POSTER ABSTRACTS

<table>
<thead>
<tr>
<th>Session Topic/Title</th>
<th>Abstract/Poster Board Numbers</th>
<th>Abstract/Poster Board Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epigenetics</td>
<td>Start # 407-End # 510</td>
<td>Genetics/Genomics Education 2303-End # 2329</td>
</tr>
<tr>
<td>Genome Structure, Variation, and Function</td>
<td>Start # 511-End # 678</td>
<td>Ethical, Legal, Social, and Policy Issues in Genetics 2330-End # 2383</td>
</tr>
<tr>
<td>Pharmacogenetics</td>
<td>Start # 679-End # 741</td>
<td>Genetic Counseling 2384-End # 2402</td>
</tr>
<tr>
<td>Complex Traits and Polygenic Disorders</td>
<td>Start # 742-End # 1119</td>
<td>Health Services Research 2403-End # 2419</td>
</tr>
<tr>
<td>Psychiatric Genetics, Neurogenetics, and Neurodegeneration</td>
<td>Start # 1120-End # 1361</td>
<td>Clinical Genetic Testing 2420-End # 2599</td>
</tr>
<tr>
<td>Bioinformatics and Genomic Technology</td>
<td>Start # 1362-End # 1721</td>
<td>Clinical Genetics and Dysmorphology 2600-End # 2785</td>
</tr>
<tr>
<td>Statistical Genetics and Genetic Epidemiology</td>
<td>Start # 1722-End # 1890</td>
<td>Prenatal, Perinatal, and Reproductive Genetics 2786-End # 2865</td>
</tr>
<tr>
<td>Evolutionary and Population Genetics</td>
<td>Start # 1891-End # 2034</td>
<td>Molecular Basis of Mendelian Disorders 2866-End # 3128</td>
</tr>
<tr>
<td>Cardiovascular Genetics</td>
<td>Start # 2035-End # 2168</td>
<td>Development 3129-End # 3157</td>
</tr>
<tr>
<td>Therapy for Genetic Disorders</td>
<td>Start # 2169-End # 2213</td>
<td>Cytogenetics 3158-End # 3223</td>
</tr>
<tr>
<td>Metabolic Disorders</td>
<td>Start # 2214-End # 2302</td>
<td>Cancer Genetics 3224-End # 3509</td>
</tr>
</tbody>
</table>

Posters should remain on the board for all three days (Sunday through Tuesday)

POSTER AUTHOR SCHEDULE

The program and abstract/poster board number next to each listing is followed by an S (Sunday), M (Monday), or T (Tuesday) to indicate the day on which authors must be present at their poster boards. Refer to the schedule below for presentation times and for the poster mounting/removal schedule.

Poster Mounting and Removal
Authors must put up and take down their posters according to the schedule below. Authors must be present at their boards based on their odd or even abstract/program/board number, and must remain at their boards for the duration of their scheduled presentation times. Posters should remain on the boards for all three days.

**Sunday, October 19**
11:00 am–1:00 pm All poster authors (Sunday, Monday, and Tuesday) place posters on boards
11:00 am–7:00 pm Posters available for general viewing
4:00 pm–6:00 pm Poster Session I (Sunday Authors Present)
4:00 pm–5:00 pm (odd poster board numbers; author must be present)
5:00 pm–6:00 pm (even poster board numbers; author must be present)

**Monday, October 20**
10:00 am–4:00 pm Posters available for general viewing
2:00 pm–4:00 pm Poster Session II (Monday Authors Present)
2:00 pm–3:00 pm (odd poster board numbers; author must be present)
3:00 pm–4:00 pm (even poster board numbers; author must be present)

**Tuesday, October 21**
10:00 am–4:00 pm Posters available for general viewing
2:00 pm–4:00 pm Poster Session III (Tuesday Authors Present)
2:00 pm–3:00 pm (odd poster board numbers; author must be present)
3:00 pm–4:00 pm (even poster board numbers; author must be present)
4:00 pm Posters closed
4:00 pm–4:15 pm All authors remove posters from boards
4:15 pm Exhibit Hall closed

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
407M

5-hydroxymethylcytosine Dysregulation in Neurodegenerative Disorders. B. Yao1, L. Lin1, C. Street1, S. Yang1, C. Slovaska1, F. Ayhan1, L. Duvic1, Z. Zalewski1, A. La Spada1, L. Ranum3, H. Chao3, D. Nelson3, P. Jin3. 1) Human Genetics, Emory University, Atlanta, GA; 2) Pediatrics and Cellular and Molecular Medicine, University of California San Diego, San Diego, CA; 3) Molecular Genetics & Microbiology, University of Florida, Gainesville, FL. 4) Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN; 5) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Neurodegenerative disorders are a heterogeneous group of chronic progressive diseases characterized by neuronal dysfunction, progressive degeneration, and progressive neuronal loss. It has been well established that epigenetic plasticity in DNA methylation-related regulatory processes influences activity-dependent gene regulation, learning and memory, and aging in the central nervous system. Recent studies that methylcytosine (mC) can be oxidized to 5-hydroxymethylcytosine (5hmC) present a particularly intriguing epigenetic regulatory paradigm in the mammalian brain. Our previous works suggest 5hmC modification at selective loci is altered in the cerebellum of FXTAS mouse model. In the present study, we systematically characterize the genome-wide 5hmC distributions in several neurodegenerative disorder mouse models with common features of ataxia, including Fragile X ataxia and tremor syndrome (FXTAS), spinocerebellar ataxia (SCA) type 1, 7, 8 and 17, together with their age-matched wildtype littermates. 1113 common loss-of-5hmC regions in all disease mouse models comparing to their wildtype littermates are identified, whereas 2182 common gain-of-5hmC regions are also generated. Gene Ontology analysis on these 5hmC differential regions indicates a strong correlation of neuronal pathways, such as Purkinje cell differentiation and cerebellum development, with common loss-of-5hmC regions. Motif identification and transcription factor analysis predict several key transcription factors in neurodevelopment and neuronal functions, including RE1-Silencing Transcription Factor (REST), Transcription Factor ETS domain 1 (Tcf4) and Foxhead box L1 (Foxl1) which could bind to DNA in a methylation-dependent manner. Altered cytosine modification at these loci thus could change their binding affinity, and affect their normal function in neurodegenerative mouse models. In addition, Approximately 15% of the 5hmC differential regions in these disease models overlap with the cerebellum enhancer regions, suggesting their potential epigenetic impact on distal regulatory regions. Our study provides the first systematic investigation of 5hmC profiles in various neurodegenerative disease mouse models, and highlights its key roles in neuronal functions. The common differential 5hmC regions in various ataxia-related disorders may sever as “hotspots” for early diagnosis and important therapeutic targets.

409M

Genome-wide Methyl-seq analysis reveals changes in hypothalamic DNA methylation patterns in response to the enhanced maternal care paradigm. D.H. Yasui1, T.Z. Baram2, A. Singh2, K.W. Dunaway2, J.M. LaSalle2. 1) Medical Microbiology, UC Davis School of Medicine, Davis, CA; 2) Anatomy and Neurobiology, UC Irvine School of Medicine, Irvine, CA; 3) Medical Microbiology and Genome Center, UC Davis School of Medicine UC Davis MIND Institute, Davis, CA.

Enhanced maternal nurturing in rat models leads to a long-lasting modulation of subsequent stress responses and altered expression of key stress-controlling genes. In the “augmented maternal care” (AMC) paradigm, enhanced mothering from postnatal days 2–9 (P9) leads to reduced corticotropin releasing hormone (Crh) expression in the hypothalamic paraventricular nucleus (PVN) which persists through life. We hypothesize that AMC epigenetically sets future gene expression levels by altering DNA methylation levels. Gene levels of Crh were assayed in AMC and control PVN by bisulfite pyrosequencing analysis. As expected, Crh methylation levels were significantly higher in AMC vs control PVN, consistent with repression of Crh expression. However, Crh methylation levels were unaffected in thalamus where no differences in gene expression were observed, suggesting that elevated PVN Crh methylation levels are due to AMC and related to Crh repression. To investigate potential additional genes contributing to the stress-resilient phenotype produced by AMC, Methyl-seq analysis was performed on AMC and control PVN to identify other differentially methylated regions (DMRs). We focused on partially methylated (>70% CpG methylation) domains (PMDs), because they comprise up to 40% of the genome in early life and are enriched in repressed genes. Thus, PMDs are likely to contain DMRs responsive to AMC. We employed a hidden Markov model trained to detect PMDs to screen for potential DMRs and identified 38 chromosomal loci, covering ~5% of the uniquely aligned rat genome. Computational analyses confirmed lower average methylation levels in 37 of the 38 identified differential PMDs, consistent with differential methylation between parental genotypes. Further analysis identified 10 enhancer regions involved in neuronal identity. Methyl-seq analysis also identified a Crh promoter region that has potentially elevated methylation levels in AMC PVN. In summary, current studies identified both marker gene (Crh) and genomewide differential PMDs. Future studies will evaluate the potential role of these DMRs in the stress-resilient phenotype of AMC and determine if these changes are shared in other stress-resistant rodent models. Current and future studies will confirm DMRs by pyrosequencing and Methyl-seq analysis and correlate these with PVN gene expression changes identified by RNA-seq. These results will provide insights into normal brain function as well as early life influences on genome wide epigenetic reprogramming of hypothalamic genes with implications for certain human for stress-related, neurodevelopmental, and psychiatric disorders.

408T

Modeling DNA methylation dynamics with approaches from phylogenetics. J. Capra1,2, D. Kostka2. 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Biomedical Informatics, Vanderbilt University, Nashville, TN; 3) Departments of Developmental Biology and Computational & Systems Biology, University of Pittsburgh, Pittsburgh, PA.

Methylation of CpG dinucleotides is a prevalent epigenetic modification that is required for proper development in vertebrates. Genome-wide DNA methylation datasets have become increasingly common, and this has enabled characterization of DNA methylation profiles from distinct stages across differentiating cellular lineages. Changes in CpG methylation are essential to cellular differentiation; however, current methods for modeling methylation dynamics do not account for the dependency structure between precursor and dependent cell types. We developed a continuous-time Markov chain approach, based on the observation that changes in methylation state over tissue differentiation can be modeled similarly to DNA nucleotide changes over evolutionary time. This model explicitly takes precursor to descendant relationships into account and enables inference of CpG methylation dynamics. To illustrate our method, we analyzed a high-resolution methylation map of the differentiation of mouse stem cells into several blood cell types. Our model can successfully infer unobserved CpG methylation states from observations at the same sites in related cell types (90% correct) and this approach more accurately reconstructs missing data than imputation based on neighboring CpGs (84% correct). Additionally, the single CpG resolution of our methylation dynamics estimates enabled us to show that DNA sequence context of CpG sites is informative about methylation dynamics across tissue differentiation. Our approach also enabled identification of genomic regions with clusters of highly dynamic CpGs. We found that enhancers, in addition to promoters and CpG islands, contain many dynamic CpG sites that are likely relevant to shifts in gene expression across the development of this lineage. Finally, we characterized genetic variation in CpG sites with a focus on mutations that disrupt the potential for CpG methylation. Our work establishes a framework for inference and modeling that is well suited to DNA methylation data, and our success suggests that other approaches used to analyze DNA nucleotide substitutions will also translate to the modeling of epigenetic phenomena.
Whole blood DNA methylation changes are associated to malignant pleural mesothelioma. E. Casalone1, S. Guarraia1, G. Fionti1, M. Betti1, D. Ferrante2, C. Di Gataetano1,2, F. Rosa1, A. Russo1,2, S. Tunes2, M. Padovan2, C. Casadio2, F. Ardissonne4, E. Ruffinì4, P.G. Betta4, R. Libinè5, R. Guaschino9, E. Piccolin11,12, D. Mirabelli11,12, C. Magnani11,12, I. Dianzan11,12, G. Matullo14,15. 1) Laboratory of Genetic Pathology, Department Health Sciences, University of Piemonte Orientale, I-28100, Novara, Italy; 2) Department of Medical Sciences, University of Turin, Turin, Italy; 3) Laboratory of Genetic Pathology, Department Health Sciences, University of Piemonte Orientale, Novara, Italy; 4) 4CPO-Piemonte and Unit of Medical Statistics and Epidemiology, Department Translational Medicine, University of Piemonte Orientale, Novara, Italy; 5) 5Thoracic Surgery Unit, University of Piemonte Orientale, Novara, Italy; 6) Chest Surgery, Department of Clinical and Biological Sciences, University of Turin, Orbassano, Italy; 7) Thoracic Surgery Unit, University of Turin, Turin, Italy; 8) Pathology Unit, Azienda Ospedaliera Nazionale SS. Antonio e Biagio e Cesare Arrigo, Alessandria, Italy; 9) Transfusion Centre, Azienda Ospedaliera Nazionale SS. Antonio e Biagio e Cesare Arrigo, I-15121, Alessandria, Italy; 10) Pneumology Unit, Santo Spirito Hospital, Casale Monferrato, Italy; 11) Unit of Cancer Epidemiology, CPO-Piemonte and University of Turin, Turin, Italy; 12) Interdepartmental Center for Studies on Asbestos and other Toxic Particulates “G. Scansetti”, University of Turin, Turin, Italy.

Malignant pleural mesothelioma (MPM) is a rare and aggressive tumor strongly associated with asbestos exposure. Its onset is usually 30–40 years after the first exposure, and it is characterized by a poor prognosis with a median survival of 12 months. Alterations in DNA methylation have been reported in several cancers, and are becoming an established hallmark of tumor. DNA methylation is frequently associated with genetic mutations but also epigenetic changes leading to gene expression modifications. The identification of MPM-specific epigenetic markers in peripheral blood might be a useful methodology for defining biomarkers for potential early detection and may define new therapeutic approaches. We found that the UHRF1 enhances motility and invasiveness in the human breast cancer cells. Chromatin immunoprecipitation (ChIP) and immunoprecipitation (IP) assays showed that the UHRF1 in regulating the progression of human breast cancer, we analyzed changes in cell morphology, migration, and invasion as well as epithelial and mesenchymal marker proteins. We found that the UHRF1 enhanced motility and invasiveness in the human breast cancer cells. Chromatin immunoprecipitations (ChIP) and immunoprecipitation (IP) assays showed that the UHRF1 protein directly bound and repressed the CDH1 through H3K9me2/3 and H3K9Ac. Inhibition of UHRF1 expression by RNA interference restored CDH1 expression in the breast cancer cells and suppressed the cell invasiveness. To further confirm our findings, we generated stable shRNA expressing and shRNA non-expressing clones of MCF-7 cells and evaluated their cell morphology and migration abilities. Our results suggested that methylation status in whole blood DNA might provide a useful biomarker for potential MPM early detection.

412T Analyses and characterization of adjacent CpG sites of TCGA BRCA

DNA methylation plays important roles in maintaining genome stability and regulating gene expression. It is fundamentally important for cell differentiation and dynamically responding to the physiological and pathologic conditions. The HumanMethylation450 BeadChip has emerged as a powerful platform to investigate DNA methylation across the human genome. That is the platform used by TCGA for the DNA methylation study. The current challenge in TCGA and other genomic initiatives is in understanding these high throughput data. DNA methylation analysis is particularly challenging, especially regarding how to combine methylation data from multiple CpG sites per gene. To address this challenge, we investigated methylation measurements as beta-values among adjacent CpG sites. We analyzed TCGA breast cancer methylation data. The latest release has HM450k data for 812 samples. An example of our analyses of these samples is summarized below. After excluding missing beta-value, we have 302559 CpG sites within 23905 genes for analysis. 21754 genes are represented by at least 2 CpG sites, with an average of 13.8 CpG sites. The variation of beta-value among the CpG sites with a gene is substantial, with standard deviation of 0.2438. The mean difference between minimum and maximum beta-value is 0.6345. We then focused on the analysis of methylation difference between the two adjacent CpG sites. As expected, we found that the beta-value difference increases as the distance between the two CpG sites increases. When the two CpG sites are within 1 kb, the difference in beta-values is usually less than 0.1. We also analyzed the beta-value differences in terms of genomic context, including sites in the promoter region, 5'UTR, first exon, gene body, and 3'UTR. We analyzed island and island shore, transcription levels, and clinical phenotypes. The insights gained from our analysis provide an effective strategy to integrate methylation and gene expression and clinical phenotypes.

413M Sports Related Concussions Induce DNA Methylation Changes in Immune Cell Trafficking and Cell Survival Pathways. H. Kim1, J. Gilli2, S. Yun3, A. Cashion4, D. Wang5, K. Merchant-Borna5, H. Lee6, J. Bazarian7. 1) SML, NINR/NIH, Bethesda, MD; 2) University of Rochester Medical Center, Rochester, NY; 3) Yotta Bioinformatics, Bethesda, MD.

Sports related concussions (SRC) are common among athletes and place them at risk for neurological and psychological dysfunction. There is considerable variation and outcome from SRC; with most athletes recovering within one week, while others may remain asymptomatic for more than one year. The purpose of this study was to determine changes in DNA methylation after the SRC in order to characterize the underlying molecular processes that accompanied recovery. Pre-season blood samples were obtained from 253 college athletes. Athletes were followed over the season and 11 sustained a SRC with subsequent blood obtained 7 days following the injury. In the SRC athletes, whole genome scale DNA methylation using Methyl Binding Domain (MBD)-seq method in peripheral blood mononuclear cells (PBMCs) was compared at baseline (pre-injury) and 7 days following the injury (sub-acute). Genomic regions that differed between baseline and the sub-acute period were determined by MACS for peak calling, PAVIS for annotation, and in-house PERL scripts provided a list of genes that significantly changed between the two time points. Ingenuity Pathway Analysis was used to determine candidate biological pathways and gene networks most related to SRC. Top canonical pathways related to SRC included: the TCA cycle II, glucocorticoid receptor signaling, as well as other immune response pathways, which includes xanthine and xanthosine. Death and survival were the primary biological function network related to SRCs. Genes with reduced DNA methylation during the sub-acute period included allograft inflammatory factor 1 (AIF1), a gene that mitigates microglial activation following TBI, which is related to inflammatory activity and cell death. The genes that increased DNA methylation contained RALBP1 and RAP1B, which reduce apoptosis following TBI, was also reduced in methylation. Reduced DNA methylation was also found in regions of genes related to protein ubiquitination including the gene SNF8. Our findings illustrate that genes related to neurological disease and psychological disorders are significantly affected by SRC, resulting in reduced methylation of genes related to cellular functions that promote recovery including immune cell trafficking and cell survival. These findings provide unique insights into molecular recovery mechanism following SRC.
414T
Genome-wide DNA methylation analyses identify loci influencing recurrent stroke risk in samples from the Vitamin Intervention for Stroke Prevention clinical trial. K.L. Keene1,2, W.M. Chen3, S.D. Turner4, A.F. Koeppe4, S.R. Williams5, M.M. Sale6, B.B. Worrall2. 1) East Carolina University, Greenville, NC; 2) University of Virginia, Charlottesville, VA.
DNA methylation is a widely accepted epigenetic factor that is important in many diseases. The folate one-carbon metabolism pathway (OCPM) pathway plays a critical role in DNA methylation through the conversion of homocysteine to methionine, wherein s-adenosyl methionine is the primary donor for de novo methyltransferase reactions. In addition to its role in DNA methylation, elevated homocysteine levels have been associated with several diseases, including cardiovascular disease and stroke. Due to the cross-talk between genes in the FOCM pathway, DNA methylation, and stroke risk, we have performed analyses of genome-wide DNA methylation status using DNA extracted from buffy coats and Illumina HumanMethylation450 BeadChip arrays, to identify regions of hypo- or hypermethylation that may influence recurrent stroke risk in 183 (105 European descent (ED) and 78 African descent (AD)) samples from the Vitamin Intervention for Stroke Prevention (VISP) clinical trial. VISP was a multi-center, double-blind, randomized, controlled clinical trial designed to determine if daily intake of a multivitamin tablet with high dose folic acid, vitamin B6 and vitamin B12 reduced recurrent cerebral infarction, nonfatal myocardial infarction (MI), or mortality. Ethnicity and gender stratified analyses identified two loci with significant (P < 10\(^{-5}\)) differential DNA methylation in ED males with (n=18) and without (n=39) a recurrent stroke, and one locus with differential methylation in ED females with (n=14) and without (n=34) a recurrent stroke. The most significant differentially methylated locus in ED males was located near the cysteine/histidine-rich 1 (CYHR1) gene (P = 4.85 \times 10^{-38}, \text{Qval} = 0.005). In ED females, the most significant locus was near the centlein, centrosomal protein (CNTLN) gene (P = 1.70 \times 10^{-7}, \text{Qval} = 0.004). No loci showed significant differential methylation in the AD samples. The CNTLN gene is located adjacent to the 9p21 locus, a region with prior genetic associations with ischemic stroke and coronary diseases. The protein encoded by the CYHR1 gene has been identified as a biomarker of good response to erythropoietin (EPO) in hemodialysis patients. Furthermore, recent studies suggest that EPO might have neuroprotective properties in the ischemic brain. Collectively, our findings could provide insight into epigenetic marks that influence recurrent stroke risk and uncover targeted treatments and regimens to alleviate recurrent stroke.

414T
Global Methylation of Fracture Risk in a Cohort of Young African Americans with Forearm Fractures. C. Sprouse1, B.T. Harmon1, H. Gordish-Dressman1, L.M. Ryan2, L.L. Tosi3, J.M. Devaney1,2. 1) Children’s national medical center, Washington, DC; 2) The School of Medicine and Health Sciences, George Washington University, Washington DC; 3) Department of Pediatric Emergency Medicine, The Johns Hopkins Hospital, Baltimore MD.
Forearm fractures account for roughly 25% of all pediatric fractures. The estimated direct costs of treating childhood forearm fractures currently exceed $2 billion dollars per year in the United States. Despite an overall trend of declining injury rates in children, multiple studies have shown that the incidence of forearm fractures among children is increasing. The reasons for this increase in pediatric forearm fracture rates remains unclear. The objective of this study was to provide new insights into the influence of methylation on fracture risk in a young African American cohort. Additionally, we hoped to identify novel epigenetic markers and elucidate novel genetic pathways, which contribute to fracture risk. Our cohort included 138 African American children ages 5 to 9 who were recruited as part of a forearm fracture study. The DNA was extracted from whole blood and analyzed for methylation changes using the Illumina HumanMethylation450 BeadChip (485,577 CpG sites). Statistical analysis was completed on 395,899 CpG sites (Partek Genomics Suite) comparing participants with forearm fractures (n=69) and age-matched controls (n=69). Associations between methylation and fracture status were examined using ANCOVA models with gender and age as covariates. After scrubbing the data and adjusting for multiple testing, 490 CpG sites were significantly different between children with forearm fractures and control participants. The top ranked CpG site by p-value (p= 1.1 \times 10^{-06}; cg23369647) was located on a CPG island of the gene SCML4, and fracture status was associated with a 58% higher beta value in children with forearm fracture compared with age matched control participants.

415M
Vitamin B deficiency and gestational programming of genes related to Alzheimer’s disease. V.C. Silva, A.L.D.A Agamme, L. Fernandez, M.T.C Muniz2, V. D’Almeida. 1) Psychobiology, Universidade Federal de São Paulo, São Paulo, Brazil; 2) Department of Biological Sciences, Universidade de Pernambuco, Recife/PE.
Fetal antecedents have been associated with increased offspring disease risk through maternal environment, placental changes, and epigenetic programming. Epigenetic regulation is crucial in the development of organisms since it is required to achieve either stable genes expression or repression at various stages of development. As methionine-homocysteine pathway supports methyl groups among other purposes, to DNA and protein methylation and this flow is sensitive to the supply of amino acids and vitamins, we investigated whether maternal vitamin B deficiency during early development alters the gene expression of App, Ps1 and Bace, which are related to Alzheimer’s disease. To this end, we submitted female mice to experimental diet one month before and during pregnancy. After 20 days, plasma homocysteine levels from deficient gestational group (DG, n=7) were approximatively 50% higher than control group (CT, n=7) (CT= 9.219 \mu mol/L, DG= 13.606 \mu mol/L, p=0.0009). The offspring’s male mice (n=6–8, per group) from control and experimental dams (CTO and DG0) were euthanized after birth, the total brain was harvested and mRNA isolated by Illumina HumanMethylation450 BeadChip arrays with Forearm Fractures. C. Sprouse1, B.T. Harmon1, H. Gordish-Dressman1, L.M. Ryan2, L.L. Tosi3, J.M. Devaney1,2. 1) Children’s national medical center, Washington, DC; 2) The School of Medicine and Health Sciences, George Washington University, Washington DC; 3) Department of Pediatric Emergency Medicine, The Johns Hopkins Hospital, Baltimore MD.
Forearm fractures account for roughly 25% of all pediatric fractures. The estimated direct costs of treating childhood forearm fractures currently exceed $2 billion dollars per year in the United States. Despite an overall trend of declining injury rates in children, multiple studies have shown that the incidence of forearm fractures among children is increasing. The reasons for this increase in pediatric forearm fracture rates remains unclear. The objective of this study was to provide new insights into the influence of methylation on fracture risk in a young African American cohort. Additionally, we hoped to identify novel epigenetic markers and elucidate novel genetic pathways, which contribute to fracture risk. Our cohort included 138 African American children ages 5 to 9 who were recruited as part of a forearm fracture study. The DNA was extracted from whole blood and analyzed for methylation changes using the Illumina HumanMethylation450 BeadChip (485,577 CpG sites). Statistical analysis was completed on 395,899 CpG sites (Partek Genomics Suite) comparing participants with forearm fractures (n=69) and age-matched controls (n=69). Associations between methylation and fracture status were examined using ANCOVA models with gender and age as covariates. After scrubbing the data and adjusting for multiple testing, 490 CpG sites were significantly different between children with forearm fractures and control participants. The top ranked CpG site by p-value (p= 1.1 \times 10^{-06}; cg23369647) was located on a CPG island of the gene SCML4, and fracture status was associated with a 58% higher beta value in children with forearm fracture compared with age matched control participants.

415M
Vitamin B deficiency and gestational programming of genes related to Alzheimer’s disease. V.C. Silva, A.L.D.A Agamme, L. Fernandez, M.T.C Muniz2, V. D’Almeida. 1) Psychobiology, Universidade Federal de São Paulo, São Paulo, Brazil; 2) Department of Biological Sciences, Universidade de Pernambuco, Recife/PE.
Fetal antecedents have been associated with increased offspring disease risk through maternal environment, placental changes, and epigenetic programming. Epigenetic regulation is crucial in the development of organisms since it is required to achieve either stable genes expression or repression at various stages of development. As methionine-homocysteine pathway supports methyl groups among other purposes, to DNA and protein methylation and this flow is sensitive to the supply of amino acids and vitamins, we investigated whether maternal vitamin B deficiency during early development alters the gene expression of App, Ps1 and Bace, which are related to Alzheimer’s disease. To this end, we submitted female mice to experimental diet one month before and during pregnancy. After 20 days, plasma homocysteine levels from deficient gestational group (DG, n=7) were approximatively 50% higher than control group (CT, n=7) (CT= 9.219 \mu mol/L, DG= 13.606 \mu mol/L, p=0.0009). The offspring’s male mice (n=6–8, per group) from control and experimental dams (CTO and DG0) were euthanized after birth, the total brain was harvested and mRNA isolated by Illumina HumanMethylation450 BeadChip arrays with Forearm Fractures. C. Sprouse1, B.T. Harmon1, H. Gordish-Dressman1, L.M. Ryan2, L.L. Tosi3, J.M. Devaney1,2. 1) Children’s national medical center, Washington, DC; 2) The School of Medicine and Health Sciences, George Washington University, Washington DC; 3) Department of Pediatric Emergency Medicine, The Johns Hopkins Hospital, Baltimore MD.
Forearm fractures account for roughly 25% of all pediatric fractures. The estimated direct costs of treating childhood forearm fractures currently exceed $2 billion dollars per year in the United States. Despite an overall trend of declining injury rates in children, multiple studies have shown that the incidence of forearm fractures among children is increasing. The reasons for this increase in pediatric forearm fracture rates remains unclear. The objective of this study was to provide new insights into the influence of methylation on fracture risk in a young African American cohort. Additionally, we hoped to identify novel epigenetic markers and elucidate novel genetic pathways, which contribute to fracture risk. Our cohort included 138 African American children ages 5 to 9 who were recruited as part of a forearm fracture study. The DNA was extracted from whole blood and analyzed for methylation changes using the Illumina HumanMethylation450 BeadChip (485,577 CpG sites). Statistical analysis was completed on 395,899 CpG sites (Partek Genomics Suite) comparing participants with forearm fractures (n=69) and age-matched controls (n=69). Associations between methylation and fracture status were examined using ANCOVA models with gender and age as covariates. After scrubbing the data and adjusting for multiple testing, 490 CpG sites were significantly different between children with forearm fractures and control participants. The top ranked CpG site by p-value (p= 1.1 \times 10^{-06}; cg23369647) was located on a CPG island of the gene SCML4, and fracture status was associated with a 58% higher beta value in children with forearm fracture compared with age matched control participants.
Whole-blood DNA methylation patterns and glycemic traits in the KORA F4 study. J. Kriebel, W. Wahl, S. Zeilinger, K. Schramm, W. Rathmann, M. Roden, A. Peters, T. Illin, M. Waldenberger, H. Prokisch, H. Grallert, C. Herder. 1) Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany; 2) Institute of Epidemiology II, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany; 3) German Center for Diabetes Research (DZD), Neuberger, Germany; 4) Institute of Human Genetics, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany; 5) Institute of Human Genetics, Technische Universitat Muenchen, Munich, Germany; 6) Institute for Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Dusseldorf, Dusseldorf, Germany; 7) Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Dusseldorf, Dusseldorf, Germany; 8) Department of Endocrinology and Diabetology, University Hospital Dusseldorf, Dusseldorf, Germany; 9) German Center for Diabetes Research (DZD), Dusseldorf, Germany; 10) Hannover Unified Biobank, Hannover Medical School, Hannover, Germany.

Recent publications indicate an involvement of DNA methylation in the development of type 2 diabetes (T2D). Thus, it is important to study the associations between genome-wide DNA methylation and blood glucose and insulin levels. The aim of the present study was to investigate the role of DNA methylation in whole blood in the development of T2D. Therefore, we performed a genome-wide association study of glycemic traits (fasting insulin, fasting and 2h glucose) in 1448 whole blood samples of the participants with available gene expression data using linear model and Benjamini-Hochberg correction for multiples testing including an area of ± 500kb around genome-wide significant signals. We additionally tested associations between DNA methylation and gene expression in a subset of 538 participants with available gene expression data using linear model and Benjamini-Hochberg correction for multiples testing including an area of ± 500kb around significant CpG sites. The CpG site cg06500161 (located in the SREBF1 gene locus) was associated with fasting glucose (B-H adjusted p = 1.7×10^-2), cg09613192 and cg2205733 (both unannotated) were associated with fasting insulin (B-H adjusted p = 1.8×10^-2 and 4.3×10^-3, respectively) and cg81024682 (located in the SREBF1 gene locus) was associated with fasting glucose, 2h glucose, and fasting insulin (B-H adjusted p = 3.0×10^-2, 1.8×10^-2, and 4.3×10^-2, respectively) after adjustment for age and sex. After adjustment for body mass index (BMI) and estimated white blood cell counts, we did not detect any genome-wide significant signals. We additionally tested associations between DNA methylation and gene expression in a subset of 538 participants with available gene expression data using linear model and Benjamini-Hochberg correction for multiples testing including an area of ± 500kb around significant CpG sites. The CpG site cg06500161 (ABCG1 gene) showed an association with ABCG1 gene expression level (B-H adjusted p = 3.0×10^-10). Our findings support previous evidence of a role of the ABCG1 gene in the regulation of glucose metabolism and suggest that the association might be modulated by epigenetic processes. Our study points towards an association between DNA methylation and T2D related traits. Analyses in insulin-responsive tissues such as adipose tissue, muscle or liver might help to elucidate the underlying mechanisms.
Distinct Patterns of DNA Methylation in Labial Salivary Gland Tissue Based on Sjögren’s Syndrome Disease Status. M.B. Cole1, X. Shao2, D. Quach2, H. Quach2, A. Baker2, L.F. Barcellos2, L.A. Criswell2. 1) Department of Physics, University of California, Berkeley, Berkeley, CA; 2) Division of Epidemiology, Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, Berkeley, CA; 3) Rosalind Russell Russell Medical Research Center for Arthritis, Department of Medicine, University of California, San Francisco, CA.

Sjögren’s Syndrome (SS, OMIM #270150) is a chronic, multisystem autoimmune disease characterized by progressive destruction of the exocrine glands, with subsequent mucosal and conjunctival dryness. A growing body of evidence indicates that epigenetic changes, in particular, altered patterns of DNA methylation, contribute to the development of this complex disease, modulating risk and severity. We report on an expanded case-control study of DNA methylation differences within labial salivary gland tissues, using biopsies sampled from 12 primary SS cases and 5 controls in the Sjögren’s International Collaborative Clinical Alliance (SICCA; http://sicca.ucsf.edu/; HHSN2682013000057C) collection. These subjects are part of a larger, 36-subject study group for which blood, gland tissue, and cell-sorted blood samples have been methylotyped (110 samples total). We generated genome-wide DNA methylation profiles using Illumina HumanMethylation450 BeadChips and further characterized full genome SNP profiles using the Illumina HumanOmni2.5-Quad platform. All methylation results were background subtracted (‘noob’ method in methylumi Biocductor package) and normalized via all sample mean normalization (ASMN) and beta-mixture quantile normalization (BMQN). Multidimensional Scaling (MDS) applied to the 360,546 CpG sites passing strict QC criteria visibly separates cases from controls within the first 2–3 components, and this clustering changes substantially with the inclusion of a separate set of 9 symptomatic SICCA controls without true SS (based on SICCA’s extensive clinical and serologic data). We demonstrate significant gene-centered mean hypomethylation across L10 and RPF2 in SS cases (FDR < 0.05). Mean methylation levels within 15 other putative SS-associated genes were similar between cases and controls. We report median methylation levels in specific BLK and KHL24 CpGs that are 15–25% hypomethylated in SS cases in addition to other significant gene-centered hypomethylation across the genome. A growing body of evidence indicates that epigenetic changes contribute to the development of this complex disease, modulating risk and severity. We report on an expanded case-control study of DNA methylation differences within labial salivary gland tissues, using biopsies sampled from 12 primary SS cases and 5 controls in the Sjögren’s International Collaborative Clinical Alliance (SICCA; http://sicca.ucsf.edu/; HHSN2682013000057C) collection. These subjects are part of a larger, 36-subject study group for which blood, gland tissue, and cell-sorted blood samples have been methylotyped (110 samples total). We generated genome-wide DNA methylation profiles using Illumina HumanMethylation450 BeadChips and further characterized full genome SNP profiles using the Illumina HumanOmni2.5-Quad platform. All methylation results were background subtracted (‘noob’ method in methylumi Biocductor package) and normalized via all sample mean normalization (ASMN) and beta-mixture quantile normalization (BMQN). Multidimensional Scaling (MDS) applied to the 360,546 CpG sites passing strict QC criteria visibly separates cases from controls within the first 2–3 components, and this clustering changes substantially with the inclusion of a separate set of 9 symptomatic SICCA controls without true SS (based on SICCA’s extensive clinical and serologic data). We demonstrate significant gene-centered mean hypomethylation across L10 and RPF2 in SS cases (FDR < 0.05). Mean methylation levels within 15 other putative SS-associated genes were similar between cases and controls. We report median methylation levels in specific BLK and KHL24 CpGs that are 15–25% hypomethylated in SS cases in addition to other significant gene-centered hypomethylation across the genome.
423M Allele-specific distribution of 5-hydroxymethylcytosine at imprinting control regions. K. Yamaizawa1-3, A. Ferguson-Smith4. 1) Pediatrics, Keio University School of Medicine, Tokyo, Japan; 2) Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK.

Covalent modification of cytosine methylation at carbon five represents a major epigenetic mark of mammalian genome and plays a pivotal role in various biological processes. Recent studies have indicated that the Tet11/12/13 translocation (Tet) family proteins can enzymatically convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). Additional processing steps and subsequent base excision repair mechanism would then result in the removal of the methylated base and its substitution with an unmethylated cytosine. However, very little is known about the distribution and function of 5hmC at differentially methylated regions (DMRs) which regulate the expression of imprinted genes. To clarify the role of 5hmC in the epigenetic landscape of genomic imprinting, we profiled the allele-specific distribution of 5hmC at representative DMRs in brain, liver, placenta and ESC cells derived from reciprocal hybrid mice. The hydroxymethylated and methylated DNA immunoprecipitations with specific antibodies to 5hmC and 5mC, respectively, were followed by quantitative PCR and pyrosequencing for SNP genotyping in hybrid crosses. We found that brain contains a substantial amount of 5hmC and that 5mC and 5hmC exist on the same allele at DMRs in our principle. Our findings raise the possibility that 5hmC could be generated on conversion of 5mC as an intermediate inactive demethylation process at DMRs. It is also noteworthy that methylation analysis with conventional bisulfite conversion could overestimate methylation levels at DMRs.

424T Obesity accelerates epigenetic aging of human liver. J.E. Hampe1, W. Erhart2, M. Brosch3, C. Roecken4, O. Ammerpohl5, C. Schafmayer5, M. Altmers6, N. Heits7, J.T. Bell8, T.D. Spector9, P. Deloukas7,8, R. Sieber10, B. Sipos11, T. Becker12, C. Roecken13, C. Schafmayer14, S. Horvath2,1, 2) Medical Department I, TU Dresden, University Hospital Dresden, Dresden, Saxony, Germany; 3) Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA; 3) Biostatistics, School of Public Health, University of California Los Angeles, Los Angeles, California, USA 90095; 4) Department of Medicine I, University Hospital Schleswig-Holstein, 24015 Kiel, Germany; 5) Department of Human Genetics, Christian-Albrechts-University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, 24015 Kiel, Germany; 6) Department of Visceral and Thoracic Surgery, University Hospital Schleswig-Holstein, 24015 Kiel, Germany; 7) Department of Twin Research and Genetics Epidemiology, Kings College London, London SE1 7EH, UK; 8) William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, EC1M 6BQ, UK; 9) Princess Al-Jawhara Al-Brahim Centre of Excellence in London School of Medicine and Dentistry, Queen Mary University of London, London SE1 7EH, UK; 10) Institute of Pathology, University Hospital Tbingen, 72074 Tbingen, Germany; 11) Institute of Pathology, University Hospital Schleswig-Holstein, 24015 Kiel, Germany.

Obesity entails an increased risk of many chronic diseases that are typically associated with older age. While it is plausible that obesity increases the biological age of some cell types, it has been difficult to test this hypothesis due to the lack of appropriate biomarkers of aging due to difficulties in the quantification of tissue age across tissues. Here, we utilize a novel biomarker of aging known as epigenetic clock to study the effect of high body mass index (BMI) on the DNA methylation ages of blood, liver, muscle and adipose tissue. The epigenetic aging is in turn may result from sexual dimorphisms in transcriptional regulation. In our previous work (Naumova, 2013), we demonstrated sex-specific differences in DNA methylation levels of the zona pellucida binding protein 2 (ZPB2) promoter that is located within chromosomal region 17q12-21, one of the best replicated GWAS regions for asthma (Mofta 2007, 2010). The differences in methylation levels between males and females were consistent with the sex-specific bias in genetic association that we found in the asthmatic families collection from Sagayenu-Lac-Saint-Jean (Naumova 2013). We next hypothesized that sex-specific differences in DNA methylation levels at the ZPB2 promoter resulted from the dosage of the sex chromosome. ZPBP2 DNA methylation was investigated in DNA samples extracted from fibroblast cell lines derived from individuals with different sex phenotypes and sex chromosome dosage: e.g. 46, XY females (sex reversal); 45, X females (Turner syndrome); 46, XY males; 46, XX females; 48, XX males, and a 48, XXXY male. No significant influence of sex phenotype or the presence of the SRY gene was detected. However, significant positive correlation was found between ZPBP2 methylation levels and the number of X-chromosomes. Individuals with one X chromosome had considerably lower methylation levels (n=9, average DNA methylation level 17.5 %) compared to individuals with one X chromosome (n=9, average DNA methylation level 29.3 %) (p = 0.005, t-test statistics). Thus, our data show that the X chromosome dosage has an effect on DNA methylation and perhaps may contribute to disease susceptibilities. Our data are consistent with the idea that sex chromosome dosage on gene expression in mice (Wijchers 2010) and DNA methylation in humans (Grafodatskaya 2013). Currently, our laboratory is working towards identifying the X-linked gene(s) that are responsible for methylation differences at autosome regions and at the asthma-associated region of chromosome 17q12-21, in particular.
Epigenome-wide DNA methylation and body mass index in monozygotic twins. J.T. Bell1, F. Gao2, C.G. Bell3, Y. Xia1, W. Yuan3, L. Roos1, P.-C. Tsai1, K. Ward1, P. DeJoukas2,4, X. Wang2, T.D. Spector1, 1) Department of Twin Research, King's College London, London, United Kingdom; 2) BGI-Shenzhen, Shenzhen, China; 3) William Harvey Research Institute, Queen Mary University of London, London, United Kingdom; 4) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom.

Obesity is a major risk factor for multiple common diseases. Genetic, environmental and recently epigenetic factors have been shown to associate with obesity, offering potential to help understand and modify disease risk. Here, we focus on identifying the epigenetic signatures of environmentally mediated risk of obesity and its impacts, by studying monozygotic twins. We performed an epigenome-wide association scan of body mass index (BMI) in 1,446 UK monozygotic twins from 723 pairs (mean age = 51.8, mean BMI = 25.9 kg/m²) from the EpiTwin project (www.epitwin.eu), part of TwinsUK. Whole blood DNA methylation was assessed using methylated DNA immuno-precipitation followed by high-throughput sequencing (MeDIP-seq), allowing for full genome-wide coverage at 10.5 million overlapping bins of size 500bp. Differentially methylated regions related to BMI were identified by comparing DNA methylation discordance to BMI discordance within twin pairs (mean absolute BMI twin-pair discordance = 4.4 kg/m², mean relative discordance = 0.2 units). The most associated signals included a region within an intron of the CDKAL1 gene (P = 3.6e-7), a GWAS obesity and type 2 diabetes susceptibility locus. Functional annotation based on gene structure, Encode data, and chromHMM chromatin states showed strong enrichment of differential methylation in poised promoters. Genomewide significant effects within poised promoters were observed in genes SUSU34 (FDR = 6e-5, P = 2e-5), and TDRD10 (FDR = 4e-6), and the SLC10A4 result validated at an overlapping CpG-site on the illumina 450k array in whole blood samples from 355 individuals. Replication is currently ongoing in 424 individuals using MeDIP-seq. Our results show epigenetic studies in monozygotic twins are a powerful tool for identifying functional genomic changes in complex traits such as obesity.

DNA-methylation and the Down Syndrome phenotype. A. Bouman1, P. Hennekam1, M.M.A.M. Mannens1, A.N.P.M. Mul1, E.J. Meijers-Heijboer1, R.C.M. Hennekam2,3,1) Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands; 2) Department of Pediatrics, Emma Children’s Hospital, Amsterdam, The Netherlands; 3) Department of Translational Genetics, Emma Children’s Hospital, Amsterdam, The Netherlands.

Down syndrome (DS) is the most frequent genetic cause of intellectual disability known today. DS is explained by the presence of an additional chromosome 21 which serves as the basis of DS-pathogenesis. However, underlying cellular and molecular processes causing the DS-phenotype remain not well understood. Several hypotheses exist regarding the implications and effects of the presence of trisomy 21. In this study we hypothesize that the presence of an extra chromosome 21 might influence the DNA-methylation pattern of (parts of) chromosome 21 and/or other loci throughout the genome. Since DNA-methylation plays an important role regarding gene transcription regulation, differentially methylated loci of specific genes can either cause increased or decreased transcription. This altered transcription may play a major role in causing several of the specific features of the DS-phenotype. White blood cell DNA-methylation patterns of 10 individuals with DS and 10 healthy (age- and sex-matched) controls were obtained using the illumina 450K Human DNA-methylation Array. DNA was bisulfite converted using the gold standard kit of Zymo®. Analysis of the data was performed using two freely available “R” packages; minfi and CHAMP. We will discuss the results of our DS methylation-study and zoom in on specific genes which show a distinct methylation patterns in DS vs. controls. Our study serves as an addition in ongoing fundamental DS-research.
431M Monoamine oxidase A (MAOA) expression level predicts alcohol consumption in Rhesus macaques. R.P. Cervera Juanes, B. Ferguson 1, B. Park 2, G. Wandef, K.A. Grant 1, B. Ferguson 1 1) Neurosciences, ONPRC, OHSU, Beaverton, OR; 2) Johns Hopkins University, Baltimore, MD; 3) 3Department of Public Health and Preventive Medicine, Oregon Health & Science University, Portland, OR

Monoamine oxidase A metabolizes the neurotransmitters serotonin, norepinephrine and dopamine, which are key regulators of behavior. Both a repeat polymorphism in the promoter of the MAOA gene (MAOA-LPR) and a history of childhood adversity have been linked to risk for alcohol use and behavioral disorders in humans and rhesus macaques. To investigate the potential genetic and epigenetic mechanisms underlying these risks, we collected blood from twelve unrelated, male rhesus macaques before and after they participated in a 12-month alcohol self-administration study. We determined MAOA-LPR genotypes, promoter CpG methylation levels and MAOA mRNA levels in all samples. We found that mRNA levels were independent of MAOA-LPR genotype. However, there was a robust association between site-specific CpG methylation level and MAOA expression. Of particular interest, our results show a strong association between MAOA mRNA level and MAOA expression by alcohol self-administration protocol. After 12 months of alcohol use, nMAOA expression was significantly reduced, and there was a moderate increase in regional and site-specific CpG methylation levels. Thus, our results indicate that blood MAOA mRNA level is more tightly associated with promoter methylation than with MAOA-LPR genotype. Moreover, since the level of MAOA expression was predictive of subsequent heavy alcohol use, CpG methylation may provide a molecular link between childhood adversity and sustained risk for alcohol use. Finally, our finding that chronic alcohol consumption further dempens gene expression suggests that reduced MAOA availability, and its downstream effects on dopamine, serotonin and epinephrine levels, may also contribute to alcohol addiction. This work was supported by NIH grants U61AA020928 and U01AA013510.


Exposure to a diabetic intrauterine environment is a strong risk factor for the development of type 2 diabetes in the offspring. DNA methylation has an important role in the epigenetic response to the environment. However, the genes which are subject to methylation in response to intrauterine diabetes exposure remain largely unknown. To discover differentially methylated genes and relevant pathways in response to intrauterine diabetes exposure, we selected, from a longitudinal study of Pima Indians, 189 individuals whose mothers had type 2 diabetes during their pregnancy and 201 individuals whose mothers did not have diabetes during their pregnancy. We assayed the offspring’s genome-wide DNA methylation in peripheral blood, collected when the participant was not diabetic, using the Illumina HM450 array. Quality control excluded the probes with a detection P value > 0.01 or having a SNP score <0.1. Follow-up analysis, including a SNP count of 395,899 methylation sites remained for association testing. The association between intrauterine diabetes exposure and methylation of CpG sites was tested in a linear model with adjustment for maternal age (at the time of the child’s birth), as well as the proportion of American Indian ancestry, age and sex of the child. We calculated the difference in methylation score (d, in SD units) in those exposed to a diabetic intrauterine environment, compared with those unexposed. Exposure to a diabetic intrauterine environment was significantly associated with methylation of CpG islands in the LHX3 (d=0.67, FDR=4.6E-4) and WNT9B (d=0.52, FDR=4.8E-2) were hypermethylated in those exposed while PPP1R3B (d=0.66, FDR=9.0E-4), ATP8B3 (d=0.56, FDR=6.7E-3) and LINC00839 (d=0.55, FDR=4.8E-2) were hypomethylated. The expression level of LHX3, a LIM1-homeodomain transcription factor which is under the control of the ISL1 transcriptional regulator, has previously been shown to be controlled by methylation in mice. In the pathway enrichment analysis of the 103 hyper-methylated and 1282 hypo-methylated genes (P<0.005) using DAVID, the sequence-specific DNA binding domain (ZnF) and the DNA repair pathway were overrepresented (P=3.7E-5 and P=1.4E-3 respectively). Of these genes, 5 SNPs in regions of interest were differentially methylated with exposure to a diabetic intrauterine environment. Validation of these findings in additional samples is ongoing.

433M DNA enrichment as a cost effective tool to examine DNA methylation at single nucleotide resolution. A. Czyz, V. Ruotti, A. Tan, D. Hill, S. Kuersten, R. Vaidyanathan. R&D Department, Epicentre an Illumina company, Madison, WI, USA

Epigenetic modifications are punctuation marks on DNA, responsible for modification of transcriptional activities. One of the most characterized epigenetic modifications is DNA methylation. DNA methylation plays a role in a variety of biological processes including embryonic development, cell differentiation, chromosome stability, and chromatin structure. Aberrant DNA methylation has been found to be associated with various intellectual disabilities, including obesity, anemia and numerous types of cancer. In mammalian cells, methylation occurs mostly at cytosines that are contained in the symmetrical dinucleotide CpG, or in an asymmetrical CHH or CHG context. Methylated nucleotides are generally grouped in promoter regions, CpG islands, and CpG shores. Epignome methylase Seq-Kit is a simple, 1-day, post-bisulfite library construction method that requires only 50–100 ng of input gDNA. However, providing enough coverage for whole-genome bisulfite libraries typically requires 60–120 Gb of sequencing, and therefore, is relatively expensive. In contrast, several strategies exist for enrichment of target sequences that are more cost effective, including probes and use of methylation-specific antibodies. The central weakness of these last methods is lack of single nucleotide resolution. But, by combining enrichment methods and DNA bisulfite treatment with the Epignome methylase Seq-Kit, we circumvent this shortcoming. To demonstrate, we took McCP2 methyl binding domain and IgG A antibodies derived against 5mcG, and used it for enrichment of beta-lymphocyte gDNA NA18508, prior to continuing with the Epignome library preparation kit. The prepared libraries were sequenced and analyzed by utilization of the commercially available BISMARK and MEDIPS bioinformatics packages. We observed significant enrichment of promoter and CpG islands, comparable to standard EpiGnome libraries without enrichment. With only 8GB (50M reads) sufficient for analysis, and thus possibility of multiplexing enriched libraries on HiSeq 2500 platform. This enrichment method provides a cost effective approach for methylation study.


As interest in studying allele-specific methylation (ASM) and it’s association with common complex diseases grows, there is a need for a resource that stores and catalogs SNPs and regions that demonstrate allele specific methylation, analogous to NCBI’s dbSNP. Additionally, as ASM is a regulatory mechanism that may be associated with hits from genome-wide association studies (GWAS), researchers need a suite of tools to help them evaluate the relationship between GWAS hits and ASM. To facilitate these investigations, we have created a new web resource called dbASM, hosted on the Genapha web server (www.genapha.ca). The aim of dbASM is twofold: 1. Curate from the literature a publicly-accessible database of known sites of ASM. 2. Provide researchers with a web-based platform of tools for exploring ASM and determining regions of interest. We will present the dbASM resource including details on the underlying database construction and data sets, in addition to the web tools and example workflows. The web tools that are currently available are: Genapha CatalogSNP Search, ASM SNP Search, SNP Counter, Methylation Plots Generation, and Sequence Viewer. GWAS Catalog SNP Search allows browsing through NHGRI’s Catalog of Published Genome-Wide Association Studies by phenotype and filtering GWAS SNPs based on their relation to suspected sites of ASM. For example, rs11742570 is associated with inflammatory bowel disease (p=2.0 E-82) and demonstrates ASM. ASM SNP Search supports finding SNPs’s based on: ASM status or interorogability; location compared to genes, a chromosome region or other SNP’s; and filtering by population minor allele frequencies and sample size. SNP Counter uses asynchronous JavaScript calls to the database to provide real-time counts of types of SNPs’s in user-selected regions of chromosome. Methylation Plots Generation is a calculates SNP correlation stratifying by genotype with CpG site methylation patterns, similar to the output of the chronicane studies (but without disease status). Using CEPH HapMap samples and genotypes and methylation assays on these same samples completed on the Illumina 27k array. Sequence Viewer displays SNPs’s in the human reference genome (based currently on GRCh37, 8.1), and displays and generates DNA methylation data for SNPs and regions of interogability via MSRE cut sites for enzymes: HpyCH4IV, Bsp1287, Bsp1287R, DpnII, HpaII and MspII. These tools are all freely available for use at: http://genapha.capture.ubc.ca/asvm/.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Chronic Fatigue Syndrome (CFS) is a complex multifactorial disease that is characterized by the presence of fatigue and other particular symptoms that cannot be mitigated by rest. CFS is a multisystem and heterogeneous disease, and the biological basis of CFS is poorly understood. Some of the most consistent differences observed in previous CFS studies are modified immune gene expression and changes in the hypothalamic-pituitary-adrenal (HPA) axis that include mild hypocortisolism and enhanced negative feedback to glucocorticoids. To determine a potential role of epigenetics in CFS, we examined the DNA methylome in peripheral blood mononuclear cells (PBMCs) isolated from female CFS patients (n = 46) and healthy controls (n = 25) using the Illumina HumanMethylation450 BeadChip array. All subjects were non-obese and did not consume medications with known immunological or epigenetic effects. A two-tiered approach (a high-confidence and a medium-confidence cutoff) was applied to examine evidence for differential DNA methylation in CFS. The high-confidence cutoff identified 578 differentially methylated sites and the medium-confidence cutoff identified an additional 13,059 differentially methylated sites. RAND-36 scores, a survey assessing quality of life, revealed that CFS patients had lower scores compared to healthy controls across most categories. The first two principal components of RAND-36 explained 85% of DNA methylation variability among subjects and discriminated between CFS and controls accurately. HPA axis response in PBMCs was explored using an in vitro dexamethasone suppression assay. An increased sensitivity to dexamethasone in CFS was observed with the appearance of two subgroups (responders and non-responders) among the CFS subjects. We identified 18 differentially methylated sites between CFS and controls among the medium confidence group that were associated with the HPA axis. DNA methylation differences also existed between responders and non-responders which underlines the potential of using DNA methylation differences in determining CFS subtypes. Our work is the first to examine DNA methylome changes in CFS and supports a role for epigenetic modifications in the increased glucocorticoid sensitivity observed in some CFS patients. These epigenomic differences could serve as potential biomarkers for future clinical research and assist in determining the biological basis of CFS.
Epigenome-wide meta-analysis of over 10,000 individuals reveals extensive perturbations in DNA methylation associated with adiposity. A.W. Drong1, S. Wahlin-Anderson2, B. Lehne3, M. Lönner4, G. Fiont5, S. Guerra6, K. Adam4, S. Kasela7, A. Richardson3, A. Dehghan5, L. Franke10, T. Esko7,11,12,13, L. Milani7, C.L. Relton14,15, J.T. Bell16, T.D. Spector16, O.H. Franco10, P. van der Harst17, C.M. Lindgren17, M.I. McCarthy18, G. Matullo6, C. Gieger19, J.S. Koornneef20, H. Grallert21, J.C. Chambers3,5,20,21, 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 3) Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 4) German Center for Diabetes Research (DZD), Neuherberg, Germany; 5) Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK; 6) Human Genetics Foundation, HuGoF, I-10126 Torino, Italy Department of Medical Sciences, University of Torino, I-10216, Torino, Italy; 7) Estonian Genome Center, University of Tartu, Tartu, Estonia; 8) MRC Integrative Epidemiology Unit, Oakfield House, University of Bristol, Bristol BS8 2BN, UK; 9) Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; 10) Departments of Cardiology and Genetics, University of Groningen, University Medical Center Groningen, Groningen & Durrell Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, University Medical Center Utrecht, UMCU, Utrecht, Netherlands; 11) Department of Social and Community Medicine, University of Bristol, Bristol, UK; 12) Department of Twin Research & Genetic Epidemiology, King’s College London, St Thomas’ Hospital Campus, Westminster Bridge Road, London SE1 7EH, UK; 13) Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA; 14) Oxford Centre for Diabetes Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 15) National Heart and Lung Institute, Imperial College London, London, London; 16) School of Medicine, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; 17) Broad Institute, Cambridge, Massachusetts, USA; 18) Oxford Centre for Diabetes Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 19) National Heart and Lung Institute, Imperial College London, London, London; 20) Ealing Hospital, St Thomas’ Hospital Campus, Westminster Bridge Road, London SE1 7EH, UK; 21) Imperial College Healthcare NHS Trust, London W12 0HS, UK.

Obesity is a global health problem and a major risk factor for type-2 diabetes and cardiovascular disease. The mechanisms underlying obesity and its complications are incompletely understood. We carried out an epigenome-wide association (EWA) study to investigate the relationship of DNA methylation with body mass index (BMI) to examine the epigenetic perturbations associated with obesity. The EWA study included four studies comprising 5,387 whole-blood samples from European (N=2,707) and South Asian (N=2,680) individuals. DNA methylation was determined by Illumina Infinium 450K Human Methylation array. Associations of DNA methylation with BMI were tested in each cohort using linear regression with adjustment for age, gender, physical activity, smoking, alcohol intake, white-blood cell estimates and technical covariates. Meta-analysis across the studies was done by METAL. Our primary EWA yielded CpG sites at 207 independent genetic loci associated with BMI at P<1×10−7, with little evidence for heterogeneity between the populations. We then carried out replication testing of the 207 sentinel CpG sites amongst 4,998 whole-blood samples from 9 independent cohorts. 187 of the 207 markers reached P<0.05 in replication testing and remained P<1×10−7 in combined analysis across all stages. A gene set enrichment analysis of the 187 identified genetic loci revealed enrichment for pathways involved in transmembrane transport and metabolic signalling (lipid metabolism, insulin signalling and PPAR signalling) at the 5% FDR level. Further, the 187 sentinel CpGs were located outside CpG islands and show a significant enrichment for “open sea” locations (p=4.4×10−5), mostly located within gene bodies (p=2.6×10−4). We overlapped the 187 sentinel CpG sites with ENCODE ChIP data and observed an enrichment for DNase hypersensitivity sites (p=2×10−5) and the activating histone mark H3K4me1 (p=3.85×10−12). We identify 187 genetic loci at which DNA methylation is associated with BMI. Our results provide new insight into potential DNA regulatory mechanisms associated with obesity and its complications.
439M

Genome-wide DNA methylation study identifies genes associated with GDF-15 levels. W.E. Ek1, A.K. Hedman2,3, E. Ingelsson2,4, U. Gyllensten1, L. Lind2, A. Johansson1,4, 1) Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 2) department of Medical Sciences, Molecular epidemiology, and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 3) Department of Medical Cell Biology, Molecular Epidemiology, Uppsala University Hospital, Uppsala, Sweden; 4) Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden.

Introduction. Growth differentiation factor 15 (GDF-15) is a member of the transforming growth factor-β (TGF-β) family with pleiotropic functions. GDF-15 is expressed in low to moderate levels in most healthy tissues and levels may increase in response to pathological stress associated with inflammation, oxidative stress or tissue damage. Overexpression of GDF-15 has been seen in various cancers and growing evidence indicate that GDF-15 levels in plasma may be a new biomarker for risk stratification and therapeutic decision making in cardiovascular disease.

Methods. We performed a genome-wide DNA methylation study to determine the association between GDF-15 and DNA methylation in 717 individuals from the Northern Sweden Population Health Study (NSPHS). DNA methylation status, in white blood cells, was determined at more than 475,000 sites distributed throughout the genome. Significant findings where replicated in 950 individuals from an independent cohort (PIVUS). We also performed gene ontology (GO) enrichment analysis to identify terms enriched for differently methylated genes in relation to GDF-15 levels.

Results. A total of 31 CpG sites, corresponding to 22 genes, were significantly associated (false discovery rate [FDR] q-values < 0.05) with GDF-15 in NSPHS. Among these, 12 sites replicated in DBSs (p < 0.05). One of the most studied microRNAs which has been associated with cancer and cardiovascular diseases. A total of 125 enriched biological processes were identified (FDR q-value < 0.05), of which 43 (34%) replicated (p < 0.05). Among the enriched biological processes, we found terms involved in developmental, biological and metabolic processes. A total of 13 enriched molecular functions were also identified (FDR q-value < 0.05), of which 5 (35%) replicated (p < 0.05). Among molecular functions, we found terms involved in protein binding and kinase activity.

Conclusion. We have shown that GDF-15 levels are correlated with DNA methylation level at numerous sites in the genome providing new leads for investigating the links between GDF-15 and cardiovascular disease. However, our results suggest that GDF-15 plays an important role in many biological processes, even in the absence of response to pathological stress.

440T

Black-white difference in regional patterns of DNA methylation: the Bogalusa Heart Study. X. Fu1, D. Sun2,3, S. Li2, F. Fernandez2, T. Chen2, Y. Liu4, Q. Lin5, W. Chen6. 1) Department of Biostatistics and Bioinformatics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA; 2) Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA; 3) Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China; 4) Microarray Core Facility, University of Texas Southwestern Medical Center, Dallas, TX, USA.

Although DNA methylation profiles vary by populations, regional patterns of DNA methylation in different populations are largely unknown. The objective of this study is to examine black-white difference in regional patterns of DNA methylation in relation to pathophysiological pathways. This study included 846 adults aged 28–51 years (594 whites and 252 blacks) from the Bogalusa Heart Study. Peripheral leukocyte DNA methylation was measured with Infinium HumanMethylation450 BeadChip. After sample and probe quality analysis and cell-type count estimation, a linear regression model with age, gender, cell-type count included as covariates was fitted to find the significantly differential methylation sites between races. To avoid bias towards high probe density region, a feature-oriented dynamic window method “lasso” was used to capture significantly differential regions. Enrichment analysis in KEGG pathways was performed with selected differential methylation-related gene sets. Besides, correlation patterns of methylation were examined by race groups, followed by a comparison with genetic LD structures. A total of 1,225 regions were captured among 48,241 sites that were found to be significantly, differentially methylated between blacks and whites. Enrichment analysis using genes located in these 1,225 regions identified 15 pathways in the KEGG database which were related to human complex diseases known to have different prevalence rates among populations. In correlation analysis, significant differences in both position and size of clusters were observed in every chromosome, especially in chromosome 2, 4, 10, 13 and 22, between blacks and whites. These correlation patterns were not all caused by genetic LD structures. In conclusion, the findings of the current study on black-white difference in regional patterns of DNA methylation has implications in understanding of potential race-specific epigenetic mechanisms and pathophysiological pathways underlying chronic diseases.
442T  
Methylation patterns are associated with chronological age in the Khomani San of South Africa. S. Gopolanyi1, C. Carja2, E. Patin4, M.S. Kobor2, H. Fraser3, M. Feldman5, A. Froment6, L. Quintana-Murci2, B. Henn1. 1 Ecology and Evolution, Stony Brook University, Stony Brook, NY, United States; 2 Department of Biology, Stanford University, Stanford, CA, United States; 3 Department of Genomics and Genetics, Human Evolutionary Genetics, Institut Pasteur, Paris, France; 4 Centre for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC, Canada; 5 Musum National d’Histoire Naturelle, Paris, France.

Changes in genomic methylation have been found to be associated with advancing age. Increased age is correlated with an overall genome-wide reduction in methylation, but also an increase in methylation of particular regions, such as CpG islands. So far, studies of methylation as it relates to aging in humans are primarily done in populations of European origin. It remains challenging to correlate effects of external factors, for example diet and smoking, which are known to be substantial.

We set out to identify CpG sites associated with aging in the Khomani San, a population of former hunter-gatherers living in the Kalahari Desert of South Africa. Previous studies have found that the KhoeSan populations are the most genetically diverse in the world and are among the few remaining societies that have not adopted an agricultural lifestyle. The Baka Pygmies are another group of African hunter-gatherers who live in the rainforests of Central Africa. Despite sharing similar subsistence strategies, and societal circumstances of poverty and malnutrition, the age structure of their populations differs greatly. It is common for Khomani San individuals to reach age of 70 or more, while Baka Pygmies rarely reach the age of 50. We investigated the methylation profiles of individuals in both groups and compared them to European methylation profiles. Saliva samples from 56 Khomani San and 36 Baka individuals were methylated on the Illumina 450K array. Khomani individuals ranged from 21 to 91 years old (median 62.5) whereas Baka individuals ranged from 5 to 59 years old (median 29.5). We performed an epigenome-wide association study (EWAS) to identify sites that were significantly associated with chronological age in the Khomani San individuals. This resulted in over 100 CpG sites. We replicated the strong signal of hypermethylation in the CpG island of the ELOVL2 gene, which was also associated with biological age. This suggests that specific CpG sites are associated with aging in diverse human populations. Our results may prove useful for forensic work as well as future epidemiological studies.


443M  
Characterizing a genomic map of 5-hydroxymethylcytosine in human brain at single base resolution through compartmental sequencing. J. Gross, G.G. Chen, A. Dialeo, R. Poujol, C. Ernst, G. Turecki, Douglas Mental Health Univ Inst, Montreal, Quebec, Canada.

The recent discovery that methylated cytosines are converted to hydroxymethylated cytosines (hmC) by the family of ten-eleven translocation (TET) enzymes has sparked significant interest in the field of epigenetics. This finding, along with that of Kriaucionis and Heintz who described the presence of hmC in purified neuronal nuclei, stimulated growing interest in research describing the genomic location of hmC. Using LM-PCR, we collected enriched regions of hmC in post-mortem human brain tissue from 24 subjects who died by natural causes, glucosylated, and digested with AbaSI, an enzyme that specifically recognizes glucosylated cytosines. Subsequently, custom biotinylated adaptors were ligated to the cleaved DNA, which was then further sheared. Standard libraries were prepared using the eQror pipeline and were sequenced on Illumina’s HiSeq 2000 sequencers. Bowtie 2.0 was used to map raw reads to the reference genome Hg19 and a custom perl script was used to determine the locations of hmC at single base pair resolution. A combination of BEDTools, R packages, and custom scripts was used to determine enriched and depleted regions in the genome, links to regulatory elements, and correlations with available gene expression data. The results present a unique characterization of hmC in human brain and provide an important reference for future research. Understanding the differences between individuals and across brain regions will provide insights into the epigenetic code and its function. There have been few known fetal-specific marker using the differences between maternal and fetal epigenetic characteristics. We analyzed methylation status of maternal blood and fetal placental DNA using microarray to find a differentially methylated region as potential epigenetic markers for PE. Methods: We performed a high-resolution tiling array analysis using a methyl-CpG binding domain-based protein (MBD) method with blood samples from normal pregnancies, placental samples from normal pregnancies, and blood samples from non-pregnancies. To validate the differentially methylated loci identified from a high-resolution tiling array analysis, bisulfite direct sequencing was carried out in four groups: 1) blood from normal pregnancies, 2) placenta from normal pregnancies, 3) blood from pregnancies with PE, and 4) placenta from pregnancies with PE. Results: Differentially methylated loci in the placenta were selected by Agilent Genomic Workbench software using an arbitrary differential ratio cut off of normalized log2 ratio > 2. We confirmed the methylation patterns of CpG sites in the 5 associated genes by selected loci located on chromosome 6, 12, and 19 in normal and PE group, all selected region showed significant hypermethylation in placenta compared with blood, respectively (P=0.004 for SOD1, P<0.0001 for DSCR3, P=0.0004 for C2CD2, P=0.0003 for UMODL1, P=0.0002 for ENST00000450830). In the normal pregnancy, all 5 selected regions were depleted in blood and enriched in placenta compared with normal pregnancies. Each DSCR3 CpG site was significantly increased in PE placental samples compared with normal placental samples (P<0.05 for all CpG sites). Conclusion: We found 5 fetal-specific hypermethylated regions. Among these genes, DSCR3 significantly alters DNA methylation and may serve as a potential candidate to be developed into epigenetic biomarkers in PE.

444T  
Global DNA hypermethylation is associated with increased myopia risk. E. Hsi1,2, K.C. Chen2, C.W. Huang3, M.L. Yu4, C.L. Liang4, S.S. Jung1. 1) Graduate Institute of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 2) Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 3) Department of Biochemistry, College of Medicine, Taipei Medical University, Taipei, Taiwan; 4) Department of Biotechnology, Kaohsiung Medical University, Kaohsiung, Taiwan.

Purpose: DNA methylation is the most extensively studied epigenetic process that modulates gene expression without changing DNA sequence and is known to be altered by environmental factors. We aimed to discover the association between global DNA methylation and high myopia and find the relevant consequences. Methods: The methylated levels of long interspersed nucleotide elements (LINE-1) are regarded as a surrogate marker of global DNA methylation. Three hundred high myopia and 300 sex-age matched control subjects were used to find the association between high myopia and LINE-1 methylation in leukocytes. We analyze the global DNA methylation among leukocytes, retina and sclera tissues in form deprivation myopia (FDM) mice model. Immunohistochemistry demonstrates the distribution of homocysteine and methionine. We use dopamine to mimic brighter environment to rescue the myopia state. Results: Subjects with high (82.1% - 91.5%) and moderate (79.6% - 82.1%) of LINE-1 methylation have 2.10 and 2.08 fold risk to have high myopia, significantly (p=0.002 and 0.001, respectively). Significantly elevated LINE-1 methylated level in blood and both retina and sclera tissues of both eyes in FDM mice. Immunohistochemical staining indicates elevated homocysteine, methionine levels in both FDM and fellow eyes. Finally, the global methylated level is reduced 2% significantly after 24 hours of dopamine stimulation. Conclusions: Our study demonstrates that LINE-1 DNA hypermethylation is associated with high myopia. Global hypermethylation and accumulated homocysteine indicate a constitutional change which is induced by FDM. Reduced global methylated level in blood may be a promising new treatment of dopamine and is a novel mechanism in treatment of myopia.

445M  
Microarray approach reveals the differentially methylated regions as potential epigenetic markers in preeclampsia. H.J. KIM1, S.Y. KIM2, S.Y. PARK1, D.J. KIM3, D.E. LEE1, J.H. LIM4, M.Y. KIM5, J.H. CHUNG2, H.M. RYU1,2. 1) Laboratory of Medical Genetics, Cheil General Hospital and Women’s Healthcare Center, Seoul, Korea; 2) Department of Obstetrics and Gynecology, Cheil General Hospital and Women’s Healthcare Center, Kwandong University College of Medicine, Seoul, Korea.

Objective: The development of preeclampsia (PE) seriously affects the health of the mother and the child, but the precise pathogenesis of PE is not well known. The placenta is considered to play a key role and DNA methylation may be associated with altered placental development and function. There have been few known fetal-specific marker using the differences between maternal and fetal epigenetic characteristics. We analyzed methylation status of maternal blood and fetal placental DNA using microarray to find a differentially methylated region as potential epigenetic markers for PE. Methods: We performed a high-resolution tiling array analysis using a methyl-CpG binding domain-based protein (MBD) method with blood samples from normal pregnancies, placental samples from normal pregnancies, and blood samples from non-pregnancies. To validate the differentially methylated loci identified from a high-resolution tiling array analysis, bisulfite direct sequencing was carried out in four groups: 1) blood from normal pregnancies, 2) placenta from normal pregnancies, 3) blood from pregnancies with PE, and 4) placenta from pregnancies with PE. Results: Differentially methylated loci in the placenta were selected by Agilent Genomic Workbench software using an arbitrary differential ratio cut off of normalized log2 ratio > 2. We confirmed the methylation patterns of CpG sites in the 5 associated genes by selected loci located on chromosome 6, 12, and 19 in normal and PE group, all selected region showed significant hypermethylation in placenta compared with blood, respectively (P=0.004 for SOD1, P<0.0001 for DSCR3, P=0.0004 for C2CD2, P=0.0003 for UMODL1, P=0.0002 for ENST00000450830). In the normal pregnancy, all 5 selected regions were depleted in blood and enriched in placenta compared with normal pregnancies. Each DSCR3 CpG site was significantly increased in PE placental samples compared with normal placental samples (P<0.05 for all CpG sites). Conclusion: We found 5 fetal-specific hypermethylated regions. Among these genes, DSCR3 significantly alters DNA methylation and may serve as a potential candidate to be developed into epigenetic biomarkers in PE.
446T
Quantification of the placental epigenetic signature of the maspin gene in maternal plasma of pregnancies complicated by small for gestational age. 
S.Y. Kim1, S. H. Kim1, S. Y. Park1, D.E. Lee1, J.H. Lim1, S.Y. Lee1,K.S. Kim1, J.Y. Han2, M.H. Kim2, D.W. Kwak2, H.M. Ryu1,2. 1) Laboratory of Medical Genetics, Chel General Hospital and Women’s Healthcare Center, Seoul, Korea; 2) Department of Obstetrics and Gynecology, Chel General Hospital and Women’s Healthcare Center, Kwandong University College of Medicine, Seoul, Korea.

Objective: Quantification of cell-free fetal DNA (cfDNA) by methylation-based DNA discrimination has been used in non-invasive monitoring of adverse pregnancy outcomes. The maspin (serpin peptidase inhibitor, clade B (ovabumin), member 5; SERPINB5) gene is hypomethylated in the placenta and completely methylated in maternal blood cells. The objective of this study was to examine the possible association between women with small for gestational age (SGA) neonates and DNA concentration in maternal plasma during first and second trimesters using tissue-specific epigenetic characteristics of the maspin gene. Methods: We performed a real-time quantitative methylation-specific PCR to quantify the concentrations of unmethylated-maspin (U-maspin) and methylated-maspin (M-maspin) in maternal plasma at 11–26 gestational weeks of women with SGA neonates (n = 55) and with appropriate for gestational age (AGA) neonates (n = 106). Results: At 11–14 gestational weeks, median U-maspin and M-maspin concentrations were significantly different in women with SGA neonates compared to women with AGA neonates. There were no significant differences in the concentrations of U-maspin and M-maspin between the two groups at 15–26 weeks of gestation. The ratio (U-maspin/(M-maspin+U-maspin))/100 in women with SGA neonates was similar to that of controls. In addition, there was a significant correlation between U-maspin and M-maspin concentration at first and second trimesters and birth weight and gestational age at delivery were not observed. Conclusion: Concentrations of U-maspin and M-maspin in maternal plasma of the first and second trimester were not compared with associated with pregnant women’s age, gender, and white-blood cell subpopulations. Our findings indicate that quantification of cfDNA in maternal plasma during early pregnancy using tissue-specific methylation changes of the maspin gene promoter is not a useful predictor of SGA pregnancies.

447M
Methylation levels in peripheral blood may reflect central mechanisms mediating GxE associations with cardiovascular outcomes. 
L.C. Kwee1, C.S. Haynes1, E.A. Grass2, I.C. Siegler1,2, B.H. Brummett2, M.A. Babacky2,3, R. Jiang1,2, A. Singh1,2, R.B. Williams2,3, E.R. Hauser1,2, W.E. Kraus1,4, S.H. Shah1,4, S.G. Gregory1. 1) Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC; 2) Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC; 3) Behavioral Medicine Research Center, Duke University Medical Center, Durham, NC; 4) Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, NC.

Previous studies have established associations between candidate genes, psychosocial factors, and cardiovascular (CV) outcomes. For example, a genetic variant (rs6318) in the serotonin receptor gene HTR2C is associated with both hypothalamic-pituitary-adrenal response to a stress recall task and increased risk of CV mortality. Also, SNPs in the TOMM40/APOE region are associated with CV outcomes, and this effect is mediated by psychosocial stress. The correlation between CV outcomes and methylation levels is mediated by psychosocial stress, which leads to the relevance of studying epigenetic mechanisms. Our study is the first to investigate the interaction of psychosocial factors and CV outcomes. DNA was obtained from 206 subjects selected from the CATHGEN biorepository of patients referred for cardiac catheterization at Duke University. We selected candidate genes previously associated with coronary heart disease endophenotypes and outcomes (HTR2C, TOMM40/APOE, ADRB2, BDNF, and EBF1) with interactions with psychosocial stress, and assayed methylation at CpG sites in these genes using the Infinium HumanMethylation450 BeadChip. After QC, data from 169 CpGs in these five regions were available for analysis. We used linear mixed models to test for association between methylation levels at each probe and an endpoint composed of all-cause mortality or incident myocardial infarction. All models were adjusted for age, sex, race, and batch. Ten CpG sites were nominally associated (p<.05) with the endpoint: six in BDNF (brain-derived neurotrophic factor; min p=8.1x10−4), and four in EBF1 (early B-cell factor 1; min p=8.5x10−4). (Due to the correlation in methylation across a gene, strict Bonferroni correction is overly conservative.) We observed consistent methylation changes, as all but one of the significant sites were methylated at slightly higher levels in controls vs. cases (Iq > 0 for all). These results indicate that it may be useful and relevant to assay peripheral blood biomarkers in the search for genetic mechanisms that mediate the central effect of psychosocial stress factors on cardiovascular events.
449M

Birth weight and DNA methylation, S. Li, D. Sun, X. Fu, C. Fernandez, T. Chen, Y. Lian, Q. Li, W. Chen. 1) Epidemiology, Tulane School of Public Health and Tropical Medicine, New Orleans, LA; 2) Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China; 3) Department of Biostatistics and Bioinformatics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA; 4) Microarray Core Facility, University of Texas Southwestern Medical Center, Dallas, TX.

Low birth weight for gestational age is a risk factor for adult cardiometabolic diseases. It has been hypothesized that intrauterine growth retardation may result in changes in DNA methylation, by which low birth weight leads to increased risk of cardiometabolic diseases. However, it is not known what specific DNA methylation changes occur as a result of intrauterine growth retardation as indicated by low birth weight. We aimed to examine the association between birth weight and DNA methylation profiles in adults. The study included 679 white and 289 black adult participants aged 28–51 years of the Bogalusa Heart Study who had birth weight data and DNA methylation data (the Infinitum HumanMethylation450 BeadChip). After quality control, beta values from 402,000 CpG sites were used as outcome variables and birth weight as an independent variable in linear regression models, adjusted for age, sex, cigarette smoking, and body mass index, separately in whites and in blacks. In whites, beta values of ten CpG sites were significantly (false discovery rate < 0.05) associated with birth weight, of which three were significantly higher for females. These significant sites are in or near genes EBF4, RAMP2, ARL4C, TBCD, PRRT1, and BA1,1, representing cellular component organization or biogenesis, developmental process, immune system process, localization, metabolic process, multicellular organism process, reproduction, and response to stimulus, in PANTHER analysis. In blacks, only one CpG site was significant (P < 0.05), with no clear functional implications. We are seeking replications in independent samples and will present updated results in the meeting. In conclusion, we have identified CpG sites whose methylation is associated with birth weight, indicating that prenatal growth environment affects DNA methylation profiles.

450T


Older maternal age during pregnancy has been associated with adverse birth outcomes, childhood cancer, type 1 diabetes, and neurodevelopmental disorders in offspring. For many of these conditions the underlying biological mechanism is unknown. One mechanism by which maternal age may affect the health of the offspring is through epigenetic modifications such as DNA methylation. Using the Norway Facial Clefts Study, a national population-based case-control study of cleft lip and cleft palate, we conducted the largest epigenome-wide association study to date investigating alterations in DNA methylation in newborns related to maternal age at delivery. The Illumina HumanMethylation450 BeadChip was used to assess whole blood DNA methylation. We included samples collected 2–3 days after delivery from 890 newborns. After data pre-processing, robust linear regression was used to identify CpG sites related to maternal age, adjusting for facial cleft status, birth weight, maternal alcohol use, maternal smoking, maternal education, parents' sex, and technical factors (batch, bisulfite conversion efficiency, birth year). Additional adjustment for five blood cell subtypes following the method developed by Houseman and colleagues was performed as a sensitivity analysis. Replication of selective findings is underway in an independent pregnancy cohort, the Norwegian Mother and Child Cohort Study, using Illumina HumanMethylation450K data from 1068 umbilical cord blood samples. In addition, using data from the Sister Study, a nationwide prospective US cohort of women with a sister with breast cancer, we are exploring whether maternal age-related methylation changes persist into adulthood. The average age of 1006 women in the Sister Study was 52.4 years. Among the 27K data and 181 women with Illumina HumanMethylation450K data generated from whole blood. Identification of the same maternal-age related DNA methylation changes in adults would suggest that these altered methylation states persist for an extended period of time, which may have implications for later health outcomes.

451M

Genetic control of the human blood methylome, J.L. McClay, A.A. Shabbalm, D.E. Adkins, G. Kumar, S. Nereida, S.L. Clark, S.E. Bergen, C.M. Hultman, P.K.E. Magnusson, P.F. Sullivan, K.A. aberg, E.J. van den Oord, Swedish Schizophrenia Consortium. 1) Center for Biomarker Research and Personalized Medicine, Virginia Commonwealth Univ, Richmond, VA; 2) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 3) Department of Medical Genetics, University of North Carolina School of Medicine, Chapel Hill, NC.

Comprehensive understanding of genetic influence on the epigenome could shed new light on human genotype-phenotype relationships. In this study, we performed whole genome association (GWA) analysis of DNA methylation in 1850 selected blood CpG sites in >1.5 million Swedish men and women. We combined results from two GWA studies resulting in 11.5 million independent SNP-methylation tag loci (meQTLs) in 697 Swedish subjects. Using SOLID next-generation sequencing (Life Technologies), we generated an average of 67.3 million methylation-enriched reads per subject to assay whole blood DNA methylation levels at >25 million autosomal CpGs. We used genome-wide association measures with 4.5 million SNPs, genotyped on Affymetrix 5.0/6.0 or Illumina Omni quad arrays, with imputation using 1000 Genomes panels. Testing used Matrix eQTL to run all >20 trillion tests in 1.5 days on 50 nodes of a computing cluster. Using stringent (1%) false discovery rate control, 15% of methylation sites showed genetic influence, while 98% of SNPs were associated with methylation at one or more sites. The high proportion of SNP effects was due to linkage disequilibrium (LD), whereby many SNPs tagged each eQTL, whereas long-range correlation between methylation sites was much less extensive. Local meQTL effects (<1Mb between SNP and methylation site) were 189,000-fold more common than distant effects. Most local effects could be explained by CpG-SNPs (SNPs altering CpG sequence), with 75% of methylation sites under local genetic influence binding a SNP with MAF >0.05% of all SNPs at >0.05% of all sites genome-wide. Local meQTL effects typically occurred outside known genes and were not enriched for several classes of genomic annotations (e.g. CpG islands, DNAse clusters). However, these sites were strongly enriched in the tissue expression atlas. Using the NHLBI GWAS catalog (www.genome.gov/gwastudies). Of the 393 methylation sites under local genetic control that overlapped the GWAS catalog, 366 encompassed a CpG-SNP. Our study confirms that genetic influence on DNA methylation is pervasive. However, the high proportion of local meQTL effects appear unlikely to affect disease susceptibility, because they typically occur outside of known functional regions. However, mutation of CpGs at critical regulatory loci may be an important etiologic mechanism for complex diseases. Our study provides a detailed map of genetic influence on the human blood methylome. Integration of meQTL and functional annotation data into gene mapping efforts could help pinpoint potentially causative mutations in large genomic regions.

452T

DNA methylation profiles of ten patients with Attention deficit hyperactivity disorder: a proof-of-principle study, C. Milani, T.V.M.M. Costa, M.M. Montenegro, G.M. Novo-Filho, E.A. Zanardo, R.L. Dutra, A.T. Dade, M. Vizzotti, V. Schuch, C.B. Mello, M. Mussatto, C.A. Kim, L.D. Kulikowski. 1) Departamento de Patologia, Laboratorio de Citogenmica, FMUSP, Sao Paulo, Brazil; 2) Unidade de Genetica, Departamento de Pediatria, Instituto da Criana, FMUSP, Sao Paulo, Brazil; 3) Departamento de Psicobiologia, Universidade Federal de Sao Paulo, Sao Paulo SP, Brazil.

Introduction: Attention deficit hyperactivity disorder (ADHD) is one of the most common childhood brain disorders and affects approximately 5.3% of children worldwide. Despite the high heritability of the disorder (estimated at 76%), no genetic marker has been consistently identified and the effects on the variability of the ADHD phenotype calls for improvements in research strategies. We performed the proof of principle study in order to characterize DNA methylation profile in peripheral blood leukocytes of individuals with ADHD, providing the first epigenome view study in this disorder. Methods: Genome-wide methylation profile was assessed using DNA extracted from blood lymphocytes of 10 patients with ADHD. We used the Illumina Infinium HumanMethylation450 BeadChip array, which allows the annotation of approximately 80,000 CpG sites per sample at single-nucleotide resolution. Results: The results of all patients showed similar methylation profile and biostatistical tools identified sites with an elevated probability to be consistently hypermethylated in chromosome 13 and 16 and hemimethylated in chromosome 18, 19 and 20. However, some patients revealed slightly different methylation profiles involved BRI3BP, RH34AL, ZNF516 and TP53TGS genes suggesting an involvement in process as apoptosis mediated by TNF, regulation of calcium ions, cell proliferation, differentiation and tumor suppression. In addition, some of the genes found to be differentially methylated showed the DNA methylation status of 96% of CpG islands located in the genome in promoter regions and non-promoter regions of genes. Profile the tissue-specific DNA methylation profiles would provide novel insights into pathogenic mechanisms, as well as help in future epigenetic therapies.
453M


1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Molecular Pathology, Institute of Biomedical and Translational Medicine, University of Tartu, Tartu, Estonia; 3) Institute of Computer Science, University of Tartu, Tartu, Estonia; 4) Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

Ageing is associated with gradual deterioration of immune response, for example, elderly people have lower responses to vaccination and higher risks for infections, autoimmune diseases and tumorigenesis. Although many blood cell types are involved, the most affected cells in this process are T-cells. The reshaping of T-cells during ageing is considered to be the most important feature of immunosenescence and causes an immune risk profile associated with higher morbidity and mortality among elderly individuals.

We analyzed the DNA methylation profiles of CD4 and CD8 T-cells and peripheral blood samples from 50 young and 50 elderly individuals to identify the epigenetic basis of these changes and to investigate their association with corresponding gene expression levels.

In DNA methylation analysis of CD4 and CD8 T-cells, we found more differentially methylated CpG sites and larger variability among CD8 T-cells, whereas a relatively low number of changes were detectable in peripheral blood samples after correction for cell subtype proportions. Age-related hypermethylation occurred primarily in the CpG islands of gene promoters with silenced transcriptional activity and was associated with the repressive histone H3K27me3 modification based on ENCODE data. More specifically, we identified a subset of functionally important genes in T-cells, with strong inverse correlation between methylation and expression levels. Among this subset, we found hypomethylation and higher expression of effector T-cell and infection-associated genes as well as hypermethylation of T-cell lineage-specific transcription factor genes in elderly individuals.

Our results suggest that in blood mononuclear cells, associated DNA methylation changes are shared among blood cells, they do not comprise all the changes that occur in specific blood-cell subsets, such as the T-cells described here. Our findings demonstrate links between age-related epigenetic changes and changes in gene expression, and these changes are relevant to the response to infections and to potential harmful consequences leading to immunosenescence.

454T


1) Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN; 2) Department of Psychiatry and Behavioral Science, Emory University, Atlanta, GA.

The epigenome sits at the interface between the external environment and the underlying genetics. There is extensive individual variation in the epigenome that can be shaped by a combination of genetic and non-genetic influences. Here we compare DNA methylation between African and European ancestral groups in two different study cohorts: (1) the CANDLE (Conditions Affecting Neurocognitive Development and Learning in Early Childhood) study, which provides genome-wide methylation (HumanMethylation27) and expression (Illumina WG-6 BeadChip) measures in newborn umbilical cord blood from African Americans (N = 67) and European Americans (N = 45) in the Memphis area, and (2) methylation measures in lymphoblastoid cell lines from Yoruba African (YRI) and CEPH European (CEU) panels of HapMap. We also evaluate the influence of maternal nutrition—specifically, plasma levels of vitamin D and folate during pregnancy—on child methylation. We define ancestry-dependent differences in methylation at ~2,300 genes (over 11% of the CpG sites we surveyed) at an FDR of 5%.

Our results show that ancestry-dependent DNA methylation patterns are remarkably stable and are likely influenced by a combination of genetic and non-genetic factors, including maternal nutrition.
Schizophrenia is a highly heritable and polygenic psychiatric disorder with largely unknown etiology. DNA methylation is an epigenetic mark that has been shown to be both heritable and influenced by environmental factors. Differential DNA methylation at specific promoter regions of schizophrenia candidate genes has previously been linked to the disease. In this study we aim to identify variation in the epigenome that is associated to schizophrenia by interrogating DNA methylation sites across the genome. We matched 450 individuals with schizophrenia and 450 healthy controls for age and gender from a relatively homogeneous Dutch population. Blood-derived whole genome DNA methylation was collected using Illumina's Human Methylation array. We applied FaST-LMM-EWASher (a linear mixed model with principal components) to identify sites associated to our phenotype of interest while accounting for potential spurious association arising from technical artifacts and cell-type heterogeneity. In addition, we modeled top sites shown to be associated to schizophrenia for presence or absence of chromatin marks using ChromMM and chromatin state calls derived from ENCODE Epigenomics Roadmap data. We identify multiple CpG loci with DNA methylation profiles associated with schizophrenia disease status (q-value < 0.05). Among these sites are regions of genes important for neuronal functioning, such as vacuolar protein sorting 52 homolog (VPS52). This gene is involved in vesicle trafficking and fusion to target membranes and is located at the classical MHC locus, a region frequently associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia for presence or absence of chromatin marks using ChromMM and chromatin state calls derived from ENCODE Epigenomics Roadmap data. We identify multiple CpG loci with DNA methylation profiles associated with schizophrenia disease status (q-value < 0.05). Among these sites are regions of genes important for neuronal functioning, such as vacuolar protein sorting 52 homolog (VPS52). This gene is involved in vesicle trafficking and fusion to target membranes and is located at the classical MHC locus, a region frequently associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia for presence or absence of chromatin marks using ChromMM and chromatin state calls derived from ENCODE Epigenomics Roadmap data. We identify multiple CpG loci with DNA methylation profiles associated with schizophrenia disease status (q-value < 0.05). Among these sites are regions of genes important for neuronal functioning, such as vacuolar protein sorting 52 homolog (VPS52). This gene is involved in vesicle trafficking and fusion to target membranes and is located at the classical MHC locus, a region frequently associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia. In addition, regions close to DNA methylation sites most associated to clinical diagnosis are significantly enriched for common variants associated to schizophrenia.
458T The Genome Wide DNA Methylation Signature of Subjects as They Enter and Exit Short Term Alcohol Treatment. R. Philibert1,2, B. Penalver3, T. White1, S. Shires1, T. Gutter1,3, F. Liesveld1, C. Eren1, N. Hollenbeck1, T. Osborn2,1, Dept Psychiatric, Univ Iowa, Iowa City, IA; 2) Behavioral Diagnostics Iowa City, IA 52240; 3) Indiana Univ, Indianapolis, IN; 4) Texas Tech University Health Sciences Lubbock, TX.

Alcoholism has a profound socio-economic and personal impact on tens of millions of individuals throughout the world. Unfortunately, our ability to quantify the chronicity and quantity of alcohol intake is often hindered by reliance on self-report while our ability to monitor treatment response is challenged by the absence of robust and reliable biomarkers. Recently, in work replicated in over 12 independent genome wide analyses, our consortium has shown that quantitative DNA methylation assessments, particularly at AHR locus cg05575921, can sensitively and specifically quantify tobacco smoking exposure. Now, in an extension of prior work with lymphoblasts, we examine lymphocyte DNA prepared from heavy, recently intoxicated subjects as they both begin and finish standard 28 day alcohol treatment using standard genome wide methylation techniques and the Infinium 450K HumanMethylationBeadChip. We show that as compared to abstinent controls, heavy alcohol intake is associated with profound widespread changes in DNA methylation with over 800 distinct loci reaching genome wide significance. In marked contrast to prior work with tobacco, there do not seem to be highly sensitive sentinel loci with the overall pattern of changes being consistent with the hypothesis that alcohol affects cellular functions by non-specific general solvent effects (i.e. changes in cytoplasmic dialect constant). Significantly, there is a marked bias for alcohol induced changes to revert as a function of abstinence. Pathway analysis of the alcohol induced changes demonstrated that key pathways characterized for whole genome analysis 450K HumanMethylationBeadChip.

460T DNA methylation profiles that distinguish rheumatoid arthritis (RA) from osteoarthritis in fibroblast-like synoviocytes can be detected in immune cells from peripheral blood. B. Rhead1, C. Holingue2, M. Cole1, X. Shao1, H. Quach2, D. Quach2, L.F. Barcellos1, L. Criswell1,1) Division of Epidemiology, Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, Berkeley, CA; 2) Rosalind Russell Medical Research Center for Arthritis, University of Medicine, University of California, San Francisco, San Francisco, CA.

RA [MIM 180300] is a chronic inflammatory disease with potential to cause substantial disability, primarily due to the erosive and deforming process in joints and the presence of extra-articular manifestations. RA is complex in etiology with contributions from both genetic and non-genetic factors. Epigenetic changes, such as altered patterns of DNA methylation, are also present in RA. Recent work by Firestein and colleagues (Nakano, 2013) compared genome-wide DNA methylation profiles in synovial fibroblasts (SFRA) from patients with RA to osteoarthritis (OA) controls and identified a set of differentially methylated genes that appear to distinguish these two forms of arthritis. Given the greater accessibility of peripheral blood compared to synovium-derived FLS, we set out to determine whether similar methylome signatures were present in immune cell subsets in peripheral blood. We generated over 400 genome-wide DNA methylation profiles for 101 women (70 cases from the UCSF RA cohort and 31 controls) using Illumina HumanMethylation450 BeadChips. Four FACS-sorted immune cell types were assayed for each individual: CD14+ monocytes, CD19+ B cells, CD4+ memory T cells, and CD4+ T cells. All samples were background subtracted using the uomi Bioconductor package and normalized using all samples to mean normalization (ASMN). All study individuals were HapMap typed for matching SNP from Illumina OmniExpress BeadChips. We excluded CpG sites near SNPs known to be common (<1%) from analysis, as well as sites with a low detection p-value, leaving 442,797 sites for analysis. We examined the top differentially hyper- and hypomethylated (n = 15 and 18, respectively) loci from the Firestein study in each immune cell type and found strong evidence for overlapping RA methylome signatures between FLS and immune cells. The most significantly differentially methylated genes in cases vs. controls include MGMT, IDH1, TTP, CABLES1, COL4A1, COL4A2, CYFIP1, EGF, FOXO1, MAP3K1, PHLPP1, PTNP14, RXRA, and TGFBR2. MGMT encodes a DNA repair enzyme and showed the strongest evidence of hypomethylation. The strongest hypermethylation was seen in MAP3K1, a serine/threonine kinase involved in some signal transduction cascades. Our investigation of DNA methylation in RA cases and controls underscores the importance of epigenetic mechanisms in RA pathogenesis and represents the largest study yet, to date, of methylome profiles derived from immune cells.
Characterizing Functional Methylocmes in Human Populations Using Novel Next-Generation Capture Sequencing Approach. X. Shao1, F. Allum1, F. Gunnard1, M.-M. Simon1, S. Busche1, M. Caron1, T. Kwan1, S. Marceau2, M. Lathrop1, A. Tchernof1, M-C. Voh1,2, T. Pastinen1, E. Grundberg1. 1) Department of Human Genetics, McGill University and Genome Quebec Innovation Centre, Montreal, QC H3A 1A5, Canada; 2) Institute of Nutrition and Functional Foods (INAF), Laval University, Quebec, QC G1V 0A6, Canada; 3) Endocrinology and Nephrology, CHU de Quebec Research Center, Quebec, QC, G1V 4G2, Canada; 4) Quebec Heart and Lung Institute, Quebec, QC, G1V 4G5, Canada.

Whole genome bisulfite sequencing (WGBS) has been widely used to detect DNA methylation at single CpG resolution. However, as only ~20% of CpG sites are variable across individuals or across human tissues, WGBS is inefficient and costly and thus not optimal for population studies of methylation variation and the impact on disease susceptibility. Meanwhile, targeted approaches such as Illumina Human Methylation 450K BeadArray or the reduced representation bisulfite sequencing are biased to gene promoter regions and cover only a minor proportion of variable CpGs. Here we introduce MethylC-Capture Sequencing (MCC-seq), a novel cost-efficient sequencing approach targeting the most informative fraction of CpGs within disease-relevant tissues. Using adipose tissue and whole blood we have designed capture panels focusing on metabolic and autoimmune diseases that target ~160Mb or ~5M dynamic CpGs/tissue mapping to regulatory elements, hypomethylated footprints or an Illumina450K site. The panels also include all SNPs and their proxies (r^2 > 0.8) linked to metabolic or autoimmune diseases (GWAS catalogue 9/2013). The approach is based on bisulfite-converted DNA libraries that are selectively enriched using the custom capture panels (Roche/Nimblegen) and sequenced on the IlluminaHiSeq2000 systems. With multiplexing up to 6-folds/lane we note on average ~62% on-target rate/sample corresponding to ~26X mean genome coverage/sample. We also find a high concordance of within samples CpG methylation on different runs (p=0.9). Methylation calls between MCC-seq and WGBS are found to be highly concordant (p=0.85), whereas lower concordance of MCC-seq and 450K (p=0.80) is observed. Moreover, our results indicate that MCC-seq is capable of simultaneous genotyping where 97% of on-target CpGs have an average concordance of 99% with genotypes from the Illumina 2.5M array. Finally, we integrated phased genotypes and the MCC-seq data to detect allele-specific methylation (ASM). Preliminary data show ~2.4% of all heterozygous SNP associated CpGs are demonstrating ASM with the highest fold enrichments in regions, most likely representing sequence-dependent effects. In addition, while imprinted loci are enriched among ASM they only account for 0.84% of detected sites. Overall, the MCC-seq is cost-efficient and provides single-base resolution of functional DNA methylene and genotype profiles suitable for large-scale disease-association studies.

Faciocapulohumeral muscular dystrophy 2 (FSHD2) testing - a UK pilot study and a clinical diagnostic service. D.J. Smith1, J.A. Whitfield2, S. O’Shea1, R. Whitington1, L. Yarram-Smith1, P.W. Lunt2, M. Williams1. 1) Bristol Genetics Laboratory, North Bristol NHS Trust, Southmead Hospital, Bristol, UK; 2) National Genetics & Genomics Education Centre, National School for Healthcare Science, Birmingham, UK.

FSHD (affecting ~1 in 20,000 individuals) is an autosomal dominant but epigenetically regulated disorder, with a characteristic pattern of progressive muscle involvement commencing in the face and shoulder-girdle. Two clinically indistinguishable forms (FSHD1 & FSHD2) differ by the molecular basis of disease, but are underpinned by hypomethylation of 4q35, leading to aberrant expression of DUX4 (a toxic transcription factor). FSHD1 (OMIM 158900) (95% cases) is caused by contraction of D4Z4 repeats, resulting in allele specific hypomethylation. FSHD2 (OMIM 158901) (~3% of cases) is contraction independent, caused by mutations in the SMCHD1 gene encoding a chromatin modulating enzyme, resulting in definitive hypo-methylation and chromatin relaxation of D4Z4 array on both 4q35 alleles. A permissive haplotype at 4q35 is required for clinical expression of both FSHD1 and 2, necessitating digenic inheritance for FSHD2. Bristol Genetics Laboratory has provided a FSHD specialist diagnostic service since 1992, and processes over 400 referrals annually. Patients negative for standard and extended deletion analysis but clinically typical for FSHD are candidates for NextGen ksequencing (4.8% diagnostic referrals).

Lemmers et al (2012) identified mutations in SMCHD1 (18p12.31) causing global hypomethylation of D4Z4. Hartweck et al (2013) identified a region showing intense hypomethylation in FSHD2 (DR1). BGL has developed a novel pyrosequencing assay for quantification of methylation using 10 statistically established sites. A pilot clinical study, with novel candidate SMCHD1 mutations found in all 9; (3 missense, 1 non-sense, 1 duplication, 2 deletion, 2 potential splice site). FSHD2 is an example of digenic inheritance, and this increases the complexity of genetic counselling, as the risk to offspring is between 25 and 50%, dependent on the haplotype of the wider family. We present a service overview, the results of this study and interesting cases highlighting the clinical utility of genetic testing, and family risks associated with this digenic disease.

Sex specific epigenetic and transcriptional responses of peripheral blood leukocytes (PBLs) to lipopolysaccharide (LPS). M. Stein1, C. Hrusch1, J. Nicodemus-Johnson1, A. Sperling2, C. Ober1. 1) Human Genetics, University of Chicago, Chicago, IL; 2) Department of Medicine, University of Chicago, Chicago, IL.

Many immune mediated diseases show striking sex differences in prevalence, yet the genetic and epigenetic factors contributing to these differences are poorly characterized. We undertook this study to identify sex-specific differences in methylation and transcriptional responses to LPS, an important activator of the innate immune response, in PBLs collected from 153 members (55% female; range 6–76 years) of a large Hutterite pedigree who previously participated in our studies of asthma. Cells were incubated with LPS+media and with media alone (control) in a closed system (TruCulture®, Myriad RBM, Austin). After 30 hours, DNA and RNA were extracted. In a pilot study of 28 school children (12 girls, 16 boys; 6–14 years old), we used the Illumina Methylation450 Beadchip for methylation and the HumanHt-12v4 for expression studies. Studies in the larger sample are underway. We defined sex-specific differentially methylated regions (DMRs) in response to LPS as CpG sites with methylation response (q-value <0.05) in one sex, but not in the other (q-value >0.2). We observed 39 sex-specific DMRs in response to LPS in girls only and 85 in boys only. 48% of sex-specific DMRs were near a gene that was detected as expressed in the arrays. Expression levels of five (8.5% of expressed genes near a DMR) were correlated with methylation levels. These genes include intriguing candidates that have previously been associated with asthma. For example, methylation levels at a CpG site in the ANXA2 gene were lower in LPS-treated samples compared to control samples in boys (P=2.15x10^-4) but remained unchanged in girls (P=0.48); ANXA2 expression was higher in LPS-treated samples compared to control samples in boys (P=1.97x10^-4) but not in girls (P=0.16). Increased expression of ANXA2 was previously associated with asthma exacerbations (PMID 10430734). Methylation levels at a CpG site in ATP2C1 were significantly higher in LPS-treated samples in boys (P=1.57x10^-4) but remained unchanged in girls (P=0.95); gene expression was also higher in LPS-treated samples in boys (P=4.91x10^-4) but not in girls (P=0.92).

Methylation levels in this gene were previously associated with a parent-of-origin effect in children with asthma (PMID 24166889). Overall, these data show that methylcytation and transcriptional responses to an important environment stimulus (LPS) can be sex-specific, potentially providing a mechanism for the sex disparities in occurrence and course of immune mediated diseases.
467M
Race-specific association between DNA methylation and body mass index: the Bogalusa Heart Study. D. Sun1,2, X. Fu3, S. Li4, Q. Li4, T. Chen4, Y. Lian4, C. Fernandez4, W. Chen4. 1) Department of Epidemiology, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA; 2) Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China; 3) Department of Biostatistics and Bioinformatics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA; 4) Microarray Core Facility, University of Texas Southwestern Medical Center, Dallas, TX, USA.

This study test the hypothesis that genome-wide DNA methylation is associated with adiposity measured as body mass index (BMI). As part of the Bogalusa Heart Study, 968 adults aged 28–51 years were recruited between 2006 and 2010. Of these individuals, 830 participants (584 whites, 246 blacks) were examined only once and used as the discovery cohort; 138 participants (95 whites, 43 blacks) were examined twice 3.2 years apart and used as the replication cohort. The discovery findings were replicated by the baseline and follow-up sample. Genome-wide methylation of peripheral leukocyte DNA samples was measured by HumanMethylation450 BeadChip and sample and probe quality analysis was performed. In the association analyses by race, adjusting for age, gender and smoking, DNA methylation levels at 17 sites in whites and at 4 sites in blacks were found to be significantly associated with BMI in discovery cohort. As shown in the table below, 2 (cg06500161 and cg00574958) of the 17 sites in whites and 1 (cg26403843) of 4 sites in blacks were significant in both baseline and follow-up replication cohorts. In conclusion, DNA methylation changes in lipid metabolism-related genes are associated with BMI values, which provides potential race-specific epigenetic mechanisms underlying the development of obesity in whites and blacks.

Effect sizes and Bonferroni-adjusted P values

<table>
<thead>
<tr>
<th>CpG</th>
<th>Gene</th>
<th>Discovery cohort</th>
<th>Baseline replication</th>
<th>Follow-up replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>P</td>
<td>β</td>
<td>P</td>
<td>β</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cg06500161</td>
<td>1.4×10⁻⁶</td>
<td>2.2×10⁻⁶</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG074958</td>
<td>9.8×10⁻⁴</td>
<td>1.1×10⁻⁴</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG26403843</td>
<td>2.6×10⁻²</td>
<td>4.0×10⁻²</td>
</tr>
</tbody>
</table>
468T
Genome-wide association scans identify differentially methylated and expressed regions related to smoking in adipose tissue. F-C. Tsai1, C. Glaistonbury3, A. Vuela1, T. Erte1, W. Yuan4, E. Demirtzakis2, P. Deloukas2, K. Small1, T.D. Spector1, J.T. Bell1. 1) Department of Twin Research & Genetic Epidemiology, King’s College London, London, United Kingdom; 2) Department of Genetic Medicine and Development, Université de Genève, Genève, Switzerland; 3) William Harvey Research Institute, Queen Mary University of London, London, United Kingdom.

Smoking is an important environmental risk factor in the development of chronic disease and cancer. Multiple epigenome-wide association scans (EWASs) for smoking have identified and replicated differentially methylated regions in blood. To explore these effects across tissues, we analysed DNA methylation profiles from the Illumina 450k array in adipose tissue samples from 349 healthy female twins, including 35 current smokers, 128 ex-smokers, and 186 non-smokers. We identified 39 CpG sites in or near 25 genes that were significantly associated with smoking at a false discovery rate of 5%. Several of the smoking-associated methylated sites, such as those in the \textit{AHRR}, \textit{F2RL3}, and \textit{CYP1A1} gene, were previously identified to be associated with smoking in blood and lung tissues. Furthermore, smoking-related regions were predominately hypo-methylated in adipose tissue, consistent with previous results. We next performed a genome-wide scan of RNA-sequencing gene expression profiles in the same 349 adipose tissue samples to identify differentially expressed exons associated with smoking. In total, 48 exons in 35 genes were associated with smoking status at a false discovery rate of 5%. Of the 35 genes, 4 genes (\textit{CYP7T1}, \textit{AHRR}, \textit{F2RL3}, and \textit{CYP1B1}) overlapped with the top methylation results. Inter-individual correlation showed strong associations between DNA methylation and exon expression at a number of genes. Thus, several of the smoking-associated loci. To our knowledge, this is the first smoking-EWAS in adipose tissue. We identified both adipose-specific and tissue-shared DNA methylation changes related to smoking, with corresponding gene expression associations with smoking status. We conclude that smoking exerts a strong effect on DNA methylation and gene expression levels across multiple tissues, and suggest that smoking should be incorporated as a covariate in future EWASs.

469M
Downstream analyses and Mendelian randomization study on methylene-wide associations with BMI reveal biological pathways underlying obesity. A. Abecasis1,2, A. Loh4, R. Cawthon12, T. Blomberg11,12, S. Wahle6, A.W. Drong5, M. Loh4, S. Zeller12,2, G. Fiorito5, S. Kasela2, R. Richmond11, A. Dehghan11, L. Fr ankle12, T. Esko11,12,13, L. Milan11, C.L. Relton11,14,15, J. Knebel15,20, H. Prokisch26, C. Herder2, A. Peters1, T. Illig15,19, M. Waldenberger14,15, O.J. Groop17,20, N. Gill20, M. Botella20, C.M. Lindgren2,21, M. McCarthy2,22, G. Mathiebu2, C. Gieger12, J.S. Kooner23,24,25, H. Grallert12,3, J.C. Chambers5,8. 1) Institute of Epidemiology II, Helmholtz Zentrum Muenchen - German Research Center for Environmental Health, Neuherberg, Germany; 3) German Center for Diabetes Research (DZD), Neuherberg, Germany; 4) Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK; 5) Welcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 6) Human Genetics Foundation, HuGeF, I-10126 Torino, Italy Department of Medical Sciences, University of Torino, I-10126, Torino, Italy; 7) Estonian Genome Center, University of Tartu, Tartu, Estonia; 8) MRC Human Genetic Epidemiology Unit, Oxford, University of Bristol, Bristol BS8 2BN, UK; 9) Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; 10) Departments of Cardiology and Genetics, University of Groningen, University Medical Centre Groningen, University & G勤奋 Center for Cardiovascular Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands; 11) Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children’s Hospital, Boston, Massachusetts, USA; 12) Program in Medical and Population Genetics, Broad Institute of Cambridge, Cambridge, Massachusetts, USA; 13) Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA; 14) Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, NE1 3BZ, Tyne and Wear, UK; 15) School of Social and Community Medicine, University of Bristol, Bristol, UK; 16) Institute of Human Genetics, Helmholtz Zentrum Muenchen - German Research Center for Environmental Health, Neuherberg, Germany; 17) Institute of Human Genetics, Technische Universitt Muenchen, Munich, Germany; 18) Institute for Clinical Diabetology, German Diabetes Center, Laibniz Center for Diabetes Research, Heinrich Heine University Dusseldorf, Dusseldorf, Germany; 19) Hannover Unifed Biobank, Medical School Hannover, Hannover, Germany; 20) Department of Twin Research & Genetic Epidemiology, King’s College London, St Thomas’ Hospital, London, SE1 7EH, UK; 21) Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA; 22) Oxford Centre for Diabetes Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom; 23) National Heart and Lung Institute, Imperial College London, London W12 0NN, UK; 24) Ealing Hospital NHS Trust, Middlesex UB1 3HW, UK; 25) Imperial College Healthcare NHS Trust, London W12 0HS, UK.

Studying body mass-related DNA methylation signatures is promising from two perspectives. First, epigenetic mechanisms might be involved in the development of obesity, and second, epigenetic regulation may partially be underlying obesity-related pathogenic processes including disturb lipid and glucose metabolism. In an epigenome-wide approach based on the Illumina Infinium 450k assay, we have identified methylation sites strongly associated with body mass index (BMI) in more than 10,000 individuals of European and South Asian ancestry. In subsamples of 3961 and 703 subjects, the identified methylation sites were integrated with single nucleotide polymorphism (SNP) and gene expression data. In addition to the observed BMI-methylation associations was explored by means of Mendelian randomization experiments using SNPs associated with BMI in published genome-wide association studies, and SNPs associated with methylation in our study as instrumental variables. Furthermore, we tested association of the identified methylation sites with incident type 2 diabetes (T2D) during a follow-up of 7.6 years on average (n = 3064), as well as with clinical traits reflecting glucose and lipid metabolism and inflammation (n = 4176). Methylation at 125 CpG sites showed strong genetic regulation (p = 4.610−6 to 1.010−360). 13 CpG sites were significantly associated with expression of genes located in cis (p = 5.510−8 to 6.110−353). Mendelian randomization experiments suggest that methylation at the majority of CpG sites might be consequential to changes in BMI. 92 methylation sites showed significant association with future risk of T2D (p = 7.110−12 with \textit{ABCG1}) or clinical traits independent of BMI, most prominently with triglycerides (p = 5.610−45 with \textit{ABCG1}), HDL cholesterol (p = 3.810−49 with \textit{ABCG1}), C-reactive protein (p = 6.510−36 with \textit{CRELD2}) and Ha1H1 (p = 1.610−12 with \textit{ABCG1}). Adjust- ment for methylation levels at the identified CpG sites resulted in a reduction of up to 68% in the association of these traits with BMI (p = 1.210−4 to 3.810−12). Ongoing analyses will determine the extent to which some of this reduction is due to mediation through methylation. The presented work provides a close examination of the CpG sites discovered in the first large meta-analysis of genome-wide DNA methylation in relation to BMI. Our findings provide new evidence for biological pathways underlying obesity and related metabolic disturbances.
47OT
Epigenome-wide association study identifies epigenetic markers for asthma and allergic disease in the MeDALL study. C.J. Xu1,2,3,4,  T.M. Kerkhof1,4,5, R.T. Lie1,2,6,  M. Arinam1,2,3,4, C. C. van Diemen1,2,3,4,  M. Bustramante1,2,  E. Morales2,  J. Sunyer6,  N. Báz7,  I. Annessi-Maeso8,  U. Gehring9,10, J.M. Antí11,  J. Bouquenet11,  P. van der Vlies1;  C.C van Diemen3,  C. Wijmenga3,4,  D.S. Postma12,  G.H. Koppelman12, 1. University of Groningen, University Medical Center Groningen, Department of Pulmonology, Groningen, The Netherlands; 2. University of Groningen, University Medical Center Groningen, GRIAC Research Institute, Groningen, The Netherlands; 3. University of Groningen, University Medical Center Groningen, Department of Environmental Health Sciences, University Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands; 4. Department of Biosciences and Nutrition, Karolinska Institutet, Sweden; 5. Sachs’ Children’s Hospital and Institute of Environmental Medicine, Karolinska Institutet, Sweden; 6. Center for Research on Environmental Epidemiology (CREAL), Barcelona, Spain; 7. Center for Genomic Regulation (CRG), Barcelona, Spain; 8. Center for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands; 9. University Hospital, Hospital Arnau de Vilanova, Department of Respiratory Diseases, Montpellier, France; 10. University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children’s Hospital Groningen, The Netherlands.

Asthma and allergic diseases are caused by a combination of genetic and environmental factors. Although a large number of genetic variants are associated with asthma and allergic risk, many of the risk factors remain unexplained. The hypothesis that epigenetics are responsible for the missing heritability and tested our hypothesis using an epigenome-wide association approach. We examined DNA methylation in asthma and allergy in 1748 blood samples collected at birth, 4 and 8 years old using a case-control design. Cases of asthma (e.g. allergic rhinitis or eczema or allergic rhinitis) were selected from 4 European birth cohorts (PIAMA, BAMBSE, INMA and EDEN), participating in the Mechanism of the Development of Allergy (MeDALL) consortium. DNA methylation profiling was performed using the Illumina Infinium HumanMethylation450K arrays. A total of 462,742 sites that measures 485,512 different CpG sites covering 96% of RefSeq genes. The target methylation analysis that is capable of interrogating ~1.5 kb regions in a single sequencing read, which theoretically covers ~90.5% of the dmCpgs identified for cleft palate only had concordant direction in genomic regions with lower CpG density and depleted for active histone marks. Nuclear hormone receptors (e.g. PPARγ, ESR1, NOX5, MIR548HD) (0.26, P = 0.03), AJAP1 (B = 0.20, P = 0.04), PARD3 (B = 0.13, P = 0.03), ERRFI1 (B = −0.13, P = 0.03), PDLIM2 (B = −0.07, P = 0.04), CSNK1D (B = −0.08, P = 0.04), and ITLN1 (B = −0.12, P = 0.04). This study identified genome-wide changes to DNA methylation profiles in individuals from women with PE and FDR value < 0.05). Of these, a number of correlated with the development and function. Functional characterisation of genes identified to be differentially methylated in diseased and non-diseased tissues is required. This research may lead to improved understanding of the genetic and environmental factors that correlate with the development of asthma and allergic diseases in childhood.

472T
Multiplexed and Quantitative DNA Methylation Analysis Using Long-Read Single-Molecule Real-Time (SMRT) Bisulfite Sequencing. Y. Yang1, R. Sebra1, I. Peter1, R.J. Desnick1, C.R. Geyer2, J.F. DeCoteau2, S.A. Scott1. 1. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 2. Cancer Stem Cell Research Group, University of Saskatchewan, Saskatoon, SK, Canada, 3. Norway Facial Clefts Study.

DNA methylation has important roles in transcriptional regulation, imprinting, X chromosome inactivation and other cellular processes, and aberrant CpG methylation is directly involved in the pathogenesis of human imprinting disorders. The many targeted methods for analysis of CpG methylation typically involve bisulfite sequencing and sequencing. Although high-throughput targeted methods have been developed to enhance sequencing depth and CpG methylation quantitation, it is not currently feasible for the capture of SMRT bisulfite sequencing four amplicons of the TUBGCP3 CpG island, which resulted in highly consistent methylation quantitation (s.d.=0.023). SMRT bisulfite sequencing was validated by comparing Cpg methylation levels to both 450K array (r=0.903±0.009) and SureSelect Methyl-Seq reads. Coupled with an optimized long-range bisulfite amplification protocol, multiplexed SMRT bisulfite sequencing is an accurate and cost-effective method for targeted CpG methylation analysis that is capable of interrogating ~1.5 kb regions in a single sequencing read, which theoretically covers ~90.5% of CpG islands in the human genome.
474T
Epigenetic mechanisms, particularly DNA methylation, are a possible link between genetic and environmental determinants of health. As the methylation undergos rearrangement in utero and is susceptible to environmental insults, it may be a mechanism explaining the developmental origins of human disease. Polybrominated diphenyl ethers (PBDEs) are flame retardant chemicals that have been widely used in consumer products over the last 40 years. In the CHAMACOS longitudinal birth cohort, which follows participants from low-income Mexican-American farmworker families in California, we have previously found that prenatal and early-life exposure to PBDEs is associated with a variety of adverse health effects, including decreased fertility, abnormal mental development, and altered thyroid function in children. Further, we have demonstrated that exposure to PBDEs is also associated with altered DNA methylation in LINE-1 repetitive elements in newborn children. Here, we report results of the analysis of site-specific DNA methylation by Illumina Infinium HumanMethylation450K BeadChips assessed in cord blood from CHAMACOS newborns and from the same children at 9 years of age. The Illumina assay simultaneously interrogates methylation at 485,512 CpG sites, in 99%; of all RefSeq Genes and 96%; of CpG Islands. Multivariate linear regression and generalized estimating equation analysis identified 8 differentially methylated positions (DMPs) associated with prenatal PBDE exposure (FDR < 0.05). These DMPs were located in 4 different genes: AMN, EPHA6, FAM150A, and SMA5D. We additionally identified clusters of CpG sites associated with prenatal PBDE exposure, which we defined as clusters that are both pairwise correlated and associated with positive environmental exposures in children. Differential cell counts were used to evaluate how white blood cell type composition modified the relationship between exposures and DNA methylation profiles at birth. New data will be presented on the potential role of altered DNA methylation in mediating the effects of environmental exposures on children's health.

475M
Comparing Statistical Methods for Differential Methylation Identification Using Bisulfite Sequencing Data. X. Yu1, S. Sun2. 1) Department of Epidemiology & Bio-statistics, Case Western Reserve University, Cleveland, OH; 2) Department of Mathematics, Texas State University, San Marcos, TX.
DNA methylation is an epigenetic modification involved in organism development and cellular differentiation. Identifying differential methylation changes helps to study genomic regions associated with diseases. Differential methylation studies on single-base-resolution has become possible with bisulfite sequencing (BS) technologies. Several methods have been developed to identify differentially methylated regions (DMRs) in BS data. In this poster, we provide a comprehensive comparison analysis of five DMR identification methods: methylKit, BSSeq, BiSeq, HMM-DN, and HMM-Fisher. We summarize the features of these methods from several analytical aspects, and compare their performances using both simulated and real BS datasets. Based on the simulated analysis results, our discoveries are summarized below. First, parameter settings may largely affect the accuracy of DMR identification. Different from the default settings, the modified parameter settings yield higher sensitivities and/or lower false positive rates. Second, all five methods show higher accuracy in the identification of simulated DMRs that are long and have small within group variations. Third, HMM-DN and HMM-Fisher yield relatively higher sensitivities and lower false positive rates than others, especially in DMRs with large variations. Moreover, in the real data analysis, five methods show low concordance, probably due to the different approaches they have used when addressing the issues in DMR identifications. In addition, we find that among the three methods (methylKit, BSSeq, and BiSeq) that involve methylation estimation, BiSeq can best present the raw methylation signals. Therefore, based on the above results, we suggest that users select DMR identification methods based on the characteristics of their data and the advantages of each method. To guarantee a higher accuracy in validation and further analysis, users may choose the identified DMRs that are relatively long and have small within group variations as a priority.

476T
Potential susceptibility factors of congenital heart disease identified by epigenome-wide association study of placenta. C. Zeng, W. Wei, J. Zhu, Z. Liu1. 1) Laboratory of Genomic Variations and Precision Biomedicine, Beijing Institute of Genomics, Chinese Academy of Sciences; 2) National office for Maternal and Child Health Surveillance, West China Second University Hospital, Sichuan University.
Congenital heart disease (CHD) is the most common birth defect and genetic, epigenetic, and environmental factors all have contributions to its occurrence. As a nutrition-exchanging organ in embryo development, placenta provides a connection between fetal and environmental factors. Epigenetic mechanisms may play an important role. In this study, we conducted an epigenome-wide association study (EWAS) in 26 placentas of CHD fetuses and 15 placentas of the healthy ones. Considering the anatomy of placenta, each sample was divided into the maternal side and the fetal side for analysis. DNA methylation profiling was performed using Illumina Methylation450K chip, and loci around SNP sites were ignored. To remove possible contaminations from the blood, we did principle component analysis (PCA) in placenta and blood samples. Two CHD and one control were then excluded due to their closeness to blood samples in PCA.

The DNA methylation of fetal side tissues differed significantly from the maternal side in PC4, which was set to be the covariant in EWAS. In association analysis using linear regression model, 44 loci of significantly differential methylation were observed (p<10−7) between cases and controls in maternal side, and their corresponding genes were enriched in alcohol metabolism (FDR = 0.001), and ALDH3A1 gene. Further, differential methylation patterns in the corresponding genes may be associated with injuries in the fetal side, especially in DMRs with large variations. More-over, in the real data analysis, five methods show low concordance, probably due to the different approaches they have used when addressing the issues in DMR identifications. In addition, we find that among the three methods (methylKit, BSSeq, and BiSeq) that involve methylation estimation, BiSeq can best present the raw methylation signals. Therefore, based on the above results, we suggest that users select DMR identification methods based on the characteristics of their data and the advantages of each method. To guarantee a higher accuracy in validation and further analysis, users may choose the identified DMRs that are relatively long and have small within group variations as a priority.
A microRNA self-regulatory network in testicular germ cell tumor. W. Chan¹, Y. Suen¹, S. Gui², L. Li¹, B. Chen¹,², 1) CUHK-Shandong University Joint Laboratory on Reproductive Genetics, School of Biomedical Sciences, Chinese University of Hong Kong, Shatin, Hong Kong SAR; 2) Department of Biological Science and Biotechnology, School of Science, Wuhan University of Technology, Wuhan, Hubei, China.

It was previously demonstrated that microRNA-199a (miR-199a) is down-regulated in testicular germ cell tumor (TGCT) partially caused by hypermethylation of its promoter. miR-199a is encoded by two loci in the human genome, namely, miR-199a-1 on chromosome (Chr) 19 and miR-199a-2 on Chr 1. Both loci encode the same miR-199a. Another microRNA, miR-214, also locates on Chr 1 and in close vicinity of miR-199a. Previous study revealed that it is co-transcribed with miR-199a-2. However, the biological significance of the co-transcription of miR-199a and miR-214 remains largely unknown. In this study, we showed that miR-199a and miR-214 were concordantly expressed in a TGCT cell line NT2 and in TGCT patient tissues. After 5-aza treatment, miR-199a and miR-214 expression was significantly increased in NT2. Silencing of DMN1 (DNA (cytosine-5)-methyltransferases) 1 with siRNA restored the expression of miR-199a and miR-214, accompanied by de-methylation of the promoters of miR-199a-1/2. Tumor protein p53 (TP53) down-regulated the expression of DMN1 in NT2 cells and overexpression of TP53 restored the expression of miR-199a and miR-214. In addition, silencing of PSD10 up-regulated the expression of TP53, while miR-214 over-expression resulted in PSD10 down-regulation and TP53 up-regulation. Collectively, our findings highlighted a miR-199a/miR-214/TP53/DMN1 self-regulatory network, which caused the down-regulation of miR-199a, miR-214, and TP53, respectively, as well as the up-regulation of DMN1. This study provides a mechanism of promoter DNA hypermethylation of miR-199a in TGCT. They also suggest a potential therapeutic approach by targeting the miR-199a/miR-214/TP53/DMN1 regulatory network in the treatment of TGCT.
482T Identification of tumor suppressor genes modulated by histone acetylation in gastric cancer. F. Wisnieski1, D.Q. Calcagno2, M.F. Leal1,2, L.C. Santos1, C.V. Gigek1, E.S. Chen1, T.B. Pontes1, S. Demachk1, R. Artigliani2, P.P. Assumpção2, L.G. Loureno2, R.R. Burbano6, M.C. Smith1. 1) Genetics Division, Department of Morphology and Genetic, Federal University of So Paulo, So Paulo, Brazil; 2) Nucleus of Research in Oncology, Joo de Barros Barreto University Hospital, Federal University of Par, Belm, Brazil; 3) Department of Orthopaedics and Traumatology, Federal University of So Paulo, So Paulo, Brazil; 4) Department of Pathology, Federal University of So Paulo, So Paulo, Brazil; 5) Department of Surgical Gastroenterology, Federal University of So Paulo, So Paulo, Brazil; 6) Human Cytogenetics Laboratory, Institute of Biological Sciences, Federal University of Par, Belm, Brazil.

Despite the fact that overall rates of gastric cancer (GC) continue to decline worldwide, the majority of patients are still diagnosed with advanced disease in Western countries. In these cases, surgical resection of the primary tumor offers limited value for a cure and has high morbidity rates. New strategies for early diagnosis and new therapeutic methods in GC continue to be explored. Epigenetic control using histone deacetylase inhibitors, such as trichostatin A (TSA), is a promising cancer therapy. This study aimed to identify genes modulated by TSA in gastric cell lines, and to evaluate the expression of selected genes in gastric resection specimens. ACP02 and ACP03 GC cell lines were treated in triplicate with 250 nM TSA for 24 hours. Differentially expressed genes (DEGs) in treated cell lines compared to controls were identified using microarray assay. The validation of selected DEGs and the assessment of their expression in 46 pairs of primary gastric adenocarcinoma and adjacent non-tumor tissues were performed using qRT-PCR. Possible associations between the expression levels of these genes and clinicopathological features were also evaluated. Microarray analysis revealed 42 DEGs (20 upregulated and 22 downregulated genes). Ten relevant functions were observed in which 2 were enriched with a significant number of genes: a) connective tissue, immune, and inflammatory diseases; b) cell cycle, drug, and lipid metabolism. The upregulated DEGs BMP8B, Bambi, and LRRC37A2 were validated and initially selected for gastric tissue analysis. Reduced transcript levels of BMP8B were found in different types of gastric carcinomas compared with adjacent non-tumor tissues (p=0.001 and p=0.006, respectively). In addition, reduced expression of LRRC37A2 was observed in early stage gastric tumors compared with adjacent non-tumor tissues (p=0.001). On the other hand, the level of expression of Bambi did not differ between gastric tumors and non-tumor samples. Our results demonstrated, for the first time, BMP8B and LRRC37A2 are possible tumor suppressor genes in gastric cancinogenesis and constitute potential targets for epigenetic therapy through histone deacetylase inhibitors.
Methylation analysis and diagnostics of Beckwith-Wiedemann syndrome in 1000 subjects. A. Ibrahim1,2, G. Kirby3, C. Hardy4, R. Dias5, L. Tee6, D. Lim7, J.N. Berg8, F. MacDonald6, P. Nightingale6, E.R.Maher1,3, A. Lim9, O. Vincent9, K. Rohde10, M.-H. Dizier1,2, J. Esparza-Gordillo1, P. Margaritte-Jeannin1,2, L. Liang6, Y.-A. Lee7, P.-F. Pignatti10, W.O.C. Cookson8, T. Pastinen9, M. Lathrop9, F. Demenais1,2, E. Bouzigon1,2,10, 1) Department of Medical Genetics, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, United Kingdom; 2) Department of Clinical Genetics, University of Dundee, Dundee, United Kingdom; 3) Centre for Rare Diseases and Personalised Medicine, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom; 4) West Midlands Regional Genetics Service, Birmingham Women’s Hospital, Birmingham, United Kingdom; 5) Wellcome Trust Clinical Research Facility, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital, Birmingham, United Kingdom.

Background - Beckwith-Wiedemann syndrome (BWS), a congenital overgrowth disorder with variable expressivity and a predisposition to tumorigenesis, is clinically diagnosed by criteria from the World Health Organization. There are no generally agreed clinical diagnostic criteria, with molecular studies commonly performed to confirm diagnosis. In particular, methylation status analysis at two 11p15.5 imprinting centres (IC1 and IC2) detects up to 80% of BWS cases (though low-level mosaicism may not be detected). In order to evaluate the relationship between the clinical presentation of suspected BWS and IC1/2 methylation abnormalities we reviewed the results of >1000 referrals for molecular diagnostic testing.

Methods and Results - 507/1091 (46.5%) referrals had a positive diagnostic test for BWS. The frequency of tumours was 3.4% in those with a molecular diagnosis of BWS. Previously reported genotype-phenotype associations with paternal uniparental disomy, IC1, and IC2 epimutation groups were confirmed and potential novel associations detected. Predictive values of previously described clinical diagnostic criteria were compared, and although there were differences in their sensitivity and specificity, receiver operating characteristic (ROC) analysis demonstrated that these were not optimal in predicting 11p15.5 methylation abnormalities. Using logistic regression, we identified clinical features with the best predictive value for a positive methylation abnormality. Furthermore, we developed a weighted scoring system (sensitivity - 75.9% and specificity - 81.8%) to prioritise patients presenting with the most common features of BWS, and ROC analysis demonstrated superior performance (area under the curve - 0.85; 95% CI: 0.83-0.87) compared to previous criteria.

Conclusion - We suggest that this novel tool will facilitate selection of patients with suspected BWS for routine diagnostic testing and so improve the diagnosis of the disorder.
Genomic imprinting is an epigenetic gene-marking phenomenon that is established in germine. The importance of genomic imprinting is to regulate gene expression in parental-specific manner. One of the new findings has shown that imprinting does not always occur similarly across all tissues. Recent studies suggest a more complicated scheme of tissue-specific promoters and epigenetic marks. We test the hypothesis that imprinted genes can be predicted by the methylation level. In case of genomic imprinting one of the parental copies is usually silenced through DNA methylation. Based on this knowledge we expect semi-methylation in imprinted genes. In order to prove this hypothesis we analysed the DNA methylation in well-known imprinted genes across the tissue panel from the same individuals. All 17 tissues from every 4 individuals were collected during the autopsy. DNA methylation analysis of the total 72 tissue samples and controls was performed with the Illumina Infinium HumanMethylation450 BeadChip in Estonian Genome Centre. We used Levene’s test for comparison of known imprinted genes with the rest of the genes captured by 450K methylation array. As a result, all imprinted genes (n=92) demonstrated less variability in the methylation level (p < 0.01) across all 17 tissues when compared to non-imprinted genes. We also visualized the CpG patterns of known imprinted genes across all tissues. Each CpG was annotated to its exact location in the genome in exon, gene body or UTR region. Visualized CpG patterns also confirmed tissue-specific nature of imprinted genes. For example, gallbladder shows medium methylation of KCNQ1DN gene as CpG sites are only partially methylated, while in isciatic nerve the CpG sites are not methylated. Using this mapping method, we have narrowed down the list of potential candidate genes to 3,000. We found that some genes meet the criteria for candidate imprinted genes in all somatic tissues, while other genes meet those criteria only in some of the tissues. As the next step we need to analyze the DNA methylation in well-known imprinted genes across the whole human genome. Our method will give a better understanding of the nature of the imprinted genes.

Characterizing the processing, localization, and function of Snord116 noncoding RNAs at the Prader-Willi locus. R. Coulson, W. Powell, D. Vasut, J. Afkarian, S. Wong, J. LaSaile, Medical Microbiology and Immunology, Genome Center, M.I.N.D Institute, University of California Davis, Davis, CA.

Prader-Willi syndrome (PWS) is a neurodevelopmental disorder caused by paternal deficiency of the imprinted 15q11-13 noncoding RNA cluster SNORD116. This repeat cluster is GC-skewed, resulting in R-loop formation, histone displacement, and chromatin decondensation specifically of the paternal allele upon neuronal transcription. SNORD116 produces two non-coding, 116HG snoRNAs, which localize to the 116HG snoRNA gene clusters of maturing neurons, and 116 host gene (116HG), which forms an RNA cloud at its paternally decondensed site of transcription. To understand the processing, localization, and functional relevance of each component of SNORD116, we engineered novel transgenic mice to test complementation of the PWS mouse model, Snord116del. Complete transgene wild-type (Ctg/WT) mice containing 27 copies of the transgenic Snord116 repeat were generated and bred to Snord116del males to produce offspring lacking paternal Snord116 but expressing the transgene (Ctg/Snord116del). RNA fluorescence in situ hybridization (FISH) analysis of brain showed that spliced 116HG localized in a distinct cloud at its decondensed transcription site and mature snoRNAs localized to the nucleolus in wild-type (WT) but not Snord116del neurons, as previously described. In Ctg/WT neurons, a single 116HG RNA cloud was detected that was significantly larger than that observed in WT neurons, suggesting that the Ctg-derived 116HG colocalizes with the endogenous 116HG RNA cloud at the paternal Snord116 locus. In addition, snoRNAs were detected in the nucleolus at a significantly higher level in Ctg/WT than in WT neurons. Despite expression of the 116HG RNA cloud, and all Ctg/Snord116del tissues, no RNA cloud or snoRNAs were detected. qRT-PCR analysis demonstrated that splicing of Snord116 was largely restricted to neuronal tissues of both Ctg/WT and Ctg/Snord116del mice, despite expression of the transgene in many tissues. These combined results suggest that processing of the 116HG transcript is dependent on neuronal factors and/or chromatin states, including the decondensed Snord116 paternal allele. Analyses of a spliced 116HG/Snord116del transgenic mouse is underway to further test the hypothesis that localization of the 116HG RNA cloud requires the decondensation of paternal Snord116 in neurons. Understanding how DNA-RNA interactions mediate the processing and localization required for Snord116 function and phenotypic rescue is critical for the development of effective PWS therapies in the future.
Random monoallelic expression in neuronal development, induced pluripotent stem cells and neural committed cells. A.R. Jeffries1,2, D.A. Uwano1, G.Cocks2, L.W. Perfect3, E.L. Dempster1, J. Mill1, J. Price1. 1 University of Exeter Medical School, Royal Devon & Exeter Hospital, Barrack Road, Exeter. 2 King’s College London, Centre for the Cellular Basis of Behaviour, The James Black Centre, London, SE5 9NU. United Kingdom.

Random monoallelic expression is the choice of which alleles to express is made at random early in development and maintained in subsequent clonal progeny. This results in cellular level heterogeneity of allele specific gene expression together with any resulting phenotypic effects. Up to 10% of autosomal genes can show random monoallelic expression yet little is known on the precise mechanisms behind the allelic choice made early in development.

We use epigenetic reprogramming to investigate the effects on previously characterized random monoallelic expressed genes. We also neutralize the resulting induced pluripotent stem cells (iPSCs) to a more committed state to further examine allelic choice.

Results: Genes which previously showed random monoallelic expression in neuronal stem cells reverted to biallelic expression after epigenetic reprogramming into iPSCs. Lineage commitment into neural stem cells followed by clonal isolation revealed a number of genes showing random monoallelic expression choice, many of which represented new gene loci. Global DNA methylation analysis was also performed using the Illumina 450k methylation beadchip. Clones undergoing monoallelic expression showed increased DNA methylation at gene promoters compared to sister clones which showed biallelic expression.

In summary, random monoallelic expression is lost during epigenetic reprogramming but re-established in a stochastic fashion when cells are committed towards a neuronal lineage. Clonal differences in allelic expression status show associated promoter level DNA methylation differences.

Somatic and genetic variations in regulatory regions revealing discordance between monozygotic twins. K. Kim1, HJ. Ban1, J. Seo1, K. Lee1, M. Yavartanoo1, SC. Kim2, K. Park3, SB. Cho4, JK. Choi2. 1 Department of Bio and Brain Engineering, KAIST, Daejeon 305–710, Republic of Korea; 2 Department of Biomedical Informatics, Center for Genome Science, National Institute of Health, KCDC, Chongchung-Buk-do 363–951, Republic of Korea; 3 Division of Molecular and Life Sciences, Hanyang University, Ansan, Gyeonggi-do 425–791, Republic of Korea; 4 Research Institute of Bioinformatics, Omics Inc., Daejeon 305–333, Republic of Korea; 5 Samsung Genome Institute, Samsung Medical Center, Seoul 135–710, Republic of Korea.

Open chromatin has important role as regulatory region in transcriptional process, and genetic variations in open chromatin have been known to contribute to discordance of gene expression and phenotype. However, the effect of somatic variation on regulatory region still remains to be elucidated. In this study, we performed FAIRE sequencing (~72X) to analyze open chromatin, array-based genotyping across 72 monozygotic (MZ) twins to identify genetic and somatic variations that can explain twin discordance in chromatin accessibility. First, we focused on the spectrum of somatic and genetic sequence variations underlying discordant open chromatin mainly at developmental level but not through natural selection. Next, we used epigenetic reprogramming to investigate the effects on previously characterized random monoallelic expressed genes. We also neutralize the resulting induced pluripotent stem cells (iPSCs) to a more committed state to further examine allelic choice.

Results: Genes which previously showed random monoallelic expression in neuronal stem cells reverted to biallelic expression after epigenetic reprogramming into iPSCs. Lineage commitment into neural stem cells followed by clonal isolation revealed a number of genes showing random monoallelic expression choice, many of which represented new gene loci. Global DNA methylation analysis was also performed using the Illumina 450k methylation beadchip. Clones undergoing monoallelic expression showed increased DNA methylation at gene promoters compared to sister clones which showed biallelic expression.

In summary, random monoallelic expression is lost during epigenetic reprogramming but re-established in a stochastic fashion when cells are committed towards a neuronal lineage. Clonal differences in allelic expression status show associated promoter level DNA methylation differences.

Osteoporotic fractures are a socio-economic burden and are increasing in incidence with the ageing of industrial societies. Since 2007, numerous genome-wide association studies (GWAS) for osteoporosis and related traits have identified multiple common genetic variants associated with BMD, however little is known concerning the role of epigenetic changes influencing BMD. The aim of this study is to investigate the association between epigenetic changes and BMD phenotypes by undertaking an epigenome-wide association study of monozygotic (MZ) and dizygotic (DZ) individual twins from EpiTwin (http://www.epitwin.eu/), a TwinsUK project (http://www.twinsuk.ac.uk/).

We have completed whole genome methylated DNA immunoprecipitation sequencing (MeDIP-seq) for 1,248 MZ twins and 320 DZ twins to quantify genome-wide methylation levels. Methylation levels for each 500 bp region (250 bp overlap) of the genome will be compared with BMD for all twins using a linear mixed effects model to account for the twin relationships. Additional effects will include age, gender, and batch. Significant regions will be differentially methylated regions (DMRs), identified to be associated with BMD through variation in methylation levels. Analysis is currently ongoing. Each twin has three measurements for BMD at total forearm, total hip, and total spine.

Identified DMRs suggest that region-specific variation of methylation levels is associated with osteoporotic phenotypes. Significant findings will be qualitatively compared to regions already identified by GWAS to influence BMD, where any overlaps would suggest an additional mechanism of influence for the region on BMD. DMRs could also identify novel drug targets to treat osteoporosis, as the majority of osteoporosis drug targets have been identified through studying the determinants of BMD.
493M Epigenome-wide association study of sexual orientation in monoy- 
gotic twins. T.C. Ngun1, W. Guo2, N.M. Ghaithaman3, K. Purkayastha4, 
F.J. Sanchez1, S. D’Andrea5, M. Zhang2,6, K. Purkayastha1, J. Duis1, 
1) Department of Human Genetics, David Geffen School of Medicine at Univer- 
sity of California Los Angeles (UCLA), Los Angeles, CA; USA Los Angeles, 
CA; 2) Bioinformatics Division and Center for Synthetic & Systems Biology, 
TNLIST, Tsinghua University, Beijing 100084, China; 3) Department of 
Pathology and Laboratory Medicine, David Geffen School of Medicine at 
UCLA, Los Angeles, CA, USA; 4) Department of Counseling Psychology. 
The University of Wisconsin-Madison, WI, USA; 5) Department of Molecular 
Cell Biology and Genetics, University of Texas at Dallas, Richardson, TX 75080, USA; 6) Department of Molecular, Cellular, 
and Developmental Biology, UCLA, Los Angeles, CA, USA.

Sexual orientation is one of the most pronounced sex differences in the 
animal kingdom. Although upwards of 95% of the general population is 
hetereosexual, a small but significant proportion of individuals (3–5%) is 
 homosexual. Previous work has consistently shown that sexual orientation 
has a strong genetic component. Male sexual orientation has been linked to 
several genomic loci including Xq28, 6p12, and 10q28, which is an imprinted 
region. As with other complex traits, environmental factors also play an 
important role. For instance, monogygotic twins show substantial levels of 
discordance for this trait. Additionally, each biological older brother a man 
has increases the probability of him being gay by 33%. Taken together, the 
evidence strongly suggests a role of epigenetic mechanisms in the regulation of 
sexual orientation. We aim to elucidate the molecular mechanisms under- 
lying sexual orientation by investigating DNA methylation patterns on a 
genome-wide basis in 34 monogygotic male twin pairs that are discordant for 
sexual orientation. The distribution of DNA methylation changes between 
sexually discordant monozygotic twins (RRBS). As a control population, we also performed RRBS on 10 monozy- 
gotic twin pairs concordant for homosexuality. Subjects were classified as 
 homosexual or heterosexual based on their sexual feelings (as measured by 
the Kinsey scale). We first consolidated nearby CpG sites into short 
regions (100–500 bp) to increase the signal to noise ratio in our data. 
A linear mixed modeling approach was used to identify regions that were 
significantly associated with sexual orientation (sexual orientation DMRs).
At a false discovery rate of 10%, we identified 136 sexual orientation DMRs. 
One of those regions was in 10q26 but there were no other overlaps between 
our data and previous linkage findings. This discrepancy was not surprising 
as prior studies primarily used families with evidence of maternal loading, 
which was not a criteria for exclusion in our study. Several regions associated with 
sexual orientation DMRs were involved in the androgen signaling path- 
way: (1) AR, which encodes the androgen receptor; (2) ZDHHC7, which 
regulates the localization of sex steroid receptors; and (3) SULT1A3, which 
is part of the alternative androgen synthesis pathway. Our findings demon- 
strate that numerous epigenetic changes are associated with sexual orienta-
 tion in humans and that our approach has the potential to identify novel 
genomes that influence this trait.

494T Genes that escape from X inactivation vary in mouse tissues. C. Dis- 
teche1,4, J. Berletch1, W. Ma4, F. Yang5, J. Shendure2, W. Noble6, X. Deng1, 
1) Dept Pathology, Univ Washington, Seattle, WA; 2) Dept Genome 
Sciences, Univ Washington, Seattle, WA; 3) Fred Hutchinson Cancer 
Research Center, Seattle, WA; 4) Dept of Medicine, Univ Washington, Seat-
tle, WA.

X chromosome inactivation (XCI) silences most genes on one X chromo-
some in females but some genes escape XCI. Surveys in cultured 
human/mouse hybrid cells and in cell lines from individuals with skewed 
XCI have shown that about 8–15% of human genes consistently escape 
XCI, 10–13% display variable levels of escape, and 10–20% vary between 
cell lines and individuals. Escape from XCI results in significant sexual 
dimorphisms in levels of gene expression, and bi-allelic expression of 
at least some escape genes is important for a normal phenotype in human 
females. Indeed, the presence of a single X chromosome (45,X) results in 
Turner syndrome. To identify escape gene in vivo and to explore molecular 
mechanisms that regulate this process we analyzed the allele-specific 
expression and chromatin structure of X-linked genes in mouse tissues and 
cells with skewed XCI and distinguishable alleles based on single nucleotide 
polymorphisms. Using a new method to estimate allele expression, we 
demonstrate a continuum between complete silencing and significant 
expression from the inactive X (Xi). Few genes (2–3%) escape XCI to a 
significant level and only a minority differs between mouse tissues, suggest-
ing stringent silencing and escape controls. Allelic profiles of DNase I hyper-
sensitivity and RNA polymerase II occupancy of genes on the Xi correlate 
with escape from XCI. Allelic binding profiles of the DNA binding protein 
CCCTC-binding factor (CTCF) in different cell types indicate that CTCF 
binding at the promoter correlates with escape. Importantly, CTCF binding at 
the boundary between escape and silenced domains may prevent the 
spreading of active escape chromatin into silenced domains.

495M Male Rett syndrome: Clinical profiling and insights into epigenomics. 
J. Duis1, S. Gupta2, S. Naird1. 1) The McKusick-Nathans Institute of Genetic 
Medicine, Johns Hopkins, Baltimore, MD; 2) Department of Neurogenetics, 
Kennedy Krieger Institute, Baltimore, MD.

Rett syndrome (RTT [MIM 312750]) is an X-linked disorder caused by 
mutations in the gene encoding methyl CpG binding protein 2 (MECP2), 
which is known to be essential in neuronal function. Once thought to be 
lethal, Rett is now extensively described. We report 7 cases of male 
Rett with special attention to the clinical profile including phenotype and 
genotype correlations. We review what is known in the literature, report 
novel mutations with insights into inheritance of these mutations, the roles 
of maternal inheritance gleaned from these cases, and how this facilitates 
global understanding of the roles epigenetics may play in this disorder.

496T Examining Escape from X Chromosome Inactivation and Sex-Differen-
tial Gene Expression. P.G. Bronson, T.R. Bhangale, K.T. Cuenne, W. Ort-
mann, R.R. Graham, T.W. Behrens. Genentech, Inc., South San Fran-
cisco, CA.

Autoimmune disease displays a predilection for females, with a sex ratio 
of up to 12:1 for systemic lupus erythematosus in women. The cause of 
this sex imbalance is unknown, but hormones are suspected to play a role. 
A gene dosage effect for the X chromosome may also be involved. Female 
embryonic cells undergo X chromosome inactivation (XCI), through CpG 
hypermethylation and chromatin remodelling initiated by the noncoding RNA 
XIST, to randomly turn all but one X into a compact, inactive state. Previous 
studies have established that ~15% of the 1,031 genes on X escape XCI.
We analyzed 195 lymphoblastoid cell lines (LCLs) from females with RNA-
Seq expression data (Gevaudan RNA Sequencing Project of 1KGP) and 
DNA sequence genotypes available (1KGP) to identify a subset of “mono-
clonal” LCLs. We defined a “monoclonal” LCL as one in which the same 
X was inactivated in ≥50% of the LCLs, as determined by informative (heterozy-
gous variants) in non-escaping genes on X. RNA-Seq reads were mapped to 
the reference genome using GSnap, which performed splice-junction-
aware split-read alignment. Using known variants in the samples, GSnap 
performed SNP-tolerant alignment and removed the reference-bias in read 
mapping. Gene expression levels were quantified in terms of RPKMs using 
Deseq. We identified 63 “monoclonal” LCLs of the 195 examined, 
and observed escape from XCI in 80 genes using informative SNPs; escape 
was defined as having >=1 heterozygote cell line with skewed XCI (≥90% 
of RNA reads coming from only one of the two alleles). From the literature 
we identified an additional 41 escaping X genes for a total of 121 genes 
that escape XCI. The Kolmogorov-Smirnov test was used to test 13,615 
genomes for expression differences between females (n=241) and males (n= 
214). Significant gene expression differences, defined as >1.2 fold difference 
and p≤5.10^-3, were observed in 13.3% of escaped genes. Expression 
differences were observed for an additional 7 X genes and 16 autosomal 
genes, including the T-cell activation RhoGTPase activating protein (an 
autoimmunity locus) and adrenomedinull (previously reported to have higher 
expression in females). Further work is underway to annotate the genes 
that escape XCI, and those that show differential expression between 
females and males. Additional datasets are also being examined. Under-
standing the mechanisms that control expression of genes on the X chromo-
some may provide new insight into autoimmunity.
497M
Placental microRNAs as potential biomarkers for noninvasive detection of trisomy 21. H. M. Ryu1,2, J. H. Lim1, H. J. Kim1, A. R. Oh1, S. Y. Kim1, D. E. Lee2, S. Y. Park1, Y. J. Han1, J. S. Choi2, K. H. Cho2. 1) Laboratory of Medical Genetics, Medical Research Institute, Cheil General Hospital and Women’s Healthcare Center, Seoul, Korea; 2) Department of Obstetrics and Gynecology, Cheil General Hospital and Women’s Healthcare Center, Kwandong University College of Medicine, Seongnam, Korea.

Objectives: The discovery of fetal nucleic acids in the plasma of pregnant women has led to the development of noninvasive prenatal tests (NIPT). The occurrence of placental microRNAs (miRNAs) in maternal blood has been described and then proposed as a particular class of molecular biomarkers for the NIPT. However, miRNAs for NIPT of fetal trisomy 21 (T21) have yet to be reported. The objective of this study was to discover a panel of placental miRNAs as potential novel biomarkers for NIPT of T21 fetuses and to predict biological functions of new biomarkers using bioinformatics tools. Method: Using microarray-based expression profiling, we compared the expression levels of miRNAs in whole blood samples from non-pregnant women, whole blood samples from pregnant women, euploid fetal placenta samples from pregnant women, and T21 fetal placenta samples from non-pregnant women. We analyzed the differentially expressed miRNAs according to the presence or absence of disease and tissue type (p value < 0.05 and two-fold expression change). Potential target genes of miRNAs were predicted using the TargetScan program. To predict their functions, the functional annotation tools provided by the WebGestalt database were used. Results: We identified 299 miRNAs which reasonably separate the whole blood from the placenta. Among identified miRNAs, 150 miRNAs up-regulated in the placenta and 149 miRNAs down-regulated. Most of up-regulated miRNAs in the placenta were members of mir-498, mir-134, and mir-127 clusters as placenta specific miRNA located on chromosome 19 and 14. Among up-regulated miRNAs in the placenta, mir-1973 and mir-3196 were highly expressed in the T21 placenta than in the euploid placenta. The two miRNAs potentially regulated 203 target genes. Interestingly, the genes were significantly associated with T21-related disorder such as congenital abnormalities and mental disorders. The genes were involved in various pathways such as ErbB signaling, calcium signaling, chemokine signaling, focal adhesion, MAPK signaling, and unfolded protein response. The miRNAs could be candidate as placenta specific as biomarkers for NIPT of T21 fetus as well as provided new insights by regulation of miRNAs into the molecular mechanisms in the placenta of T21 fetus.

498T
Detection of in vivo G-quadruplex structure of the ANXA5 promoter that contributes to the recurrent pregnancy loss. H. Inagaki1, S. Ota2, H. Miyamura3, M. Tsutsumi2, T. Kato1, H. Niszawa1, I. Yanagihara1, H. Kurahashi1. 1) Molecular Genetics, ICMS, Tottori Health University, Toyooka, Aichi, Japan; 2) Obstetrics and Gynecology, Fujita Health University School of Medicine, Toyoake Aichi, Japan; 3) Developmental Medicine, Osaka Medical center and Research Institute for Maternal and Health, Izumi, Osaka, Japan.

Recent findings have highlighted the possibility that polymorphisms within the annexin A5 gene (ANXA5) promoter contribute to the etiology of recurrent pregnancy loss. However, the underlying mechanisms are unknown. An M2 haplotype of the polymorphism confers a high risk of onset for recurrent pregnancy loss. The M2 haplotype shows lower promoter activity and is more likely to have acquired differences in expression during evolution compared with constitutive BAE genes (\(P_{\text{RNA}}\times10^{-8}, P_{\text{protein}}<10^{-5}\)). We show that ANXA5 genes are more likely to have evQTLs (expression variability quantitative trait loci) than BAE genes (\(P<10^{-8}\)). Finally, we observe that fetal brain MAE genes are significantly associated with T21-related disorder such as congenital abnormalities and mental disorders. The genes were involved in various pathways such as ErbB signaling, calcium signaling, chemokine signaling, focal adhesion, MAPK signaling, and unfolded protein response. The miRNAs could be candidate as placenta specific as biomarkers for NIPT of T21 fetus as well as provided new insights by regulation of miRNAs into the molecular mechanisms in the placenta of T21 fetus.

499M
An epigenetically regulated expression-variable class of genes depleted in neurodevelopmental CNVs. A. Gimelbrant1,2, L. A. Weiss1,3,4. 1) Cancer Biology, Dana-Farber Cancer Institute, Boston, MA., United States; 2) Genetics, Harvard Medical School, Boston, MA, United States; 3) Psychiatry, University of California San Francisco, San Francisco, CA, United States; 4) Institute for Human Genetics, University of California San Francisco, San Francisco, CA, United States.

We recently discovered that >3,000 human autosomal genes can be expressed from a single allele in one cell, and from the other allele -or both- in a neighboring cell. This epigenetic phenomenon is known as variable monoallelic expression (MAE). MAE is a mitotically stable process generalizing a somatic mosaic of MAE and BAE cells in a given tissue, with mRNA expression levels higher in BAE cells compared with MAE cells for the same gene. Little is known about functional consequences of MAE. We hypothesize that MAE increases expression level variability in the population, with potential impact on phenotypic variation.

In this study, we used a chromatin signature to identify MAE genes in lymphoblastoid cell lines and in human fetal brain tissue. We assessed available data to compare expression variation in MAE genes to BAE genes at three distinct scales. At the trans-species level, we evaluated gene expression differences between humans and non-human primates. In the human population, we used genetic mapping of expression variability data to identify genes with polymorphisms influencing expression variation. Finally, we provide an analysis of pathogenic and non-pathogenic neurodevelopmental copy number variant (CNV) data in humans.

We find that MAE genes show higher interspecies variation in expression level at the mRNA and protein level and are more likely to have acquired differences in expression during evolution compared with constitutive BAE genes (\(P_{\text{RNA}}<3\times10^{-8}, P_{\text{protein}}<4\times10^{-5}\)). We show that MAE genes are more likely to have evQTLs (expression variability quantitative trait loci) than BAE genes (\(P<10^{-8}\)). Finally, we observe that fetal brain MAE genes are significantly associated with T21-related disorder such as congenital abnormalities and mental disorders. The genes were involved in various pathways such as ErbB signaling, calcium signaling, chemokine signaling, focal adhesion, MAPK signaling, and unfolded protein response. The miRNAs could be candidate as placenta specific as biomarkers for NIPT of T21 fetus as well as provided new insights by regulation of miRNAs into the molecular mechanisms in the placenta of T21 fetus.

500T
Defining the role of CGGBP1 protein in FMR1 gene expression. M. Goracci, S. Lanni, F. Falumbo, G. Mancano, P. Chiurazzi, E. Tabolacci, G. Venti. Institute of Medical Genetics, Catholic University, Rome, Italy.

Fragile X syndrome (FXS) is the most common heritable form of cognitive impairment and is caused by the expansion over 200 repeats and subsequent methylation of the CGG triplets at the 5’ UTR of the FMR1 gene, leading to FMR1 silencing. Specific proteins that could recruit components of the silencing machinery we investigated the role of CGGBP1 in FMR1 gene transcription. CGGBP1 is a highly conserved protein which binds specifically unmethylated CGG tracts. The role of CGGBP1 on FMR1 transcription is yet to be defined. Sequencing analysis and expression studies through quantitative PCR of CGGBP1 were performed in cell lines with different allele expansions (wild-type WT, premutation, methylated full mutations FXS and unmethylated full mutation UFM), demonstrating no differences between them. ChIP assays showed that CGGBP1 binds unmethylated CGG triplets of the FMR1 gene proportionally to the length of the repeats. We also observed that CGGBP1 binds unmethylated CGG triplets of the FMR1 gene. CGGBP1 is not a direct regulator of FMR1 transcription. Support by Telethon Onlus, FRAXA Foundation and Italian Association for fragile X syndrome.
Conclusions: In the current study, RNA-Seq analysis revealed remarkable RNA-editing in the retina of CNV mouse model. RNA editing could play a substantial role in retina inflammation during CNV.
Dasgupta therapeutic strategy for individuals with incurable mitochondrial disease. fibroblasts with high percentage of mtDNA A3243G mutation. These mito-A3243G, resulted in an increased import miR-181c and miR-423-5p from mitochondria. These miRNAs were also increased in MELAS-associated mitochondria compared with wild type mitochondria. RT-qPCR results confirmed these miRNAs were richer in 43B mitochondria than in HSI mitochondria. These miRNAs were also increased in h16 fibroblast mitochondria. However, these miRNAs were unchanged in their cellular. These findings suggested the MELAS mutation, mtDNA A3243G mutation in myopia by inhibiting the underlying scleral changes. Our findings show that scleral miRNAs can be manipulated in culture to miR-103 (1.5 fold, p=0.007). The cytotoxicity assay findings did not reveal of COL1A1, the most prevalent protein in the scleral ECM, undergoes mRNAs. The most prevalent protein in the scleral ECM. col1a1 (at least 2-fold, min p=5.9×10^-5), and up-regulated after inhibition of miR-98 and miR-103) and mRNA (let-7c, let-7e, miR-98 and miR-103) expression. Four biological replicates were assessed for each target and data normalized to the negative scramble controls. QPCRs were performed using Taqman® assays. GAPDH, POLR2A, RNU44 and RNU6B served as housekeeping genes, and the delta CT method was used for data analyses. A live/dead viability/ cytotoxicity assay (Life Technologies) was used for assessing toxicity of the treatments. Results: All four miRNAs were tested and inhibited as expected with their respective inhibitors (at least 40-fold, min p=0.001). We also investigated the potential of altering COL1A1 gene expression levels via selective manipulation of miRNA activity in scleral fibroblast cultures. Methods: Human scleral fibroblasts from donor eyes were grown in culture via selective manipulation of miRNA activity in scleral fibroblast cultures. Regulation of COL1A1 in myopia by inhibiting the underlying scleral changes. Scleral thinning and extracellular matrix (ECM) loss are classic features of high/pathological myopia caused by excessive axial ocular elongation of the eye. miRNAs represent potential therapeutic targets for myopia control by promoting/enhancing ECM deposition in the sclera. Since COL1A1 is the most prevalent protein in the scleral ECM, undergoes degradation and loss in myopia and miRNAs are known to regulate COL1A1, we investigated the potential of altering COL1A1 gene expression levels using standard methods. miRNA inhibitors directed at let-7c, let-7e, miR-98 or miR-103 were introduced into the cells via electroporation. Cells were lysed after 48 hours, followed by RNA extraction and reverse transcription using the Taqman® Cells to CT kits for studying miRNA (let-7c, let-7e, miR-98 and miR-103) and mRNA (COL1A1) expression. Four biological replicates were assessed for each target and data normalized to the negative scramble controls. QPCRs were performed using Taqman® assays. GAPDH, POLR2A, RNU44 and RNU6B served as housekeeping genes, and the delta CT method was used for data analyses. A live/dead viability/cytotoxicity assay (Life Technologies) was used for assessing toxicity of the treatments. Results: All four miRNAs were tested and inhibited as expected with their respective inhibitors (at least 40-fold, min p=0.001). We also investigated the potential of altering COL1A1 gene expression levels via selective manipulation of miRNA activity in scleral fibroblast cultures. Methods: Human scleral fibroblasts from donor eyes were grown in culture via selective manipulation of miRNA activity in scleral fibroblast cultures. Regulation of COL1A1 in myopia by inhibiting the underlying scleral changes. Scleral thinning and extracellular matrix (ECM) loss are classic features of high/pathological myopia caused by excessive axial ocular elongation of the eye. miRNAs represent potential therapeutic targets for myopia control by promoting/enhancing ECM deposition in the sclera. Since COL1A1 is the most prevalent protein in the scleral ECM, undergoes degradation and loss in myopia and miRNAs are known to regulate COL1A1, we investigated the potential of altering COL1A1 gene expression levels using standard methods. miRNA inhibitors directed at let-7c, let-7e, miR-98 or miR-103 were introduced into the cells via electroporation. Cells were lysed after 48 hours, followed by RNA extraction and reverse transcription using the Taqman® Cells to CT kits for studying miRNA (let-7c, let-7e, miR-98 and miR-103) and mRNA (COL1A1) expression. Four biological replicates were assessed for each target and data normalized to the negative scramble controls. QPCRs were performed using Taqman® assays. GAPDH, POLR2A, RNU44 and RNU6B served as housekeeping genes, and the delta CT method was used for data analyses. A live/dead viability/cytotoxicity assay (Life Technologies) was used for assessing toxicity of the treatments. Results: All four miRNAs were tested and inhibited as expected with their respective inhibitors (at least 40-fold, min p=0.001). We also investigated the potential of altering COL1A1 gene expression levels via selective manipulation of miRNA activity in scleral fibroblast cultures. Methods: Human scleral fibroblasts from donor eyes were grown in culture via selective manipulation of miRNA activity in scleral fibroblast cultures. Regulation of COL1A1 in myopia by inhibiting the underlying scleral changes. Scleral thinning and extracellular matrix (ECM) loss are classic features of high/pathological myopia caused by excessive axial ocular elongation of the eye. miRNAs represent potential therapeutic targets for myopia control by promoting/enhancing ECM deposition in the sclera. Since COL1A1 is the most prevalent protein in the scleral ECM, undergoes degradation and loss in myopia and miRNAs are known to regulate COL1A1, we investigated the potential of altering COL1A1 gene expression levels using standard methods. miRNA inhibitors directed at let-7c, let-7e, miR-98 or miR-103 were introduced into the cells via electroporation. Cells were lysed after 48 hours, followed by RNA extraction and reverse transcription using the Taqman® Cells to CT kits for studying miRNA (let-7c, let-7e, miR-98 and miR-103) and mRNA (COL1A1) expression. Four biological replicates were assessed for each target and data normalized to the negative scramble controls. QPCRs were performed using Taqman® assays. GAPDH, POLR2A, RNU44 and RNU6B served as housekeeping genes, and the delta CT method was used for data analyses. A live/dead viability/cytotoxicity assay (Life Technologies) was used for assessing toxicity of the treatments. Results: All four miRNAs were tested and inhibited as expected with their respective inhibitors (at least 40-fold, min p=0.001). We also investigated the potential of altering COL1A1 gene expression levels via selective manipulation of miRNA activity in scleral fibroblast cultures. Methods: Human scleral fibroblasts from donor eyes were grown in culture via selective manipulation of miRNA activity in scleral fibroblast cultures. Regulation of COL1A1 in myopia by inhibiting the underlying scleral changes. Scleral thinning and extracellular matrix (ECM) loss are classic features of high/pathological myopia caused by excessive axial ocular elongation of the eye. miRNAs represent potential therapeutic targets for myopia control by promoting/enhancing ECM deposition in the sclera. Since COL1A1 is the most prevalent protein in the scleral ECM, undergoes degradation and loss in myopia and miRNAs are known to regulate COL1A1, we investigated the potential of altering COL1A1 gene expression levels using standard methods. miRNA inhibitors directed at let-7c, let-7e, miR-98 or miR-103 were introduced into the cells via electroporation. Cells were lysed after 48 hours, followed by RNA extraction and reverse transcription using the Taqman® Cells to CT kits for studying miRNA (let-7c, let-7e, miR-98 and miR-103) and mRNA (COL1A1) expression. Four biological replicates were assessed for each target and data normalized to the negative scramble controls. QPCRs were performed using Taqman® assays. GAPDH, POLR2A, RNU44 and RNU6B served as housekeeping genes, and the delta CT method was used for data analyses. A live/dead viability/cytotoxicity assay (Life Technologies) was used for assessing toxicity of the treatments. Results: All four miRNAs were tested and inhibited as expected with their respective inhibitors (at least 40-fold, min p=0.001).

Olfactory neuroepithelium (ONE) represents a site where neural stem cells can be recovered (NSC), besides being a readily accessible site. They can be cultured by sphere assay, forming spherical cluster with multipotent and progenitor cells in suspension, with ability to differentiate. These features make it a model for studying molecular and cellular neural processes. Mainly, neural stem cells have been used to study the mechanisms that regulate the self-renewal and differentiation of NSCs, in order to be used as a tool to investigate neural processes.

508T Epigenetic factors regulating DUX4 expression in muscle cells. J. Balog1, P.E. Thissen2, K.R. Strassheim3, Y.D. Krom1, A. de Jong1, R.J.L. Lemmers1, P.J. van der Vliet1, R. Tawil1, S.J. Tapscott4, S.M. van der Maarel1. 1) Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Neuromuscular Disease Unit, Department of Neurology, University of Rochester Medical Center, Rochester, New York, United States of America; 3) Division of Human Genetics, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America.

Facioscapulohumeral dystrophy (FSHD) is a progressive, dominantly inherited myopathy affecting 1:1,400,000–20,000 individuals. A key factor in the pathomechanism of FSHD is DUX4, a double-homeobox gene whose transcription factor which misexpression causes muscle cell death. Sporadic expression of DUX4 in myonuclei is preceded by chromatin relaxation of the D4Z4 macrosatellite repeat array that contains a copy of the DUX4 gene in every single D4Z4 unit. These epigenetic changes in D4Z4 are a consequence of either a contraction of the D4Z4 repeat array in ~95% of cases (FSHD1), or of mutations in SMCHD1 (Structural Maintenance of Chromosomes Flexible Hinge Domain Containing) in most other cases (FSHD2). We aimed to determine the epigenetic requirement of DUX4 transcript expression in primary human myoblasts and myotubes. First we established the dynamics of DUX4 expression during myogenesis in a series of control (n=6), FSHD1 (n=4) and FSHD2 (n=6) human primary myoblast and myotube samples. DUX4 transcripts were detected in FSHD1 and FSHD2 myotubes and a robust increase of DUX4 transcript levels was observed in FSHD myotubes. To understand the epigenetic changes underlying transcriptional upregulation of DUX4 we quantitatively measured DNA methylation and activating histone modifications (H3K4me2, H3K4me3 and H3K27ac) or repressive histone modifications (H3K9me3 and H3K27me3) at D4Z4 in the same samples DNA methylation levels were lower in patient myoblasts and myotubes, but no significant change was found during myogenensis suggesting that increased levels of DUX4 in myotubes do not correlate with changes in DNA methylation. Activating histone modifications did correlate with transcription, with higher levels in patient samples compared to controls. We could not confirm the reported loss of H3K9me3 at D4Z4 in patient myotube samples. Furthermore downregulation of known chromatin modifiers involved in H3K27me3 and H3K9me3 did not induce DUX4 transcription in control myoblasts and myotubes while depletion of SMCHD1 resulted in transcriptional derepression of DUX4. Interestingly the level of DUX4 was lower in control PRS compared to patient FSHD1 or FSHD2 myoblasts but subsequently decreased during myogenic differentiation possibly contributing to the leaky expression of DUX4. We conclude that SMCHD1 is a central chromatin regulator at D4Z4 controlling DUX4 expression in FSHD1 and FSHD2.

509M Integrative analysis reveals enhanced regulatory effects of human large intergenic noncoding RNAs in lung adenocarcinoma. X. Kong1, Y. Zhou1, J. Huang1, J. Jiang1, J. Huang1, P. Zhang1, Y. Zhu1, Y. Shi1, L. Hu1. 1) Molec Gen, Inst Hilf Sci, Shanghai, China; 2) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America; 3) State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China.

Although there is an accumulating appreciation of the key roles that large intergenic noncoding RNAs (lincRNAs) play in diverse cellular processes, our knowledge of how lincRNAs function in cancer remains sparse. Here, we present a comprehensive landscape of RNA-seq transcriptome profiles of lung adenocarcinomas and their paired normal counterparts from 72 South Korean patients to unravel the mysteries of lincRNA functions and mechanisms. We found asymmetric expression patterns of lincRNAs between cancer and adjacent paired normal tissues, with an increased breadth and quantity of lincRNAs during cancer progression. Consistent with previous findings, lincRNAs were typically coexpressed with their neighboring genes, albeit to an extent similar to that of pairs of neighboring protein-coding genes. By applying a mathematical model based on a mathematical model based on a mathematical model, we identified a set of gene expression, we distinguished an additional subset of lincRNAs termed "regulatory lincRNAs," representing their dominant roles in gene regulation. The number of regulatory lincRNAs was significantly higher in cancerous compared to normal tissues, with greater enrichment of target protein-coding genes associated with cell growth and differentiation processes observed in cancer tissues. Moreover, dozens of lincRNAs switched from non-regulators in normal tissues to regulators in tumor tissues. Our integrated analysis reveals enhanced regulatory effects of lincRNAs in lung adenocarcinoma and provides a resource for the study of regulatory lincRNAs that facilitate tumorigenesis and may serve as new targets for clinical therapy.
A Neuroepigenomic Model of the Fetal Alcohol Exposure Spectrum. 

Maternal alcohol consumption during pregnancy causes a continuum of heterogeneous disorders termed Fetal Alcohol Spectrum Disorders (FASD). Patients affected with FASD show life-long defects, particularly affecting the central nervous system and its complex traits. Our group has developed an animal model using C57BL/6J (B6) mice and genome-wide molecular technologies. The results show that exposure of alcohol during neurodevelopment in B6 mice causes behavioral disabilities matching FASD patients in resulting offspring. Further, the resulting mice show changes in brain gene expression as well as epigenetic marks, including DNA methylation, multiple histone modifications, and ncRNA expression. Interestingly, the genes affected by the epigenetic marks last a lifetime. Also, the genes affected participate in the critical neural processes of apoptosis, neurodevelopment, cellular identity, cell-cell interaction, and signalling. Using genome-wide arrays to interrogate the cytosine methylation and ncRNA expression within the whole brain and hippocampus of adult mice prenatally exposed to alcohol, we have identified an epigenomic footprint that is characterized by alterations to imprinted regions of the genomes, controlled by DNA methylation, that encode multiple developmentally important non-coding RNAs (ncRNAs) and are regulated by CCCTC-binding factor (CTCF), a zinc finger protein. These processes are developmentally integral to the outcome of larger scale cortical brain structure formation, since the events of both pre- and post-natal neurodevelopment are highly dependent on the epigenotype as well as the experience and environment of the differentiating cells at the molecular level. Taken together, our models suggest that ethanol can create significant long-term changes to the molecular mechanisms that create and maintain an epigenetic landscape that is essential to normal brain function and future neurodevelopment. The proposed model suggests a potential for effective attenuation of disease endophenotypes, given the plastic nature of the epigenome in response to enriched postnatal environments.
511S
Allele-aware ChIP-seq alignments identify allelic differences in transcription factor binding at disease-associated loci. M.L. Buchkovich1, K.E. Eklund1, Q. Duan1, L. y1,2,3, K.L. M1, T.S. Furey1,4,5. 1) Department of Genetics, University of North Carolina, Chapel Hill, NC; 2) Department of Biostatistics, University of North Carolina, Chapel Hill, NC; 3) Department of Computer Science, University of North Carolina, Chapel Hill, NC; 4) Department of Biology, University of North Carolina, Chapel Hill, NC.
Allelic differences in protein-DNA binding can influence gene transcription at disease-associated loci. These differences can be detected as allelic imbalance, or enrichment of one allele, in ChIP-seq data. Allele-aware alignments that use genotyping data during mapping more accurately align sequence reads to heterozygous sites than traditional, genotype-independent alignments, and are important for detecting imbalances. We used an allele-aware aligner, GSNAP, and heterozygous sites from complete (sequencing-based), partial (HumanM2Duo array + imputation), or no sample-specific genotype data to map CREB1 ChIP-seq reads from GM12878, a human lymphoblastoid cell line. In the case of no genotype data, common variants (1000 Genomes; EUR MAF>0.05) were considered heterozygous sites. Both CREB1, a transcription-activating protein, and B-lymphocytes are known to play a role in inflammation. Using complete genotype data and requiring a minimum of 5 sequences containing each allele, we detected 204 sites of allelic imbalance (binomial p<0.01). We correctly identified 161 and 146 of these imbalances using partial and no genotypes, respectively, in addition to 52 and 47 imbalances not detected using complete genotypes. Two sites of allelic imbalance, rs2382818 and rs713875, are in linkage disequilibrium (1000 Genomes EUR: r2>0.8) with inflammatory bowel disease (IBD)-associated variants, and are predicted to be strong enhancers of disease. The imbalance at rs2382818 was only detected in allele-aware alignments (chromHMM) in GM12878 cells. While the imbalance at rs713875 was detected in the genotype-independent and all three allele-aware alignments, the imbalance at rs2382818 was only detected in allele-aware alignments using complete or partial genotypes. We performed electrophoretic mobility shift assays for these two variants with purified CREB1 protein and observed allelic differences in CREB1 binding of both variants in the directions consistent with the predicted imbalances. Together, these data suggest that differential binding of CREB1 to alleles of these variants may influence IBD by disrupting the transcription of one or more genes. This study highlights our ability to create allele-aware alignments in the absence of completely genotyped samples and the utility of these alignments in identifying sites with allelic differences in protein binding at disease-associated loci.

512M
Functional genomics approaches reveal novel role of a key hematopoietic transcription factor in blood disorders. M. Bysnka-Bishop1, D. Vandeommen1, A.E. Campbell2, PR. Arc2, Y. Yao2, CA. Keller2, BM. Giardini3, P. Gadue2, FF. Costa4, RL. Nemiroff5, GA. Blobel6, DL. French7, RC. Hardison8, MJ. Weiss9, ST. Chour10. 1) The Pennsylvania State University, Center for Comparative Genomics and Bioinformatics, Department of Biochemistry and Molecular Biology, University Park, PA; 2) The Children’s Hospital of Philadelphia, Division of Hematology, Philadelphia, PA; 3) The Children’s Hospital of Philadelphia, Department of Pathology and Laboratory Medicine, Philadelphia, PA; 4) University of Campinas, Department of Internal Medicine, School of Medicine, Campinas, Brazil; 5) The University of Pennsylvania School of Medicine, Department of Obstetrics and Gynecology, Philadelphia, PA.
GATA1 (MIM #305371) is a hematopoietic transcription factor important for the development of erythroid and megakaryocytic lineages. Mutations in the human hematopoietic cell expresses a full-length form of GATA1 and a short form, GATA1s, that lacks the N-terminal transactivation domain (NAD). Germline GATA1 mutations that result in exclusive production of GATA1s (referred to as “GATA1s mutations”) were identified in patients with congenital hypoplastic anemia and neutropenia (MIM #300835). In individuals with Down syndrome (MIM #190685), somatic GATA1s mutations lead to transient myeloproliferative disease (TMD) (MIM #190685) and acute megakaryocytic leukemia (AMKL) (MIM #190685). We investigated the function of the human GATA1s by creating and analyzing induced pluripotent stem cells (iPScs) from somatic tissues of patients with Down syndrome-associated TMD, as well as congenital macrocytic anemia due to a germline GATA1s mutation. Microarrays, transcriptome profiles of iPScs, and protein profiling of cell lines that express upregulated by GATA1s were enriched for megakaryocytic and myeloid genes, while the downregulated genes were enriched for erythroid genes. Furthermore, compared to controls, erythropoiesis by GATA1s iPScs was markedly reduced, while megakaryopoiesis and myelopoiesis were enhanced. Single cell expression profiling of 91 selected hematopoietic genes revealed small, but significant differences in gene expression distributions for 40 genes between wtGATA1 and GATA1s cells. This suggests that small changes in expression of many genes contribute to the GATA1s-mediated change in lineage bias from erythroid to myelo-megakaryocytic lineage fate. To provide a mechanistic explanation of the observed phenotype, we performed ChIP-seq experiments to study differences in binding patterns between full length GATA1 and GATA1s. ChIP-seq results indicate that GATA1s binding is specifically impaired at erythroid target genes, but normal at megakaryocytic genes, implicating the N-terminus as a chromatin occupancy factor selective for erythroid cells.

513T
Identifying rare, non-coding DNA variants in Systemic Lupus Erythematosus. S.J. White1, S. Cantili1, F.J. Rossello2, D. Bellucchio3, E.F. Morand4. 1) MIM-PHI Institute of Medical Research, Monash University, Clayton, Victoria, Australia; 2) Victorian Bioinformatics Centre, Monash University, Clayton, Victoria, Australia; 3) Agilent Technologies, Mulgrave, Victoria, Australia; 4) Southern Clinical School, Monash University, Clayton, Victoria, Australia.
Systemic lupus erythematosus (SLE) is a debilitating multisystem autoimmune disease. Patients with SLE suffer a marked loss of life expectancy, and severe morbidity, due to autoimmunemediated inflammation of multiple organs. A classic hallmark of SLE is the presence of autoantibodies. This is a clear biological link between SLE and antibody-producing B-lymphocytes (B-cells), strengthened by the clinical efficacy of anti-B cell therapies in certain patients. Genetic studies have associated SLE susceptibility with a number of Single Nucleotide Polymorphisms (SNPs), many of which are found in or near genes with B-cell function. However, the known SLE SNPs together only account for a small percentage of the genetic contribution to SLE, and as most SNPs are located in non-coding DNA, the mechanism of their effect is often unclear. Several SLE-associated SNPs are located in such DNA elements that control gene regulation, with sequence changes disrupting the binding of specific transcription factors. In 2012 the ENCODE consortium published papers outlining genome-wide studies of open chromatin in a range of cell types. These studies demonstrated that DNasel hypersensitive sites (DHS) strongly correlate with known and predicted regulatory sequences. In addition, they are enriched for known and predicted transcription factor binding sites (TFBS). This provides a clear link between sequence variants and functional activity, and rare variants in B-cell regulatory loci (defined by open chromatin in B-cells) would explain genetic susceptibility to SLE in a way not detectable by GWAS. We have designed a custom capture approach using an Agilent SureSelectXT2 Custom Library to screen 49K B-cell DHS in genomic DNA from SLE cases. This covers ~17Mb of genomic sequence, with up to 16 samples being pooled in a single lane on the Illumina HiSeq 2500. We are currently analysing this unique sequence data from a cohort of SLE cases, and plan to perform functional validation using reporter constructs and changes in DNasel-sensitivity.
514S

Coordinated Regulatory Variation Associated with Maternal Glycemic Traits Regulates HKDC1 Expression. T.E. Reddy, C. Guo, A.E. Ludvik, M.G. Hayes, L.L. Armstrong, D.M. Schollens, C.D. Brown, B.T. Layden, W.L. Lowe. 1) Institute for Genome Science & Policy, Duke University Medical School, Durham, NC; 2) Department of Biostatistics & Bioinformatics, Duke University Medical School, Durham NC; 3) Division of Endocrinology, Metabolism and Molecular Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; 4) Department of Preventive Medicine, Division of Biostatistics, Northwestern University Feinberg School of Medicine, Chicago, IL; 5) Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 6) Jesse Brown Veterans Affairs Medical Center, Chicago, IL.

Genome wide association studies (GWAS) to understand the genetics of complex human traits strongly suggest that non-coding genetic variation is a major contributor to disease risk. One hypothesis is that variation in gene regulation may contribute to altered expression of disease-relevant genes. As one example, in our previous GWAS, we identified SNPs in 10q22 near the putative hexokinase gene HKDC1 that were associated with two hour glucose levels in an oral glucose tolerance test administered at ~28 weeks gestation during pregnancy. None of the associated SNPs were within the protein-coding regions of HKDC1, suggesting that the SNPs instead could have a regulatory effect on HKDC1. To address that hypothesis, we have comprehensively mapped regulatory variation in the 10q22 region associated with maternal glucose levels. We used open chromatin data from the ENCODE project and the Roadmap Epigenome Consortium to define regulatory elements, and measured regulatory function of each element across four different human ancestry groups using reporter assays constructed from DNA from the 1,000 Genomes Project. To disrupt linkage within regulatory elements and to better assay rare and personal variants, we complemented the set of natural haplotypes using site-directed mutagenesis. With this approach, we saturated the common and rare variation in regulatory elements across this gestational hyperglycemia locus. These assays revealed a surprising amount of regulatory variation controlling expression of HKDC1. The effects of regulatory variants were strongly and significantly coordinated across risk haplotypes. Furthermore, when cloned into a luciferase reporter assay, multiple regulatory elements acted independently. These results suggest that regulatory variants with a coordinated effect across a large haplotype contribute to a reduction in HKDC1 expression that is greater than would be expected for a single causal regulatory variant. We confirmed the reporter assay results using ChIP-seq analysis and observed extensive monoallelic expression of genes in all autosomes in this study. Together, these results provide evidence for a novel model of disease in which coordinated regulatory effects across a risk haplotype have a magnified effect on endogenous gene expression, leading to subtle but significant effects on phenotype.

515M


Lipid and lipoprotein levels influence risk of cardiovascular disease; however, studies have traditionally focused on the assessment of plasma cholesterol and triglyceride levels. The plasma lipidome consists of many thousands of lipid metabolites, with large structural diversity. As part of the San Antonio Family Heart Study, lipidomic profiles (quantification of 319 lipid species) were generated from 1,240 Mexican American individuals from 30 extended families using mass spectrometry. Individual lipid measures were correlated with gene expression profiles for 20,413 genes from PBMCs, and promoter variants (cis-eQTL) were identified that were strongly associated with both expression and individual lipid profiles. Three variants were further characterized. Variants rs88736 within 1000 bp of the TSS 5' of the TMC4 gene and rs887316 in the proximal promoter of the RAD9α gene, were associated with a phosphatidylinositol lipid levels (p = 1.0×10−4 and 4.2×10−4). The rs606458 variant in the SFI1 gene promoter was associated with levels of a phosphatidylinosine levels (p = 1.7×10−6). We tested the two alleles of each variant in luciferase reporter assays introducing promoter-containing plasmids in K562 cells. All variants resulted in differential activity in the reporter assay, suggesting altered gene expression for each allele. However, ChIP-seq data from ENCODE did not reveal the likely DNA-binding protein mediating the observed effect. Therefore, we adapted our novel methodology, Hybridization Capture of Chromatin-Associated Proteins for Proteomics (HyCCAPP) to isolate the desired promoter regions from crosslinked cells using hybridization capture with complementary oligonucleotide probes, as initially developed for the analysis of genomic regions in yeast. The methodology was developed to isolate individual chromatin regions after protein-DNA crosslinking, and use mass spectrometry to characterize and quantify the bound proteins. In our adaptation of this method, a promoter was captured with over 3% efficiency, and the isolated material was amenable to proteomic analysis. HyCCAPP now allows the effective characterization of DNA-binding proteins for promoter variants that show differential activity in our reporter assays, and will help elucidate the underlying mechanisms by which the variants identified in our cohort modulate gene expression and plasma lipid levels.

516T


Trisomy 21 is a model disorder of altered gene expression. Several studies have addressed the transcriptome differences between normal and affected individuals; however all of these studies suffer from the presence of “noise” due to gene expression variation among different individuals. We have previously used a pair of monozygotic twins discordant for trisomy 21 in order to study the global dysregulation of gene expression ( Nature: 508: 7496,2014 ). The majority of previous studies focused on aneuploidies were conducted on cultured cell populations or tissues, but studies focusing on gene and allelic expression behaviour at the single cell level are lacking. In this study we explore the allele specific expression in Trisomy 21 using transcriptome studies in single cells. We have used 40 normal cells and 48 trisomic cells from the fibroblasts of the monozygotic twins discordant for trisomy 21 and compared the ASE (allele specific expression), and their transcriptional metrics in these two cell groups. Remarkably we have observed extensive monoallelic expression of genes in all autosomes including on chromosome 21. In addition a series of samples from mosaic trisomy 21, trisomy 13 and trisomy 18 are in different stages of investigation. These studies in single cells will provide a fundamental understanding of the gene expression dysregulation and allele specific expression in aneuploidies and may contribute to the understanding of the phenotypic heterogeneity of these syndromes.
517S
Analysis of long-range interactions in primary human cells reveals CFTR new regulatory elements. S. Moisan1,2, S. Berlivet3,4, J. Dosseh3, C. Férec5, J.-S. Masquelier1, INSEMM U1076 1Laboratoire de Génétique, Génomique Fonctionnelle et Biotechnologies, Brest, France; 2) Université de Bretagne Occidentale, Faculté de Médecine et des Sciences de la Santé, Brest, France; 3) CHU Brest, Hôpital Morvan, Laboratoire de Génétique Moléculaire, Brest, France; 4) Unité de Génétique et Développement, Institut de Recherches Cliniques de Montréal, Québec, Canada; 5) I4178 Department of Biochemistry and Goodman Cancer Research Center, McGill University, Montréal, Québec, Canada.

The cystic fibrosis transmembrane conductance regulator (CFTR) gene was identified in 1989. Twenty five years later, the regulatory mechanisms controlling its complex expression are still not fully understood. Although, 1970 mutations have been identified, many cases of cystic fibrosis or CFTR Related Disorders remain in still of unknown genetic cause. The promoter, which binds transcription factors and drives some aspects of CFTR expression, cannot alone account for tissue specific control. This implicates distinct cis- or trans-acting elements in cell-type-specific regulation of CFTR expression. The aim of our project is to study long-range regulatory mechanisms of the CFTR gene. Interestingly, the majority of the human genome is composed of non coding DNA whose function has not been thoroughly investigated. A significant number of conserved non-coding sequences (CNCs) are found in gene-poor regions, these large intergenic regions must have kept a function throughout evolution. We first developed the Chromosome Conformation Captures (3C) approach to map these potential regulatory elements which could interact specifically with the CFTR gene by three-dimensional folding mechanism. Subsequently, we enhanced our analyses with a high throughput adaptation of 3C: the 3C-resolution Captures (3C) technology. This approach allows the analysis of millions chromatin interactions. Thus, we have analyzed the spatial organization of a ~790kb region, comprising the CFTR gene, with 145 SC primers. Interactions between this locus and the promoter of another gene, UBE2L3, which is the ubiquitin-conjugating enzyme, can be monitored with the Ion PGM™, in primary epithelial cells, which express the gene and primary skin fibroblasts, which do not express the gene. We compare chromatin conformation in order to identify uncharacterized regulatory elements that act specifically in CFTR-expression. Our approach is validated by the identification of previously characterized regulatory elements. Moreover, we identify novel chromatin contacts of the CFTR promoter with chromosomal regions, which could potentially be involved in CFTR gene expression regulation. Thanks to 3C and 3C-derived analyses, we could identify new possible mutations far from the gene, which may lead to its dysfunction by modifying the chromatin conformation. These analyses will be pursued on patients affected by cystic fibrosis or by CFTR Related Disorders, in whom either a single mutation or none was found in the CFTR coding sequence.

518M

Cis-regulatory elements (CREs) are found in the non-coding DNA of the genome and act as binding sites for transcription factors and drives some aspects of gene expression. In vertebrates, CREs can be either upstream or downstream of the transcription start site. CREs are found in the promoter, intron, or in the first intron of the gene. CREs can be found in regions that are not actively transcribed, and they are more common in gene-poor regions, these large intergenic regions must have kept a function throughout evolution. We first developed the Chromosome Conformation Captures (3C) approach to map these potential regulatory elements which could interact specifically with the CFTR gene by three-dimensional folding mechanism. Subsequently, we enhanced our analyses with a high throughput adaptation of 3C: the 3C-resolution Captures (3C) technology. This approach allows the analysis of millions chromatin interactions. Thus, we have analyzed the spatial organization of a ~790kb region, comprising the CFTR gene, with 145 SC primers. Interactions between this locus and the promoter of another gene, UBE2L3, which is the ubiquitin-conjugating enzyme, can be monitored with the Ion PGM™, in primary epithelial cells, which express the gene and primary skin fibroblasts, which do not express the gene. We compare chromatin conformation in order to identify uncharacterized regulatory elements that act specifically in CFTR-expression. Our approach is validated by the identification of previously characterized regulatory elements. Moreover, we identify novel chromatin contacts of the CFTR promoter with chromosomal regions, which could potentially be involved in CFTR gene expression regulation. Thanks to 3C and 3C-derived analyses, we could identify new possible mutations far from the gene, which may lead to its dysfunction by modifying the chromatin conformation. These analyses will be pursued on patients affected by cystic fibrosis or by CFTR Related Disorders, in whom either a single mutation or none was found in the CFTR coding sequence.

519T
Systemic Lupus Erythematosus-associated functional variants influence the gene expression of UBE2L3 through the regulation of promoter and enhancer activities. S. Wang1, B. Ho2, P. Gaffney1,2. 1) Arthritis and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, OK; 2) Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK.

Genetic variants in the UBE2L3 region have been associated with multiple autoimmune diseases including Systemic Lupus Erythematosus (SLE), Crohn’s disease, celiac disease, and rheumatoid arthritis. We previously reported an SLE-associated functional risk haplotype in the UBE2L3 region that led to significant increases in both mRNA expression of UBE2L3 and protein expression of UBE2L7. However, the mechanism by which the risk haplotype influences UBE2L3 expression has not yet been described. In this report, we employed a combination of approaches including EMSA, Luciferase reporter assay, Western blotting, and RT-qPCR to evaluate the functional potential of SLE-associated variants in the UBE2L3 region and to characterize the mechanism by which these functional variants influence UBE2L3 gene expression. Through the use of EMSA, we observed functional evidence for two variants in and around the UBE2L3 promoter region. A functional variant (rs140490), located150 bps upstream of the UBE2L3 transcription start site increases nuclear protein binding (P=0.0031). Luciferase reporter assays demonstrated that the risk allele of this functional variant increases UBE2L3 promoter activity (P=0.0038) in EBV-transformed B cells. In contrast, another functional variant (rs2266959), located in the first intron of UBE2L3, reduces nuclear protein binding (P=0.0014). Luciferase reporter activity assays demonstrated that rs2266959 falls within an enhancer element, the risk allele of which leads to a significant increase in the enhancer activity (P=0.018). The variant rs140490 increases binding of nuclear proteins to the promoter, while the variant rs2266959 reduces binding of nuclear proteins to the enhancer. However, both functional variants lead to significantly increased transcriptional activity of the promoter and the enhancer, which implies that both variants may be responsible for the gene expression phenotype produced by the risk haplotype of UBE2L3. Further mechanistic characterization of the functional variants of UBE2L3 is required to understand the role of this risk haplotype in autoimmune disease pathogenesis.
520S
Comparative genomics and abstraction of DNA sequences allow the identification of cis-regulatory signatures. P.EM. Guimaraes1, M.A. Almeida2, T.J.P. Sobreira3, H.A. Castillo4, J.X. Neto5, P.S.L. Oliveira1, 1 Laboratório de Evolução de Genes e Genomas, INEGB, UFU, Patos de Minas, MG, Brazil; 2 AT&T GCC, Department of Genetics, Texas Biomedical Research Institute, San Antonio TX USA; 3 Laboratório de Bioinformática, LNBio, CNPEN, Campinas, SP, Brazil; 4 Laboratório de Modificação do Genoma, LNBio, CNPEN, Campinas, SP, Brazil.

Gene expression fine regulation is achieved by the complex interaction between transcription factors (TF) and their respective binding sites (TFBS). Identifying these short DNA motifs is a major current goal in modern molecular biology. However, the efficiency of current prediction algorithms is limited due to the sequence variation between TFBS and the complexity of transcription regulatory networks. As an endeavor to modify this scenario, here we present and validate a novel computational pipeline to improve the prediction of functional TFBS. This pipeline uses a "phylogenetic footprint" filter and converts predicted TFBS to an alphabet, which describes evolutionary-functionally related families of TF. The "phylogenetic footprint" works as a selection step for orthogonal sequences. Further, TFBS were predicted in syntenic blocks, for each species evaluated, and converted into strings of families of functional related transcription factors. The family clustering reduces TFBS variability and prediction redundancy. Finally, only TFBS families predicted at the same gene and relative position and present in the set of selected species were considered. We carried our routine in order to identify functional cis-regulatory elements acting in aldha2 gene, searching for syntenic sequences in a 30 kb window flanking the coding sequence of the mouse gene, including all intronic sequences, comprising a total of 138,191 bases. Orthologous segments were identified, identifying gene family in rat and chicken genomes. A total of 1,318 predicted TFBS families were considered constrained by our "spatial and evolutionary filter", reducing in 99.6% the original universe of 382,750 predicted elements. Despite the reduction in TFBS, it could one access side effect: the lost of functional binding sites. To evaluate the efficiency of our routine we compared the final set of predicted TFBS to those previously experimentally identified as real regulators of aldha2 gene. All validated TFBS were represented in our set, indicating that we have the considerable reduction of predicted TFBS we still clamped the real ones. Taken together, our data suggest that the routine can be applied to high-throughput analyses and help to draw more focused experimental validation protocols. The same strategy can be applied to identification of mRNA seed pairing, splicing enhancer elements and others regulatory signatures and structural elements.

521M
Detecting gene-by-environment interactions using allele specific expression. D.A. Knowles1, S.B. Montgomery3, A. Battle2. 1 Stanford University, Stanford, CA; 2 Johns Hopkins University, Baltimore, MD.

The impact of environment and lifestyle on human health is dramatic, with major risk factors including substance use, pollution, diet and exercise. However, the interaction between environment and individual genetic background is poorly understood; detecting gene-by-environment interactions is both statistically and computationally challenging. Using RNA sequencing of primary tissue (whole blood) from 922 individuals in the DGN cohort, and help to draw more focused experimental validation protocols. The same strategy can be applied to identification of mRNA seed pairing, splicing enhancer elements and others regulatory signatures and structural elements.

522T

RNA is the first intermediary between genetic variants and their downstream effects on human disease. Recent work has shown that SNPs altering gene expression as opposed to coding SNPs may be the major drivers of complex phenotypic variation. Therefore, characterization of how regulatory SNPs (eQTLs) change between tissues, environments, and populations is central to understanding of the genetic basis of human disease. In this work we leverage a unique data set of expression from 19,360 genes collected in CD4+ T cells and CD14+ monocytes, from 183 Europeans (Eur), 91 African Americans (AfAm), and 74 Asians (Asn) genotyped at 700K SNPs. We also profiled 415 genes in CD14+/CD16- dendritic cells after exposure to LPS, influenza, and LPS. We estimate components of heritability for genotyped SNPs within and between contexts across all genes; determining cis-genetic heritability for each pair of populations for each gene. We estimate the h^2^ cис in (Eur,AfAm,Asn) at (0.03,0.041,0.031) for CD4s, and (0.048,0.053,0.045) for CD14s (e.s. < 0.001), observing higher h^2^ cис for CD14s than CD4s and an ordering of h^2^ cис across populations with the same gene relative position and present in the set of selected species (AfAm > Eur > Asn). The cross cell type genetic correlations ρ g, pop = (0.353, 0.251,0.244), implying different sets of eQTLs between cell types. Upon exposure to LPS, flu and IFNγ, h^2^ cис estimates for Eur CD14 increased by 0.06, 0.053, and 0.08, with similar patterns in AfAm and Asn. In contrast to the cross population, the cross exposure correlations were dramatically higher with ρ g, exp = (0.93,0.542,0.779). The higher value of ρ g, exp than ρ g, cis coupled with the increases in h^2^ cис are consistent with a model where upon exposure, the expression of a gene expression with this effect. In CD4-CD14, across populations we observe ρ g, pop(Eur,AfAm)=0.56, ρ g, pop(Eur,Asn)= 0.52, and ρ g, cis(AfAm,Asn)=0.40. The cross population genetic correlations are higher than cross tissue and lower than cross exposure correlations ρ g, exp. This can be explained by population genetic distance. In all, our results suggest cross tissue analysis will uncover distinct regulatory regions while stimulation will increase the power to detect eQTLs.

523S

Human cytomegalovirus (HCMV), is a clinically important and ubiquitous herpesvirus that infects humans worldwide. HCMV is a DNA virus which replicates in epithelial cells and neurons. The virus infects individuals early in life and persists throughout life in the host, as it is essential for viral growth and might serve as a potential target for combat HCMV infection. RNase P complexed with external guide sequence (EGS) represents a novel nucleic-acid-based gene interference approach to modulate gene expression and is used in this study to explore a potential therapeutic approach for inhibiting UL49 gene expression and replication of HCMV. A functional EGS RNA was constructed to target the UL49 mRNA. The EGS RNA was shown to be able to direct human RNase P to cleave the UL49 mRNA, qualitatively and quantitatively, in vivo. A reduction of 68% in the mRNA and protein expression levels and a reduction of 1000-fold in viral growth were observed in human cells that expressed the functional EGS. Further studies of the function of UL49 gene indicate that the reduction of UL49 gene expression has no effect on HCMV gene replication and gene expression of IE2 and qB. However, it leads to the down-regulation of the gene expression UL99, which is highly immunogenic and play roles in the assembly and egress of virus particles. This observation has provided us a valuable new clue to understand the role of UL49 gene during HCMV lytic infections.
524M Deep genetic, transcriptomic and epigenetic maps of human blood elements. L. Chen1,2, S. Busche3, A. Datta4, O. Delleaneau1, K. Downes1, S. Ecker1, H. Ketelsen1, M. Kostadima2, D.S. Paulus5, OBIO-BLUEPRINT CONSORTIUM6, 1) Human Genetics, Wellcome Trust Sanger Institute, Cambridge, Cambridge, United Kingdom; 2) Department of Haematology, University of Cambridge, NHS Blood and Transplant, Cambridge Biomedical Centre, Cambridge, Cambridge, United Kingdom; 3) McGill Epigenome Mapping Centre, McGill University, Montreal, Quebec, Canada; 4) European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom; 5) Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands; 6) UCL Cancer Institute, University College London, London, United Kingdom; 7) http://www.blueprint-epigenome.eu/.

Haematopoiesis, the formation of blood elements, has been studied as a paradigm of stem cell biology and development. The BLUEPRINT project (http://www.blueprint-epigenome.eu/) is generating deep genetic, transcriptomic and epigenetic maps of over 30 human blood cell types and their malignant counterparts that will be available to the scientific community. A first stream of research has focused on cataloguing the transcriptional and alternative splicing changes in rare blood progenitor cells. We performed high-throughput RNA transcriptome analysis of eight highly-purified cell populations (HSC, hematopoietic stem cells; MPP, multipotent progenitor cell; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte macrophage progenitor; MEP, megakaryocyte erythrocyte progenitor; EB, erythroblasts and MK, megakaryocytes). We identified 17,881 novel splice junctions that were specific to these progenitor cell populations; (2) 20,924 high-confidence alternative-splicing junctions, where approximately 17% were cell-type specific and assigned categories of genes such as RNA splicing and transcription factors; (3) two differentially spliced transcription factors that were important in megakaryopoiesis by using in vitro experiments. These results suggest that alternative splicing is a crucial post-transcriptional regulatory mechanism in hematopoietic cell function. A second stream of research has focused on the effect of common sequence variation on the epigenetic landscape using detailed and integrated analyses. In three peripheral blood cell types, monocytes, neutrophils and naïve CD4 T cells, we have generated data for whole-genome methylation and RNA-seq, ChiP-seq of two histone modifications and DNA methylation by 450K arrays in an initial set of 48 healthy individuals with European-ancestry. Here we describe the first epigenetic quantitative trait loci (QTLD) analyses of this integrated data set to further understand the effects of genotype-phenotype correlations via epigenetic mechanisms in these immune cells.

525T Analysis of monoallelic expression in human individual cells revealed novel imprinting loci. J. Santon1, OBO. BLUEPRINT CONSORTIUM2, 1) Human Genetics, Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Department of Haematology, University of Cambridge, NHS Blood and Transplant, Cambridge Biomedical Centre, Cambridge, Cambridge, United Kingdom; 3) McGill Epigenome Mapping Centre, McGill University, Montreal, Quebec, Canada; 4) European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom; 5) McGill University, Montreal, Quebec, Canada; 6) European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom; 7) McGill Epigenome Mapping Centre, McGill University, Montreal, Quebec, Canada; 8) Structural and Computational Biology Group, Spanish National Cancer Research Center (CNIO), Madrid, Spain; 7) Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands; 8) UCL Cancer Institute, University College London, London, United Kingdom; 9) http://www.blueprint-epigenome.eu/.

The aim of our study is to identify novel imprinting loci in human hematopoietic cells. Here we describe the first epigenetic quantitative trait loci (QTL) analyses of this integrated data set to further understand the effects of genotype-phenotype correlations via epigenetic mechanisms in these immune cells.

526S Somatic Rhodopsin transcriptional repression by artificial DNA-binding proteins targeted to cis-regulatory elements. S. Bott1, E. Marrocco1, N. de Frisco1, F. Curioni1, M. Sofia1, M.L. Baccò1, S. Rossi1, F. Simonelli1, E.M. Surace1, 1) Tigern (Telethon Institute of Genetics and medicine), Napoli, Italy; 2) Department of Ophthalmology, Seconda Università degli Studi di Napoli, Napoli, Italy; 3) Dipartimento di Scienze Mediche Veterinaria, Università di Bologna, Bologna, Italy; 4) Dept. of Translational Medicine, “Federico II” University, Napoli, Italy.

Artificial DNA-binding proteins are becoming a novel and versatile tool in development therapeutics. The ZF protein (ZFPs) modular structure enable both the specific assembly of multiple ZFPs to generate a DNA binding domain (DBD) with different target specificities and the use of various effector domains including transcriptional regulation domain or nucleases for either genomic regulation or editing. We showed that a selected artificial transcriptional repressor (Zinc-Finger & repressor, ZF6-FR) delivered by AAV2-8 vector down regulates selectively mutated human RHO expression levels, ameliorating the retinal phenotype in a mouse model of autosomal dominant Retinitis Pigmentosa (adRP). We speculate that a protein, which contains exclusively the DBD (ZF6-DBD) domain without the effector domain (KRAB), may confer a higher therapeutic efficiency and tolerability genome-wide (off target) than ZF6-KRAB. Therefore, we removed from the construct the KRAB domain. The adRP mouse model eyes treated with AAV8-CMV-ZF6-DBD demonstrate robust and consistent recovery of the ERG a-wave and b-wave responses along a wide range of luminance in both scotopic and photopic conditions when compared to EGFP treated eyes. In addition, ZF6-DBD treated eyes show significant higher responses compared to AAV8-CMV-ZF6-KRAB. Based on sequence identity between the human and the porcine RHO genes, we generated a CRISPR-Cas9 mediated gene editing construct in the porcine retina. To assess "pure" transcriptional effects and not possible secondary degeneration owed to rhodopsin knock-down, we performed an early sacrifice. Fifteen days after vector administration the retina regions transduced were harvested and RT-PCR studies were performed. Transcriptional levels of Rho showed that retinas that received ZF6-DBD demonstrate more than 40%; of transcriptional repression of Rho. Furthermore, next generation sequencing (NGS, RNA-seq) was used to assess the differential transcriptional profiling of AAV8-CMV-ZF6 and AAV8-CMV-ZF6-KRAB treated eyes compared to controls. The dataset generated show specific down-regulation of Rho expression. Ongoing experiments are showing remarkably similar results with Transcription activator-like effector (TALE) technology. In conclusion these data support the use of artificial DNA-binding proteins as therapeutics.

527M Epigenetic promoter silencing in Friedreich ataxia is dependent on the length of the GAA triplet repeat mutation. Y. Chutake1, C. Lam1, W. Costello1, M. Anderson1, S. Bidichandani1, 1) Department of Pediatrics, Univ Oklahoma HSC, Oklahoma City, OK; 2) Department of Biochemistry, Univ Oklahoma HSC, Oklahoma City, OK; 3) Department of Biostatistics & Epidemiology, Univ Oklahoma HSC, Oklahoma City, OK; 4) Dept. of Translational Medicine, “Federico II” University, Napoli, Italy.

Friedreich ataxia (FRDA) is an autosomal recessive disease caused by an expanded GAA triplet-repeat (GAA-TR) mutation in intron 1 of the FXN gene. Most patients are homozygous for expanded alleles containing 100–200 repeats with a consensus repeat length of 120 repeats. Both transcriptional deficiency and phenotypic severity are significantly correlated with the length of the shorter of the two expanded alleles. We show that repressive chromatin extends from the expanded GAA-TR in intron 1 to the transcriptional start site (FXN-TSS). Using a chromatin accessibility assay and a high-resolution nucleosome occupancy assay, we found that the FXN-TSS, which is normally in a nucleosome-depleted region is rendered inaccessible by altered nucleosome positioning in FRDA. Consistent with the altered epigenetic landscape, the FXN gene promoter was found to be in a transcriptionally non-permissive state in FRDA. Both metabolic labeling of nascent transcripts and an unbiased whole transcriptome analysis revealed a severe deficiency of transcriptional initiation in FRDA. FXN transcriptional deficiency, measured by qRT-PCR in patient cell lines bearing a range of expanded alleles (200–1122), was significantly correlated with the length of the shorter of the two expanded alleles. Importantly, this was noted both upstream (R²=0.84, p=0.014) and downstream (R²=0.89, p=0.002) of the expanded GAA-TR repeat length. Correlation coefficients are significantly related to repeat length. A bilinear regression model revealed that length-dependence =0.84; p=0.014) and downstream (R²=0.89, p=0.002) of the expanded GAA-TR repeat length. Correlation coefficients are significantly related to repeat length. A bilinear regression model revealed that length-dependence was significantly weaker when the shorter of the two expanded alleles contained <400 triplets. Direct measurement of FXN promoter function in patients with expanded alleles in vitro, using a 4×400 bp promoter construct, demonstrated that the shorter of the two expanded alleles revealed a significantly greater deficiency in individuals with longer GAA-TR alleles. Thus, deficient transcriptional initiation caused by epigenetic silencing of the FXN promoter is a major cause of FXN transcriptional deficiency in FRDA, and it is dependent on the length of the expanded GAA-TR mutation.
528T
Characterizing the predictive features of regulatory genetic variants. N.M. Ioannidis1, N. Larson2, Y. Zhang2, A. Battie1, S. Montgomery1, S.N. Thibodeau1, W. Sieh1, A.S. Whittemore1, C.D. Bustamante1, 1) Stanford University School of Medicine, Stanford, CA; 2) Mayo Clinic, Rochester, MN.

Whole-genome and whole-exome sequencing technologies are increasingly enabling studies of genetic variation in large numbers of healthy and diseased individuals; however, interpreting the clinical significance of the many genetic variants identified in these studies remains a critical challenge. One mechanism by which genetic variants can achieve clinical significance—even those that do not affect protein structure, such as noncoding, intronic, and synonymous variants—is by regulating gene expression. Here we use a machine learning approach to identify genomic features predictive of genetic variants that are associated with gene expression regulation. We test a large set of features including sequence conservation, overlapping functional elements from ENCODE and Ensembl, and position relative to the nearest transcription start site and splice site. We identify predictive features by training random forest models on known cis-expression quantitative trait loci (cis-eQTLs) in lymphoblastoid cell lines (LCLs). These models achieve an area under the receiver operating characteristic curve (AUC) of between 0.75 and 0.9 depending on the stringency with which cis-eQTLs are defined. We also characterize the tissue-specificity of the genomic features of regulatory genetic variants by comparing the predictive performance of our random forest on cis-eQTLs in LCLs versus cis-eQTLs found in normal prostate tissue from an independent sample of around 500 prostate cancer patients. We anticipate that our results will be useful in future studies of regulatory variation and for prioritizing the likely clinical significance of rare and novel genetic variants identified in large-scale clinical sequencing studies.

529S
Characterization of enhancer gene interactions using DNasel and gene expression data across 110 cell types. P. Kheradpour1, 2, M. Kellis1, 2, Roadmap Epigenomics Consortium. 1) Massachusetts Institute of Technology, Cambridge, MA; 2) Broad Institute, Cambridge, MA.

Recent efforts to characterize diseases through genome-wide association studies and annotate the genome using ChIP-seq experiments have led to a dramatic increase in putative functional genomic regions. Nevertheless, many of the implicated loci have fallen outside coding regions and are thought to be regulatory in nature, efforts to link these regions to their target genes, thereby permitting a better understanding of their importance, has lagged considerably. Generally, experimental linking techniques only permit the interrogation of a small number of specific regions or produce a genome-wide linking at very low resolution.

We utilize DNasel hypersensitivity and expression data from 110 cell types produced by the ENCODE and Roadmap Epigenomics projects to produce linking confidences between hypersensitive regions and nearby genes. We find that high confidence links are supported by independent datasets such as eQTL annotations. Moreover, we find that after controlling for distance several factors, including the orientation of the upstream gene and the presence of an interleaving CTCF, greatly affect the linking confidence. In addition to predicting links between genes and nearby putative enhancers, we are interested in answering fundamental biological questions such as the number of enhancers per gene and where these enhancers fall relative to a gene’s transcription start site. We address these questions through a careful analysis of the distribution of confidence scores for each gene and by performing a scaling analysis on the number of cell types.

530M

Imbalance in allele-specific expression (ASE) of homologous copies of a gene may result from heterozygous SNP cis-acting-eQTLs. Identifying these regulatory variants provides insight into the dynamics of the transcriptome and may highlight genetic variation relevant to complex traits. RNA sequencing is a powerful tool for quantifying ASE by directly capturing mRNA transcript sequence information. It is important to characterize such regulatory variation in a tissue-specific manner, as most publicly available eQTL results are from lymphoblastoid cells and tissue-specific regulatory variation may go undetected. We recently conducted a genome-wide allele-specific eQTL (aseQTL) study on 471 samples of normal prostate tissue (from patients with prostate/bladder cancer) by generating gene expression (RNA-seq) and chromosomally-phased genotype (Illumina Infinium 2.5M) datasets. We additionally imputed common (MAF >1%) exonic SNPs using 1000 Genomes data to increase the proportion of ASE-informative reads, resulting in ~2.7% of all aligned reads being allele-specific. From ~1.6M QC-passed SNPs and 16,607 expressed genes, we evaluated 7,435,266 aseQTL (SNP x gene) associations, restricting our analyses to common autosomal SNPs within 500kb of gene transcription start/stop sites. We tested aseQTL associations by jointly modeling the total and ASE gene read counts, and inconsistent results (e.g., trans-acting or imprinting effects) were excluded. We identified 393,046 significant aseQTLs (FDR = 0.05), with 10,870 genes corresponding to ≥1 putative aseQTL. From the set of most significant aseQTLs per gene, 172 SNPs corresponded to multiple genes. Functional annotation for these SNPs was enriched for ENCODE enhancer histone marks, suggesting evidence of cis co-regulation. The effect size estimates were inversely correlated with MAF (Spearman ρ = −0.38), indicating purifying selection pressure on high-impact regulatory variation. Large-effect aseQTLs were highly concentrated near the TSSs of the regulated genes and tended to co-localize with ENCODE promoter histone marks. Our findings also highlight dynamic expression in genes preferentially expressed in prostate cancer. We find that 110 of the 122 prostate cancer aseQTLs are specific to prostate tissue samples. Further investigation into the prostate-specific gene expression database within our expressed gene set, 81 were significantly associated with aseQTLs. Overall, our results demonstrate widespread and complex cis-acting functional variation in the prostate transcriptome.
Uncovering SMCHD1 regulation of the Protocadherin Cluster. A.G. Mason1, J. Balog1, R.J.L. Lemmers1, R.C. Silsker2, B.T. Heijmans2, R. Taye2, S.J. Tascott3, S.M. van der Maarel3, 1) Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; 3) Netherlands Consortium for Healthy Aging, PO Box 9600, Leiden 2300 RC, The Netherlands; 4) NeuroMuscular Disease Unit, Department of Neurology, University of Rochester Medical Center, Rochester, NY, USA; 5) Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

Faciocapulohumeral muscular dystrophy (FSHD) is a rare disease affecting 1:15,000–20,000 individuals and is clinically characterized by progressive weakness and wasting of the facial and upper extremity muscles. In FSHD individuals, due to the contraction of the D4Z4 macrosatellite repeat array on chromosome 4 (FSHD1), or due to SMCHD1 mutations (FSHD2), the somatic repression of the retrogene DUX4 is incomplete leading to the varied expression of the germline transcription factor in muscle fibers. SMCHD1 encodes a protein that regulates chromatin repression at different loci, and it has been shown to be important for the regulation of the protocadherin (Pcdh) gene cluster in mice. Protocadherins are important for cell adhesion and are predominantly, but not exclusively, expressed in the nervous system. PCDH genes create cell-surface diversity, are monoozalicly expressed in neurons, and are important for neuronal connectivity, self-recognition, and self-avoidance. Studies in SMCHD1 knockout mouse embryos show that the Pcdh cluster is in fact dysregulated. We are interested in studying the role that SMCHD1 has on the PCDH cluster in a neuronal cell context as well as in FSHD muscle cells to fully decipher the normal function of SMCHD1. To investigate the regulation effects that SMCHD1 has on the PCDH cluster, we created a series of knockdown lines in human and mouse neural cells. CRISPR/Cas9 mediated genome editing was performed to delete PCDH genes in vitro and in vivo. We performed ChIP-seq and ChIP-qPCR experiments being done to interrogate the regulation of the PCDH gene cluster in FSHD patient primary muscle cells and, as a reference, in neuronal SK-N-SH cells. Genomic segments identified in each cell line were further characterized using luciferase assays to approximate a peripheral nerve motor neuron. Luciferase assays were performed to identify genomic segments with at least a 5-fold increase in activity over a control vector. This revealed 26 novel genomic segments with enhancer activity in at least one cell line. Of the 32 genomic segments, nine showed a significant difference in luciferase activity between the major and minor alleles. Currently, we are further characterizing a subset of prioritized rSNPs by deleting the genomic segment in cultured cells using CRISPR/Cas9 technology followed by luciferase assays. The identified gene clusters regulated by these alterations are being investigated. Any SNP associated with altered gene expression will represent candidate modifiers of peripheral nerve-related diseases.

Sex-specific Association of Alu Retrotransposons with Gene Expression. S. Linker1, D. Hedges2, 1) HIHG, University of Miami, Miami, FL; 2) The Ohio State University, Columbus, OH.

Transcriptional regulation is a highly coordinated process necessary for the proper function of an organism. Although many mediators of gene expression have been identified (e.g., CpG methylation, promoter accessibility), much of the repetitive content of the genome remains unexamined for its direct association with expression. Retrotransposons (RT) have the potential to alter gene regulation. However, to date there have been limited studies analyzing that association of RTs with alterations in gene expression in humans. RTs are a major cause of variation both between individuals as well as within an individual, due to somatic retrotransposition. This indicates that if RTs are indeed modifiers of gene expression, then they can influence phenotypic variability among as well as within humans. It is therefore essential for disease studies to identify if RTs have an effect on gene expression.

The research presented here examines the association of RTs with differential gene expression. Our results suggest that the presence of a RT within a gene is indeed associated with decreased expression of the same gene. More importantly, this effect of RTs on gene expression was sex-specific, and largely driven by effects in males. Lastly, the genes found to contain variations in RTs are enriched for Autism susceptibility genes, which are also enriched for differences in expression between the sexes. These results provide evidence for RTs in mediating sex-specific effects of human gene expression.
MicroRNA Expression, APOE and TOMM40. L. M. Bekris, Y. Shao, M. Shaw, B. McCue, J.B. Leverenz. 1) Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH; 2) Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University; 3) Lou Ruvo Center for Brain Health, Cleveland Clinic, Cleveland, OH.

MicroRNAs (miRNAs) play an essential role in post-transcriptional gene regulation in the brain. Differences in brain miRNA expression between Alzheimer’s disease (AD) and cognitively normal controls have been reported. Genetic variants located at the APOE locus are the strongest risk factors for late-onset AD described to date. A complex regulatory structure exists at the APOE locus that includes putative regulatory sites within the TOMM40 gene. Little is known about the influence of miRNA post-transcriptional regulation at this locus. The aim of this investigation was to demonstrate that brain miRNA correlate with APOE and TOMM40 brain expression differentially according to brain region or the presence of AD. In addition, it was demonstrated, using bioinformatics tools, that a subset differentially expressed miRNAs that correlated with APOE or TOMM40 expression are predicted to target APOE or TOMM40. To achieve this aim, miRNA and miRNA quantitative RT-PCR were used to measure miRNA expression in post-mortem brain from AD (n=21) or cognitively normal age-matched control (n=21) hippocampus (HP) and cerebellum (CB). Quantitative RT-PCR was used to measure APOE and TOMM40 mRNA. Western blotting was used to measure brain APOE and Tomm40 protein. Linear regression was used to determine if APOE and TOMM40 miRNA expression correlate with protein expression in AD (n=8) compared to cognitively normal age-matched controls (n=8) or HP compared to CB. Linear regression analyses were used to determine if miRNA qRT-PCR levels differentially correlate with TOMM40 or APOE expression in AD (n=6) compared cognitively normal age-matched controls (n=6). Post-mortem brain mRNA were identified that significantly correlate with AD or control brain APOE or TOMM40 expression levels. In addition, a subset of miRNA were predicted to target APOE or TOMM40. These results suggest that an intricate APOE locus regulatory structure may be further fine-tuned by miRNA post-transcriptional modulation according to brain region or disease status.

538S

MicroRNAs (miRNAs) are 21–23-nucleotide-long small non-coding RNAs that are processed from longer transcripts forming stem-loop structure by RNA-induced silencing complex (RISC) and functions as mediators in gene silencing. MicroRNAs play essential roles in gene regulation by inhibiting translation of messenger RNAs (mRNAs), and by digestion of mRNAs, or by RNA interference, during various vital function and phenomenon. Heat shock is an external stress, and cells that are subjected to heat shock immediately respond to such an environmental stress for protecting themselves and for maintenance of homeostasis. A major response to heat shock is the expression of a certain set of proteins, referred to as heat-shock proteins, and heat-shock proteins play important roles in cells under heat stress conditions. A recent study (Mol Cell, 39, 292-299, 2010) suggested that HSP70 and HSP47 interacted with the AAO2 protein that is a major component of RISC, and would participate in loading of small RNAs into RISCs, leading us to the possibility that there may be some sort of relationship between heat shock and gene silencing. In this study, we investigated gene silencing mediated by miRNAs in mammalian cells exposed to a mild heat hyperthermia (40 °C) by means of miRNA activity assay using a luciferase reporter gene as well as miRNA expression analysis using a DNA microarray. Our data indicated that the gene silencing activities involving miRNAs were enhanced without increasing in their expression levels under such a heat-stress condition. Additionally, the gene silencing activity appeared to be independent of the cytoprotective action involving heat shock proteins that are immediately activated in heat-shocked cells and that function as molecular chaperons for restoring heat-denatured proteins to normal proteins. Our findings point out the possibility that gene silencing involving endogenous miRNAs might play a subsidiary role in heat-shocked cells for an aggressive inhibition of the expression of heat-denatured proteins.
539M The functional role of IncRNAs in breast tissue-specific gene regulation. E. Wagner, Y. Hao, Y. Liu, C. He, 1, 2) Dept. of Epidemiology, Indiana University School of Public Health, Indianapolis, IN; 2) Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN; 3) Dept. of Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN.

Long non-coding RNAs (lncRNAs) are a class of RNAs that are more than 200nt but do not code for protein. Increasing evidence suggested that IncRNAs play a key role in gene regulation. However, few studies have investigated the role of lncRNAs in gene regulation at genome-wide level, especially in target tissue. In order to determine the functional role of lncRNAs in tissue-specific gene regulation, we have conducted a genome-wide association study of co-expression of lncRNAs and mRNAs in 119 breast tumors and 23 normal breast tissue samples using RNA-seq data (2x100nt). After QC and normalization, the expression levels of 4,777 known lincRNAs and 21,128 mRNA were estimated. We used linear regression to examine the association between each mRNA and lncRNA expression, adjusting for age, race, tissue type, and sequencing batch. Multiple testing was corrected using Bonferroni correction. We considered two possible underlying mechanisms: 1) differential co-expression mechanism in which the association between lncRNA and mRNA expression differs in normal and tumor breast tissue; 2) dose-response co-expression mechanism in which there is an association between lncRNA and mRNA expression and the association is similar in normal and tumor breast tissue. Accordingly, we tested an interaction term between IncRNAs and tissue type in regression model in order to detect the differential co-expression. We found that 694 pairs of IncRNA-mRNA (comprising of 524 lncRNAs and 76 mRNAs) are differentially expressed in tumor and normal breast tissue. In interaction<1x10−12, Genes such as TFES3, DEDD, and DNL1 were differentially co-expressed with multiple lncRNAs. These genes are involved in chromosomal translocation, apoptosis, and cellular movement, functions implicated in breast cancer. For dose-response co-expression, we found 804 pairs of IncRNA-mRNA (comprising of 13 IncRNAs and 706 mRNAs) are similarly co-expressed in tumor and normal breast tissue. To identify cancer-related co-expression, we further restricted the results by differentially expressed IncRNAs and mRNAs in tumor and normal breast tissue and identify genes including GABARAP1, DCTPP1, KIFC1, and MAG1FE1. Ingenuity pathway analysis revealed that the identified genes are primarily clustered in inflammatory response, cell death and survival, and cellular movement. Our study suggested that IncRNAs play a key role in regulation of genes that are important for breast cancer initiation and progression.


Small non-coding RNAs (ncRNAs) comprise many classes of distinct RNAs with diverse and unknown functions. Thus, a complete and accurate catalog of ncRNA genes is essential for elucidating the non-coding repertoire. Here we report the first genome-wide study to explore the tissue specificity properties along with other characteristics of ncRNAs such as cleavage specificity and secondary structure. We collected publicly available small-RNA sequencing data from adipose, B-cell, blood, brain, colon, heart, kidney, liver, lung and skin samples from 275 human normal, non-diseased subjects. We first performed stringent data quality control steps to account for biological and inter-laboratory variation. We then quantified the expression values of all tissues and identified 2215 scRNAs, 83% (85% in GENCODE) of which belong to the mRNA, tRNA, snRNA, snoRNA, and rRNA classes. Overall, we find 33% of scRNAs to be expressed in at least one tissue type, with no tissue expressing more than 46% of all possible scRNAs. Distinctions of expression levels were very similar across tissues, with miRNA having the highest absolute expression levels, followed by snoRNA C/D box, H/ACA box and tRNA. Clustering of overall expression levels of these tissues mimics the stem cell potency groups: liver and lung (endothelial), brain and skin (ectoderm), heart and kidney (mesoderm). We then analyzed the tissue specificity of scRNAs using Shannon entropy introduced by Schug et al. Unexpectedly, the tRNA (58%), scRNA (51%), and snRNA (48%) classes were highly tissue specific even though they are generally thought to perform housekeeping cellular functions. Over half of the tissue-specific scRNAs were identified in B cells, similar to tRNA-derived miRNAs reported recently, and some of the scRNAs are residing in the same compartment as lincRNAs, which may suggest why scRNAs have been referred to as ubiquitous miRNAs. Taken together, these results suggest that scRNAs typically thought to function in basic biological processes may work through different mechanisms depending on tissue type.

541S A tissue-specific lincRNA in the TRAF1-C5 risk locus as a putative cis-regulator bridging genetics and disease. T. Messemaker, R.B. Marques, T.W. Huizinga, A. Adriaans, A.M. Bakker, A.W.G. Berendsen, N. Daha, H.M.M. Mikkers, F. Kurreeman, 1, 2) Depart. of Rheumatology, Leiden University Medical Centre, Leiden, Netherlands; 2) Department of Molecular Cell Biology, Leiden University Medical Centre, Leiden, Netherlands; 3) Department of Human Genetics, Leiden University Medical Centre, Leiden, Netherlands.

In the last decade genome wide association studies (GWAS) have identified genetic polymorphisms that associate with Rheumatoid arthritis (RA). However, the way these genetic regions contribute to disease remains ill defined. We previously identified the TRAF1-C5 locus as a predisposing risk factor to the development of RA. In the present study we investigated functional consequences of this risk locus. Using expression quantitative trait loci (eQTL) datasets, we observed an association between the risk allele and expression of TRAF1 and C5 at the mRNA level in various blood-related cell types. As part of an underlying mechanism we identified a novel large non-coding RNA intergenic of TRAF1 and C5 (C5T1-lincRNA). The lincRNA is transcribed in the same orientation as TRAF1 and C5 by RNA polymerase II, is highly transcribed in liver, and its expression is rapidly induced in different immune cells by specific immune stimuli. Expression of C5T1-lincRNA correlated with either C5 or TRAF1 expression in a tissue specific manner. In addition, knockdown of the intergenic transcript in a hepatocyte cell line resulted in decreased C5 levels. Together our data support the involvement of a novel lincRNA in regulating C5 and TRAF1 expression. We propose that this lincRNA, which is fully located within the associated region, is responsible for the RA associated altered RNA levels of TRAF1 and C5 and plays a role in RA susceptibility.


The approximately three billion base pairs of the human DNA represent a storage devise encoding information for hundreds of thousands of processes that can go on within and outside of a cell. This information is partially revealed in distinguishable transcripts that are potentially derived from 12 billion nucleotides, considering strandedness and the diploid nature of the most genomes. Results stemming from the efforts to catalogue and analyze the RNA products made by many cells in the Encyclopedia of DNA Elements (ENCODE), in human, fly-worm (modENCODE) and mouse ENCODE projects have contributed to help shed light on both the functional content, how this transcriptional information is organized and its evolutionary conservation. Currently for the human genome, there are annotated a total of 196,520 transcripts (41 % coding) within 57,820genic regions (35% coding) based on v19 in Genencode. Analyses of both the regulatory and transcriptome data sets of phase 2 of the human ENCODE projects underscore several under-appreciated lessons concerning the conservation and novel functions of long and short non-coding (nc)RNAs. These lessons include: a) specific short ncRNA classes are enriched in micro-vesicles and used for inter-cellular communication b) the conservation of long ncRNAs between mouse and human of several transcriptional features that appear to be independent of the degree of sequence similarity and c) a multiplicity of cap modifications present on short ncRNAs. These and other insights drawn from the multiple ENCODE and other genome-wide data sets assist in understanding what is often seen as dauntingly complex but elegantly organized genome and continues to prompt a reconsideration of the definition of a gene.
The landscape of retrotransposon expression in human lung adenocarcinomas. A. Biton1, B. Xue2, D. Risso3, T. Speed4,4, L. He2, Biton. 1) Department of Medicine, Lung Biology Center, University of California, San Francisco, San Francisco, California 94110 USA; 2) Department of Molecular and Cell Biology, Division of Cellular and Developmental Biology, University of California at Berkeley, Berkeley, CA 94720, USA; 3) Department of Statistics, University of California, Berkeley, CA 94720, USA; 4) Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville VIC 3052, Australia.

Nearly half of the human genome consists of repetitive elements. The largest class of mammalian repetitive elements are retrotransposons, a class of transposable elements that can duplicate in the genome through RNA intermediates using a “copy and paste” mechanism. Retrotransposons are considered a major source of genetic variation and de novo insertion of retrotransposons in the genome can introduce mutations and modify gene splicing and expression. Normally silenced in adult somatic cells, aberrant retrotransposon activation has been associated with cancer. For example, using exome sequencing data, de novo insertions of L1 elements have been detected in tumor suppressor genes of various cancer types (Helman E. 2014 Genome Res). In this study, we are interested in detecting potential aberrant retrotransposon expression in cancer genomes and determining their potential effect on proximal genes.

We studied the expression of retrotransposons (LINEs, LTRs, and SINEs) in RNA-Seq data from tumor/normal pairs of lung adenocarcinoma from public database including The Cancer Genome Atlas. The analysis of the retrotransposon expression was performed at two levels. First, at the level of the consensus sequence, by mapping the reads to the consensus sequences listed in RepBase, to crudely evaluate the abundance of each retrotransposon class. Second, at the locus-specific level, by using the RepeatMasker annotation in the genome and investigating the potential effect of the retrotransposons in regulating the adjacent genes.

We observed that lung adenocarcinomas have distinct retrotransposon expression profiles compared to the normal samples. Indeed we found that tumor and normal samples could be distinguished solely based on their retrotransposon expression profiles. Interestingly, in addition to LINEs and SINEs, we observed aberrant expression of LTRs (including but not limited to HERV18 or HERVK). Using the locus-centric approach, we found that derepression of some retrotransposons could be involved in inhibition of their adjacent gene. We also observed examples of concurrence of exonization of the retrotransposon and differential exon-usage within the same gene. Altogether, these results suggest that retrotransposons may play roles in cancer development.

Overexpression of the FMR1 mRNA in premutation carriers is isoform specific. C. Yrigollen1, D. Pretto1, J. Eid1, H-T. Tang2, E. Loomis1,3, C. Raske4, B. Durbin-Johnson5, P. Hagerman1,4, F. Tassone1,4 1) Department of Biochemistry and Molecular Medicine, University of California Davis, Sacramento, CA; 2) Pacific Biosciences, Inc., Menlo Park, CA; 3) Department of Public Health Sciences, University of California Davis, School of Medicine, Davis, CA; 4) MIND Institute, University of California Davis Medical Center, Sacramento, CA.

Premutation carriers of the FMR1 gene are at risk to develop multiple fragile X associated disorders, including the neurodegenerative disorder fragile X-associated tremor/ataxia syndrome (FXTAS). Elevated FMR1 mRNA levels have been observed in premutation carriers; whether all FMR1 isoforms are differentially expressed or only specific ones, is currently unknown.

Thus, using single molecule real time (SMRT®) sequencing and qRT-PCR we have determined which of the potential transcribed FMR1 isoforms are expressed and, more importantly, determined if all or only some of the transcripts are overexpressed in premutation carriers compared to normal controls.

The SMRT® sequencing approach has provided us complete cDNA sequence reads from exon 1 through exon 9 and exon 9 - exon 17 within a single FMR1 transcript. Our results showed that out of the 24 predicted isoforms in human, only 16 FMR1 isoforms are expressed in peripheral blood mononuclear cells, brain, and primary fibroblast cells. Interestingly, although the isoforms missing exon 12 were the most abundant in both premutation carriers and controls in all three tissues, only two were differentially expressed in carriers.

Specifically, out of the 16 observed isoforms, Iso10 and Iso10b, both with exons 12 and 14 spliced, and differing only in the splicing acceptor site used at exon 17, were significantly overexpressed.

These findings demonstrate that the differential expression of FMR1 isoforms observed in premutation carriers is isoform specific. The loss of exon 12 results in a shortened unstructured variable loop between two β sheets in the FMRP KH2 domain while the splicing of exon 14 results in the loss of a nuclear export signal and introduce a frameshift that alters the amino acid sequence from exon 15 - exon 17, giving rise to truncated proteins.

Why FMR1 alternative splicing regulation is compromised in premutations is not yet understood, nor the functional consequences that result from abundance of Iso10 and Iso10b, but these changes in expression of specific isoforms are likely to play a relevant role for understanding the pathogenesis of FXTAS and of the FMR1 associated disorders.
545M

High throughput identification of exonic sequence variation that exhibits allelic imbalance in RNA splicing. R. Soemedi
1,2,4, T. Esko
1,2,4, A. Metspalu
1,2
1) Center for Computational Molecular Biology, Brown University, Providence, RI; 2) Molecular and Cellular Biology and Biochemistry, Brown University, Providence, RI; 3) Department of Computer Science, Brown University, Providence, RI; 4) Center for Biomedical Engineering, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

Introduction: The rate of discovery of causative mutations for human diseases has accelerated in recent years, but a significant proportion of published causality lacks direct evidence of pathogenicity of the observed variants. Recent studies have predicted that -15–30% of exonic variants can alter RNA splicing, which may result in exon skipping and aberrant splice site usage; both can potentially lead to more deleterious changes due to frameshift. We have developed High Throughput Splicing Assays (HTSAs) to test tens of thousands of wild-type/wild-type pairs of splice substrates simultaneously for the presence of allelic imbalance in RNA splicing.

Methods: We designed a library of 200-mer oligonucleotides consisting of 5,192 pairs of mutant and wild-type version of exons (median exon length = 76 nt) and their flanking introns derived from reported missense and nonsense mutations at Human Gene Mutation Database (n=1064) and common SNPs (n=228). HTSA in vivo was done by incorporating the library into three-exons splicing constructs under CMV promoter that were subsequently transfected into HeLa cells. HTSA in vitro was performed in HeLa nuclear extract using in vitro-transcribed library containing two-exons splicing constructs. Original pool of substrates and spliced RNAs were extracted, made into Illumina libraries and subjected to parallel sequencing in HiSeq2500.

Results: We observed allelic imbalance in RNA splicing in 1,036 (26%) of the 3,930 tested SNPs that were tested in both HTSA in vivo and in vitro (>1.5 fold-change and P < 0.05 by two-sided Fisher’s exact test, fold-change). We further identified alternative splice site usage and intron retention in mutant RNAs, that is, in mutant sequences exclusively 37% (n=75) and 49% (n=187) mutant/wild-type pairs of the library substrates, respectively (P < 0.001 by two-sided Fisher’s exact test). The resulting datasets were made into ‘executive summary’ plots for the annotation and comparison of multiple potential splicing sites.

Conclusions: Our study facilitates a better characterization of exonic sequence variation of potential deleterious consequence resulting from aberrant splicing. Our findings indicate the urgency of integrating splice-altering variant hypothesis in characterizing causal variants in human disease. Machine-learning approaches can be further applied to generate a superior prediction tool for identifying variants that lead to gene processing defects.

546T

Impact of microRNA binding site polymorphisms on gene expression variation. T. Amño1, U. Vida1, T. Elbers1, E. Metspalu1,2
1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Department of Bio-technology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

Functional interpretation of GWAS-identified loci presents a major challenge since only a small fraction of variants has a direct effect on protein-coding regions. MicroRNAs (miRNA) bind to specific motifs called miRNA response elements (MREs) in the 3’-UTR of mRNAs, repressing the activity of their target mRNAs. We used SNP data to analyze the miRNA conservation and increasing number of known human miRNAs and SNPs in MREs provides an opportunity to systematically investigate the potential impact of common genetic variants on regulatory interactions between miRNAs and their target mRNAs.

We hypothesized that polymorphisms within the MREs that disrupt or generate a binding site can affect the expression of target genes, leading to allelic-expression specific modulation. cis-eQTLs (variants associated with gene expression changes) from the recently published joint pairs of splicing individuals) were retrieved from bloodetQTLbrowser (http://genenetwork.nl/bloodetqtlbrowser/). SNPs and proxies (R²=1) were mapped to the 3’UTR of the cis-eQTL-affecting genes. Putative effect of the SNP on the binding of miRNA was estimated using PolymiRTS v2 database. In total, 4829 individual SNPs and proxies were found in the putative MREs of 2166 REFSEQ transcripts. Out of 2578 miRNAs present in miBase v20, 2573 were predicted to be influenced by at least one SNP. Altogether, 24 858 SNP-miRNA associations were reported. To restrict the analysis to specific miRNAs, a list of 123 miRNAs was identified using 11 publicly available small RNA sequencing datasets using a threshold of 10 reads in at least 6 studies. In addition, to add confidence to miRNA binding site prediction, MRE SNPs were filtered against 36 independent AGO2, AGO3 and AGO4 PAR-CLIP and HITS-CLIP datasets from STARBase database. The filtering using a consensus list of blood-specific miRNAs revealed a set of 1128 SNP-miRNA associations, consisting of 123 miRNAs, 771 SNPs and 614 transcripts. After filtering against AGO-binding-guided mutation-conservative set of 331 SNP-miRNA associations remaining, consisting of 101 miRNAs, 206 SNPs and 192 transcripts. We suggest that our analysis enables identification of miRNA-associated eQTLs for future functional studies and helps to prioritize genetic variants that are important in complex diseases or traits.

547S

High-Resolution Genomic Analysis of Human Mitochondrial RNA Sequence Variation Reveals Genetic Determinants of Post-transcriptional Modification and Interplay with the Nuclear Genome. A. Hodgkinson1, Y. Idaghdour1, E. Gbeha1, J.-C. Grenier1, E. Hip-Ki1, V. Bruat2, J.-P. Goulet3, T. de Maillard1,2, P. Awadalla1,2
1) CHU Sainte-Justine Research Centre, Department of Pediatrics, Faculty of Medicine, Université de Montreal, 3175 Chemin de la Côte-Sainte-Catherine, Montreal, Quebec H3T 1C5, Canada; 2) CARTaGENE, 3333 Queen Mary Road, Office 493, Montreal, Quebec H3V 1A2, Canada; 3) Department of Biology, Division of Science and Mathematics, New York University Abu Dhabi, Post Office Box 129198, Abu Dhabi, United Arab Emirates.

Mutations in the mitochondrial genome have been associated with a wide range of diseases, as well as fundamental biological processes such as ageing. Previous studies have found that sequence variation across mitochondrial genomes within a single individual (heteroplasmy) is reasonably rare, however little is known about the extent and distribution of sequence variation in the mitochondrial transcriptome. By ultra-deeply sequencing mitochondrial RNA from the whole blood of ~1000 individuals at an unprecedented level of coverage (>8000X) from the CARTaGENE project, we find a remarkable level of variation across individuals, as well as sites that show consistent patterns of post-transcriptional modification. Using a genome-wide association study we find that the rate of post-transcriptional modification at functionally important sites in mitochondrial transfer RNAs is under strong genetic control, largely driven by a missense mutation in MRPP3 that explains ~22% of the variance. Furthermore, we find nuclear genetic variants that are associated with modifications in other mitochondrial genes, and explore the relationship between mitochondrial post-transcriptional modification and genetic risk of complex diseases, such as neurodegeneration. The identification of a major nuclear genetic determinant of post-transcriptional modification in mitochondria, develop our understanding on the level of cross-talk between the nuclear and mitochondrial genomes, and suggest that IRNA post-transcriptional modification may impact cellular energy production.

548M

RCARE: RNA-Sequence Comparison and Annotation for RNA Editing. SY. LEE1,2, JG. Joung1, CH. Park3, JH. Park4, JH. Kim5
1) Seoul National University Biomedical Informatics (SNUBI) and Systems Biomedical Informatics Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 110799, Korea, BS; 2) Seoul National University Biomedical Informatics (SNUBI) and Systems Biomedical Informatics Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 110799, Korea, MS; 3) Seoul National University Biomedical Informatics (SNUBI) and Systems Biomedical Informatics Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 110799, Korea, PhD; 4) Seoul National University Biomedical Informatics (SNUBI) and Systems Biomedical Informatics Research Center, Division of Biomedical Informatics, Seoul National University College of Medicine, Seoul 110799, Korea, BS; 5) University College of Medicine, Seoul 110799, Korea, MD, PhD, MS.

Posttranscriptional sequence modification of transcripts through RNA editing is an important mechanism for regulating protein function and is correlated with human disease phenotypes. The identification of RNA-editing sites is a fundamental step in the study of RNA editing. Determining the location of condition-specific RNA-editing sites and elucidating their functional roles will help toward understanding various biological phenomena that are mediated by RNA editing. In the present study, it is proposed that RCARE (RNA-sequence comparison and annotation for RNA editing), which searches, annotates, and visualizes RNA-editing sites using thousands of previously known editing sites, can be used for comparative analyses between multiple samples. This a new tool, RNA-sequence comparison and annotation for RNA editing (RCARE), which determines condition-dependent RNA-editing sites, provides rich and systematic annotations, and delivers ‘executive summary’ plots for the annotation and comparison of multiple samples. RCARE allows for the analysis of multiple RNA-Seq data through a user-friendly web interface. RCARE is freely available at http://www.snuibi.org/software/rcare/.
The RNA Editing Landscape in Acute Myeloid Leukemia. E. Meduri1,2, B. Huntly1,2. 1) Department Of Haematology, Cambridge Institute For Medical Research, University Of Cambridge, Cambridge, United Kingdom; 2) Wellcome Trust—Medical Research Council Cambridge Stem Cell Institute, Cambridge, United Kingdom.

Recent next generation sequencing efforts have annotated the role of somatic DNA mutations in the generation of malignancies such as acute myeloid leukemia (AML). However, growing evidence points to epigenetic variation as a further and perhaps equal contributor to the malignant phenotype. We present pairwise transcriptomic and whole genome/exome sequencing (WGS/WES) on 172 AML samples from 33 patients was performed, using a stringent pipeline to identify RNA editing events (REE). Surprisingly, we observed a marked heterogeneity in the number of editing events per sample, which ranged from <1000 to >13,000 events and known proven candidates and outcome markers in AML, such as cytoge- genetic and mutational status and overall survival.

Global identification of binding sites for the splicing regulatory factors SRSF5 and hnRNP1A. G.H. Bruun1, T.K. Doktor1, S. Briner1, A. Masuda2, B. Palsiaini1, A.R. Krainer2, K. Ohno3, B.S. Andersen1. 1) Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark; 2) Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan; 3) Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA.

RNA splicing is a fundamental process required for correct splicing of almost all protein coding genes. It is critically dependent on numerous splicing-regulatory factors (SRFs) throughout all tissues. The consensus expression and activity of the splicing regulatory factors (SRFs) that function by binding to the SRSFs. Thus, mutations/SNPs that affect SREs, may disrupt splicing and a rapidly increasing number of human diseases are now known to be caused by aberrant splicing. Unfortunately, specific binding motifs are not precisely defined for even the best characterized and most abundant SRPs and will assess the functional consequences of the REE in these genes including alterations of protein coding, RNA abundance and RNA splicing. We are also comparing these tumor specific patterns to normal tissue patterns of RNA editing and are correlating the patterns of REE with known prognostic and outcome markers in AML, such as cytoge-


Background: Alternative splicing expands the repertoire of gene functions and is a signature for different cell populations. Here we characterize the transcriptome of human bone marrow subpopulations including progenitor cells to understand their contribution to homeostasis and pathological conditions such as atherosclerosis and tumor metastasis. To obtain full-length transcript structures, we used long read technologies such as Pacific Biosciences (PACBio) and Oxford Nanopore (ONT). Full-length transcript consensus sequences were obtained for the PacBio data using the RS_IsoSeq protocol from genome-guided assembly. Full-length transcript consensus sequences for genome-guided assembly. Full-length transcript consensus sequences for genome-guided assembly. Full-length transcript consensus sequences were obtained for the PacBio data using the RS_IsoSeq protocol from PacBio's SMRT Analysis software. Quantitation for each sample was done independently for each sequencing platform using Sailfish to obtain the TPM (transcripts per million) using k-mer matching. Results: PacBio's long read sequencing technology is capable of sequencing full-length transcripts up to 10 kb and reveals heretofore-unseen isoform diversity and complexity within hematopoietic cell subpopulations. A comparison of sequencing depth revealed a 50-fold difference between the expected and de novo transcript assembly with short read, second-generation sequencing reveals that, while short reads provide precision in determining portions of isoform structure and supporting larger 5' and 3' UTR regions, it fails in providing a complete structure especially when multiple isoforms are present at the same locus. Increased breadth of isoform complexity is revealed by long reads that permits further elaboration of full isoform diversity and specific isoform abundance within each separate cell population. Sorting the distribution of major and minor isoforms reveals a cell population-specific bias, particularly on RNA editing and shows how tissue specificity and diversity are modulated by alternative splicing.
553S  Analysis of the expression levels of chitinase-like proteins, Ym1, Ym2 and breast regression protein-39, in mouse tissues. M. Ohno, Y. Kida, M. Sakaguchi, Y. Sugahara. Dept Applied Chemistry, Kogakuin Univ, Hachioji, Tokyo, Japan.

Mice produce chitinase-like proteins (CLPs), which are structurally homologous to chitinases but lack the ability to degrade chitin. The CLPs belong to glycosyl hydrolases family 18. Mice express primarily three CLPs, including Ym1, Ym2 and breast regression protein-39 (BRP-39). Recently, CLPs have attracted considerable attention due to their increased expression in a number of pathological conditions, including asthma, allergies, rheumatoid arthritis and malignant tumors. Little is known, however, about the significance of CLPs increased expression levels during pathophysiological states. Because CLPs lack chitinolytic activity and detectable functions, their biochemical properties have only been partially defined. The quantification of Ym1, Ym2 and BRP-39 individually is important steps in gaining insight into the in vivo regulation of the CLPs. In this study, we established quantitative real-time PCR systems to quantify the expression of Ym1, Ym2 and BRP-39 individually and compare their expression levels with those of mammalian chitinases and reference genes in mouse tissues. We found that Ym1 and BRP-39 mRNAs were expressed at a high level in the mouse lung, whereas Ym2 mRNA was the most abundant CLP in the stomach, followed by lung. Furthermore, the expression levels of Ym1 and BRP-39 in the mouse lung were higher than those of two active chitinases.


Mice and humans express two active chitinases, chitotriosidase and acidic mammalian chitinases. They also produce chitinase-like proteins (CLPs), which lack any detectable functions. Mice express Ym1, Ym2 and breast regression protein-39 (BRP-39), whereas humans produce YKL-40, the human homologue of BRP-39, but do not synthesize homologues of Ym1 and Ym2. The chitinases and CLPs exhibit sequence homology to bacterial chitinases and belong to glycosyl hydrolases family 18. YKL-40 levels have been shown to be elevated in patients with rheumatoid arthritis, atherosclerosis, asthma, allergies, and cancers. Here, we established and validated quantitative real-time PCR systems to quantify the expression of YKL-40 and compared their expression levels with those of mammalian chitinases and reference genes using the same scale in human tissues. We found that YKL-40 mRNA was widely expressed in human tissues. The highest levels of YKL-40 mRNA were detected in human liver, followed by the kidney and lung tissues. Then we quantified and compared the expression levels in mouse and human tissues using the human-mouse hybrid standard DNA. Our results indicate that BRP-39/YKL-40 mRNA was expressed at high levels in the mouse lung but was weakly expressed in normal human lung. There were prominent differences in BRP-39/YKL-40 expression in the lungs of humans and mice. In contrast, YKL-40 mRNA expression in human liver tissue was higher than BRP-39 expression in mouse liver.

555T  Transcriptome Profiling of Gamma Secretase Inhibition in Breast and Ovarian Cancer Cells. A. Manmernien1,2, H. Ahonen1,2, H. Peltonen2,3, M. Hiltunen4, A. Haapasalo5, M. Heinäniemi6, M. Nykter2, V. Kytölä2, V-M. Kosma2, 1) School of Medicine, Institute of Clinical Medicine, Pathology & Forensic Med, University of Eastern Finland, Kuopio, Finland; 2) Imaging Center, Clinical Pathology, Kuopio University Hospital, Kuopio, Finland; 3) School of Medicine, Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland; 4) A. I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland; 5) Institute of Biomedical Technology, University of Tampere, Tampere, Finland.

Gamma secretase (GS) is a transmembrane enzyme that acts on over 90 substrates, many of which are involved in signaling networks dysregulated in cancer and Alzheimer’s disease. GS inhibitors are known to inhibit cancer growth, and this is largely attributed to their ability to disrupt Notch signaling. It is, however, likely that the other substrates also play a role in the anticancer effect of GS inhibition. Here we report a whole transcriptome sequencing study assessing the effect of DAPT, a GS inhibitor, on gene expression in MCF7, MDA-MB-453 and OVCAR3 cells. The RNA-sequencing generated 73 Gb of raw sequence data. Our data shows that GS inhibition (GSI) by DAPT causes cell line specific and global alterations in gene expression in MCF7, MDA-MB-453 and OVCAR-3 cells. The analysis (edgeR) yielded 15, 6 and 36 differentially expressed (DE) genes for MCF7, MDA-MB-453 and OVCAR-3 cells (FDR = 1%). None of the genes were shared by all cell lines. Next, DE was analyzed between grouped treated and untreated samples. A total of 61 genes were identified to be differentially expressed, 30 of them downregulated and 31 upregulated (FDR = 1%). Notably, NOTCH1 was included in the downregulated group, and NOTCH1 and HES1 are reported to regulate ribosomal and histone protein coding genes. The most prominent biological effects were downregulation of Notch signaling pathway and upregulation of ribosomal and histone proteins. In order to explore the biological effect of GSI, the list of DE genes was analyzed for KEGG pathway enrichment. Upregulated genes were enriched in Ribo-some (q=2.25E-09) and Systemic lupus erythematosus (SLE) (q=0.011) pathways. The SLE pathway was enriched due to its pathway containing histone genes. As expected, the downregulated genes were found enriched in Notch signaling pathway, which included EP300, NOTCH1, CREBBP, HES1 and NCOA2 (q=0.0046). Further, Connectivity Map analysis applied to the GSI signature identified 16 compounds (FDR = 1%) suggesting that the gene expression signature of GSI shares significant similarity with CDK inhibitors and apoptosis inducers. Overall, our results suggest that the effects of GSI are mediated through multiple mechanisms, implying that the cancer growth inhibiting properties of gamma secretase inhibitors are not exclusively due to inhibition of Notch signaling.


Pancreatic islets of Langerhans consist of several endocrine cell types. Previous studies of sorted cell populations have provided limited insight into the cellular diversity within islets. Here, we investigate the cellular heterogeneity of human islet cells by performing single cell (SC) RNA-seq. We studied islet cell transcriptomes using islets from 2 healthy organ donors and 2 individuals with type 2 diabetes (T2D). RNA-seq was performed in the whole islet (bulk) and 149 SCs from the healthy donors (respectively 68 SCs and 81 SCs) and 126 from T2D (respectively 63 SCs and 63 SCs). On average we obtained 36.6 millions reads per SC. Clustering analysis allowed us to identify the two most abundant cell types, one expressing high levels of insulin (INS) and the other glucagon (GCG); two additional but less abundant cell populations were each expressing the other major islet hormones, either somatostatin (SST) or pancreatic polypeptide (PPY). In particular we were able to identify in the samples of the two normal individuals the following gene subsets 4 SCs: 92 INS, 12 CGG, 2 SST and 4 PPY SCs. For the two T2D individuals we found respectively 31 INS, 26 CGG, 1 SST and 0 PPY SCs. Interestingly we noticed that the previously assumed mutually exclusive expression of INS and GCG in beta and alpha cells respectively is not universally observed since many cells expressing high levels of both INS and GCG suggesting an intermediate alpha-beta cell phenotype; 16 from the two normal individuals and 37 for the T2D SCs were expressing INS and GCG. This novel cell subpopulation was more abundant in the patients than in the normal individuals (Fisher p-value < 0.0003). Currently we are performing a single cell RNA-seq of sorted beta SCs from the same individuals to assess the variability within this specific cell subpopulation. In addition, comparison with known eQTLs and allele-specific expression analysis is ongoing in order to identify T2D specific gene expression regulation. This study of single cells is expected to provide a better understanding of the different cell types of the pancreatic islets in in health and disease.
557M Analysis of Genome-Wide RNA-Sequencing Data Reveals Age of the CEPH/Utah (CEU) Lymphoblastoid Cell Lines Systematically Biases Gene Expression Profiles. Y. Yuan, L. Tian, D.S. Lu, S.H. Xu. Chinese Academy of Sciences and Max Planck Society (CAS-MPG) Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China.

In human, Lymphoblastoid cell lines (LCLs) from the CEPH/CEU (Centre d’Etude du Polymorphisme Humain - Utah) family resource have been widely used for examining the genetics of gene expression levels. However, we noted that CEU/CEPH cell lines were collected and transformed approximately thirty years ago, much earlier than the other cell lines from the pertaining individuals, which we suspected could potentially affect gene expression, data analysis and interpretation of results. In this study, by analyzing RNA sequencing data of CEU and the other three European populations, we systematically examined and evaluated the potential confounding effect of LCL age on gene expression levels and patterns. Our results indicated that gene expression level and pattern of CEU cell lines have been biased by the older age of CEU cell lines. Interestingly, most of CEU-specific expressions are associated with functions such as cell proliferation which are more likely due to older age of cell lines rather intrinsic characters of the population. We suggested the results should be carefully explained when CEU LCLs were used for transcriptomic data analysis in future studies.

558T Transcriptome analyzes in human corneas derived from keratoconus and control individuals from Poland. M. Gajeczka1,2, J.A. Karolak1, M. Kabza1, D.M. Nowak1, P. Polakowski1, J.P. Szaflik1, Kabza1, D.M. Nowak1, P. Polakowski1, J.P. Szaflik1. 1) Department of Genetics and Pharmaceutical Microbiology, Poznan University of Medical Sciences, Poznan, Poland; 2) Institute of Human Genetics, Polish Academy of Sciences, Poznan; 3) Department of Ophthalmology II, Medical Faculty, Medical University of Warsaw.

Keratoconus (KTCN, OMIM 148300) is a degenerating, usually bilateral disorder of the eye, characterized by progressive stromal thinning resulting in conical shape of the cornea. Here we present a pilot keratoconus study based on transcriptome analyzes in human corneas derived from keratoconus and control individuals from Poland. A total of 96 individuals from Polish population are included into this study: 48 keratoconus patients and 48 individuals without keratoconus phenotype. Detailed clinical evaluation in each individual involved in the study was completed. Protocol for RNA extraction from human cornea was established. Total RNA from corneas for RNaseq and small RNA analyzes was extracted. Illumina HiScan system is applied in transcriptome analyzes in this study. Transcriptome analyzes were initiated. Preliminary results of transcriptome analyzes in human corneas will be presented. Since genetic factors involved in keratoconus etiology are in majority unrecognizable, the newest available technologies are necessary to be involved in further keratoconus research. Support: National Science Centre in Poland, Grant no. 2012/05/E/NZ5/02127.

559S Defining tissue compartment gene expression dysregulation in asthma by multi-tissue whole transcriptome sequencing. A. Wesołowska-Anderson1, R. Davidson2, C. Urbanek3, G. Sajó4, C. Eng3, J. Rodríguez-Santana4, M. Canto5, E.G. Burchard1,2, M.A. Seibold1,2. 1) Integrated Center for Genes, Environment and Health, National Jewish Health, Denver, CO; 2) Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO; 3) Department of Medicine, University of California, San Francisco, CA; 4) Centro de Neurologia Pediatrica, San Juan, PR; 5) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA; 6) Department of Pediatrics, National Jewish Health, Denver, CO.

The extent to which gene expression in asthmatics is dysregulated among systemic immune (blood) tissue relative to respiratory airway tissues is unclear. Additionally, although 22 genome-wide association study (GWAS) asthma risk loci have been identified, expression of 161 genes within these loci (+/−200 kb of the reported markers) in the blood and airways has not been systematically examined nor has their differential expression in asthma disease. To investigate blood vs. airway gene expression patterns in asthma we performed whole transcriptome (WT) RNA-seq on whole blood and brushed nasal airway epithelium from Puerto Rican children with asthma (n=10) and controls (n=10); and whole blood and brushed bronchial airway epithelium from adults with asthma (n=12) and healthy controls (n=8). Whole transcriptome RNA-seq was performed in a 2x100bp paired end fashion using Illumina TrueSeq methodology. RNA-seq data was analyzed using the Tuxedo analysis pipeline. We identified 440, 347, and 70 genes differentially expressed among asthmatics in the bronchial, nasal, and blood WT data, respectively. There was substantial overlap (65 genes) in upper (nasal) and lower (bronchial) airway differential expression. In contrast, we found less overlap of blood differentially expressed genes with both the lower (14 genes) and upper airways (9 genes). Among the 161 GWAS genes, 72.3% (99 genes) were expressed in all tissues, 85.1% (137 genes) were expressed in at least one tissue, 23.4% (32 genes) were expressed only in airway tissues, and 2.2% (3 genes) were only expressed in the blood. There was a significant enrichment for GWAS genes among genes differentially expressed in the airway (17 genes, p=9.4×10−6), in contrast only 1 GWAS gene was among the genes differentially expressed in blood of asthmatics (p=0.38). Among differentially expressed GWAS genes were IL1RL1, which encodes for the IL-33 receptor, signaling of which is known to drive airway Th2 inflammation, and IKZF3, a chromatin remodeling factor important for B lymphocyte differentiation. Our results indicate that asthma disease is associated with a larger number of airway (both upper and lower) vs. blood gene expression changes. Additionally, we find largely unique groups of differentially expressed genes in the asthmatic airway vs. blood. Our examination of GWAS loci genes suggests that risk conferred by multiple asthma GWAS genes may operate through modulation of gene expression, specifically in the airway.

Alopecia areata (AA) is a common autoimmune disease manifesting with hair loss that ranges from spontaneously resolving patches (patchy AA, AAP) to lifelong total scalp (alopecia totalis, AT) or total body (alopecia universalis, AU) involvement. AA is a highly prevalent autoimmune disease in which the hair follicle is attacked by cytotoxic T lymphocytes. Obstacles to identifying treatments for AA include limited understanding of disease pathogenesis, as well as the molecular signatures of affected lesions. In order to identify gene expression signatures that correlate with disease, we conducted a comprehensive expression analysis of affected scalp skin of patients with AA. Scalp biopsies from approximately 20 subjects with AAP, 20 subjects with either AT or AU, and 20 healthy control subjects were obtained from lesional/perilesional scalp skin of AA patients and normal scalp skin of healthy controls. Additionally, non-lesional samples were taken from patients with AAP. RNA was extracted from these samples and microarray analysis was performed on Affymetrix microarray data in order to identify genes differentially expressed among the three sample groups. Using Ingenuity Systems’ IPA to interrogate genes upregulated in AT/AU and AAP, we identified two strong gene expression signatures in total skin from human AA, namely, an IFN-g response (IFN) signature, including IFN-gamma and IFN-inducible chemokines, a cytotoxic T cell (CTL) signature, including CD8 and granzymes, implicating these as the dominant inflammatory cells in AA pathogenesis. We identified a hair keratin (KER) signature using IPA, and the GO term Epidermis T cell (CTL) signature, including CD8 and granzymes, implicating these as the dominant inflammatory cells in AA pathogenesis. We identified a hair keratin (KER) signature using IPA, and the GO term Epidermis

561T  Dynamic Transcriptome Changes in Cell-free Synovial Fluid following Meniscal Injury Suggests the Potential for Early Intervention. D.V. Vanci1, J. Rampersaud2, A. Mehta3, S.N. Page1, B.P. Lesniak1, J.M. Vance2, M.A. Pericak-Vance2, L.D. Kaplan1, L. Wang2, 1) University of Miami, Miller School of Medicine, Miami, FL; 2) John P. Hussman Institute for Human Genomics,University of Miami, Miller School of Medicine, Miami, FL; 3) UHealth Sports Performance and Wellness Institute, University of Miami, Miller School of Medicine, Miami, FL.

Meniscus injury is one of the leading causes of osteoarthritis (OA), often leading to knee joint replacement. Osteoarthritic changes in articular cartilage have been seen as early as 2 years post-meniscal damage. We hypothesize that molecular changes are initiated within the knee capsule at the time of meniscus injury and long before the manifestation of OA. Identification of those early changes via a minimally invasive procedure could facilitate early intervention. Towards this end, we characterize transcriptome in synovial fluid (SV) at different time points (1 to 7 months) following a meniscal injury. Eight male patients ages 24–48 years with meniscal injury diagnosed by magnetic resonance imaging (MRI) were ascertainment. SV was collected from the affected knee joint by needle aspiration at the time of meniscectomy, prior to initial incision. Total RNA was extracted from cell-free SV and sequenced on Illumina HiSeq2000. The RNA-Seq generated 18 million to 59 million reads for each sample. The STAR program was used to align the reads against human reference genome to map RNA species. Aligned reads were normalized by estimating effective library size. The normalized gene expression is expressed as count per million (CPM). The EdgeR program was used to perform clustering and differential gene expression analysis. In total, 764 RNAs (including coding, noncoding RNA and miRNAs) were present in at least one sample. Clustering analysis demonstrated that samples with longer injury duration (LID) and shorter injury duration (SID) formed distinctive clusters. Differential gene expression analysis revealed that 65 RNAs were downregulated and 78 RNAs up-regulated in either rare (FDR<0.05) in 4 individuals with SID compared to 4 individuals with LID. The most differentially expressed gene is SLC2A9 (P=2.1x10^-12, FDR=9.6x10^-9), which is barely detectable in SID but is abundant in LID samples. SLC2A9 is a sugar transporter in meniscal and articular chondrocytes and can be induced by proinflammatory cytokines, e.g. IL-1beta. SLC2A9 is believed to play a role in the development and survival of chondrocytes in cartilage matrices. Our data demonstrate dynamic transcriptome changes in SV following meniscal injury and suggest the utility of using RNA biomarkers in SV to monitor disease progression and guide early intervention.

562S  Defining the transcriptional landscape of microRNAs in human peripheral blood. G. Hemmrich-Stanisak1, Z. Du3, M. Huebenthal2, M. Paulsen2, J. Hartwig2, P. Rosenstiel1, D. Kabelitz2, A. Franke1. 1) Inst. of Clinical Molecular Biology, University of Kiel, Kiel, Germany; 2) Institute of Immunology, University of Kiel, Kiel, Germany.

Next generation sequencing of small RNAs (sRNAseq) is a widely used application to microRNA-based diagnostic or prognostic marker profiles for conditions like diseases. However, it remains a challenging task as what PCR-based sRNAseq protocols suffer from the limitation that measurement of a particular miRNA is not independent from other miRNAs. The introduction of biases towards certain miRNAs during sequencing library construction leads to discrepancies in the overall abundance of sequenced miRNAs. Particularly in complex tissues such as peripheral blood these effects can be dramatic. Due to the differential miRNA expression in different blood cell types and the uneven distribution of cell types based on different conditions e.g. diseases, age, gender, etc. it is difficult to distinguish between true positive, false positive but also false negative signals. To enable correction for the above-described confounding effects and to understand in general how miRNAs are distributed in different cell populations in whole blood samples as drawn in clinical practice, we sought to investigate the transcriptional profiles of miRNAs in seven distinct peripheral blood cell populations of 40 healthy individuals using sRNAseq. To complete the picture, we also examined circulating miRNAs from serum and investigated the miRNA content of isolated blood-borne exosomes. We used magnetic activated cell sorting (MACS) to purify CD56+ (NK cells), CD19+ (B cells), CD8+ (T cells), CD4+ (Th cells), CD14+ (Monocytes), CD15+ (Neutrophils) and CD235a+ (Erythrocytes) cell populations to gain comprehensive knowledge about the miRNA transcriptome of most blood cell types. The results indicate distinct cell type specific miRNA profiles that accurately reflect the developmental relationships between the different cell lineages. A substantial fraction of miRNAs is expressed in subsets or exclusively in one single cell type. Interestingly, Erythrocytes have a distinct miRNA signature containing more than 100 different miRNA species - a fact to be taken into account when investigating miRNA transcriptomes of whole blood samples. Clustering analysis revealed distinct groups of co-regulated miRNAs that in turn show close relationships in their respective validated target genes and the corresponding biological pathways. The here presented data will help to interpret miRNA-based marker profiles more carefully, taking into account technically and naturally occurring confounding effects.
563M Systematic Analysis of Age and Sex Effect Identified Different Behaviors between Coding and Non-coding Genes and Two Age-dependent Patterns of Gene Expression. M. Narahara1, Y. Tabara2, T. Kwaguchi², F. Matsuda3, R. Yamada1. 1) Statistical Genetics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; 2) Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan.

Sexual dimorphism and aging lead to differences in a wide range of phenotypes (e.g., disease prevalence and courses). Elucidating the molecular mechanisms underlying the differences would provide new insights for complex diseases. We analyzed gene expression microarray data of human peripheral blood samples obtained from 298 individuals to characterize sexually dimorphic genes and aging genes. This study provides new insights particularly into differences between the autosomes and X chromosome, and between coding and non-coding RNAs.

We identified 2,570 (2,071) sex-biased transcripts (genes), 1,672 male-biased and 898 female-biased transcripts, by FDR<5%. Autosomal transcripts were more often male-biased, whereas X-specific transcripts were more often female-biased (P=8.9E-07). We estimated that ~7.05% of genes at X-specific regions were incompletely inactivated. Female-biased transcripts on the X chromosome were significantly enriched at the short arm compared to the long arm (P=0.0003). Such a difference was not evident for male-biased genes and ncRNAs.

We assessed aging effects by a LOESS regression, and, by FDR<5%, identified 5,254 (3,891) transcripts (genes) for females but no transcripts for males. Subsequently, we analyzed the 5,254 age-dependent transcripts with hierarchical clustering, which identified two clusters. 38% of aging transcripts were classified into the first cluster that showed decreasing tendencies with age, and enriched by ncRNAs. 62% were classified into the second cluster that showed increasing tendencies with age, and enriched by mRNAs.

Lastly, we analyzed age-sex interaction. We chose the most significant transcripts by FDR<30%, for which we conducted hierarchical clustering, and identified two clusters. Notably, in both clusters, female-male differences were largest at the youngest age and became smaller as age.

Our findings indicate the existence of a large number of sex-biased genes and aging genes for human whole blood samples. We analyzed these genes in details such as comparison among chromosomal regions and RNA types, and clustering analysis.

564T Whole Transcriptome Sequencing in Alzheimer’s Disease Reveals Gene Expression Differences Related to TGFβ1 Pathway. K. Nho1,2, V. Ramanan1,3, S. Riachy1, K. Kuan4, H. Gao2, Y. Liu2, T. Foroud1, H. Edenberg1, A. Saykin2,3,4,5. 1) Radiology and Imaging Sciences, Indiana University, Indianapolis, IN; 2) Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN, USA; 3) Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA; 4) Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA; 5) Department of Neurology, Indiana University School of Medicine, Indianapolis, IN, USA.

**Background:** Human brain function depends on the precise regulation of gene expression. Gene expression analysis (transcriptome profiling) has become a major focus of neurodegenerative disease research. Prior studies demonstrated that abnormal gene expression patterns may contribute to the onset and progression of Alzheimer’s disease (AD).

**Methods:** We performed whole transcriptome analysis of blood-derived RNA to identify blood-based biomarkers in AD. We performed RNA-Seq on 37 participants (7 healthy control (HC), 8 subjective cognitive decline (SCD), 5 early mild cognitive impairment (EMCI), 10 late MCI (LMCI), 7 AD) from the extensively studied IMAS (Indiana Memory and Aging Study) cohort. The libraries were sequenced on SOLID 5500XL and the resulting 75 nt reads were mapped to human reference genome (hg19) and a splice-junction library, respectively, using BFAST. Read counts were calculated using bamtools from NGSUtils [Breuse MR, 2013]. After normalization, we identified genes that are differentially expressed between any of the diagnosis groups (stages of AD development) using the edgeR package. Benjamini and Hochberg’s algorithm was used to correct for the false discovery rate (FDR). We performed pathway analysis with enrichment of association in order to identify biological pathways over-represented, with threshold for input defined as nominal gene-level significance (uncorrected p<0.05) using MetaCore.

**Results:** We identified 31 differentially expressed genes between any of the diagnosis groups within FDR-corrected p<0.05. These included TGFβ1 (FDR-corrected p = 3.93 x 10^-14) and SERPING1 (FDR-corrected p = 3.58 x 10^-7). The deficiency of TGFβ1 signaling has been shown to increase both Aβ accumulation and neurodegenerative processes in AD models. SERPING1 encodes a protein that inhibits the inflammatory process. These significantly differentially expressed genes suggest possible early stage up-regulation. A pathway analysis of all differentially expressed 1,022 genes with uncorrected p<0.05 revealed enrichment in 10 functional pathways within FDR-corrected p<0.05 including multiple pathways related to cell cycle, immune activation, and apoptosis.

**Conclusions:** Whole transcriptome sequencing analysis of blood in AD reveals differentially expressed genes related to immune activation including TGFβ1. These data add to the growing understanding of the genetics and pathobiology of AD that may lead to novel therapeutic targets for AD.


Multipotent stromal cells (MSCs) are known for their distinctive ability to differentiate into mesenchymal cell lineages such as adipocytes, osteoblasts, and chondrocytes. They can be isolated from numerous tissue sources including bone marrow. Because of their differential potential and secretion of many growth factor(s), MSCs are thought to exhibit inherent qualities of regeneration and immune suppression. Based on their characteristic features, MSCs are seen as advantageous for the field of regenerative medicine in treating a variety of injuries and disorders, in addition to graft versus host disease. Since the percentage of MSCs derived from bone marrow is low, MSCs must be cultured for several cell passages to obtain sufficient cell numbers for a desired cell-based therapy. However, after several rounds of passaging, we have shown previously that the quality of these cells declines as demonstrated by decreases in cell proliferation, increases in cell size, reduced multipotent differentiation potential and differences in gene expression. In this study, we wish to identify molecular markers of proliferation of MSCs that can predict the cell population quality. Human MSCs derived from the bone marrow of 8 different donors were grown under identical conditions and total RNA was harvested at cell passages 3, 5, and 7. The proliferative potential was measured for each donor/passage using two different assays, percent confluency at 96 hours and percentage of Edu positive cells after 6 hours in culture. Total RNA was hybridized on a two color microarray for each donor/passage and gene expression data was correlated with both cell proliferation assay results. Using a random forest and a multiplicity adjustment, 28 genes were identified as statistically significant and highly correlated (0.72 ≤ r ≤ -0.73) with cell proliferation. When the significant gene lists were analyzed by Ingenuity Pathway Analysis software, these genes were involved in the top scoring networks (p<0.05) of cellular growth and proliferation, cellular development, and cell cycle. Thus, we have identified novel gene markers that are indicative of MSC proliferative quality and may be used to rapidly assess a population of MSCs.
Transcriptome profiling of patterns of RNA expression could be a powerful approach to identify networks of genes that play a role in disease. We investigate the recent genetic markers that could be associated to chromosomal instability in patients affected by a Bloom's syndrome. Methodology: We performed gene expression profiling using high throughput sequencing (deep-sequencing RNAseq) using the Illumina's HiSeq 2500 platform followed by differential gene expression analysis of samples derived from two patients affected by Bloom's syndrome and three unaffected controls. The data analysis was generated using specialized softwares for preprocessing (CASA V 1.8.2 e Seyclean), mapping and alignment (Bowiez2 v1.0 e Samtools v0.1.18) and differential expression analysis (HTSeq-count v0.6.4p2). Results: The RNAseq assay revealed the precise location of transcription limits, with resolution of a single nucleotide and high level of efficiency showing high genetic complexity. It was possible the mapping of 22,334 genes, which 11,938 had minimum coverage required for analysis of differential expression. Of these, 399 were presented differentially expressed, being 216 up regulated in the group with Bloom syndrome and 183 up regulated in the control group. Unexpected all 216 genes of Bloom syndrome are connected to immunological systems. Conclusions: Our results suggested that gene expression network in Bloom's syndrome could interfere in the regulation of the pathways associated with the immunological systems regulation probably caused by disturbance of DNA repair mechanisms. The study of the transcriptome using next generation sequencing is a competent approach with presenting a good cost-coverage. Thus, enable construction of a genic interaction network leading to new possibilities to investigate the pathogenesis of Bloom's syndrome just as well in another chromosomal instability syndromes.

Transcriptome Applied to a Bloom's Syndrome: Immunological Insights. M.M. Montenegro1,2, G.M. Novo-Filho1,2, E.A. Zanardo1,2, R.L. Dutra1,2, A.T. Dias1,2, T.V.M. Costa1,2, A.M. Nascimento1,2, F.B. Piazzone1,2, C. Milani1,2, S.C.S. Andrade1,2, G. Gasparini1,2, C.A. Kim1,2, L.D. Kulikowski1,2. 1) Department of Pathology, Cytogenomic Laboratory / LIM 03, HC-FMUSP, São Paulo, Brazil; 2) Genetic Unity, Department of Pediatrics, Children Institute, HC-FMUSP, São Paulo, Brazil; 3) Department of Zootechny, EASALQ-USP, Piracicaba, Brazil.


A major problem in studying neurogenetic syndromes is obtaining live neurons from people with these disorders. We have taken a novel approach to study gene expression in neurons from Angelman or Dusp15q syndrome subjects using dental pulp stem cells (DPSC). Shed teeth were collected to generate DPSC lines from 3 AS deletion, 3 idic(15) and three neotypical subjects. RNAseq was performed on DPSC and DPSC derived neurons. We identified 20 genes in AS, 120 genes in idic(15) and 3 genes in both groups (DPT, MED12L and AKR1C7) that were significantly different from control DPSC-neurons. Copy number correlated to gene expression for most genes across the 15q11.2-q13.1 critical region. Two thirds of the genes differentially regulated in idic(15) were down-regulated compared to controls including the transcription factors FOXO1 and HAND2, while in AS the genes did not show a clear trend except in the 15q region. Pathway analysis revealed increased cytokine activity related genes including four genes that increased in idic(15) samples: CCL7, CCL2, MMP1 and MMP3, while steroid hormone biosynthesis genes were slightly enriched in both idic(15) and AS neurons. Overall there were more dramatic changes in gene expression in the idic(15) duplication than AS deletion cell lines, perhaps because the mechanism of AS may be through protein targeting by UBE3A. Nonetheless, the finding of a significant increase in both HERC2 and UBE3A expression in idic(15) the one consistent mRNA increase in these two genes in AS vs deletion neurons may impact the AS phenotype, at least in deletion cases.


Identification of viral RNA in human samples is a useful tool for human pathology research and clinical diagnostics. However, most methods for detection of viral RNA, such as qPCR and microarray hybridization, are highly targeted, provide no sequence information, or fail to capture information regarding the broad range of viral RNA abundance reflective of viral load or transcriptional activity. Previously, we have demonstrated the utility of using biotinylated capture probes to enable generation of quality RNA sequencing (RNA-Seq) libraries from highly degraded formalin-fixed, paraffin-embedded (FFPE) samples. This capture approach provides a quantitative view of the human coding transcriptome by enriching for next-generation sequencing (NGS) RNA-Seq library fragments derived from exo-nuclear sequence over those derived from intergenic or ribosomal RNA, and is able to capture coding RNA that constitutes much less than 1% of the total transcriptome. Using this same concept of sequence enrichment, we are designing an assay to capture NGS library fragments derived from viral sequences of interest from the background of human host sequences. Initial test panels target some important human pathogens, such as human papilloma virus (HPV), enterovirus, hepatitis, and West Nile virus. This development aims to facilitate RNA-Seq for both quantitation of viral RNA in relation to the human host and the ability to collect sequence information useful for the discovery of novel viral variants.

569M Identifying conserved genomic responses to inflammation in vascular endothelial cells. L. Antounian1,2, A. Medina-Rivera3, J. Dennis1, L. Rapkin1, F. Gagnon4, M.D. Wilson1,2. 1) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 2) Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; 3) CIHR-STAGE Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; 4) Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada.

Vascular endothelial cells are involved in the inflammatory response associated with many diseases. In response to cytokines like tumor necrosis factor alpha (TNFα), a potent pro-inflammatory signaling molecule, NFκB and JNK signaling pathways become activated and regulate the genomic response to infection. Transcription factors (TF) and epigenetic modifications control the dynamic regulatory landscape of cells. It is not known if the epigenetic regulation of the inflammatory response is conserved in mammalian species, many of which serve as models for human disease. To understand the epigenetic regulation of vascular endothelial cells in the context of inflammation, we are studying regulatory regions of the genome. We are using chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) to identify epigenetic modifications and TF binding sites: H3K27ac, a mark of active enhancers; H3K36me3, a mark of active gene bodies; H3K4me2, a mark of active and latent enhancers and promoters; and two TFs cJun and RelA that are activated by the signaling cascades JNK and NFKB, respectively. To perform a controlled comparison of endothelial cell inflammatory gene regulation, we are culturing primary endothelial cells isolated from the aorta of four mammalian species - human, mouse, rat, and cow - with and without treatment of TNFα. Using multiple sequence alignments, we are comparing the location of epigenetic modifications and TF binding between species. Since evolutionary conservation of epigenetic modifications or TF binding sites has been shown to enrich for tissue-specific and process-specific functions, we aim to identify conserved and species-specific regulatory elements. Preliminary results show that 16% of enhancers are conserved between human and rat, while 5% of cJun binding sites are conserved for aortic endothelial cells grown under basal conditions. We find that 74% of conserved cJun sites (N=995) and 84% of conserved H3K27ac sites (N=14,188) are shared with human umbilical vein endothelial cells. This suggests that conservation of regulatory regions is important for vascular endothelial cell identity. Evolutionarily conserved regulatory regions will be used to prioritize single nucleotide polymorphisms in regulatory regions related to disease.
570T Transcriptional and Epigenetic Landscape of Megakaryocytes derived from Human Induced Pluripotent Stem Cells. L.J. Vasquez1, TM. Moreau2, Exence de Monclere1, N. Sorone1, N. Sowada6, M. Kousi1, A. Medeira1, H. Thiele2, F. Lepri3, W. Lenzock3, A. Radiconi2, T.L. Schwarzenberg1, D.J. Morris-Rosenahl4, L. Wenzeck1, N. Katsanis5, H. Thiele5. 1) Center of Human Disease Modeling, Duke University, Durham, NC; 2) Geneva University Hospital, Department of Medical Genetics, Geneva, Switzerland; 3) Sanger Institute, Cambridge, United Kingdom; 4) Scripps Research Institute, La Jolla, CA; 5) Medical Genetics, University of Ulm, Ulm, Germany; 6) University of Limerick, Limerick, Ireland. DNA methylation was probed genome wide using Illumina 450K arrays. Next we interrogated whether hiPSC-MKs were equivalent to MKs derived from neonatal cord blood (CB) or adult peripheral blood (PB). Finally, we compared regulation of genes related to platelet biogenesis in the formers such as CD9 and VWF with incomplete epigenetic remodelling of the hiPSC genome. Indeed, our study provides mechanistic insights in the reported phenotypic and functional differences between fetal, neonatal and adult cells. Developmental differences in the regulation of TFIIIB, which recruits Pol III to target genes. Disease-associated SNPs in BRF1 are linked to human pathogenesis, such as alveolar capillary dysplasia (ACD) and pulmonary venolobar syndrome (PVOS). BRF1 is a multifunctional factor that regulates the transcription and activation of genes involved in alveolarization of the lung. The presence of functional polymorphisms in BRF1 and their association with human pathologies suggest a role for BRF1 as the cause of this syndrome, a finding supported by the recapitulation of key neyrodevelopmental phenotypes upon suppression of BRF1 in zebrafish embryos. BRF1 associates with BD1 and TP to form the transcriptional complex for genes involved in alveolarization. Therefore, the lack of function or the absence of BRF1 expression in the zebrafish model is associated with lung morphogenesis defects. Furthermore, BRF1 knockdown led to increased hypoxia signaling. Here we describe an autosomal recessive disorder characterized by intellectual disability and cerebellar hypoplasia as well as facial dysmorphic features, short stature, microcephaly, dental anomalies, and scoliosis. Whole exome sequencing revealed biallelic missense alterations of BRF1 in the patient with ACD associated with susceptibility of GSD. Methods: Genomic DNA was isolated from whole blood samples of 234 female patients with GSD and 399 gallstone-free female controls. HLA-G +1537A/C, 14-bp insertion/deletion, and +3142G polymorphisms were genotyped using PCR-RFLP or Pre-Developed TaqMan Allelic Discrimination Assay. Genotype, allele, and carrier frequencies of the three HLA-G SNPs were determined by direct counting. Hardy-Weinberg equilibrium was assessed for each SNP in both control and study groups using Haploview 4.2. Results: HLA-G +3142G genotype showed association with susceptibility of GSD (OR = 0.66, Pc = 0.03). Allele +3142G also shows the same result (OR = 0.75, Pc = 0.02). The other two polymorphisms show the same frequencies with controls. Conclusions: Based on our findings, HLA-G +3142G genotype as protective in female patients with GSD. Our results need to be replicated and verified in another large-sized cohorts or other ethnics.

571S BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. P.L. Tan1, F. Hög2, N. Sowada6, M. Kousi1, A. Medeira1, H. Thiele2, F. Lepri3, W. Lenzock3, A. Radiconi2, T.L. Schwarzenberg1, D.J. Morris-Rosenahl4, L. Wenzeck1, N. Katsanis5, H. Thiele5. 1) Center of Human Disease Modeling, Duke University, Durham, NC; 2) Geneva University Hospital, Department of Medical Genetics, Geneva, Switzerland; 3) Sanger Institute, Cambridge, United Kingdom; 4) Scripps Research Institute, La Jolla, CA; 5) Medical Genetics, University of Ulm, Ulm, Germany; 6) University of Limerick, Limerick, Ireland. RNA polymerase III (Pol III) synthesizes RNAs and other small non-coding RNAs to regulate protein synthesis. Dysregulation of Pol III transcription has been linked to cancer, and germline mutations in genes encoding Pol III subunits or RNA processing factors cause neurogenetic disorders in humans, including alveolar capillary dysplasia and pulmonary venolobar syndrome. Here we describe an autosomal recessive disorder characterized by intellectual disability and cerebellar hypoplasia as well as facial dysmorphic features, short stature, microcephaly, dental anomalies, and scoliosis. Whole exome sequencing revealed biallelic missense alterations of BRF1 in the patient with ACD associated with susceptibility of GSD. Methods: Genomic DNA was isolated from whole blood samples of 234 female patients with GSD and 399 gallstone-free female controls. HLA-G +1537A/C, 14-bp insertion/deletion, and +3142G polymorphisms were genotyped using PCR-RFLP or Pre-Developed TaqMan Allelic Discrimination Assay. Genotype, allele, and carrier frequencies of the three HLA-G SNPs were determined by direct counting. Hardy-Weinberg equilibrium was assessed for each SNP in both control and study groups using Haploview 4.2. Results: HLA-G +3142G genotype showed association with susceptibility of GSD (OR = 0.66, Pc = 0.03). Allele +3142G also shows the same result (OR = 0.75, Pc = 0.02). The other two polymorphisms show the same frequencies with controls. Conclusions: Based on our findings, HLA-G +3142G genotype as protective in female patients with GSD. Our results need to be replicated and verified in another large-sized cohorts or other ethnics.

572M Association study of the Human Leukocyte Antigen-G and gallstone disease in Han Chinese. H. Yang1,2, S. Shiff5,4,5, Y. Lee1,4,6,7,8,1, Medical Research, Mackay Memorial Hospital, New Taipei City, Taiwan; 2) Mackay Junior College of Medicine, Nursing and Management, Taipei, Taiwan; 3) Division of Gastroenterology, Department of Internal Medicine, Mackay Memorial Hospital, Taipei, Taiwan; 4) Department of Medicine, Mackay Medical College, New Taipei City, Taiwan; 5) Department of Pediatrics, Mackay Memorial Hospital, Taipei, Taiwan; 6) Department of Pediatrics, Taipei Medical University, Taipei, Taiwan; 7) Institute of Biomedical Sciences, Mackay Medical College, New Taipei City, Taiwan. Background and Goal: Gallbladder stone induce inflammatory responses and affect extra-hepatic bile ducts called gallstone disease (GSD). The pathology and environmental risk factors of GSD are well documented. However, genetic polymorphisms of immune or inflammatory regulation in GSD remain unclear. Human leukocyte antigen (HLA)-G was defined as non-classical Major histocompatibility complex class I based on relative low polymorphisms. HLA-G proteins have seven different isoforms from alternative splicing and have immunosuppressive properties. HLA-G expression is able to confer resistance to Nature killer cell- or cytotoxic T lymphocyte-mediated destruction. HLA-G exerts its inhibitory effect broadly as this molecule has been found to interact with numerous inhibitory receptors on a variety of immune cells. Although only a handful of studies have been performed that evaluate a possible role for HLA-G-variant conditions of the bowel, we believe it prudent to include this material to increase risk for the development of malignancy in the setting of inflammatory conditions of the chronic gastrointestinal and inflammatory diseases, such as GSD. In this study, we investigated whether polymorphisms of the HLA-G gene were associated with GSD susceptibility. Methods: A total of 347 HLA-G was isolated from whole blood samples of 234 female patients with GSD and 399 gallstone-free female controls. HLA-G +1537A/C, 14-bp insertion/deletion, and +3142G polymorphisms were genotyped using PCR-RFLP or Pre-Developed TaqMan Allelic Discrimination Assay. Genotype, allele, and carrier frequencies of the three HLA-G SNPs were determined by direct counting. Hardy-Weinberg equilibrium was assessed for each SNP in both control and study groups using Haploview 4.2. Results: HLA-G +3142G genotype showed association with susceptibility of GSD (OR = 0.66, Pc = 0.03). Allele +3142G also shows the same result (OR = 0.75, Pc = 0.02). The other two polymorphisms show the same frequencies with controls. Conclusions: Based on our findings, HLA-G +3142G genotype as protective in female patients with GSD. Our results need to be replicated and verified in another large-sized cohorts or other ethnics.

573T Structural Variation Analysis Using Nanochannel Genome Mapping to Predict Novel Gene Function and Induction of Pluripotency of Human Fibroblasts. E.H. Cho1, H. Dai1, A. Pang1, R. Williams2, K. Nazor2, K. Bhutani3, N. Schork2, J. Loring2. 1) BioNano Genomics, Inc., San Diego, CA; 2) Center for Regenerative Medicine, The Scripps Research Institute, La Jolla, CA; 3) Hu-Gen Bio, J. Craig Venter Institute, La Jolla, CA. Four key genes, POUSF1, SOX2, KLF4, and MYC, are commonly used for reprogramming human fibroblasts into stem cells. Induced pluripotency of human fibroblasts into stem cells currently relies on three different transfection methods of these genes: retroviral, Sendai virus, and mRNA. The latter two methods are non-integrating methods, while the former method integrates into the genome. Retroviral transfection to induce pluripotency is currently the most efficient method, but due to the potential damaging effects of viral integration into the chromosomes, it is thought to be of high risk for future clinical applications. To investigate whether structural changes result during the process of induction of pluripotency in stem cell genomes we used nanochannel genome mapping technology from BioNano Genomics to compare structural differences among cell line derived by the three reprogramming methods and the parental genome. By producing de novo genome map assemblies of each genome, comparing to a reference genome, and cross comparing all four assemblies, we determined sample-specific SV calls for each of the three methods. After automated assembly and comparison, followed by manual verification, 14 sample-specific SVs were found in the retrovirus-induced sample, 12 in the Sendai virus induced sample, and 8 in the mRNA induced sample, all with respect to the Hgt9 human reference genome. Examination of this dataset provided no evidence that any of the methods induce significant changes in the genome. Further study of a female patient Graves OR = 0.66, Pc = 0.03). Allele +3142G also shows the same result (OR = 0.75, Pc = 0.02). The other two polymorphisms show the same frequencies with controls. Conclusions: Based on our findings, HLA-G +3142G genotype as protective in female patients with GSD. Our results need to be replicated and verified in another large-sized cohorts or other ethnics.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
574S
An exercise filled lifestyle may alter the gut metagenome exposed to polychlorinated biphenyl. E. Rampersaud1, M. Veerapan1, M. Toborek1. 1) John P. Hussman Inst Human Genomic, University of Miami Miller School of Medicine, Miami, FL; 2) University of Miami.

The gut microbiome comprises of multiple phyla of microbes which responds to environmental stressors. Moreover, the gut microbiome composition has been implied in the catabolism of drugs and other xenobiotic compounds such as polychlorinated biphenyl (PCB). PCBs are persistent organic pollutants which are carcinogenic. Our group has previously shown that exercise can attenuate changes conferred by oral PCB in the murine gut microbiome. However, the role of the gut microbiome in exercise-based PCB degradation remains unclear. We propose to study the composition and functional characteristics of the gut metagenome upon the exposure to PCB and exercise. 32 mice were used in 4 study groups comprising of sedentary and exercised mice which were either exposed to PCB or not (PCB delivery vehicle only). All mice were voluntarily exercised or remained sedentary for 5 weeks and were treated with oral gavage of relevant PCB congeners or vehicle. Mouse feces were collected before gavage and 2 days after gavage. Collected fecal material was subjected to total DNA extraction and subsequently sequenced using the Illumina HiSeq 2000. All raw reads were assembled, mapped and analyzed using MG-RAST v3.3.6s and MatR v1.0.0. All samples were uploaded and analyzed successfully with MG-RAST. Our analyses revealed significant differences between sedentary and exercised mice gut microbiome – observably, an increase in bacteriodes in exercised mice (p < 0.001). The bacteriodes increase is more apparent upon treatment with PCB in exercised mice but a notable decrease in microbiome diversity based on the α-diversity value. No observable microbiome clustering was found between groups. Based on preliminary functional analyses, PCB exposure increases presence of microbes involved in information storage and metabolic proteins (p < 0.05). Our results indicate and replicate that exercise attenuates the effects oral PCB exposure phenotypically and metabolically (we were not able to alter immune systems and the observed alteration likely indicates a possible change to immune responses responsible for PCB degradation. Further studies are required to elucidate specific pathways and mechanisms involved in this process. Through this study, we postulate that exercise can be used to remediate populations exposed to oral PCBs.

575M

Africans harbor the greatest levels of human genetic variation, and differences in diet are likely to have produced distinct selection pressures resulting in genetic adaptations. An additional important source of variation that can influence both health and disease is the gut microbiome. To better understand how genetic adaptations, culture, and diet interact, we analyzed the gut microbiomes from ethnically diverse rural Africans and urban European Americans. In our study we examined natural variation in gut microbiota in a population of about 100 inbred strains of mice from the Hybrid Mouse Diversity Panel (HMDP) and a population of a 200 men from the Metabolic Syndrome in Men (METSIM) study. Profiling of the gut microbiota in the HMDP (n= 599 mice) revealed that the genetic background strongly influences gut microbiota composition and that the gut microbiome composition explains a considerable portion of the variation in obesity traits. There was insufficient power to estimate heritability in the human sample, nevertheless associations with traits were observed. We also observed evidence of strain-specific shifts in gut microbiota after feeding mice with obesity, high fat/high sucrose (HF/HFS) diet, suggesting a strong effect of host genetics on the plasticity of gut microbiota. Genome-wide association of the panel identified 7 significant loci in mouse genome associated with relative abundances of the families Clostridiaceae, Bacteroidaceae, and Lactobacillaceae. Integrating datasets from multiple intermediate phenotypes (clinical parameters, metabolite levels and gene expression) provided additional support for relationships between specific microbial taxa and metabolic traits. Together, we detected significant correlations between Roseburia intestinalis and metabolic phenotypes in the human study. In addition, we showed similar microbial affinities in human and mice by inferring microbial co-occurrence networks. Understanding how genetic, culture, and diet impact the key factors in the missing heritability of complex diseases and the integration of microbiota data with different intermediate phenotypes may allow us to obtain valuable insights into diseases-associated metabolic pathways.

576T
Interactions between host genetics and diet affect gut microbiota and influence metabolic traits in mouse and human. E. Org1, J. Wha2, Y. Blum3, B. Emert3, E. Y. Kang4, B. Parks5, R. Knight3, T. Drake5, M. Laakso6, E. Eskir7, A. J. Luisi5,7, 1) Department of Medicine, David Geffen School of Medicine, UCLA, Los Angeles, CA; 2) Computer Science Department, UCLA, Los Angeles, CA; 3) Department of Chemistry and Biochemistry, University of Colorado Boulder, CO; 4) Howard Hughes Medical Institute, Boulder, CO; 5) Department of Pathophysiology and Laboratory Medicine, UCLA, CA; 6) Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland; 7) Department of Human Genetics, David Geffen School of Medicine, UCLA, Los Angeles, CA.

An increasing body of evidence suggest that host genetics as well as the diet affects the composition of gut microbiota. Moreover, it is clear that shifts in microbial communities can lead to cardiovascular and metabolic diseases. In our study we examined natural variation in gut microbiota in a population of about 100 inbred strains of mice from the Hybrid Mouse Diversity Panel (HMDP) and a population of a 200 men from the Metabolic Syndrome in Men (METSIM) study. Profiling of the gut microbiota in the HMDP (n= 599 mice) revealed that the genetic background strongly influences gut microbiota composition and that the gut microbiome composition explains a considerable portion of the variation in obesity traits. There was insufficient power to estimate heritability in the human sample, nevertheless associations with traits were observed. We also observed evidence of strain-specific shifts in gut microbiota after feeding mice with obesity, high fat/high sucrose (HF/HFS) diet, suggesting a strong effect of host genetics on the plasticity of gut microbiota. Genome-wide association of the panel identified 7 significant loci in mouse genome associated with relative abundances of the families Clostridiaceae, Bacteroidaceae, and Lactobacillaceae. Integrating datasets from multiple intermediate phenotypes (clinical parameters, metabolite levels and gene expression) provided additional support for relationships between specific microbial taxa and metabolic traits. Together, we detected significant correlations between Roseburia intestinalis and metabolic phenotypes in the human study. In addition, we showed similar microbial affinities in human and mice by inferring microbial co-occurrence networks. Understanding how genetic, culture, and diet impact the key factors in the missing heritability of complex diseases and the integration of microbiota data with different intermediate phenotypes may allow us to obtain valuable insights into diseases-associated metabolic pathways.

577S
Drosophila fragile X mental retardation protein is associated with chromatin and regulates replication stress-induced DNA damage response. Y. Cheng1, W. Zhang2, Y. Li2, Z. Chen2, D. Chen2, P. Jin1. 1) Human Genetics, Emory University, Atlanta, GA, USA; 2) State Key Laboratory of Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, P.R. China.

Fragile X syndrome, a common form of inherited mental retardation, is caused by the loss of the fragile X mental retardation protein (FMRP). As a selective RNA-binding protein, FMRP has been shown predominately localized in cytoplasmic regulatin the translational control. However, it is known that a small portion of FMRP is present in nucleus and its nuclear function has been elusive. We have found that Drosophila dFmr1 is required for replication stress-induced H2Av phosphorylation in DNA damage response (DDR). Replication stress could induce the expression of dFmr1 and promote the nuclear accumulation of dFmr1. We show that dFmr1 is associated with chromatin in a domain-specific manner upon the stimulation of replication stress, which is essential for its ability to induce the phosphorylation of H2Av. Furthermore, we have performed ChIP-seq analyses using dFMR1 antibody and identified specific dFMR1 binding sites in S2 cells. By characterizing the distributions of dFMR1 between control and UV exposure conditions, we found that the sites bound by dFMR1 upon UV treatments significantly overlap with replication origins identified previously. These results together reveal an unexpected nuclear role of FMRP in DDR, and uncover a feed-forward mechanism by which dFmr1 and early DDR induced by replication stress incrcorately regulate each other, thereby synergistically triggers the activity of DDR signaling cascade.
578M Impact of Structural Variants on the three-dimensional multi-scale chromatin conformation of the human genome. D. Piewczynski1,2, J. Kim2, C. Zhang2, A. Malhotra2. 1) ICM, University of Warsaw, Warsaw, mazovia, Poland; 2) The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA; 3) The Centre for Innovative Research, Medical University of Białystok, Białystok, podlasie, Poland.

In this study we investigate the causative relation between copy number variants (CNV) and three dimensional chromatin structure of various cell types. Spatial arrangement of chromosomes can play a very important role in many biological processes. By rendering distant genomic regions to a close spatial proximity it can precipitate long-range interactions between genes and transcription factors. Recently developed Chromosome Conformation Capture (3/4/5C) techniques allow us to quantify interaction frequency between genomic regions and infer their spatial arrangement. On the other hand recent studies such as 1000 Genomes Project suggest that Human genomes are very variable, not only in terms of single mutations, but also large scale structural variants. Our three-dimensional simulation method uses homo-polymer modeling techniques, yet is done separately at different scales: whole nucleus, chromosomes, 10mb regions modeled as random walk, 1mb giant loops structures, 100kb rosettes and random loops, 10kb local neighborhoods. For all spatial scales we use the same evolutionary algorithm that constructs an ensemble of chromatin conformations based on a given experimental contact map containing interaction frequencies. Moreover, in order to overcome limitations of homo-polymer models, we map various local genomic features such as active genes, open chromatin marks, GC ratio and others onto 1D chromatin chain. We also consider structural constraints: resistance to stretching and bending, which accommodate for biophysical chromatin fiber properties. As evolutionary operators we employ three mutation operators (single bead displacement, fragment displacement, fragment rotation) as well as a crossover operator (splitting two structures at a random position and gluing them cross-wise), which are chosen with predefined probabilities. In our study we combine the data being generated by various studies including the 1000 Genomes Project regarding structural variants type, size and location, with their spatial localization in three-dimensional model of the nucleus. We use public HiC maps of the resources from ENCODE project in order to further map the locations of genomic regions. We hypothesize that the shape of chromatin fiber around deletion, duplication, inversion and mobile element insertions is strongly correlated with the type of CNV, suggesting a causative link.

579T Higher order chromatin structure and CFTR gene regulation: roles of cis-regulatory elements and CTCF/cohesin complex. R. Yang, N. Gosalia, J. Kershner, S.H. Leir, A. Harris. Lurie Children’s Research Center and Department of Pediatrics, Northwestern University, Feinberg School of Medicine, Chicago, IL.

The cystic fibrosis transmembrane conductance regulator (CFTR) gene encodes a cAMP-activated chloride ion channel, which when mutated, causes the genetic disease cystic fibrosis. CFTR expression is highly tissue-specific and it has been shown that cis-elements play an important role in the regulation of CFTR. Some of these elements are located far (>100kb) from the gene promoter and regulate gene expression through chromosome looping which establishes physical interactions. In order to better understand the molecular mechanism of CFTR gene regulation and look for additional distal regulatory elements we used 4C-seq (chromosome conformation capture combined with high-throughput sequencing). This technique can identify and quantify interactions between selected genomic sites (viewpoints) and rest of the genome. 4C-seq experiments were performed in human intestinal and airway cell lines that use different cis-regulatory elements, and also in primary human airway and epididymis cell cultures. Viewpoints were used at the CFTR promoter, at intronic and distal enhancers, and at CTCF-binding insulators flanking the gene. The data correlate well with previous 3C data we generated for the CFTR locus and also reveal potential novel cis-regulatory elements. Moreover, results confirmed the cell type-specific properties of these chromatin interactions. We subsequently performed studies in which we 1) manipulated the levels of the architectural proteins CTCF and cohesin, or 2) removed specific cis-regulatory elements from the endogenous CFTR gene by CRISPR/Cas9-mediated targeting and examined the impact on 3D locus structure and CFTR expression. These data provide substantial insight into higher order chromatin structure mediated by both cis-regulatory elements and the CTCF/cohesin complex and their critical role in CFTR gene regulation.


In contrast with model organisms, human cell lines have not historically been readily compatible with mutagenesis screens to identify genes involved in specific cellular phenotypes. Recent advances in genome editing technologies - specifically, with the CRISPR/Cas9 system - have enabled efficient genome-wide recessive genetic screening in haploid human cell lines (Wang et al. Science 2014). However, taking advantage of haploid genetics limits the choice of a cellular model primarily to KBM-7, a near-haploid chronic myelogenous leukemia cell line, and its haploid derivative, HAP1. While both have been used successfully in previous studies, HAP1 has not been well characterized. HAP1 was derived in a failed attempt to induce pluripotency and has several notable distinctions from KBM-7, including the fact that it is adherent and that it no longer maintains a second copy of chromosome 8 (Carette et al. Nature 2011), but little else has been reported regarding its cellular state. To facilitate further efforts to dissect the genetic architecture of human phenotypes in these haploid cellular models, we have generated RNA-seq and ATAC-seq data from the HAP1 cell line. For comparison, we are also generating equivalent datasets from other myeloid cell lines, including the parental line (KBM-7), and a lumphoblastoid cell line (GM12878) that has been extensively characterized in previous studies. Beyond serving as a reference dataset for further experimentation, characterization of the nuclear architecture of a haploid cell line may offer insights into fundamental human biology. Both the RNA-seq and ATAC-seq datasets allow for hierarchical clustering of the samples to better understand how related their cellular phenotypes are. Furthermore, genes differentially expressed based on lineage will be used to guide closer examination of the results may also yield a more nuanced understanding of gene regulation in human cells.

581M A Computational Pipeline for Characterization, Identification, and Significance Analysis of the Somatic Copy Number Variations from Genome Sequencing Datasets. A. Harmanci1, A. Serin Harmanci2, M. Gunel2, M.B. Gerstein1. 1) Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06527, USA; 2) Departments of Neurosurgery and Genetics, Yale Program in Brain Tumor Research, Yale School of Medicine, New Haven, CT 06510, USA.

Large scale genomic aberrations, namely the structural variations, are important driving factors for tumorigenesis and tumor evolution. Among these variants, the unbalanced variants, the copy number variants, are significant because they may change the dosage of the tumor suppressor or oncogenes and drive the formation of tumor. We are presenting a complete pipeline for identification and significance analysis for somatic CNVs using genome sequencing (exome and whole genome) datasets. Our pipeline takes multiple datasets as input and uses a signal processing based segmentation method to identify CNVs for each sample. Following CNV identification, pipeline utilizes a genome wide test to assign a significance of recurrence for each copy number variant segment. In addition, our pipeline estimates the ploidy, tumor purity and clonal fraction for each segment, generates the closest subclonal events and computes significance for these segments. We compare our pipeline with several other CNV identification tools and present benchmarking results. We also applied our pipeline on several exome sequencing datasets of patients with brain tumor and present significantly altered clonal and subclonal events including tumor suppressor and oncogenes. In addition to coding elements, we also generate significantly altered non-coding elements.
582T

Method for classifying candidate structural variants into true positives and false positives. H. Parikh1, J. Zook1, M. Pratt2, G. Bartha2, M. Eberle3, M. Sait1. 1) Genome Scale Measurements Group, Biosystems and Biomatics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA; 2) Personalis Inc., 1350 Willow Road, Suite 202, Menlo Park, CA 94025, USA; 3) Illumina Inc., San Diego, CA 92122, USA.

The human genome contains genomic variants ranging in size from single nucleotide polymorphisms to large structural variants (SVs). SVs include variations such as novel sequence insertions, deletions, inversions, mobile-element insertions, tandem duplications, interspersed duplications and translocations. SVs have long been implicated in phenotypic diversity and human diseases; however, identifying SVs in a whole genome has proven elusive. Recent advances in next-generation sequencing technologies have facilitated the analysis of SVs in unprecedented detail. However, clinical adoption of human genome sequencing requires methods with known accuracy of SV calls across genome. Hence, we developed methods to make highly confident SV and non-SV calls for NA12878, the pilot genome for the Genome in a Bottle Consortium, for which we recently published high-confidence small variants and homozygous reference sites. Because numerous methods have been developed to find candidate SVs, we decided to develop methods to look for evidence of these SVs in mapped sequencing reads from multiple sequencing technologies. We collected 9 different Gold Standard SV call sets totaling 12,738 deletions that were derived from pedigree sequencing, microarrays, and other validation methods. We annotated these candidate SV sites with parameters such as mean coverage, mean insert-size, numbers of discordant pairs, numbers of soft clipped reads, mean mapping quality score, and numbers of heterozygous and homozygous SNP genotype calls from 3 sequencing technologies (Illumina, Moleculo, and PacBio). In addition, we annotated randomly generated regions to understand characteristics of non-SV regions. Graphical visualization of the annotation parameters has shown clear distinction between true positive and false positive SVs. A key advantage of the proposed method is its simplicity and flexibility to generate various annotation parameters from aligned sequence data based on different sequencing datasets from the same genome. This allows integration of multiple sequencing datasets to identify high-confidence SV and non-SV calls that can be used as a benchmark to assess false positive and false negative rates. We are currently creating a classification method based on the annotation parameters to generate both high-confidence SV calls and high-confidence non-SV calls for genomes selected as reference materials by the Genome in a Bottle Consortium.

583S

Genomic CNVs can cause sudden infant death syndrome (SIDS). A. Pfeufer1,2, M. Arnold3, M. Cohen4, S. Hersms5, D. Dörn6, T.A. Plötz7, I. Sinicina8, E.A. Mitchell9, M. Donner10, D.T. Mage11, E.R. Behr12, T.W. Mühleisen13, S. Cichon4, T. Meiling12, A. Peters14, H.W. Mewes1,2, M. Klintschar5, T. Bajanowski12, M. Vennemann1,2, 1) Helmholtz Zentrum München, Neuherberg, Germany; 2) TU München, München, Germany; 3) University of Sheffild, Sheffild, UK; 4) University of Basel, Basel, Switzerland; 5) MHH Hannover, Hannover, Germany; 6) LMU München, München, Germany; 7) University of Auckland, Auckland, New Zealand; 8) Dupont Haskell Global Centers for Health & Environmental Sciences, Newark, USA; 9) Biomolecular Core Laboratory, AI duPont Hospital for Children, Wilmington, USA; 10) St. Georges University of London, London, UK; 11) Forschungszentrum Jülich, Jülich, Germany; 12) Universität Essen, Essen, Germany; 13) Universität Münster, Münster, Germany.

The contribution of monogenic disorders to sudden infant death syndrome (SIDS) is established (most commonly LQTs and MCAD deficiency). In contrast the importance of complex genetic predispositions is less clear. Also copy number variations (CNV) involving critical genes may predispose to SIDS. Patients: We performed genome-wide SNP and CNV genotyping of 368 SIDS cases using the Illumina HumanHap 660x3 quad array. 320 cases originated from the multi-center German study on sudden infant death syndrome (GeSID) and 48 cases were recruited from the Sheffield SIDS study in the U.K. As controls we used 823 population based individuals from the KORA cohort. Methods: In a GWAS design we compared SIDS cases with population controls for genome-wide SNP association. As age matching is not feasible in SIDS we performed sex matching of controls and geographical adjustment by multidimensional scaling. We included Chr.X markers as SIDS exhibits 2:1 male:female sex bias indicating possible genetic risk factors on the sex-chromosomes. Markers were imputed to 4.8 Mio SNPs using the haplotype backbone provided by the 1000 Genomes project. Autosomal SNP-markers were analyzed using an additive model adjusted for sex. Markers on Chr. X were analyzed stratified by sex and then meta-analyzed. In another line of analysis we called genome-wide CNVs in the SIDS cases using the CRLMM algorithm. Results: The GWAS analysis showed no significant association beyond genome-wide significance level (p<10-6) neither with autoimmune markers nor with markers on Chr. X. CNV analysis revealed deletions in five cases: One case had a 5.2 Mb deletion in 5q3. another case had a 2.2 Mb deletion in 18p1 and a third case had a 370 kb deletion in 5q2. Two further unrelated cases had nearly identical 400 kb deletions in 18p1. None of these CNVs involved genes previously identified to be involved with SIDS or Mendelian diseases involved with premature death. Conclusion: Our GWAS was powered to detect common CNP association signals with OR≥1.8 and MAF≥0.2 with ≥95% power. The failure to detect such associations in the human genome is well in accordance with the expectation of strong evolutionary selection against any common genetic SIDS risk factors. The detection of CNVs in 1.4% of cases (5/368) indicates a small but significant contribution of structural genomic mutations to SIDS etiology.
584M

In vivo dissection, causality identification, and network analysis of copy number variants associated with autism spectrum disorders. C. Gozio1,2,1. S. Moon1, J.P. Willer1, M.A. Savidan1, S.C. Brodsky2, T. Blumen-tha2, A. Ragavendran2, M.E. Talkowski4, N. Katsanis1. 1) Center for Human Disease Modeling, Duke University Medical Center, Durham, NC, USA; 2) Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA; 3) Department of Neurology, Massachusetts General Hospital, Boston, MA, USA; 4) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA; Harvard Medical School, Boston, MA, USA.

Copy number variants (CNVs) are frequent lesions involved in both rare and complex human traits. This has raised the challenge of identifying which genes within a CNV drive clinical traits. We have shown previously how the combinatorial use of surrogate phenotypes in zebrafish embryos and genomic studies can help dissect CNVs on 1q21.1, 8q24.3, and 16p11.2. However, although successful in identifying primary drivers, these data also indicated that each of the primary dosage-sensitive transcripts within each CNV were not sufficient to drive the CNV-associated pathology. We therefore extended our study to model cis-eQTPs within CNVs and to determine global transcriptional changes and dissect likely contributory transcripts. We focused on the 1q21.1 CNV, encompassing a minimal nine-gene region, deletion of which is associated with microcephaly, while the reciprocal duplication is associated with macrocephaly. Systematic overexpression and suppression of all genes in the CNV showed that dose perturbation of the chromodomains-helicase-DNA-binding protein CHD1L gave significant head size changes: overexpression of human CHD1L mRNA led to macroce- phy, while suppression of chd1l lead to a significant decrease in head size. However, given that a) human genetic data suggested that CHD1L might not be sufficient to explain causality of this CNV, we asked whether cis-CNV and trans-CNV genetic interactions might contribute to pathogene-sis. To answer the first question, we systematically tested the pairwise interaction of CHD1L with each of the other 8 genes within the CNV, through which we identified several epistatic partners that modulate the expressivity of micro- and macrocephaly. For the second question, we generated RNA-seq data from heads of control, CHD1L RNA- and chd1l MO-injected embryos, and identified 700 differentially expressed transcripts; strikingly, the top 20 genes in this cohort have been implicated in the genetic causality of autism. Epistatic analysis of candidates whose dosage is perturbed reciprocally by the CNV showed strong complementation for a subset of genes, indicating that their interaction is likely related to the observed neurodevelopmental phenotypes. The combination of CNV dissec- tion with in vivo genetic interaction studies highlights the genetic complexity of genomic lesions and informs functional networks relevant to ASD patho-mechanism.

585T

Increased relative mitochondrial DNA content in Keratocystic Tumors. A. Kondkar1, K. Abu-Amor1,2, T.A. Azad1, T. Sultan1, H. Kalantari1, A. Al-Muawad1. 1) Department of Ophthalmology, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 2) anterior Segment Unit, Department of Ophthalmology, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 3) Department of Ophthalmology, College of Medicine, Univers-ity of Florida, Jacksonville, Florida, USA.

Mitochondrial DNA (mtDNA) is extremely prone to oxidative stress. To provide evidence for the possible association of oxidative stress with Keratocystus, we estimated the changes in relative mitochondrial DNA (mtDNA) content. The study included 119 patients with Keratocystus and 208 controls matched for gender, ethnicity and systemic disease status. We selected controls with higher age group than those of the patients as the mtDNA copy number trend to increase with age. The age mean (SD) was 26.4 (7.6) and 54.5 (14.4) y for patients and controls respectively. Within this cohort, mtDNA was estimated by real-time quantitative polymerase chain reaction (qPCR) method using ND1 as a mtDNA gene and human globulin (HGB; also known as cytoglobin gene - CYGB) as the reference single-copy nuclear gene. The mean relative mtDNA content was found to be significantly higher in patients with Keratocystus (1.20 copies) than the normal control subjects (1.04 copies) [mean difference = 0.16 copies, 95% confidence interval (CI) = 0.07 to 0.25; P = 0.0004]. Subjects with high mtDNA content (>1.259 copies i.e., greater than 75th percentile) were found to be at an increased risk of the disease (odds ratio = 2.62, 95% CI = 1.40 to 4.89; P = 0.0025). High mtDNA content appears to be associated with increased risk of Keratocystus. Increased mtDNA content in Keratocystus patients may imply a response to oxidative stress, possibly in part because of mitochondrial respiratory chain defects.

586S

Structural haplotypes of the human amylase locus and their relationship to obesity. C.L. Usher1, R.E. Handsaker1,2, T. Esko2,2, J.E. Moon1,2, A. Matsgaluri1, J.N. Hirschhorn1,2, S.A. McCarthy1,2. 1) Genetics, Harvard Medical School, Boston, MA; 2) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA; 3) Estonian Genome Center, University of Tartu, Tartu, Estonia.

The three human amylase genes (AMY2B, AMY2A, and AMY1) encode enzymes that digest starch into sugar. These genes reside in a complex genomic locus characterized by inversions, deletions, and tandem duplications. Only a handful of the structural haplotypes are known, yet there is evidence of other, unknown haplotypes segregating in populations.

We first identified these structural haplotypes, as well as their population frequencies and evolutionary history, by tracking their transmission in a hundred father-mother-offspring trios using droplet digital PCR (ddPCR) and sequencing read-depth analysis. We found 10 structural haplotypes, each with different combinations and copy numbers of AMY2B, AMY2A, and AMY1. These structures help define a long history of recurrent mutation and provide clues for more accurate copy number genotyping at this locus.

The presence of higher AMY1 copy numbers in populations with high starch diets has led to speculation that AMY1 copy number contributes to obesity. Supporting this idea, a recent study reported that AMY1 copy num-ber has the largest effect yet observed for common variation on obesity. To investigate the amylase locus’s effect on BMI, we used the advanced meth-ods that we had already developed for amylase to genotype copy number for all amylase genes in 500 Estonians from each of the BMI extremes (BMI < 22 and BMI > 33). This design had >99% power to detect effects one-tenth as strong as those reported in the recent study. As expected, we observed strong associations between obesity and the copy number of any amylase gene. However, we found no association between obesity and the copy number of any amylase gene.

587T

Constitutional chromothripsis: A novel phenomenon in congenital dis-orders. A. Alhariri1, M. Pronold1, A. Slavolivé1,2, J. Shieh1,2. 1) Department of Pediatrics, Division of Medical Genetics, UCSF, San Francisco, CA; 2) Department of Pathology and Laboratory Medicine, UCSF, San Francisco, CA; 3) Institute of Human Genetics, UCSF, San Francisco, CA.

Chromothripsis (chromosomal shattering) is characterized by complex multiple genomic rearrangements recently observed in cancer genomes and complex human traits. This has raised the challenge of identifying which genomic locus characterized by inversions, deletions, and tandem duplications. Only a handful of the structural haplotypes are known, yet there is evidence of other, unknown haplotypes segregating in populations. A novel phenomenon in congenital disorders, it will be important to recognize this phenomenon.

In this study, we investigated the amylase locus’s effect on BMI, we used the advanced methods that we had already developed for amylase to genotype copy number for all amylase genes in 500 Estonians from each of the BMI extremes (BMI < 22 and BMI > 33). This design had >99% power to detect effects one-tenth as strong as those reported in the recent study. As expected, we observed strong associations between obesity and the copy number of any amylase gene. However, we found no association between obesity and the copy number of any amylase gene.

These studies help define a long history of recurrent mutation and provide clues for more accurate copy number genotyping at this locus.

We first identified these structural haplotypes, as well as their population frequencies and evolutionary history, by tracking their transmission in a hundred father-mother-offspring trios using droplet digital PCR (ddPCR) and sequencing read-depth analysis. We found 10 structural haplotypes, each with different combinations and copy numbers of AMY2B, AMY2A, and AMY1. These structures help define a long history of recurrent mutation and provide clues for more accurate copy number genotyping at this locus.

The presence of higher AMY1 copy numbers in populations with high starch diets has led to speculation that AMY1 copy number contributes to obesity. Supporting this idea, a recent study reported that AMY1 copy number has the largest effect yet observed for common variation on obesity. To investigate the amylase locus’s effect on BMI, we used the advanced methods that we had already developed for amylase to genotype copy number for all amylase genes in 500 Estonians from each of the BMI extremes (BMI < 22 and BMI > 33). This design had >99% power to detect effects one-tenth as strong as those reported in the recent study. As expected, we observed strong associations between obesity and the copy number of any amylase gene. However, we found no association between obesity and the copy number of any amylase gene.
588T End-point zygosity and CNV determination from crude samples. C. Liu, S. Dong. MBS R & D, Life Technologies, south san francisco, CA. End-point zygosity and CNV determination based on PCR Rn/Rm reading is expected to reduce the cost and increase throughput as compared to current real time Cq-based measurement. There is an immediate need by AgBio customers to screen seed zygosity and copy number of transgenes through end-point reading. Moreover, Agbio customers expect to determine zygosity and CNV using crude samples in order to stream-line their workflow. For end-point quantification, the major challenge is to control the saturation of PCR to maintain the segregation of end-point intensity according to the input copy number of the target transgenes in reference to a control gene. We have tried a few approaches such as Asymmetric PCR, ARCS (Amplification Ratio Control System) PCR, Dual Tailing PCR, pyrophosphate removal and controlling plateau of PCR by running a third PCR in the background. Among the tested approaches, controlled plateau of PCR (CoP'ed PCR) showed best separation of copy number through end-point PCR reading. We further optimized CoP'ed PCR conditions using purified DNA and crude plant and blood samples. By CoP'ed PCR, we are able to do zygosity with crude plant samples to satisfy the immediate needs for AgBio customers. Furthermore, since this approach improves the sensitivity for the copy number separation for not only end-point but also real-time PCR for crude blood samples, in the future, it can be used for CNV analysis directly from human blood, such as prenatal chromosome copy number variation detection since the percentage of fetal DNA in the maternal blood sample is only 10%-15%. It is impossible to detect the copy number abnormality through traditional real time or end-point PCR.

589S Sequencing the genomes of single cells. P. Ribaux, C. Borel, F. Santoni, E. Falconnet, S.E. Antonarakis. Dept Genetic Medicine, Univ Geneva Medical School, Geneva, Switzerland. Whole-genome amplification and next-generation sequencing advances enable investigation of somatic structural and nucleotide variation to single-cell resolution. The ultimate goals of our study are (i) to identify disease-associated somatic mutations and (ii) to uncover the extent of low-abundance DNA variations in individual cancer cells in order to underlie mechanisms of tumor evolution. Because of the technical challenge of detecting and analyzing genomic heterogeneity among single cells, we first analyzed individual cells in culture and tested the robustness of our experimental workflow. We choose the K562 cells, a human immortalized myelogenous leukemia line and F-121, a human primary Trisomy 21 fibroblast cell line. We used the C1 Single Cell Auto Prep System (Fluidigm) to capture hundreds of individual cells and to generate high quality of individual amplified DNA. So far, 96 barcoded whole-exome libraries were sequenced at deep coverage (PE, 100bp). Variant calls (CNVs and SNVs) were generated with an in-house analysis pipeline. Here, we will discuss the amplification uniformity, the detectable fraction of the exome and the level of DNA contamination. By comparing single cells and bulk of cells datasets, we will assess the percentage of allelic drop out for each single exome based on the heterozygous SNVs. High quality single-cell genome sequence will greatly enhance the genetic analysis of somatic genomic disorders. C.B. and P.R. contributed equally.

590M Deletions of Regulatory Boundaries are Associated with Congenital Disease. M. Spielmann1,2, J. Ibn Salem1,2,3, D. Kohler1, M.I. Love4,5, H.R. Chung4,6, N. Huang6, M.E. Hurles6, M. Haendel7, N.L. Washington8, D. Smedley7, C.J. Mungall9, S.E. Lewis9, C.E. Ott5, S. Bauer1, P. Schoeld1,10, S. Mundlos1,2, P.N. Robinson1,2,3,4, 1) Institute for Medical and Human Genetics, Charite-Universitätsmedizin Berlin, Berlin, Germany; 2) Max Planck Institute for Molecular Genetics, Ihnestr. 63, 14195 Berlin, Germany; 3) Department of Mathematics and Computer Science, Free University Berlin, Takustr. 9, 14195 Berlin, Germany; 4) International Max Planck Research School for Computational Biology and Scientific Computing, Ihnestr. 6373, 14195 Berlin, Germany; 5) Otto Warburg Laboratory, Ihnestr. 63/73, 14195 Berlin, Germany; 6) Wellcome Trust Sanger Institute, CB10 1SA Hinxton, UK; 7) Oregon Health & Science University, Department of Medical Informatics & Clinical Epidemiology, 7239 Portland, OR, USA; 8) Lawrence Berkeley National Laboratory, Mail Stop 84R0171, 54720 Berkeley, CA, USA; 9) The Jackson Laboratory, 04609 Bar Harbor, ME, USA; 10) University at Cambridge, Department of Physiology, Development and Neuroscience, Downing Street, CB2 3EG Cambridge, UK. Recent data from genome-wide chromosome confirmation capture analysis (Hi-C) indicate that the human genome is divided into conserved mega-base-sized local chromatin self-interacting regions called topological domains. These topological domains form the regulatory backbone of the genome and are separated by regulatory boundary elements or barriers. Copy-number variations (CNVs) can potentially alter the topological domain architecture by deleting or duplicating the barriers and thereby allowing enhancers from neighboring domains to ectopically activate genes causing misexpression and disease, a mutational mechanism that has recently been termed "enhancer adoption". In this study the Human Phenotype Ontology (HPO) database was used to relate the phenotypes of 922 deletion cases recorded in the DECIPHER database to monogenic diseases associated with genes in or adjacent to the deletions. We sought combinations of tissue-specific enhancers and genes adjacent to the deletion and associated with phenotypes in corresponding tissue, whereby the phenotype matched that observed in the deletion. We compared this computationally with a gene-dosage pathmechanism that attempts to explain the deletion phenotype based on enhancer insufficiency and genes located within the deletions. Up to 11.8% of the deletions could be best explained by enhancer adoption or a combination of enhancer adoption and gene-dosage effects. Our results suggest that enhancer adoption caused by deletions of regulatory boundaries can contribute to a subset of syndromes and should thus be taken into account for their medical interpretation.

591T Sensitive and efficient analysis of somatic mosaicism using genome-wide SNP array data. M. Volpato1,2,3,4, K. Heffelfinger1,2,5, M. C. Smedley1,2, S. Shih1,2, J. Chung1,2,3,4, 1) Max Planck Institute for Molecular Genetics, Ihnestr. 63, 14195 Berlin, Germany; 2) Max Planck Institute for Molecular Genetics, Ihnestr. 63/73, 14195 Berlin, Germany; 3) Otto Warburg Laboratory, Ihnestr. 63/73, 14195 Berlin, Germany; 4) Wellcome Trust Sanger Institute, CB10 1SA Hinxton, UK; 5) Oregon Health & Science University, Department of Medical Informatics & Clinical Epidemiology, 7239 Portland, OR, USA; 6) Lawrence Berkeley National Laboratory, Mail Stop 84R0171, 54720 Berkeley, CA, USA; 7) The Jackson Laboratory, 04609 Bar Harbor, ME, USA; 8) University at Cambridge, Department of Physiology, Development and Neuroscience, Downing Street, CB2 3EG Cambridge, UK. The relevance of somatic mutation to human phenotype. Among the challenges to practical survey of somatic mutations is that detection methods must be at once sensitive (since mutations may exist in only a small fraction of sampled cells), efficient (since the rate of mosaicism is low and many samples must be analyzed), and agnostic (since mutation locations are not known a priori). hapLOH, a computational algorithm we developed, provides sensitive, efficient, and agnostic profiling of somatic chromosomal imbalance using best-guess haplotype estimates and whole-genome SNP array data. We used hapLOH to characterize mosaicism in 31,223 individuals from 10 studies conducted as part of the GENEVA Consortium. We called 1,134 mosaic mutations (ranging in size from 42 Kb to 146 Mb [all of chromosome 8]) in 895 individuals, which is a substantially higher rate than the original published analysis of these data (Laurie et al., Nat. Gen., 2012). The results validate recent predictions that the extent of somatic mosaicism has been underestimated due to lack of sensitivity for low-frequency mutations, and confirm the value of using haplotype information for sensitive detection of allelic imbalance. We describe the locations of recurrent somatic mutations, which are likely to contain genes important for proliferation, and copy number classifications for events with sufficient log R ratio evidence. Since hapLOH is sensitive to very low-frequency events, more than 90% of the observed calls cannot be classified even though the signal of imbalance is strong. Sample-specific empirical null simulations and comparisons of observed B allele frequency and log R ratio shifts to predicted values suggest the rate of false positives is low. Incidentally, we found that depressed phase concordance can distinguish inherited from somatic mutations.
Monozygotic Twin Pairs: CNV and sequence concordance. A. Abdellau"1, E. Eh"1, J.J. Hottenga"2, Z. Weber"1, H. Mbabak"1, G. Willemsen"1, T. van Beijsterveldt"1, A. Brookes"3, J.J. Hudson"1, E.C.J. de Geus"1, K. Ye"1, P.E. Slagboom"1, G.E. Davies"1, D.I. Boomsma"1.

1) Biological Psychology, VU University Amsterdam, Amsterdam, Noord Holland, Netherlands; 2) Avera Institute for Human Genetics, Avera McKennan Hospital & University Health Center, Sioux Falls, SD, USA; 3) Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA.

Overall, our conclusions would not be accessible to those measurement in that study may also have been subject to a similar systematic error. This CNV as a quasi-continuous variation independent of

The large putative de novo CNVs detected with microarray data were not present in the subsample that had whole-genome sequence data available. The large putative de novo CNVs detected with microarray data were not in the sample that had whole-genome sequence data available. We do expect whole-genome sequence data to allow us to search for smaller de novo CNVs that cannot be detected with micro-array data.


We have now also defined less common haplotypes containing even numbers of AMY1 and AMY2A/2B, which are associated with CNVs of pancreatic amylase genes. Most published works have measured CNVs from 15q11.2 to 1.2 for a first series of qPCR validation experiments. Two out of eleven post-twinning de novo CNVs were validated with qPCR in the same twin pair. This 13-year old twin pair did not show large phenotypic differences. The remaining putative de novo CNVs from 15q11.2 were found significantly more often in older twins, suggesting that these were capturing larger CNVs. The large putative de novo CNVs detected with microarray data were not present in the subsample that had whole-genome sequence data available. We do expect whole-genome sequence data to allow us to search for smaller de novo CNVs that cannot be detected with micro-array data.

Inverted repeats mediate complex genomic rearrangements including quadruplication. C.R. Beck1, C.M.B. Carvalho1,2, L. Banser2, T. Gambin3, D. Stubbolato, B. Yuan1, K. Sterle2, S.M. McCahan3,4, M. Henneke1, P. Seemann1, G.M. Hobson1,2, J.R. Lupski1,2,6,7.

1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Centro de Pesquisas René Rachou - PUCRS, Belo Horizonte, MG, Brazil; 3) Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE; 4) Jefferson Medical College, Thomas Jefferson University, Philadelphia, PA; 5) University Medical Center Göttingen, Georg August University, Department of Pediatrics and Adolescent Medicine, Division of Pediatric Neurology, Germany; 6) Department of Pediatric Neurology, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic; 7) University of Delaware, Department of Biological Sciences, Newark, DE; 8) Department of Pediatrics and Human Genome Sciences, Novartis, Children’s Hospital, Houston, TX.

Inverted repeats (IRs) predispose loci to inversions and can lead to copy number alterations by mediating complex genomic rearrangements (CGRs). Clustering of breakpoints in cases of duplication and triplication of the dosage sensitive PLP1 gene occurs within three sets of distal (telomeric), nested inverted repeats. These CGRs result in Pelizaeus-Merzbacher disease (PMD, MIM #312080), and approximately 80% of PMD individuals have amplification of PLP1. We systematically investigated the region encompassing PLP1 in both unaffected and affected individuals to identify the IRs responsible for the CGRs and the mechanism of rearrangement. In phenotypically normal individuals, we identified a copy number neutral polymorphism between two IRs that tends to be passed on to affected individuals. Breakpoint analysis suggested that the inversion is prevalent in the population, with a frequency of approximately 40%. Furthermore, integrating these data with sequence information for the locus showed that inversion alleles were present on two haplotypes in the families of affected patients. Therefore, there is a common polymorphic inversion at the locus that appears to be recurrent. We have assembled a cohort of 17 individuals affected with PMD who have breakpoints clustering within the distal IRs. The 17 patients all have CGRs, and duplication-inverson triplication-duplication rearrangements were observed in 16 individuals. Through genome-wide linkage analysis and linkage disequilibrium mapping, Southern blotting, quantitative PCR, and cloning methodologies, we determined that the IRs responsible for the inversion are implicated in these rearrangements. Breakpoint junction analysis has underscored the presence of point mutations close to CGRs that are repulsive in origin and the presence of additional template switches at 5’/3’ or ~30% of these events. Finally, in one patient, a CGR resulting in duplication of PLP1 and quadruplication of a proximal genomic region was identified. In this patient, junction sequencing revealed that the mechanism of formation of quadruplication involved imperfect resolution of a candidate breakpoint junction flanked by microhomology-mediated break-induced replication (MMEBR) model. Thus, detailed analysis of the PLP1 locus has provided experimental evidence for another mechanism by which inverted repeats affect genomic change.

Inverted repeats mediate complex genomic rearrangements including quadruplication. C.R. Beck1, C.M.B. Carvalho1,2, L. Banser2, T. Gambin3, D. Stubbolato, B. Yuan1, K. Sterle2, S.M. McCahan3,4, M. Henneke1, P. Seemann1, G.M. Hobson1,2, J.R. Lupski1,2,6,7.

1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Centro de Pesquisas René Rachou - PUCRS, Belo Horizonte, MG, Brazil; 3) Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE; 4) Jefferson Medical College, Thomas Jefferson University, Philadelphia, PA; 5) University Medical Center Göttingen, Georg August University, Department of Pediatrics and Adolescent Medicine, Division of Pediatric Neurology, Germany; 6) Department of Pediatric Neurology, 2nd Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic; 7) University of Delaware, Department of Biological Sciences, Newark, DE; 8) Department of Pediatrics and Human Genome Sciences, Novartis, Children’s Hospital, Houston, TX.

Inverted repeats (IRs) predispose loci to inversions and can lead to copy number alterations by mediating complex genomic rearrangements (CGRs). Clustering of breakpoints in cases of duplication and triplication of the dosage sensitive PLP1 gene occurs within three sets of distal (telomeric), nested inverted repeats. These CGRs result in Pelizaeus-Merzbacher disease (PMD, MIM #312080), and approximately 80% of PMD individuals have amplification of PLP1. We systematically investigated the region encompassing PLP1 in both unaffected and affected individuals to identify the IRs responsible for the CGRs and the mechanism of rearrangement. In phenotypically normal individuals, we identified a copy number neutral polymorphism between two IRs that tends to be passed on to affected individuals. Breakpoint analysis suggested that the inversion is prevalent in the population, with a frequency of approximately 40%. Furthermore, integrating these data with sequence information for the locus showed that inversion alleles were present on two haplotypes in the families of affected patients. Therefore, there is a common polymorphic inversion at the locus that appears to be recurrent. We have assembled a cohort of 17 individuals affected with PMD who have breakpoints clustering within the distal IRs. The 17 patients all have CGRs, and duplication-inverson triplication-duplication rearrangements were observed in 16 individuals. Through genome-wide linkage analysis and linkage disequilibrium mapping, Southern blotting, quantitative PCR, and cloning methodologies, we determined that the IRs responsible for the inversion are implicated in these rearrangements. Breakpoint junction analysis has underscored the presence of point mutations close to CGRs that are repulsive in origin and the presence of additional template switches at 5’/3’ or ~30% of these events. Finally, in one patient, a CGR resulting in duplication of PLP1 and quadruplication of a proximal genomic region was identified. In this patient, junction sequencing revealed that the mechanism of formation of quadruplication involved imperfect resolution of a candidate breakpoint junction flanked by microhomology-mediated break-induced replication (MMEBR) model. Thus, detailed analysis of the PLP1 locus has provided experimental evidence for another mechanism by which inverted repeats affect genomic change.
596M

The 600kb 16p11.2 duplication and deletion (BP4-BP5) are among the most frequent copy number variants, and the duplication is associated with neurodevelopmental disorders, including autism spectrum disorder and schizophrenia, and a decrease in head circumference and BMI. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number. The 16p11.2 locus, like duplications of other genomic regions, shows highly variable expressivity and incomplete penetrance. To decipher the complex effects of the 16p11.2 duplication on cognitive, behavioral, medical, and anthropometric traits, we studied a cohort of 270 duplication carriers as well as a group of 184 deletion carriers. As full length and truncated (69% of full length amino acids) transcripts. Sequencing and RT-PCR analysis revealed that one such CNV breakpoint led to transcript truncation but successful RNA expression in sperm. This suggests that the regulation of expression and behavior of spanxb is complex and that the regulation of expression and behavior is not entirely explained by differences in copy number.
599M

Identifying Population-Specific Structural Variation in Human Blood Group Genes. K. Fox1, T. Shafer1, D. Crosslin1, M. Delaney2, J. Johnson1, D. Nickerson1. 1) School of Medicine, Genome Sciences, University of Washington, Seattle, WA, USA; 2) Puget Sound Blood Center, Research Institute, Seattle WA, USA; 3) University of Washington, Department of Laboratory Medicine, Seattle, WA, USA; 4) University of Washington, Department of Medicine, Seattle, WA, USA.

Blood type variation through hemagglutination is the clinical gold standard. Assigning units of blood based on serological typing can sometimes fail to identify less common blood type variants, risking exposure of the recipient to clinically significant allo-antigens. Blood types reflect genetic variation at blood group genes, which harbor low frequency variants encoding hundreds of blood types not detected by conventional blood typing. High throughput sequencing approaches, such as exon sequencing, have the potential to assign high resolution blood types based on DNA sequence data. RH (RHD, RHCE) and MNS (GYP A, GYPB, GYPE) blood group genes contain common; structural variants (SVs) that result from unequal crossing-over or gene conversion events. However, the influence of SV characterization on blood typing for cross-matching has yet to be fully explored.

Alioasensialization is most common in chronically transfused patients, particularly in sickle cell disease (SCD) which affects 1 in 500 African-Americans. SCD patients are exposed to multiple blood donors over their lifetimes, and are at high risk (>20%) of developing significant alloantibodies. The National Heart, Lung, and Blood Institute (NHLBI) funded the Exome Sequencing Project (NHLBI-ESP), and, through this effort, subjects of African-American (n=1,714) and European ancestry (n=3,405) with heart, lung, and blood-related phenotypes were exome sequenced. Using these data, we present read-depth-based SV calls, including 70 unique duplications and 66 unique deletions in six human blood group genes (ABO, RHD, RHCE, GYP A, GYPB, GYPE) from 5,119 exomes with >20X coverage. We confirmed known deletions (Rh D-, MNS GPB null) and have identified novel SV in the ABO blood group gene (deletion of ABO exons 5–7; predicting blood type O). These analyses highlight the need for further studies not only to characterize the diversity of SV in these genes, but also to explore the phenotypic effects of such variation in the context of transfusion medicine.

600T

Characterization of Copy Number Variation and Loss of Heterozygosity Using high resolution SNP MicroArray -The Miami Experience. G. Ghafar1, B.J. Iliagan2, C. Hung3, S.A. Hosseini4, P. Berkner5, R. Yusupov2, L. Brenton1, M. Rodiguez1, B. Johnson1, O.A. Bodamer1. 1) Division of Clinical and Translational Genetics, Dr. John T. Macdonald Foundation, Department of Human Genetics, University of Miami, Miller School of Medicine, Miami, FL 33136, USA; 2) Memorial Regional Hospital, 1150 N 35th Ave # 490, Hollywood, FL, 33021, USA.

Background: MicroArray analysis is considered the first-tier clinical diagnostic test for individuals with unexplained developmental delay/intellectual disability (DD/ID), autism spectrum disorders (ASD), and/or multiple congenital anomalies (MCA). In addition to clarification of chromosome deletions and duplications; MicroArray provides information regarding long contiguous stretches of homozygosity (LCSH). Here we report the results of MicroArray analysis from patients with DD, ASD, and MCA referred to the Clinical Molecular Genetic Laboratory (CMGDL), at the University of Miami from June 2012 to May 2014. Methods: The Illumina HumanOmn1-Quad or CytoSNP-850K BeadChip (illumina,CA) array platforms were used to determine the presence of Copy Number Variations (CNVs) and Loss of Heterozygosity (LOH) in our study population. Results: Pathogenic copy number variations were detected in 19/150 (12.6%) patients ranging from 77kb to whole chromosome gain or loss. Pathogenic CNVs were detected in the following chromosomes: chr. 16 in 5 patients (16p11.2;16p12.2; 16q23.1q23.2); X chr in 5 patients, including a mosaic Turner syndrome (chromosome 10); 17 in 4 patients, including chr. 17 (p11.2;17q25.3); 18 in 2 patients (18q11.22q11.23; 18q12.1q12); Y chr in 5 (5q13.2; 6p11.32); chr. 18 (18p11.32q23). LOH was detected in 18/150 (12.0%) patients ranging from 6 Mb to 600 Mb. 83 (53.5%) patients had no CNVs or LOH spanning any genes or genes known to be associated with disease. Their results were generally skewed towards CNVs. In 30 patients, we have identified at least one CNV of uncertain significance. Conclusion: These results demonstrate the utility of SNP genotyping data for the detection of clinically significant abnormalities including CNVs, mosaicism or LOH which may harbor recessive mutations. Interpretation of LCHS and CNVs remains challenging particularly without detailed clinical information.

601S

The role of the genomic architecture of transposable elements in the formation of copy number variants: evidence from one schizophrenia family. G. Guiffanti1, S. Gauld2, P. DeCrescenzo1, H. Mangalam3, A. Rodriquez3, R. Maddur4, C. Pato2, F. Macciardi1. 1) Department of Psychiatry Division of Epidemiology & Division of Child and Adolescent Psychiatry Columbia University/NYPSI New York, NY; 2) Istituto Superiore di Sanita’, Rome, Italy; 3) OIT, University of California, Irvine, Irvine, CA; 4) MCS Argonne National Laboratory Computation Institute, University of Chicago; 5) Department of Psychiatry and the Behavioral Sciences Keck School of Medicine, USC, Los Angeles, CA; 6) Dept of Psychiatry & Human Behavior, University of California, Irvine, Irvine, CA.

Many CNVs are flanked by transposable elements (TE). It is emerging that the genomic architecture of TE might predispose certain regions to the mutational mechanisms that result in CNVs. The breakpoints of CNVs are frequently observed within the repetitive sequences of TEs, including SINE, LINE, and LTR, which suggested that regions of the genome that possess repeat sequences are more prone to rearrangements arising independently in different individuals. We sought to explore the role of TEs flanking de novo and inherited CNVs in a family of four siblings, three schizophrenia probands and one unaffected sibling, descending from a schizophrenia affected mother and an unaffected father. This family is part of the Genomic Psychiatric Cohort (GPC) sample (Pato et al, 2013). We used ForestSV to assess CNVs from whole genome sequences (~40X). CNV deletions were included in the analysis if the prediction confidence score was >0.7, which keep the error rate below 5%. First, we determined pattern of transmission of CNV deletions from the affected mother and the unaffected father to three probands and sibling; second, we estimated the prevalence of reference TEs (hg19) in 1,000 bp regions upstream and downstream the breakpoints of de novo and inherited CNVs. The number of CNVs per individual ranged from 2,020 to 2,131 unique deletions. We discovered a total of 3,780 inherited autosomal deletions, including 657 transmitted to probands and 101 to the unaffected sibling, respectively. 575 deletions were found in one or more siblings but not in the parents and were classified as de novo. Of these, 327 de novo deletions were unique to the probands, 61 to the sibling and 191 were shared by siblings, both affected and unaffected. Structurally, reciprocal deletions comprised 37% of the de novo CNV breakpoints; 36% were flanked with Alus, 18% with LINEs and 65% with LTR at either the 5’ or 3’ . Our survey of regions at the breakpoints of CNVs deletions confirm the pattern of genomic instability introduced by the repeat sequences of TE and support their role in the mechanisms that lead to copy number variants formation.

602M

De novo Germline Variants from WGS of Autism Spectrum Disorder Trios. M. Gurral1, W. Brandler1, D. Mathotra1, J. Estabillo1, T. Gadomski2, A. Watson1, T. Chapman1, D. Antin1, T. Solomon1, A. Moyzis1, A. Bhandari1, L. Wong1, C. Corsello1, N. Akshoomoff2, L. lakoucheva1, E. Couchesnes2, J. Sebat1. 1) Psychiatry, University of California, San Diego, La Jolla, CA, USA; 2) Neuroscience, University of California, Irvine, CA, USA; 3) Neurology, University of California, San Diego, La Jolla, CA, USA; 4) Psychiatry, University of California, San Diego, La Jolla, CA, USA.

Autism spectrum disorder is a genetic anomaly with diverse footprints. Based on the recent clinical estimates, it impacts more than 1% of the children around the globe. In last few years, there has been considerable effort to delineate the genetic basis of the disorder through genome sequencing projects. Despite finding several rare de novo genetic variants from exome and whole genome sequence studies, many rare variants remain to be discovered. Here, we report rare de novo genetic variants through the whole genome sequencing of 150 CNVs from 150 subjects comprising of 47 probands, 19 healthy siblings, their 46 parent-pairs, and a healthy family trio contributed by total of 47 families. All de novo single nucleotide variants (SNVs) and copy number variations (CNVs) were detected through the machine learning algorithm, supplied with bioinformatically processed data for 67 trios. Our findings include three novel de novo nonsense mutations, five de novo deletions and one duplication, all six CNVs are > 10 kb. These CNVs, five de novo deletions and a single duplication, have been validated through Affy6 and Illumina array platforms. The three nonsense mutations observed in probands are likely pathogenic. All the four de novo deletions observed in probands are rare and encompass exons. Largest de novo SV in our samples is ~500 kb duplication observed in a proband spanning through several genes. Taken together, it is clear that the burden of de novo nonsense SNVs and CNVs in ID/ASD is high and reiterates the need to determine the role of the parent of origin for about 30% of total de novo SNVs and one forth of the CNVs are contributed by mother.
Copy number variation (CNV) is a major class of genomic variation in humans. Many CNVs are benign and constitute normal variation while others have pathogenic effects. Detecting CNVs genome-wide at high resolution and determining their nature (i.e., deletions or duplications) is important for both clinical diagnostics and biomedical research. At present, sequencing-based CNV detection cannot affordably compete with array-based methods. Clinical cytogenetics relies heavily on arrays for detecting pathogenic CNVs. Array-based genome-wide CNV detection platforms are also used to efficiently and cost-effectively study genome structure in large cohorts and to track genome structural stability for quality control in cell cultures, particularly in stem cell model systems. Array-based CNV calls also provide orthogonal validation of sequencing-based calls. In this study, we characterize the relative performances of all currently available commercial high-density oligonucleotide arrays for genome-wide CNV detection. This work is an extension of our earlier study [PMID: 22140474] using a previous generation of arrays. Here, we collected CNV data from 15 different array designs: two Affymetrix, four Agilent, and nine Illumina designs. The extensively characterized HapMap/1000 Genomes sample, NA12878, was hybridized to each array in two technical replicates. Two sets of CNVs were then called from each array.

To address these questions, we first developed novel ways to use increasingly abundant whole genome sequencing data to identify mCNV loci and the copy number alleles that segregate at these loci. We applied these methods to 849 genomes sequenced in Phase 1 of the 1000 Genomes Project at low coverage and to hundreds of other deeply sequenced genomes, creating a locus, allele and haplotype map for human duplication CNV. We typed 3,926 duplication CNVs of which 1,432 appear to exhibit three or more alleles. We also developed new molecular-biological assays to evaluate genotype accuracy at these loci. The discrete integer copy number determinations from these assays are highly (99%) concordant with our computational calls from sequencing data across all 849 genomes.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
607S
SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints. H. Kehrer-Sawatzki1, J. Vogt1, K. Bengesser1, S. Bammert1, K.B.E. Claes2, K. Wimmer3, V.F. Maunter4, R. van Minkelen5, E. Legius6, H. Brems6, M. Upadhyaya7, C. Lazarro7, J. Högel7, T. Rosenbaum7, L. Messiaen7,8,9, D.N. Cooper1. 1) Human Genetics, University of Ulm, Ulm, Germany; 2) Centre for Medical Genetics, Ghent University Hospital, Belgium; 3) Division of Human Genetics, Medical University Innsbruck, Austria; 4) Department of Neurology, University Hospital Hamburg Eppendorf, Germany; 5) Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands; 6) Department of Human Genetics, KU Leuven, Belgium; 7) Institute of Medical Genetics, School of Medicine, Cardiff University, UK; 8) Molecular Diagnostics Unit, Hereditary Cancer Program, Catalán Institute of Oncology (ICO-IDIBELL), L’Hospital de Llo- bregat, Barcelona, Spain; 9) Department of Pediatrics, Duisburg General Hospital, Germany; 10) Medical Genomics Laboratory, Department of Genetics, University of Alabama at Birmingham, Alabama, USA.

Background: Genomic disorders are caused by copy number changes that may exhibit recurrent breakpoints processed by nonallelic homologous recombination. However, region-specific disease-associated copy number changes have also been observed which exhibit non-recurrent breakpoints. The mechanisms underlying these non-recurrent copy number changes have thus far not been fully elucidated. Results: We analyze large NF1 deletions with non-recurrent breakpoints as a model to investigate the full spectrum of causative mechanisms, and observe that they are mediated by various DNA double strand break repair mechanisms, as well as aberrant replication. Further, two of the 17 NF deletions with non-recurrent breakpoints, identified in unrelated patients, occur in association with the concomitant insertion of SINE/variable number of tandem repeats (Alu) SVA retrotransposons at the deletion breakpoints. The respective breakpoints are refractory to analysis by standard breakpoint-spanning PCRs and are only identified by means of optimized PCR protocols designed to amplify across GC-rich sequences.

Conclusion: De novo large deletions encompassing many hundreds of kilobases could constitute a novel and as yet under-recognized mechanism underlying large-scale copy number changes in the human genome.

607T

Genome mapping in nanoarrays (BioNano Genomics) represents a new single-molecule platform complementary to short-read sequencing for genome assembly and structural variation analysis. Extremely long molecules of hundreds of kilobases fluoresce sequentially labeled at sequence motifs and elongated in nanochannels enable direct interrogation of genome structure at a high resolution. The high throughput of the BioNano Irys system has made possible, for the first time, rapid analysis of multiple genomes and cross-sample comparison to identify genome structural variation at high resolution. To date, we have de novo assembled more than 20 normal and diseased human genomes and analyzed their structural variation content. Our genome map assemblies cover at least 90% of non-N-base portions of the genome and also extend into subcentromeric and subtelomeric regions of the genome.

Here, we present results from extensive analysis of an Asian genome and a CEPH trio. We detected hundreds of large structural variants per genome and haplotype differences in these genomes. In the YH genome, we found 708 insertions/deletions and 17 inversions larger than 1 kb. Without considering 59 SVs that overlap with N-base gaps in hg19, 609 out of 666 (90%) are supported by orthogonal experimental methods (resequencing- and/or fosmid assembly-based) or historical evidence in public databases. For the CEPH trio, we identified novel and previously reported structural variants consistent with Mendelian inheritance. We also used publicly available public sequence read data to confirm and refine our SV calls.

Overall, our genome map assemblies provide valuable structural information otherwise difficult or impossible to decipher with short-read sequencing data alone.

610S
Estimating the parental haplotype source of germline-transmitted de novo duplications. Y. Liu1,2, S. Vattathil1,2, L. Huang3, X. Xia4, G.E. Davies5,6, E.A. Ethel7, J.J. Hettenga8, A. Abdeliaou9, S. Rucinski3, S. Aru2, D. Boomsmma10, T.H. Beat9,11, P. Sche10,11,2. 1) Program in Biostatistics Bioinformatics and Systems Biology, The University of Texas Graduate School of Biomedical Science, Houston, TX, USA; 2) Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; 3) Human & Molecular Genetics Program, The University of Texas Graduate School of Biomedical Sciences, Houston, TX, USA; 4) Center for Genomic Medicine and Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Hanover, NH, USA; 5) Avera Institute for Human Genetics, Sioux Falls, SD, USA; 6) Biological Psychology, VU University, Amsterdam, The Netherlands; 7) Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA; 8) Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA; 9) Department of Genetics, The University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.

De novo germline-transmitted duplications arise during gamete formation (in mitosis) but also de novo duplications are generated as a result of mutations in the germline that can be inherited by a single individual. Here we present a new computational approach to interpret the genetic origin of de novo duplications by using population genetic data from family-based genome-wide association (GWA) studies. Specifically, we first use trios to identify de novo duplications in offspring. We then check for consistency among plausible combinations of transmitted haplotypes and observed SNP array data (genotypes implied by the “B allele” frequencies) for the offspring and their parental origin. For each scenario described above, we are attempting to leverage these datasets to resolve the question of which haplotype, derived from the same parental chromosome (situation 2 above). We are currently attempting to strengthen our conclusions with additional analyses of monzygotic twins from the other study and by ruling out sources of artifact, such as somatic mutation.
Our results suggest that rare CNVs account for a substantial portion ofiqs defined and an alteration of females' fecundity. In addition to CNV size, these

The impact of rare large CNVs has most conclusively been observed in clinical cohorts and it is unclear how these variants affect health in general populations. To investigate the burden of rare large CNVs in the general population, we analyzed the Estonian Genome Centre cohort. It is a longitudinal, prospective, population biobank encompassing 5% of the Estonian adult population and linked to comprehensive personal, educational, medical and daily life data. Within a subset of 7877 individuals, we identified 65 carriers of known genomic disorder lesions, equivalent to a prevalence of 0.8% in the general adult population. Their phenotypes are significantly enriched for GO processes involved in neurodevelopment, ≥11 genes, respectively. Concordantly, we found that the frequency score, but not presence of aohnolog, correlated with ID prevalence. Our results suggest that rare CNVs account for a substantial portion of the population variability of educational attainment. Carriers of syndromic genomics defined by rare CNVs and by non-recurrent rare CNVs, respectively, reference cohorts should not readily be considered as healthy carriers. Rare large CNVs represent a significant public health issue as they impact life quality on multiple levels including medical and socioeconomic metrics.

Autism Spectrum Disorder (ASD) is a highly heterogeneous neurodevelopmental condition with many contributing risk genes and loci. Genome wide microarray and sequencing studies have implicated numerous genes, for example members of the NFκB and SHANK families, PTCCH1, CHD8, and SCN2A as risk candidates in ASD. Mutations in voltage gated sodium channel SCN2A, many of which de novo in origin, have previously been associated with several forms of epilepsy. In recent whole exome and genome sequencing data in ASD, SCN2A has emerged as one of only a small number of genes with multiple independent de novo loss of function mutations, making it one of the best new candidate genes. Through our own ASD whole genome sequencing project, we have uncovered a complex structural variant at SCN2A in a family quartet comprised of parents and two affected female offspring, initially, a 1.7kb single exon (exon 18) deletion was discovered as a de novo variant in both children thought to have arisen by germline mosaicism. However, further analysis revealed that the deletion breakpoint can also be detected in a blood DNA sample from the father, despite his having two copies of exon 18. Using detailed molecular analyses, we have shown that in the father there is a complex rearrangement such that he carries one intact allele and in addition to the transmitted deletion, he also possesses a third structure carrying a neighbouring intronic deletion. We are pursuing a number of different technical approaches to resolve the nature of this variant and the mechanism by which it has arisen. We are also assessing mRNA transcripts from all family members to investigate the functional impact of the variant structures. Resolving this complex variant will aid in the interpretation of the exonic deletion in the affected individuals in this family and also demonstrates the importance of detailed genomic data in understanding variants. Our findings have further implications in studying models of genetic disease as other complex similar events could give rise to cryptically inherited variants misinterpreted as de novo events. Furthermore, this example highlights the power of whole genome sequencing to reveal variants not detectable by other technologies.

616S Genome-wide copy number variants analysis identifies deletion variants associated with ankylosing spondylitis. S. Yim1, S. Jung1, T. Kim2, S. Shim3, Y. Chung1. 1) The Catholic University of Korea, College of Medicine, Seoul, South Korea; 2) Hanyang University Hospital for Rheumatic Diseases, Seoul, South Korea; 3) Chungnam National University Hospital, Daejeon, South Korea.

We aimed to discover ankylosing spondylitis (AS)-associated copy number variants (CNVs) in Korean subjects and their synergistic roles in the development of AS. We performed genome-wide association study (GWAS) in 309 AS patients and 309 controls using Agilent CNV microarray. AS-associated CNV regions (CNVRs) were replicated in two independent sets (625 cases and 891 controls in total) by qPCR and deletion-typing PCR. In the CNV-GWAS, we identified 227 CNVRs significantly associated with AS. Of the candidate CNVRs, nine were successfully replicated in the first replication; 1q32.2 (HHAT), 1p34.2 (BMP8A), 2q31.2 (PRKRA), 6p21.32 (HLA-DPB1), 11q22.1 (CNTN5), 13q13.1 (EEF1DP3), 14q24.2 (RGS6), 16p13.3 and 17p13.1 (CHD8). Given the smaller sample size, we performed a second replication in 12 T1D cases and 2,206 families with T1D-affected offspring. In our initial analyses using a standard approach with PennCNV software, we were also replicated that the ImmunoChip-wide case-control odds ratio (OR) for rDEls and rDUPS was highly sensitive to NC thresholds and procedures. Analysis using an optimal QC pipeline, with the highest validated accuracy, suggested no overall CNV burden for T1D, but some association for rDEls longer than 500 kb (OR=1.67, p=0.003), noting that ImmunoChip genome coverage is mostly concentrated in known autoimmune regions. No specific CNV regions (CNVRs) showed association at a corrected p-value threshold, although CNVR frequencies were lower than expected (most less than 0.1%), significantly reducing statistical power. Transmission disequilibrium tests (TDT) run on the family dataset provided no support for CNVRs contributing to familial clustering of T1D. We present an R-package, plumbCNV, that provides an automated approach for QD and detection of rare CNVs, which can facilitate equivalent analyses of large-scale SNP array datasets in any disease. A power analysis indicating sample sizes required to obtain convincing evidence for single locus tests is also provided.


Copy number variations (CNVs) have been proposed as a possible source of ‘missing heritability’ in complex human diseases. Several studies of type 1 diabetes (T1D) have found null associations with common copy number polymorphisms (CNPs), but CNVs of low frequency and high penetrance could still play a role. We used the Log-R-Ratio (LRR) intensity data from a dense single nucleotide polymorphism (SNP) array, ImmunoChip, to detect rare CNV deletions (rDEls) and rare CNV duplications (rDUPS) in 6,808 T1D cases, 9,954 controls, and 2,206 families with T1D-affected offspring. Our initial analyses using a standard approach with PennCNV software were affected by biases that led to detection of CNVs subsequently confirmed by quantitative polymerase chain reaction (qPCR) to be false positives. We developed a series of quality control (QC) tests that were demonstrated to identify and correct probe intensity biases. Statistical testing and validation showed that the ImmunoChip-wide case-control odds ratio (OR) for rDEls and rDUPS was highly sensitive to QC thresholds and procedures. Analysis using an optimal QC pipeline, with the highest validated accuracy, suggested no overall CNV burden for T1D, but some association for rDEls longer than 500 kb (OR=1.67, p=0.003), noting that ImmunoChip genome coverage is mostly concentrated in known autoimmune regions. No specific CNV regions (CNVRs) showed association at a corrected p-value threshold, although CNVR frequencies were lower than expected (most less than 0.1%), significantly reducing statistical power. Transmission disequilibrium tests (TDT) run on the family dataset provided no support for CNVRs contributing to familial clustering of T1D. We present an R-package, plumbCNV, that provides an automated approach for QD and detection of rare CNVs, which can facilitate equivalent analyses of large-scale SNP array datasets in any disease. A power analysis indicating sample sizes required to obtain convincing evidence for single locus tests is also provided.
619S Genetic variation in introns that flank alternatively spliced exons: A new way to look for disease-related variants. A. Neininger, W.C. Steward, D.A. Greenberg. Battelle Center for Mathematical Medicine, Nationwide Children’s Hospital, Columbus, OH, 700 Children’s Drive; The Ohio State Department of Pediatrics, Columbus, OH, 700 Children’s Drive; The Ohio State Department of Statistics, Columbus, OH, 1958 Neil Ave.

Simple exon mutations are not the only kinds of genetic variants that can lead to inherited disease. Intronic variation is also important, but the relatively high degree of variability in introns makes the identification of specific disease-related alleles difficult. For example, the gene BRD2 is related to epilepsy susceptibility and it contains no exonic disease-related mutations, but certain sequences of an intron are associated with disease. This intron is highly variable and it also contains an alternatively spliced (AS) exon. This led us to ask: Are introns surrounding AS exons more variable than other introns? AS exons are found in approximately 75%-90% of genes, and 1/3 of these exons can introduce a premature termination codon (PTC). Therefore, understanding the variability that regulates alternative splicing could allow us to identify intronic sequences that are related to disease. Using sequence data from 1000 Genomes, we first compared the frequency of common variants in the introns flanking AS and regularly spliced exons. We found statistically significant differences between introns that flank AS exons compared to those that flank regularly spliced exons. This finding suggests that this variation influences regulation of splicing. We also subdivided the introns flanking AS exons into two groups: those that introduce a PTC leading to nonsense-mediated decay, and those that do not. We then explored a method that evaluates different intronic sequences with respect to their new ones. Our approach to evaluate the PTC is to use a footprint from the analysis of 46 epilepsy patient haplotypes and 48 control haplotypes to identify the free energy. Our results suggest that free energy is not only a footprint in the loci that AS exons appeared able to predict which sequences exist in the population and which do not. Moreover, sequence similarity alone does not appear to be associated with disease. We did not observe any statistical differences between these two groups. In conclusion, approaches that account for differences in the physical gene structure may give us a better way to understand intronic sequence variation, and consequently, a better way to identify disease-related “alleles”.


Acidic mammalian chitinases (AMCase) have been shown to be closely associated with asthma in mouse model, allergic inflammation, and food processing. AMCase is expressed in mice and humans. The amino acid sequence of human AMCase shares 81% identity with that of mouse counterpart. There are several variants with the amino acid substitutions in human AMCase. Recent genetic association analyses of AMCase haplotypes for asthma revealed significant associations between the variant haplotype and complex diseases. It is shown that the free energy of the hypervariable BRD2 intron that flanks the AS exon appeared able to predict which sequences exist in the population and which do not. Moreover, sequence similarity alone does not appear to be associated with disease. We did not observe any statistical differences between these two groups. In conclusion, approaches that account for differences in the physical gene structure may give us a better way to understand intronic sequence variation, and consequently, a better way to identify disease-related “alleles”.


The 1000 Genomes Project provides an essential reference catalog of human variation with more than 60 million variants ranging from single nucleotide polymorphisms to structural variant events including inversions and duplications. Also provided are global allele frequencies and genotypes for 2535 individuals from 26 different populations across Europe, Africa, South Asia and South America, which enables many other projects to better interpret their results. Primary uses for the 1000 Genomes data include imputation panels to create whole genome variant sets from exome or array-based genotypes; as filters of “normal” or shared variation in rare disease or cancer sequencing projects; and to explore demography and selection in human populations. The 1000 Genomes Project is now drawing to a close. Here we describe plans to maintain the resource in order to ensure it remains the valuable data set it is today by providing long-term support for the 1000 Genomes Project resource. For example, we will continue to host both the FTP site (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp) and the project website (http://www.1000genomes.org) to ensure the community can access both the raw data and the documentation about the project. We will also create a stable version of the 1000 Genomes Browser (http://browser.1000genomes.org) based on the project’s final data release. This project specific Ensembl-based browser displays all of the 1000 Genomes variants as soon as possible and will use the GRCh37 assembly of the human reference genome. We will also maintain the existing tools and incorporate new ones to assist users in accessing the data they desire. Our most popular tools are the Data Slicer—that allows users to select genomic subsections of our alignment (BAM) and variant (VCF) files and thus download just the piece of the file they need—and the Variation Pattern Finder, which allows users to discover patterns of shared variation in a specific region of the genome. Other tools include the VCF to PED converter, which allows users to generate PLINK format files from remotely hosted VCF files and the recently introduced the Allele Frequency Calculator that will calculate allele frequencies in bulk for specific subpopulations from our VCF files.


Genome-wide association studies (GWAS) have been instrumental to identify genetic variants associated with complex traits. However, risk variants tend to fall in non-coding regions and likely affect gene regulatory mechanisms. Even with increasing regulatory annotations provided by ENCODE and others, predicting the impact of a sequence change on gene regulation remains a challenge. Identification and functional characterization of sequence variants in regulatory elements is crucial for understanding the molecular determinants of complex traits. Here, we integrate DNA sequence with DNase footprinting data to predict the impact of a sequence change on transcription factor binding. Applying this approach to 653 Dnase-seq samples, we identified 3,831,862 regulatory variants predicted to affect active regulatory elements for a panel of 1,891 transcription factor motifs. We observe that these variants are enriched for having a MAF < 1%, are more likely to be in enhancer regions, and tend to affect factors active in < 5 cell-types. Using our newly developed method, QuASAR (Quantitative Allele-Specific Analysis of Reads), we examined the data for variants exhibiting allele-specific binding (ASB). We identified 3,217 binding variants within footprints that are significantly enriched (20% FDR). Overall, we estimate that 56% of our predicted regulatory sites in footprints show an ASB signal. To assess the effect these variants may have on complex phenotypes, we examined their association with complex traits and observed that ASB-SNPs are enriched 1.22-fold for complex traits variants (from GWAS). Using results from meta-studies for lipid levels and height, we identified variants whose binding sites were enriched for associated SNPs. Factors for lipid levels include GR and HNF4a, as well as regulators of immune function such as CREG, PU.1, IRF-1 and IRF-2. Factors for height include stem cell and developmental regulation factors such as GATA and OAF. Additionally, these annotations can be very useful for fine-mapping. For example, because rs532436 (a genetic loci associated with an increase in LDL) is in a footprint for the factor USF, the annotation increases the likelihood from 40% to 70% that this is the functional SNP in the loci. These results show that integration of footprint annotations into GWAS meta-study results improves the identification of likely causal SNPs and provides a putative mechanism by which the phenotype is affected.
The success of whole exome sequencing diagnosis in a large cohort of patients with Mendelian disorders. T. Roscioli1,2,2, L. Evans3,4, M.J. Cowley5, K. Ying6, Y. Zhu7, N. Walsh1, K. Mowat7, A. Hackett2, E.P. Kirk1, R. Sachdev1, M.L. Freckmann1, M. Lipke1, M. Buckley2, M. Dinger6.

1) Department of Medical Genetics, Sydney Children's Hospital, Sydney, NSW, Australia; 2) School of Women's and Children's Health, University of New South Wales, Randwick, NSW, Australia; 3) Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, NSW, Australia; 4) Newcastle GOLD Service, Hunter Genetics, Waratah, NSW, Australia; 5) SEALS laboratory, Prince of Wales Hospital, Sydney, Australia.

Molecular diagnosis in patients with Mendelian disorders has been improved significantly by advances in next generation sequencing (NGS). NGS has resulted in enhanced mutation identification, genetic management and in some cases, the creation of novel therapies. Mendelian disorders are among the most amenable to diagnosis through whole exome sequencing (WES) as the coding 2% of the genome is highly enriched for disease-causing mutations. WES was applied to a cohort of 52 patients from 37 families who were selected from clinical genetics clinics in New South Wales and for whom no molecular diagnosis was known. This cohort is phenotypically heterogeneous, however the majority of patients presented with intellectual disability (62%) consistent with population frequency and diagnostic importance. Other diagnoses included skeletal dysplasias (14%), retinitis pigmentosa (8%), haematological disorders (8%), seizures (3%), metabolic conditions (3%), and dysmorphic syndromes (3%). DNA from annotated exons and splice sites, untranslated regions, and the mitochondrial genome were captured with a Nextera extended exome kit on Illumina HiSeq 2500 and sequenced with paired end reads at 30x coverage. Additional content including SNPs targeting African specific diseases and a likely novel disease gene for intellectual disability involved in neurite growth in a consanguineous family.


1) EMBL-EBI, Hinxton, Cambridge, United Kingdom; 2) Centre for Research Computing, EMBL-EBI, Hinxton, Cambridge, United Kingdom; 3) CRMF, Staronice 2, 05-973, Poland; 4) Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland; 5) 4Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain; 6) Interdisciplinary Centre for Mathematical and Computational Modeling, Warsaw University, Warsaw, Poland

The European Genome-phenome Archive (EGA) is now a joint service from the European Bioinformatics Institute (EBI), UK and the Centre for Genomic Regulation (CRG), Spain. This service provides a permanent archive for all types of genetic, -omics and phenotypic data that has been consented for use in biomedical research, but where the data access is managed through an application and secured encrypted data delivery process to limit risks to the privacy of research participants. Accepted submissions include raw data from genome sequence, transcriptome, epigenome or other -omics data. The EGA is also stores called variants, genotypes, study summary statistics and associated sample phenotypes. Access to the data is managed by appropriate data access committees (DACs) that will approve access based on requests that meet the data-use and patient consent rules governing their studies. Here we describe how EGA service will scale for future submission, data archiving and dissemination including key improvements in phenotypic query of studies as well as the provision of new workflows to upload updated data and generate key performance indicators. Accepted submissions are stored in a federated archive for all types of genetic, -omics and phenotypic data that has been consented for use in biomedical research, but where the data access is managed through an application and secure encrypted data delivery process to limit risks to the privacy of research participants.
Integrative Japanese Genome Variation Database from the cohort study of Tohoku Medical Megabank Organization (ToMMo). Y. Yamaguchi-Kabata, Y. Kawai, T. Mimori, F. Katsuoka, N. Narai, K. Koijima, J. Danjoh, S. Saito, X. Pan, J. Yokozawa, R. Saito, Y. Sato, K. Tsuda, Y. Kuroki, K. Kinoshita, J. Yasuda, M. Yamamoto, M. Nagasaki, Dept. Integrative Genomics, Tohoku Medical Megabank Organization, Tohoku University, Sendai, Miyagi, Japan. Tohoku University Tohoku Medical Megabank Organization (ToMMo) was founded to establish an advanced medical system to foster the reconstruction from the Great East Japan Earthquake. The organization will develop a biobank that combines medical and genome information during the process of rebuilding the community medical system and supporting health and welfare in the Tohoku area. We have started the prospective cohort study in the region, with genome analyses, to establish the advanced, personalized medicine based on the individuals genomic data. Although reports of common variants with their frequencies are accumulating for each population, many of low-frequency variants remain undetected or lack of their frequencies. Therefore, cataloging genomic variants from whole-genome sequencing and estimation of variant frequency are necessary for foundation of genomic medicine for each population. To make a reference panel of genomic variation of Japanese, we have sequenced whole genomes of about 1,000 cohort participants, and collected single nucleotide variants (SNVs). Integrative Japanese Genome Variation Database provides data of genomic variations obtained by whole-genome sequencing of Japanese individuals who participate the genome cohort study of ToMMo. The current release provides SNV data obtained from the individuals from the cohort. The first release contains data of SNVs (about 5 millions on autosomes) that exist at least 5.0 % frequency in the samples. This browser can be used to search those SNVs and get their information such as frequency. The database and variants data would be useful for genome analysis and strategy of genomic populations for local populations.

Complete haplotypes of the human light chain immunoglobulin loci from a hydatidiform mole BAC library provide insights into locus-specific signatures of genetic diversity. K. Metz, Sternberg 1,2 , C.T. Wat son 3,4 , T.A. Graves-Lindsay 1, R. Warren 1, M. Maig 1, J. Schein 3, R.A. Holt 1, R.K. Wilson 1, E.E. Eichler 3,4, F. Breden 1, 1) The Genome Institute, Washington University, St. Louis, MO; 2) Department of Genome Sciences, University of Washington, Seattle, WA; 3) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 4) Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada; 5) Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada; 6) Howard Hughes Medical Institute, Seattle, WA. Immunoglobulin gene (IG) loci and their associated germline variation are critical for pathogen-mediated immune response, but establishing complete sequence is problematic because of large-scale duplications and the fact that many genomic resources are established from somatically rearranged material. Here, we present a complete haplotype from the essentially haploid hydatidiform mole, CHM1. We sequenced a set of tiling BAC clones from CHM1 across the light chain IG loci, kappa (IGK) and lambda (IGL), to generate single haplotype representations of these regions. The IGL haplotype is 1.25 Mb of contiguous sequence from nine clones with four novel V gene and one novel C gene alleles. There is also an 11.9 kb insertion that does not contain any functional genes. The IGK haplotype is comprised of a 644kb proximal and a 466kb distal contig separated by a gap that is also present in the GRCH37 reference genome sequence. Our effort added an additional 49kb of unique sequence that extended into the gap. The IGK haplotype contains six novel V gene and one novel J gene alleles. In addition, we identified a 16.7kb region with increased sequence identity between the proximal and distal IGK contigs with signatures of interlocus gene conversion. In two instances, we observed the presence of functional alleles that had been previously classified as either "distal" or "proximal" alleles residing at loci in the alternate location. Due to the large inverted duplications, over 80% of the IGK variable (IGKV) locus consists of segmental duplications with >95% sequence identity suggesting that, unlike in the heavy chain locus (IGH), the duplications may be responsible for sequence homogenization rather than diversity. When we compared nucleotide and structural diversity among the light chain and heavy chain haplotypes, we find a three to six fold enrichment of diversity in the IGH locus compared to the light chain loci, that supports the theory that the heavy chain is more important in determining antigenic specificity.

Defining Variation Sensitive Regions in Genes Associated with Hearing Loss. A.N. Abou Tayoun 1, C.A. Casat 2, D.M. Jordan 1, A.L. Muirhead 1, K.A. Lafferty 1, A.L. Hernandez 1, J. Shen 5, M.S. Lebo 1,2, S.R. Sunyaev 3, H.L. Rehm 1,2, S.S. Amr 1, 1) Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, MA; 2) Division of Genetics, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA; 3) Division of Pathology, Harvard Medical School, Cambridge, MA; 4) Authors with equal contribution. The classification of novel variants is a major challenge facing clinical genetics laboratories. Because most novel variants do not have functional, genetic or population data to support clinical classification, in silico approaches are commonly used to prioritize candidate disease-causing variants. We have previously shown that a systematic evaluation of gene-disease associations can largely eliminate unnecessary interpretation of variants in genes with weak disease association. Applying this approach to 145 genes included in hearing loss panels, we found 54 genes (37%) with insufficient association to hearing loss that may be excluded from testing panels. Here, we extend this approach to the domain level in the remaining 91 genes with sufficient evidence for disease association, to test the hypothesis that systematic evaluation of domain-disease association can improve variant interpretation. Using Pfram predicted domain boundaries (http://pfam.xfam.org), the frequency of variation in the general population (Exome Sequencing Project, N=6,503), and clinically classified variants from ClinVar (http://www.ncbi.nlm.nih.gov/clinvar), we determine the domain level tolerance to variation within each gene. We statistically identify regions that are most sensitive to functional variation in the general population and also most commonly impacted in symptomatic individuals. This approach leverages the fact that many genes contain multiple domains in each gene. Our data show that this approach has utility in finding regions of greatest interest in predicting variant pathogenicity in known disease genes. For example, almost all pathogenic missense variation is confined to the ion channel domain of the KCNQ4 gene and the DNA binding domains of the PAX3 and SOX10 genes. These domains represent 15–46% of the total coding region in their respective genes, and are almost devoid of functional variation in the general population. In contrast, regions outside these domains had little or no pathogenic missense variants in symptomatic individuals. This knowledge helps define variation sensitive regions in hearing loss genes that can be used to prioritize candidate disease variants, increasing the sensitivity and specificity of variant assessment within these genes.
Deep targeted sequencing of SLE associated LD blocks reveals multiple putatively functional variations in strong LD with SLE GWA SNPs: A haplotype based assessment of disease risk. P. Raj1, QZ. Li1, D. Karp1, E. Rafi1, C. Liang1, B. Wakekald1, K. Viswanathan1, I. dozmarov1, N. Olsen2, J.A. James3, J.A. Kelly2, B. Lauwerys4, PK. Gregersen5, BP. Tsao6, P. Gaffney7, EK. Wakeland7. 1) Immunology, UT Southwestern Med Center, Dallas, TX. 5323 Harry Hines Blvd. Dallas, Texas, 75390; 2) Division of Rheumatology, Penn State Hershey Medical Center, 500 University Drive, Hershey PA 17033; 3) Oklahoma Medical Research Foundation, 825 North East 13th Street, Oklahoma City, Oklahoma 73104, USA; 4) Service de Rhumatologie, SSS/IREC/RUMA, Cliniques Universitaires Saint-Luc & Université catholique de Louvain Avenue Hippocrate 10, bte B2.5390, 1200 BRUXELLES, Belgium; 5) Feinstein Institute for Medical Research in Manhasset, New York; 6) Division of Rheumatology, University of California Los Angeles, Los Angeles, California, United States of America. A strong genetic predisposition has been observed in susceptibility to SLE. Genome-wide association analyses (GWAS) have identified more than 30 SLE risk loci in humans thus far, and ongoing analyses by the SLE immunochip consortium will probably define even more. We performed deep targeted sequencing of the SLE associated LD blocks with aim to develop a genomic strategy for the functional characterization of all the genetic risk alleles in SLE. For this, we sequenced ~4.4 Mb of the human genome in 773 SLE patients and 576 controls using a custom target enrichment sequencing strategy. We called 146K pass quality variations with an average of 107-fold coverage of the entire linkage disequilibrium (LD) blocks at more than 70 potential disease risk loci, thus providing detailed sequence information for all of the identified SLE risk loci. We obtained 98 percent matches between sequencing and immunochip genotypes on subset of sample. The common (MAF>0.05) genetic variants identified in this dataset were used to assemble a list of potentially functional variations in both coding and non-coding (from ENCODE) regions within these LD blocks, thus allowing the disease associations within SLE risk loci to be integrated with functional properties. Then, haplotype were formed including known GWAS and potentially functional variations in strong LD with it. We found multiple functional variations in strong LD with SLE GWA SNP in many cases of SLE risk gene. Further we assembled a list of all genic hotspots, which could be used to identify new candidate protein encoding genes. The sensitivity of this intra-human method. Interestingly, genes that are associ-ated with childhood disease outcome also demonstrate a significant relationship to dN/dS, and we have performed some validation using a secondary exome dataset. By analyzing all X-chromosome genes, we identified novel candidate genes with early mortality, and several of these have been verified by recent exome and gene identification studies. Furthermore, we demonstrate intragenic localized patterns of variants that suggest pathogenic hotspots, which could be used to identify new candidate protein domains linked to diseases. Our results suggest intragenic localization analysis is a valuable tool to help prioritize novel disease genes in sequence interpretation.

Identifying candidate genes and domains for X-linked diseases using population exome data. X. Ge1,2, P. Kwok3,4, J. Shieh1,2. 1) Division of Medical Genetics, Department of Pediatrics, University of California, San Francisco, San Francisco, CA; 2) Institute for Human Genetics, University of California, San Francisco, San Francisco, CA; 3) Department of Dermatology, University of California, San Francisco, San Francisco, CA; 4) Cardiovascular Research Institute, University of California, San Francisco, San Francisco, CA. Many new human disease genes can be identified through high-throughput sequencing. Yet variant interpretation for the large amounts of genomic data remains a major challenge. There is an abundance of variation of uncertain significance, and it is difficult to interpret variants in genes that lack disease annotation. As clinically-significant disease genes may be subject to negative selection, we predicted these genes might have a paucity of nonsynonymous variation in the population. The aim of this study was to develop methods to predict pathogenicity using population exome data. We analyzed and integrated human X-chromosome bulk variant data from six thousand individuals (the NHLBI Exome Sequencing Project) to assess intra-human ratios of substitution rates at non-synonymous and synonymous sites (dN/dS) as a potential measure for gene-based pathogenicity. Indeed, we find that the dN/dS ratio is significantly lower in OMIM disease genes, supporting the sensitivity of this intra-human method. Interestingly, genes that are associ-ated with childhood disease outcome also demonstrate a significant relation-ship to dN/dS, and we have performed some validation using a secondary exome dataset. By analyzing all X-chromosome genes, we identified new candidate genes with early mortality, and several of these have been verified by recent exome and gene identification studies. Furthermore, we demonstrate intragenic localized patterns of variants that suggest pathogenic hotspots, which could be used to identify new candidate protein domains linked to diseases.
633T
A15924G mt-tRNAT Gene Mutation is not the Primary Cause of Mitochondrial Myopathies. A. Cakiris1,2, B. Kara1, B. Ustek3,1, Genetics, Institute for Experimental Medicine, Istanbul, Turkey; 2) Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey; 3) Pediatric Neurology, Kocaeli University Medical Faculty, Kocaeli, Turkey.
Mitochondrion is a essential organelles present in all nucleated mammalian cells. Mitochondrial genome has a very high mutation rate at 10~17 folds higher than the nuclear genome. Those mtDNA mutations may occur within proteins, tRNA or tRNA genes. But more than half of mutations that are disease related are located in mt-tRNA genes. Some of these are polymorphic and the others are pathogenic mutations. In order to analyze mitochondrial DNA mutations, the entire mitochondrial genome was amplified in two overlapping polymerase chain reactions. Mitochondrial DNA was deep sequenced by next generation sequencing technology. The blood or muscle tissue of the 267 patients diagnosed with undefined mitochondrial myopathy and 330 healthy individuals have been analyzed. A15924G mt-tRNA mutation was detected to be homoplasmic in 12 patients and 6 healthy control samples. A15924G mt-tRNA mutation can alter the free energy of the thermodynamic ensemble for the secondary structure of wild type TRNT and affect the structure of tRNAs. This mutation was previously shown to be associated with the LIMM disease as a pathogenic mt-TRNT mutation. In another study, this mutation has been replicated as a common mtDNA polymorphism and not the primary cause of LIMM. In our study, this mutation was found in 12 patients, 7 of which belongs to haplogroup I, 2 to F1b1, 1 to U7b, and 1 to T2h haplogroup. Among the healthy individuals, this haplogroup distribution is as follows: 4 haplogroup I, 1 U7b, and one to haplogroup F1b1. According to the above results, the A15924G mutation in both patients and healthy individuals is considered as a haplogroup marker. These results suggest that the A15924G mutation is a common mtDNA polymorphism, rather than a pathogenic mtDNA mutation.

634S
A Mouse Mutagenesis Scheme to Isolate Lethal X-Linked Recessive Mutations. F.J. Probst1, R.R. Corrigan2, M.N. Bainbridge3, S.N. Jhangiani3, H.V. Doeblinger4, J. Hu5, R.A. Gibbs6, 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Neuroscience, Baylor College of Medicine, Houston, TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.
Large-scale mutagenesis of the mouse has yielded numerous new autosomal mutants, but X-linked recessive phenotypes have been virtually nonexistent. We have devised a two-generation mouse mutagenesis screen to capture lethal and sub-lethal X-linked recessive mutations. C57BL/6J (B6) male mice are mutagenized with the chemical “supermutagen” ENU. These mice are then mated to 129S6/SvEvTac (129) females to generate G1 progeny. The females from this cross will be X(B6*)X(129) (i.e., they will have inherited a mutated B6 X chromosome from their father and a 129 X chromosome from their mother). These G1 animals are then bred to 129 males to generate G2 progeny. In theory, the X(B6*)X(129) G1 female from each pedigree will transmit her X(B6*) chromosome to half of her male progeny and her X(129) chromosome to the other half. When meiotic “crossing over” is considered, this means that, for each X-linked allele, about 1/2 of the G2 males will inherit the B6 copy and 1/2 will inherit the 129 copy. However, if the X(B6*) chromosome carries an X-linked recessive lethal mutation, then there will be no male mice with B6 alleles at the markers flanking the mutation, since all of these mice will not survive to weaning. DNA from 10 G2 male progeny of each G1 female are therefore combined to create a single “pool” of DNAs for each pedigree. Each pool is then run on an Illumina Mouse Medium Density Linkage Panel. If the female carries a lethal X-linked recessive mutation, the male mice that inherited this mutation will have died prior to weaning. Genotyping of the B6 alleles in the pools for the markers around the mutation. Though some drift is to be expected in the data, the chance of a particular region of the X chromosome appearing to be homozygous for the 129 alleles is less than 1/1,000, and additional mice can be bred from potential mutant G1 carriers in order to confirm the result. DNA from critical G1 female animals is then subjected to exome capture via a Roche Nimblegen EZ Library, and the products are analyzed by next-generation sequencing on an Illumina apparatus. Causative mutations are confirmed via linkage analysis and additional phenotyping and breeding studies. Many of the resulting mutants are likely to be models of human X-linked diseases, thus providing new insights into the cause and pathophysiology of X-linked diseases.

635M
The relative impact of DNA mutation and RNA editing as sources of somatic sequence variation in the transcriptomes of normal adult tissues. D. Omdal1, N. Huang1, P. Kheradpour2,3, M. Kellis2,3, K. Ardiles2,3, The GTEx Consortium. 1) Department of Genetics, Washington University School of Medicine, St Louis, MO; 2) MIT Computer Science and Artificial Intelligence Lab, Cambridge, MA; 3) Broad Institute, Cambridge, MA.
The full extent and origins of somatic mosaicism in humans are unknown but could have profound implications for our understanding of human health and disease. We have developed a statistical framework for inferring DNA mutations from multi-tissue RNA sequencing data, which considers sequencing error, sampling error, RNA editing and allele-specific expression as alternative explanations for apparent sequence differences among transcripts from the same locus. We have applied this framework to the GTEx Consortium data to detect RNA editing and somatic mutation from 1,514 non-cancerous tissues obtained from 175 donors. We identified 174 putative somatic mutations across 77 donors (min/donor=0; max = 8). In contrast, we observed some evidence for RNA editing in one or more tissues at an average of 1,733 sites per donor (min, 669 sites; max, 6,522). Thus, RNA editing appears to dominate over somatic mutation as a process creating high-frequency somatic sequence variation in the transcriptomes of normal cells. We constructed an expression fingerprint combining ADAR1, ADAR2B1, and ADAR2B2 that is strongly correlated with the average extent of RNA editing across individual tissues (R=0.88 in brain tissues, p<6.7 x 10^-4 and R=0.76 in other tissues, p<4.7 x 10^-4). Like RNA editing by the ADAR family of enzymes, somatic mutations can be “programmed” to some extent by directed enzymatic activity. V(D)J recombination, catalyzed by RAG1 and RAG2, and somatic hypermutation, catalyzed by AID, act on the variable regions T-cell receptor and immunoglobulin genes. Of the 174 somatic mutation calls, 86 (49%) occur in the immunoglobulin genes, and 165 is the single tissue with the most antibody mutations (24/86, 28%). Parity motivated by this observation, we have performed a post hoc study of OC on an additional somatic mutation calls by read depth and manual curation, but definitive evidence will be provided by experimental validation of the entire callset, which is ongoing. Of the non-antibody mutation calls, 18 (20%) are observed across multiple tissues in the same person, 30 (34%) are observed in multiple people, and 73 (83%) are derived from transcript sequences, many of the mutations are likely to be functional, and we interpret their origin and impact by investigating their sequence context, relation to cancer genes, and medical history of the donors.

636T
Distinct variation in the LILRB3 and LILRA6 genes encoding a myeloid inhibitory and activating receptor pair. A. Bashirova1, 2, C. O’hUigin3, M. Carrington1, 2, 1) Leidos Biomedical Research, Frederick, MD; 2) Ragon Institute of MGH, MIT, and Harvard, Boston, MA.
The leukocyte immunoglobulin-like receptor (LILR)B3 and LILRA6 genes encode homologous inhibitory and activating orphan receptors. Both genes are expressed in monocytes and exhibit a strikingly high level of polymorphism at the amino acid level. We have demonstrated recently that LILRα (but not LILRB3) can be present in variable copy numbers in the genome, probably due to crossovers between the two genes. In a cohort of 228 healthy white individuals, we observed that 64% had two copies of LILRA6 per diploid genome, while 32% had more than two and 4% had only one LILRA6 copy. We have now characterized SNPs in the two genes in a subset of individuals with two copies of LILRA6 (N=91). Across seven exons encoding the signal peptide and extracellular D1-D4 domains, we have identified 46 and 38 amino acid changing SNPs in LILRB3 and LILRA6, respectively. Among these, 35 SNPs were located at identical positions within the two genes. While minor allele frequencies (MAFs) of more than 10% were observed in both genes, MAFs in LILRA6 were greater than 10%, there were substantial differences in MAFs of many SNPs between LILRB3 and LILRA6. In exons encoding the signal peptide and the D3/D4 domains, some SNPs had MAFs of <2% in one gene and up to 40% in the other gene. These differences were also reflected in distinct patterns of linkage disequilibrium (LD) between pairs of SNPs in the two genes, with LILRA6 demonstrating a higher degree of LD. The dataset on LILRB3/A6 SNPs will be useful in further functional and evolutionary characterization of this highly polymorphic locus, which may be involved in regulation of immune responses in disease pathogenesis.
637S
Finding effectively neutral sequence in the presence of coding and noncoding conserved elements. A.E. Woerner1,2, K.R. Veeramah1, M.F. Hammer1, 1 Computer Science, University of Arizona, Tucson, AZ; 2 Department of Ecology and Evolution, Stony Brook, NY; 3 Arizona Research Laboratories, University of Arizona, Tucson, AZ.

Early studies that have attempted to utilize putatively neutral loci in the genome for demographic inference have focused on pseudogenes, introns, and repetitive elements such as Alus. More recent studies have attempted to reduce the impact of positive and negative selection as well as background selection and genetic drift by finding loci that are either in regions of low gene density and low recombination rate, or are at a maximal distance from the nearest gene in genetic units. However, nucleotide diversity may not only be reduced near genes, but also near noncoding conserved elements. Because ~5% of primate genomes is composed of evolutionarily conserved elements, with ~2/3 of these elements being noncoding, skews in diversity in regions in linkage with such elements could have large effects on genomic patterns of nucleotide variation. Here, we show that the combination of multiple conserved noncoding elements has more of an influence on nucleotide diversity than the distance to the nearest gene. Using an approach that models the combined effects of linkage to coding and noncoding elements, we predict levels of polymorphism on both the autosomes and the X chromosome. Based on our predictions we identify loci that we consider to be effectively neutral, i.e., regions whose expected diversity is at least 99% of our inferred neutral rate. Using the relationship between diversity and the local recombination rate as a proxy for the effects of directional selection on neutral sites, we show that unlike other loci in the genome, our effectively neutral loci show no correlation between the local recombination rate and diversity. We conclude that "neutral" sequence is extremely rare in the genome. Finally, we use these neutral regions to estimate the relative effective population size of the X chromosome vs. the autosomes in African and non-African human populations.

638M
The associations of multiple genes with systemic sclerosis by next generation sequence technology. H. Li, X. Guo, S. Assassin, J. Revelle, M. Mayes, X. Zhou. University of Texas - Houston Medical School, Houston, TX.

Systemic sclerosis (SSc) is a rare and complex immune-mediated disease characterized by vasculopathy, fibrosis of skin and internal organs, and the presence of autoantibodies. According to genome-wide association studies, the HLA region contains the strongest SSc-associated loci. Classical HLA class II genes (DPB1, DQB1 and DRB1) have been clearly associated with SSc. However, the non-classical HLA genes have not been systematically investigated. Here we investigated the HLA genes, PSORS1C1, CCHCR1, MICA, MICB, NFKB1L1, NOTCH4, C6orf10, BTN2L, TAP2 and TNFAIP3 in association with SSc by using next generation sequencing technology (NGS, Ion torrent personal genome machine). We sequenced all exons for the association with SSc by using next generation sequence technology.

639T
Evolutionary constraint and disease associations of post-translational signaling sites in human genomes. J. Reimand, G.D. Bader. The Donnelly Centre, University of Toronto, Toronto, Ontario, Canada.

Interpreting genome variation in the context of phenotype and molecular function, distinguishing disease variants, and assessing personal risk are central challenges of biomedical genomics, hindered by predominantly rare inter-individual variation. Potentially functional protein-coding variants found from genome sequencing and genome-wide association studies are often ranked by evolutionary conservation and population frequency. However it is also likely that certain intrinsic protein features carry special functional significance. We hypothesized that protein regions involved in post-translational modifications (PTMs) are functional hotspots of human genomes. PTMs are biochemical alterations of amino acids that extend the functional repertoire of proteins, conducted by enzymes that recognize substrate proteins through short linear motifs (SLMs) in protein sequence. Currently we have public experimental data on ~130,000 PTM sites in human proteins. The centrality of PTMs in biological processes suggests the functional importance of underlying genome sequence. Here we found that PTM-associated protein regions representing ~11% of protein sequence follow a specific constraint in the human population. This is distinct from major sources of variation such as conservation, recombination, or GC content. PTM regions are also enriched in disease and cancer mutations, further emphasizing the deleteriousness of their variation. PTM constraint is exemplified in chromosomal context, gene signatures of most human tissues, and across a thousand processes and pathways. PTM regions include ~250,000 protein residues whose substitution would disrupt SLMs bound by PTM enzymes. Intriguingly these residues are enriched in disease mutations, confirming the implication that "neutral" sequence is extremely rare in the genome. Finally, we use these neutral regions to estimate the relative effective population size of the X chromosome vs. the autosomes in African and non-African human populations.
Genetic analysis of dendritic cell responses to influenza using RNA sequencing reveals novel genotype by stimulation effects on alternative splicing. (1, 2) J. Chen3,4, C. Edleistein5, P. Bray6, C. Shaw7, 1) Baylor College of Medicine, Houston, TX; 2) Thomas Jefferson University, Philadelphia, PA.

Human platelets are responsible for coagulation physiology and maintenance of haemostatic balance. Abnormal platelet reactivity is associated with various disease states such as bleeding disorders, atherosclerosis, occlusive or thrombotic cardiovascular disorders, inflammation and cancer, resulting in significant morbidity and mortality. The Platelet RNA and eXpression-1 study was designed to investigate significant heritable inter-individual variation in platelet functional properties. This study profiled platelet mRNA and microRNA expression as well as genotype of 154 healthy human subjects. To study the regulatory landscape of the human platelet, we conducted expression quantitative trait locus (eQTL) analysis and identified a large number of loci significantly associated with gene expression levels in cis and trans. The comprehensive approach to cataloging all platelet eQTLs can be of great help in future studies, such as the prioritization of trait-associated loci implicated in genome-wide association studies.
643S
Ttn as a likely causal gene for QTL of alcohol preference on mouse chromosome 2. L. Wang1, Y. Jiao1, Y. Huang1, B. Bennett1, R.W. Williams1, D. Li3, H. Zhao4,5, A. Gelment4,6,7,8, H.R. Franzon9,10, L.A. Farrow1,11, W. Gu1,11. 1) Orthopaedic Surgery, University of Tennessee Health Science Center, Memphis, TN; 2) Pharmacology, University of Colorado Denver, Aurora, CO; 3) Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, TN; 4) Psychiatry, School of Medicine, Yale University, New Haven, CT; 5) Epidemiology and Public Health, School of Medicine, Yale University, New Haven, CT; 6) Genetics, School of Medicine, Yale University, New Haven, CT; 7) VA Connecticut Healthcare Center, West Haven, CT; 8) Neurobiology, School of Medicine, Yale University, New Haven, CT; 9) Psychiatry, School of Medicine, University of Pennsylvania, Philadelphia, PA; 10) VISN 4 MIRECC, Philadelphia VAMC, Philadelphia, PA; 11) Departments of Medicine (Biometric Genetics), Neurology, Ophthalmology, Genetics & Genomics, Biostatistics, and Epidemiology, Schools of Medicine and Public Health, Boston University, Boston, MA.

Many quantitative trait loci (QTL), influencing mouse model phenotypes for alcoholism, have been mapped genetically. However, the gene(s) comprising the QTL (QTG) are largely unknown. In previous work, Dr. Bennett and colleagues created congenic strains carrying the DBA/2IBG (D2) region for alcohol preference (AP) on chromosome 2, on a C57BL/6IBG (B6) background. Subsequently, interval specific congenic recombinant strains (ISCRS), in which the full D2 QTL region was broken into smaller, partially overlapping regions of introgression, were generated and tested. With information from two ISCRS, the QTL has been mapped onto mouse chromosome 2 (Chr2) in a region of 3.4 Mb by using C57BL/6J (B6) x DBA/2J (D2) recombinant inbred (RI) strains as well as by using F2 populations. Several candidate genes, Gad1, ApoG3, Atf2, Sod2 and Sod9, have been evaluated but none of them is confirmed for a definitive role in the regulation of the QTL of AP on Chr2. We have been searching candidate gene for this QTL, intensively by using an integrative approach including 1) bioinformatics tools to search potential function relevant genes of alcohol preference within the QTL region; 2) searching for single nucleotide polymorphisms (SNPs) within the exons of every gene between B6 and D2 in the QTL region; 3) conducting real time PCR to examine the differentially expressed genes between B6 and D2 in the murine kidney, and 4) analysis of association of candidate gene within human population. Titin (Ttn) is known as a giant muscle protein expressed in the cardiac and skeletal muscles. However, its expression level in the tongue is known to be higher than that in the mouse kidney, here we investigated if Ttn plays a role in the regulation of AP. Our data indicated that 1) the expression level of Ttn in the less AP congenic strains is significantly higher than that in B6; 2) the expression of a Ttn probe in the BxD RI strains is negatively correlated to that of AP; 3) One SNP is in up- and the other is in down-stream of Ttn. The alcohol consumption of the B6 genotype is significantly higher than that of D2 genotype in the BxD RI strains, based on data from multiple reports using two-bottle of choice; and 4) the polymorphism of TTN in human population is highly associated with alcoholism. We conclude that Ttn is a likely causal gene for the QTL on Chr2 for the AP.

644M
Cross-population Meta-analysis of eQTLs: Fine-mapping and Functional Study. X. Wen1, G. Moyer-Brailean2, F. Luca3, R. Pique-Regi4, R. Ngu4, O. S. Franckowiak5. 1) Dept Biostatistics, Univ. of Michigan, Ann Arbor, MI; 2) Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI.

Data sets collected from multiple ethnic groups provide great potentials to understand the genetic basis of gene regulations. Nevertheless, the joint analysis of multiple-population eQTL data presents great challenges, mainly due to varying LD patterns and heterogeneity of genetic effects presented in multiple populations. In this study, we develop a set of statistical and computational tools to tackle the problem of cross-population meta-analysis of eQTLs, primarily focusing on the applications of fine-mapping and functional study. Our approach holds three distinct advantages over the existing methodologies:

1. Our method effectively identifies eQTL signals that show consistent genetic effects across populations while accounting for the potential heterogeneity based on a powerful Bayesian meta-analysis framework (Wen and Stephens, 2014 AOAS; Flutre et al. 2013 PLoS Genetics).
2. Our approach performs efficient multiple SNP fine-mapping analysis across multiple populations allowing varying LD patterns.
3. Building on the fine-mapping results, our method enables high-resolution functional analysis of eQTL signals accounting for LD.

We apply our methods to re-analyze the GEUVADIS data consisting of five population groups. Furthermore, we use base-pair resolution annotation obtained with an improved CENTIPEDE method using ENCODE DNase-seq data. Our preliminary results show that:

1. Meta-analysis is more powerful and robust in identifying eQTL signals. We can identify 6,555 genes harboring at least one eQTL at FDR 5% level, in their cis-regions (eGenes) using the joint meta-analysis approach. In comparison, separate subgroup analysis identifies 1,803 (TOSI), 2078 (GBR), 2100 (FIN), 960 (CEU) and 1042 (YRI) eGenes.
2. Many genes containing cis-eQTLs are shared in their cis-regions. With the meta-analysis powered fine-mapping approach, we are able to confidently identify 674 genes containing 2 or more independent cis-eQTL signals and 67 genes containing 3 or more signals.
3. Binding variants are significantly enriched in the identified population consistent eQTL signals (p-value < 1.6 × 10-11).

645T
Linking systems genetics and co-expression analysis to elucidate diabetic kidney disease regulatory networks. T. Leak1, C. Komorosky1, V. Nair1, H. Huang1, B. Keller1, A. Randolph1, R. Nelson2, M. Kretzler1, T. Werner2,3,4. 1) Internal Med-Nephrology, Univ of Michigan, Ann Arbor, MI; 2) National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona; 3) Genomatix Software GmbH, Munich, Germany.

Expression quantitative trait loci (eQTL) analysis associates genetic and transcriptomic data, but does not provide insight into how eQTLs influence sets of genes in complex diseases. Using diabetic kidney disease (DKD) as a case study, genome-wide SNPs, protocol biopsy gene expression and clinical data were obtained from 65 Pima Indians. Genome-wide systems genetics was paralleled with weighted co-expression framework analysis (WGCNA) was performed on gene expression profiles from microdissected glomeruli. Integration of glomerular eQTL analysis and WGCNA was carried out to identify transcriptional coregulation-modules. Matrix eQTL was employed to test allele dosage effect on gene expression levels. An in silico promoter analysis approach was used to identify common transcription factor binding sites (TFBS) frameworks shared between the promoters of cis-eQTL transcripts and potentially within DL blocks of associated expression (eSNPs). Multivariate linear regression was then used on the module-eQTLs, promoter gene matches, TF belonging to TFBS families represented in the framework and additive genotypic model, to test for association with indices of renal function and diabetes traits. A total of 13,179 transcripts were parsed into 37 co-expression modules and Matrix eQTL resulted in 3,112 cis-eQTL gene pairs. To establish a mechanistic link between correlated cis-eQTL modules, a module with a similar number of transcripts and eSNPs was analyzed via an in silico promoter analysis approach (dark magenta module). Framework analysis of tag-eSNP LD block sequences with kidney associated promoters revealed a four-enzyme framework (ETS-SP1-PAX2-PAX5-HOMF; 5.64 fold over presented) that matched 2 of the 3 eQTLs (TAF9, HPGD). The LD block spans TAF9, RAD17 and MARVELD2, two of which (TAF9 and RAD17) have matches of the framework in their promoters. A genome-wide screen for all human promoters (N=101,492 revealed only 17 matches. Four of these genes, ETS TFs and rs4640041 (located in the same LD block and an eSNP for RAD17) showed suggestive evidence of association with decreased renal function (P ≤ 0.000001). Our integrative framework potentially highlights how eSNPs may influence the expression of several genes linked in a LD block with renal function. Preliminary data indicate that other eQTLs may also act on gene sets employing similar strategies, i.e TFframeworks that could coordinate transcriptional regulation, all of which is linked with DKD.
646S  
1) Health Studies, Univ Chicago, Chicago, IL; 2) UChicago Research Bangladesh, Dhaka, Bangladesh; 3) University of North Carolina, Lineberger Comprehensive Cancer Center, Chapel Hill, NC.  
A large fraction of human genes are regulated by genetic variation. However, our understanding of the mechanisms by which genetic variants influence transcription is incomplete. A more in-depth understanding of eQTLs (expression quantitative trait loci) can be gained by identifying cellular features that mediate the effects of genetic variants on gene expression. In this work, we use Sobel tests of mediation to determine (1) if trans-eQTL associations are mediated by transcripts in cis with the SNP showing the trans-association and (2) if cis-eQTL associations are mediated by local chromatin structure (measured as DNA methylation). Using data on 1,799 Bangladeshi individuals with genome-wide SNP and array-based expression data, we demonstrate that among 197 independent trans-eQTLs (FDR< 0.05), 76 were also cis-eQTLs, and evidence of mediation was detected for 55 of these 76 trans-eQTLs (Sobel P <0.01). Using a subset of these participants (n=377) with genome-wide DNA methylation data (Illumina 450k), we detect mediation by cis-CpG methylation status for 22% of 2,598 observed cis-eQTLs, suggesting that alteration of chromatin features is a common mechanism by which genetic variation influences gene expression. In these analyses, “partial” mediation, rather than “full” mediation, is most often observed, and using simulated data, we show that this result is expected. More specifically, in the presence of mediator measurement error or potential LD between mediator and causal variants, a statistically detected mediation will be observed as partial mediation. Our results also show that eQTL associations can become significantly stronger after adjusting for a potential mediator, a counterintuitive observation. Using simulated data, we demonstrate that this phenomenon can occur when the true mediator is not the true mediator, but is correlated with the true (unmeasured) mediator. In this work we demonstrate mediation for a large fraction of cis- and trans-eQTLs; however, our ability to detect mediation would be enhanced by the use of additional mediating features, such as histone modifications. Mediation analysis will be useful for validation and discovery of eQTLs and is a valuable tool for characterizing the biological mechanisms underlying eQTLs.

648T  
Analysis of genetic and environmental determinants of gene expression. F. Luca1, C. Harvey1, G.O. Davis1, G. Moyerbrailean1, D. Watzla1, X.W. Wen1, R. Pique-Regi1. 1) Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI; 2) Department of Biostatistics, University of Michigan, Ann Arbor, MI.  
The effect of genetic variants on a molecular pathway, and ultimately on the individual's phenotype, is likely modulated by "environmental" factors. However, it is generally difficult to determine in which tissues and conditions genetic variants may have a functional impact. We denote the functional genetic variants that show cellular environment-specific effects as GxE expression quantitative trait loci (GxE-eQTLs). Achieving a better understanding of the mechanisms underlying GxE-eQTLs is a critical step in understanding the link between genotype and complex phenotypes. To identify and characterize GxE-eQTLs we have established a new two-step and cost-effective experimental approach. In the first step, we identify global changes in gene expression using low-coverage sequencing of pools of highly multiplexed samples. In the second step, we select a subset of samples for deep sequencing and allele-specific analysis. For the first step, we generated 1056 RNA-seq libraries in pools of 96 spanning 250 cellular environments across 5 cell-types (3 individuals), and 50 different treatments (including hormones, dietary components, environmental contaminants and metal ions). Relevant GO categories were enriched in the observed global gene expression changes (e.g. immune response for Dexamethasone, ion homeostasis for Zinc). We then analyzed allele specific expression (ASE) using a novel method (Quasar) that allows for joint genotyping and allele specific analysis on RNA-seq data. Across 21 cellular environments we discovered 2469 instances of ASE (FDR<10%), corresponding to 1402 ASE genes. Using Bayesian model across treatments within cell types, we observed that 95.9% ASE signals are shared (95% credible interval [93.4%-99.3%]) and their effect sizes are highly concordant (posterior correlation coefficient 0.9). This is highly consistent with previous analysis of condition-specific eQTLs. Nevertheless, we still identified 270 loci with a Bayes factor supporting GxE (200 sites treatment-specific and 70 sites control-specific GxE-eQTLs). We are now in the process of expanding the ASE analysis to 32 additional cellular environments. Our results constitute a comprehensive catalog of GxE-eQTLs and we anticipate that it will contribute to the discovery and understanding of GxE interactions underlying complex traits.

647M  
Characterizing the genetic basis of innate immune response in TLR4-activated human monocytes. S. Kim1, B. Pütz2, M. Bechheim3, J. Becker1, M. Nöthen1, B. Müller-Myhsok1, V. Hornung2, J. Schumacher1. 1) Institute of Human Genetics, University of Bonn, Bonn, Germany; 2) Institute of Molecular Medicine, University of Bonn, Bonn, Germany; 3) Statistical Genetics, Max-Planck-Institute of Psychiatry, Munich, Germany.  
Toll-like receptors (TLRs) play a key role in innate immunity. Apart from their function in host defense, dysregulation in TLR-signaling can confer risk to autoimmune diseases, septic shock or cancer. Despite major advancements in our understanding of how the innate immune system recognizes pathogens, the genetic basis for differences in innate immune responses is only poorly defined. This study was aimed to characterize the genetic basis of variation in gene expression in TLR4-stimulated human monocytes. For this purpose we isolated monocytes of 137 individuals and stimulated them with lipopolysaccharide (LPS) to activate Toll-like receptor 4 (TLR4). From these donors, we performed transcriptome profiling and genome-wide SNP-genotyping. Comparing unstimulated versus TLR4-stimulated monocytes revealed expression quantitative trait loci (eQTLs) that are unique to TLR4 stimulation. Among these, we show that SNPs conferring risk to primary biliary cirrhosis (PBC), inflammatory bowel disease (IBD) and celiac disease are immune response eQTLs for novel candidate genes, bringing new insights into the pathophysiology of these disorders in the context of TLR4-activation.
469S


Although genome-wide association studies (GWAS) have implicated thousands of common variants in various human diseases, our ability to localize association signals to individual causal alleles - and thereby elucidate mechan-isms of pathogenesis - remains poor. This is partially due to our inadequate understanding of the effects of non-coding variation, which includes the majority of disease associations. To address this, we extended the use of the massively parallel reporter assay (MPRA) as a tool for the identification of non-coding causal alleles. We introduced modifications to MPRA allowing us to empirically measure the regulatory potential of 29,000 single nucleotide polymorphisms (SNPs) and the differential allelic effects at those sites in a single experimental setup. To evaluate the assay, we tested all variants in perfect linkage disequilibrium with the highest-ranking SNP for each of 3,992 eQTLs found in the 1000 Genomes lymphoblastoid cell lines (CLLS). In addition, we performed more comprehensive testing at eQTL peaks overlapping 163 variants from the NHGRI GWAS catalog. Overall, we successfully interrogated >95% of the targeted loci and detected thousands of regions that enhance expression activity. The identified regulatory loci are highly enriched for DNAse hypersensitivity sites, as well as for histone marks predicting enhancer and promoter activity. At the same time, a substantial proportion of regulatory sequence remains unaccounted for by current mark-ers. We identified putative causal alleles for hundreds of eQTL peaks and show strong correlation with traditional luciferase reporter assays. We will also discuss the detailed several instances where the causal allele localizes to known disease-associated loci. Having experimentally identified regulatory alleles in two separate 1000 genome cell lines, we evaluated the same regions in a hepatocyte cell line, thus providing direct evidence of hundreds of tissue-specific regulatory elements. This work demonstrates the applica-tion of MPRA to functional genetics, allowing us to pinpoint causal alleles and investigate the molecular mechanisms underlying disease/trait associations.

650M

Long intergenic non-coding RNA eQTLs are enriched for complex trait-associated SNPs and do not distally regulate the expression of protein-coding genes. I. McDowell1, C. Guo2, G. Tex Consortium3, A. Pali4, C. Brown1, T. Reddy5, B. Engelhardt6. 1) Computational Biology & Bioinformatics, Duke University, Durham, NC; 2) University Program in Genetics and Genomics, Duke University, Durham, NC; 3) Biostatistics & Bioinformatics, Duke University, Durham, NC; 4) Center for Human Genetics Research, MIT, Cambridge, MA; 5) Department of Genetics, University of Pennsylvania, Philadelphia, PA.

Long intergenic non-coding RNAs (lincRNAs) participate in a variety of biological processes including mammalian X-chromosome inactivation, somatic maintenance, age-related, and cancer progression. Expression quantitative trait loci (eQTL) studies have proved effective in characterizing the genetic regulation of protein-coding genes but have largely neglected lincRNA genes for a number of reasons. Here, we performed a genome-wide association study to identify regulatory variants associated with lincRNA and protein-coding gene expression in the Genotype-Tissue Expression (GTEX) pilot data. Genotypic variation at 10 M single nucleotide polymorphisms (SNPs) including SNPs imputed from over 3.5 M quality-filtered SNPs was associated with the expression of 4,000 lincRNA genes and 18,000 protein-coding genes as quantified by RNA-seq. We performed association mapping in four tissue types independently by an additive effects Bayesian regression model and across all four tissues jointly using a multi-tissue Bayesian regression model. We identified cis-lincRNA-eQTLs (linc-eQTLs) associated with 1,566 lincRNA genes (36%) and cis-protein-coding eQTLs (pc-eQTLs) associated with 9,794 protein-coding genes (54%) at a 5% FDR. We found substantial sharing of linc-eQTLs and pc-eQTLs across tissues, suggesting that the regulation of both lincRNA and protein-coding genes relies on regulatory elements that are conserved across tissues. Nevertheless, linc-eQTLs were more tissue-specific overall than pc-eQTLs after controlling for tissue-specific expression levels, suggesting greater tissue-specificity of linc-eQTLs relative to pc-eQTLs. While both linc-eQTLs and pc-eQTLs were enriched for enhancer linkages to associate SNPs, the patterns for linc-eQTLs and pc-eQTLs were different, with linc-eQTLs more enriched for DNase hypersensitivity sites, as well as for histone marks that enhance expression activity. The identified regulatory loci are highly enriched for DNAse hypersensitivity sites, as well as for histone marks predicting enhancer and promoter activity. At the same time, a substantial proportion of regulatory sequence remains unaccounted for by current markers. We identified putative causal alleles for hundreds of eQTL peaks and show strong correlation with traditional luciferase reporter assays. We will also discuss the detailed several instances where the causal allele localizes to known disease-associated loci. Having experimentally identified regulatory alleles in two separate 1000 genome cell lines, we evaluated the same regions in a hepatocyte cell line, thus providing direct evidence of hundreds of tissue-specific regulatory elements. This work demonstrates the application of MPRA to functional genetics, allowing us to pinpoint causal alleles and investigate the molecular mechanisms underlying disease/trait associations.
Uncovering expression variability and eQTLs on the X chromosome. K. Kukurba1,2, A. Battle3, S.M. Montgomery1,3. 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Pathology, Stanford University, Stanford, CA; 3) Department of Computer Science, John Hopkins University, Baltimore, MD.

The X chromosome plays a distinctive role in genotype-phenotype relationships. However, existing genetic associations and eQTL studies have entirely disregarded it due to the special challenges of conducting analyses on the X chromosome; in particular, accounting for properties including random X-inactivation, dosage compensation, and distinct modes of inheritance relative to autosomes. To deal with these challenges and to better identify genotype-phenotype patterns on the X chromosome, we compared the utility of different statistical models for eQTL discovery on the X chromosome. By applying well-performing models, we studied X chromosome eQTL within the Depression Genes and Networks (DGN) study cohort of 922 genotyped individuals with RNA-sequencing data from whole blood. In the male and female populations, we detected a high proportion of shared eQTLs (85%) of X-chromosome eQTLs are shared between sexes as well as identified sex-specific eQTLs on the X chromosome. We further characterized the properties of these eQTLs by comparing effect sizes and evidence of selective constraint between the X chromosome and autosomes, specifically evaluating variants in lower frequency ranges. In addition, we investigated the contributions of sex to overall expression variation across the transcriptome. We discovered that females have higher expression variance on the autosomes. Conversely, we found that males have higher expression variance on the X chromosome than females, a pattern expected in males where hemizygosity results in more extreme effects arising from genetic variation. Lastly, we examined the phenomenon of random X-inactivation in females across different ages. We found that older individuals typically have more allelic skewing on the X chromosome compared to younger individuals; an observation which may be explained by clonal selection in the hematopoietic stem cell population. Together, this work advances our understanding of how sex and eQTLs shape human expression variation on the X chromosome contributing to sexual dimorphism.

Expression Quantitative Trait Loci (eQTL) describe a fundamental aspect of biology: inherited allelic variation that impacts gene expression. Universal eQTL (uQTL) describes contexts where the same eQTL relationship, indicated by repeated identification of an association between genotype changes in one genomic region and expression of the same gene, is almost always present regardless of the tissue analyzed or ancestry of the sample population. These uQTL are inherently interesting because they indicate a strong connection between the regulation of gene expression and the genetics that influence expression, and because they represent cases of super-replication that can be leveraged to infer the genetic causes and impacts of eQTL. While the existence of eQTL has been indirectly acknowledged there has not been a targeted effort to discover uQTL.

Making use of an eQTL detection pipeline developed in our lab capable of independently testing billions of genotypic-expression associations in a few hours, we analyzed >35 publicly available and new datasets, including HapMap lymphoblastoid cell lines and cancer/healthy tissues in the TCGA and of independently testing billions of genotype-expression associations in a study cohort of 922 genotyped individuals with RNA-sequencing data from whole blood. In the male and female populations, we detected a high proportion of shared eQTLs (85%) of X-chromosome eQTLs are shared between sexes as well as identified sex-specific eQTLs on the X chromosome. We further characterized the properties of these eQTLs by comparing effect sizes and evidence of selective constraint between the X chromosome and autosomes, specifically evaluating variants in lower frequency ranges. In addition, we investigated the contributions of sex to overall expression variation across the transcriptome. We discovered that females have higher expression variance on the autosomes. Conversely, we found that males have higher expression variance on the X chromosome than females, a pattern expected in males where hemizygosity results in more extreme effects arising from genetic variation. Lastly, we examined the phenomenon of random X-inactivation in females across different ages. We found that older individuals typically have more allelic skewing on the X chromosome compared to younger individuals; an observation which may be explained by clonal selection in the hematopoietic stem cell population. Together, this work advances our understanding of how sex and eQTLs shape human expression variation on the X chromosome contributing to sexual dimorphism.

Universal eQTL: discovery and replication across cell type and population. S.A. Shenoy1, T.L. Vincent1, R.G. Crystal1, J.G. Mezey1,2. 1) Department of Genetic Medicine, Weill Cornell Medical College, New York, NY; 2) Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY.

Sexual dimorphism is observed across a variety of complex phenotypes in model organisms and humans, including immune function in healthy individuals, as well as autoimmune disorders and neuropsychiatric disorders, among others. In model organisms, and to a lesser extent in humans, sex-specific quantitative trait loci have been identified for sexually dimorphic traits. To extend our understanding of the contribution of sex to the genetic basis of human immune function and dysfunction, we performed a sex-aware expression quantitative trait locus (eQTL) study of highly purified CD4+ T cells and monocytes, representing adaptive and innate immunity, in a cohort of 212 healthy individuals, as part of the Immunological Monitoring (ImmVar) Project. We estimate that of genes with significant cis-eQTLs in each sex, less than 10% are unique to one sex, with the degree of sex bias approximately equal in each of the two cell types. Where associations are shared across sexes, we see strong correlation between effect sizes of male- and female eQTLs (r2=0.94, CD4; r2=0.67, CD4+), suggesting little in the way of sex-specific regulation of alleles. We observe bi-directional regulation in cases where cis-eQTLs are shared across sexes but show opposite direction of allelic effects across sexes. We do, however, observe cis-eQTL allelic direction changes across cell-types, as has been reported previously. Although the degree of sex-bias between males and females is similar across the cell-types, we observe tissue-specificity in the genes with sex-biased eQTLs. Functional analysis of sex-specific eQTLs within cell types identifies different categories of significant enrichment, with sex-specific eQTLs enriched in antigen and immune function, and mitogen and oxidative stress in CD4. We identified genetic variants previously associated with complex traits in human diseases as cataloged in the NHGRI catalog of genome-wide association studies that are also significant sex-specific eQTLs in our datasets. These include variants associated with diseases that exhibit sexually dimorphic characteristics, including Crohn’s disease, bipolar disorder, and rheumatoid arthritis. Our results characterizing the sex-specific genetic architecture of the transcriptome in these immune cells will be compared and contrasted to our parallel analyses of the same cell-types in populations of different ancestry, as well as in tissues of the Genotype-Tissue Expression (GTEx) Project.
656M

Genome-wide Identification of microRNA Expression Quantitative Trait Loci in the Framingham Heart Study. T. Huan¹,², J. Rong³, C. Liu¹,², X. Zhang¹,², K. Tannirne⁴, R. Johanes¹,², B. Chen¹,², J. Murabito¹, C. Yao¹,², P. Courchesne¹,², P. Munson⁵, C. O’Donnell¹,², N. Cox⁵, A. Johnson¹,², M. Larson¹,², D. Levy¹,², J. Freedman². 1) The National Heart, Lung, and Blood Institute’s Framingham Heart Study, 73 Mt. Wayte Avenue, Framingham, MA, USA; 2) The Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA; 3) Department of Mathematics and Statistics, Boston University, Boston, MA, USA; 4) Department of Medicine, University of Massachusetts Medical School, Worcester, MA, USA; 5) Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, MD, USA; 6) Department of Human Genetics, University of Chicago, Chicago, IL, USA.

Introduction: Identification of miRNA expression quantitative trait loci (miR-eQTL) may yield insights into regulatory mechanisms of miRNA transcription, and thus may help elucidate important roles of miRNAs as mediators in complex traits. Several published miR-eQTL studies were based on modest sample sizes (n<200). These studies revealed a few cis-miR-eQTLs, but uncertainty persists regarding the number of miR-eQTLs and their relations to regulatory elements in the human genome. Methods: We analyzed the associations of approximately 10 million 1000 Genomes Project imputed SNPs (at minor allele frequency >0.01) with whole blood miRNA expression levels measured by Taqmen qPCR using a Fluidigm platform in 5239 Framingham Heart Study (FHS) participants. Associated SNPs residing within 1Mb of the mature miRNA (cis) and those occurring more than 1 Mb away (trans) were identified separately using an additive regression model adjusted for age, sex, and family structure. Replicability of the identified cis and trans miR-eQTLs was tested by splitting the overall samples 1:1 into pedigree independent sets. Results: 280 miRNAs were expressed in >200 individuals and were used for analysis. At a false discovery rate <0.1, we identified 5269 cis miR-eQTLs (representing 982 loci at linkage disequilibrium r2<0.7) for 76 mature miRNAs. cis miR-eQTLs showed high concordance among both pedigree independent sets (the concordance rate is 50%). We found that most cis miR-eQTLs were located upstream of their associated mRNA (57%) and intragenic (40%), 5% were intergenic miRNAs. 59% of cis miR-eQTLs were located 300–500kb upstream of their associated intergenic miRNAs, suggesting that distal regulatory elements may affect interindividual variability in miRNA expression levels. We also found that cis miR-eQTL SNPs were highly enriched for cis mRNA eQTLs and regulatory SNPs (e.g., SNPs in promoter, enhancer, and transcription factor binding sites annotated by ENCODE). By cross-linking cis miR-eQTL SNPs with GWAS SNPs and by linking cis miR-eQTL miRNAs with differentially expressed miRNAs for complex traits we identified 245 cis miR-eQTLs that are associated in GWAS with complex traits. Several of the cis miR-eQTL miRNAs displayed differential expression in relation to the corresponding GWAS trait in the FHS. Conclusion: Our study provides a roadmap for understanding the genetic basis of miRNA expression, and sheds light on miRNA involvement in a variety of complex traits.

657T

Dissecting the genetic regulation of exosome RNA cargo in a large family. E.K. Tsang¹,², X. Li²,³, V. Anaya³, K.J. Karczewski¹, D.A. Knowles³, K.S. Smith²,³, S.B. Montgomery²,³. 1) Biomedical Informatics Program, Stanford University, Stanford, CA, United States; 2) Department of Pathology, Stanford University, Stanford, CA, United States; 3) Department of Computer Science, Stanford University, Stanford, CA, United States; 4) Department of Genetics, Stanford University, Stanford, CA, United States; 5) Co-first authors; 6) Co-senior authors.

Many studies have characterized genetic variation affecting gene expression through mapping of expression quantitative trait loci (eQTLs) in a variety of isolated cell types and conditions. Meanwhile, little attention has been paid to the impact of such genetic variation on intercellular communication. Exosomas are one mechanism through which cells transfer information. They are small extracellular vesicles that package biomolecules, including RNA and protein, from their cell of origin, and can deliver this cargo to influence neighboring cells. Exosomes are produced by many different cell types and are thought to be involved in the progression of certain diseases, including cancer, as well as being important for maintaining homeostasis. However, while exosomes may be relevant to several biological processes, the regulation of exosome contents remains poorly understood. In this work, we sequenced the small RNA transcriptomes of lymphoblastoid cell lines and their exosomes from a seventeen-member three-generational family. We investigated which transcripts were specifically exported in exosomes or retained in cells through differential expression analyses. By combining publicly available whole genome sequence data for the seventeen individuals with our RNA sequence data, we mapped eQTLs in both cells and exosomes and evaluated the amount of sharing between the two groups. While eQTLs of different cell types and conditions were identified, between 13% and 40% of genes were differentially expressed between cells and exosomes. This finding suggests that the RNA contents of exosomes are actively regulated and are not simply a reflection of the cellular transcriptome. We also searched for motifs enriched in the transcripts more highly expressed in exosomes than in cells to find evidence of export signals. To characterize the effect of genetic variability on exosome contents, we performed the first eQTL study in exosomes. We found evidence of sharing of small RNA expression between cells and exosomes. We also identified cell- and exosome-specific eQTLs. Since exosomes are mediators of intercellular communication, our characterization of eQTL in exosomes highlights the impact of regulatory variants beyond the level of individual cells.

658S

Analysis of hypothalamic transcriptome and proteome in 100 strains of mice on high fat diet. Y. Hasin¹, A. Khan², V.A. Petyuk³, B.W. Parks¹, C.D. Rau¹, C. Pan³, P.D. Peiowski³, R.D. Smith⁴, A.J. Lusis⁴, D.J. Smith⁴. 1) David Geffen School of Medicine, Cardiology, UCLA, Los Angeles, CA; 2) Department of Molecular & Medical Pharmacology, University of California, Los Angeles; 3) Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington.

Regulation of body weight and appetite are multifaceted processes, in which hypothalamic nuclei play a pivotal role. We here report the first integrated transcriptome and proteome analysis in the mouse hypothalamus, using paired end RNA sequencing and LC-MS Proteomics Analysis in 100 inbred mouse strains from the Hybrid Mouse Diversity Panel (HMDP). HMDP strains allow high resolution mapping and we were able to identify thousands of new transcripts and isoform variants as well as mapping 283 pQTLs. We are currently constructing gene co-expression network from these data, using linear and non linear correlation measurements. In addition these mice were extensively phenotyped for various of metabolic phenotypes, such as plasma metabolites and lipids, which allows us to associate gene and protein expression with particular phenotypes.
659M  
Quantifying Context Specificity of Gene Regulation using Predicted Gene Expression Levels.  S.V. Mozaffari1, E.T. Garnav1, 2, K. Aquino-Michaels2, N.J. Cox3, H.K. Im5, 1 Committee of Genetics, Genomics, and Systems Biology, University of Chicago, Chicago, IL, USA; 2 Department of Medicine, Section of Genetic Medicine, University of Chicago, Chicago, IL, USA; 3 Department of Human Genetics, University of Chicago, Chicago, IL, USA; 4 Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA; 5 Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA.

Genetic regulation of gene expression is likely to play an important role in the etiology of complex traits as indicated by the enrichment of expression-associated variants (eQTLs) and DNase hypersensitivity sites among trait-associated variants. Tissue specificity and other context dependent regulation of the transcriptome is an area of current interest. We propose a method to quantify the degree of context specificity by using prediction models trained and tested in different contexts. For example, we examined the context specificity of gene regulation by quantifying the ability to predict gene expression levels using models trained in whole blood, tibial nerve, and muscle, and tested on expression levels assayed in lymphoblastoid cell lines. We generated predictive models of gene expression levels using single variant regression coefficients. SNPs significantly associated with gene expression levels were kept in an additive model with p-value threshold of 0.05 for variants in the vicinity of the gene and 0.05 for variants that were in linkage disequilibrium. We used expression data from whole blood, tibial nerve, and muscle from the GTEx pilot data for training and independent lymphoblastoid cell lines (LCL) from 1000 Genomes as test set. When using whole blood models, we found that for 8.5% of the genes, the predicted levels were significantly associated with the observed levels in LCLs with FDR <5%. For 15% of the genes, correlation was better than 10%. ERAP2 was the best predicted gene with correlation between predicted and observed of 0.87. When using models trained with tibial nerve and muscle data, 7.5% and 7.4% of the genes showed correlation with FDR less than 5%. Performance is slightly inferior to whole blood indicating more similarity between LCL and whole blood than with tibial nerve or muscle. As expected, we found that genes with higher heritability, higher expression levels, and higher variability tended to be better predicted. In conclusion, we present a novel prediction-based method to quantify context specificity of gene regulation and find that a substantial portion of the regulatory mechanism is likely to be common across tissues.

660T  
Genetic architecture of the transcriptome of four tissues in a twin cohort.  A. Bull1; A. Vihuela1; A.A. Brown1; M. Davies1; K. Small1; R. Durbin2; T.D. Specter2; E.T. Dermitzakis1.

1) Genetic Medicine and Development, University of Geneva, Geneva, Switzerland; 2) Wellcome Trust Sanger Institute, Hinxton, United Kingdom.

Understanding the genetic architecture of gene expression is an intermediate step to understand the genetic architecture of complex diseases. To explore the underlying causes of gene expression variation, we aimed at characterizing the mRNA fraction of the transcriptome of ~400 female MZ and DZ twin pairs from the TwinUK cohort in four tissues: fat, skin, blood and lymphoblastoid-cell lines (LCLs) [2330 samples in total] and used genotype information imputed into the TwinUK cohort to perform subsequent analysis. Using a linear regression approach with SNPs in a 1Mb window each side of the TSS for each gene and tissue we identified 9166 significant cis-eQTLs in fat, 9551 in LCLs, 8731 in skin and 3531 in blood (1% FDR). But cis-eQTLs are only a small part of the genetic effects that affect gene expression. By exploiting the twin structure of our sample, we disected the proportions of gene expression variation which is due to different genetic and non-genetic causes. We observed that, on average, common cis-eQTL only explained about 15% of the variance in gene expression while the majority of the genetic effects that affect gene expression were due to genetic variants in cis (mainly rare variants or common variants with small effects) explained about 30% of heritability. The remaining 50% of the heritability was explained by genetic variants in trans. Moreover, we observed that between 70% and 80% of the SNPs that were cis-eQTL in one tissue have a similar effect in another cis-eQTL signaling another tissue. Genetic variation may also affect gene expression by modifying mRNA splicing processes. We calculated the association of cis SNPs and the frequencies of exon-exon links per individual and identified 15,1566 and 41,064 splicing QTLs (eQTLs) per tissue. We observed that between 35% to 60% of the QTLs are shared among tissues, highlighting the importance of tissue specific effects in alternative splicing. Finally, we looked for genetic variants that have a cis effect in one tissue and, at the same time, a downstream effect, in trans, in another tissue. We defined a downstream signal as a QTL that regulates expression in cis (gene expression) and answered. For example: How replicable are eQTL, especially trans-acting or distant eQTL? Which eQTL act in a tissue-specific manner, and which act across tissues? In order to provide more insight into gene regulation, we used a cross-platform and cross-tissue comparison of son of eQTL signals using data from two gene expression platforms, RNA-seq and expression microarrays, in a genetically-characterized pedigree of vervet monkeys (Chlorocebus aethiops sabaeus) obtained from the Vervet Monkey Colony (VRC). For eQTL mapping of steady-state transcription levels, we used two data sets: 1) RNA sequencing (RNAseq) in a matched set of five tissues (blood, skin fibroblasts, caudate, adrenal, and pituitary) from each of 35 monkeys (a total of 175 samples), and 2) microarray gene expression data, in peripheral blood samples from 327 monkeys. The gene expression data was inverse-normal transformed. We used genotype data from the VRC genotype mapping set consisting of 500K common autosomal SNPs that was generated based on whole genome sequencing (WGS) from the VRC monkeys. For association analysis between genotypes and gene expression levels, we used the EMMAX statistical test, using age, sex, batch, and pregnancy status as covariates. Among the 5,539 gene expression probes that passed our filter for high quality probes, 3,134 probes (57%) were heritable at FDR<0.01 (SOLAR). Among the heritable probes, we identified 101 significant eQTLs, 24 significant eQTLs in cis and 37 probes in trans (i.e., on a different chromosome than a probe). The trans signals include two eQTL hotspot regions where five or more probes (gene expression) were mapped (on chromosomes 5 and 9). Forty percent of the transcripts that were regulated by eQTL in trans signals were also regulated by cis-eQTL. We then ran our prediction models on the RNAseq data at the significance level of p<0.05. Despite a small sample size with the RNAseq data, we identified 59–238 genes per tissue that were regulated by eQTL (p<5e-08). That includes genes regulated across all four tissues, LRPC3 and 31 genes in three tissues, and 23 genes in two tissues.

The emergence of high throughput sequencing (HTS) technologies has coincided with the development of advanced genomic reference populations, including the mouse Diversity Outbred (DO) heterogeneous stock. The application of HTS to genetically diverse mapping populations has the potential to provide nucleotide resolution of causal variants underlying phenotypic differences. However the increase in information content comes at the cost of increased analytical complexity. We have developed novel methods and software to exploit the high genetic diversity and heterogeneous diploid genomes of the DO to yield new layers of information and inform fine mapping of phenotypic and expression QTL (eQTL). Importantly, these methods can be readily applied to any human expression dataset where genotyping data is available.

RNA-seq alignment to individualized diploid genomes yields direct, accurate estimates of allele specific expression (ASE), and improves eQTL identification and resolution. To illustrate this, we profiled the liver transcriptomes of 450 DO mice by RNA-seq and estimated gene, isoform, and allele expression. We correlated gene expression differences to genetic variation and identified 9,000 local eQTL (eQTL) and 900 distant eQTL, showing that most of the variation in transcript abundance derives from segregating local genetic variation. Allelic expression differences confirmed that cis-acting mechanisms underlie most local eQTL and DO allele estimates correlate well with gene expression in livers from the eight DO founder strains. Most eQTL appear biallelic suggestive of a single causal variant, however complex 3- and 4- allele patterns are observed. Cis-eQTL with allelic expression patterns that deviate from the known strain ancestry are most amenable to fine mapping, and in some cases provide evidence of a causal variant. Distal eQTL with large effects are rare in the adult liver transcriptome, however we have developed conditioning methods to amplify trans effects and test candidate regulatory genes in the interval.

Segnature software for constructing individualized diploid genomes and gene annotation files are available at github.com/jaxcs/Segnature. Analytical pipelines described above can be found at do.jax.org. All eQTL results are available as an interactive web application at do.jax.org.
667S
Meta-Analysis of Liver eQTL Studies and Cross-tissue eQTL Compari-
sion using GTEx Data. E.L. Seiser1, K. Xia2, F.A. Wright1, F. Innocenti1.
1) Division of Pharmacotherapy and Experimental Therapeutics, The
University of North Carolina at Chapel Hill, Chapel Hill, NC; 2) Department
of Biostatistics, The University of North Carolina at Chapel Hill, Chapel Hill, NC;
3) Department of Statistics, North Carolina State University, Raleigh, NC.
Identification of liver-specific expression quantitative trait loci (eQTLs) may
aid in determining genes and genetic variants relevant to the pathogenesis
of complex liver-associated diseases and the biological mechanisms under-
lying drug pharmacology by providing a functional link between genetic
variation and expression levels. Previous liver eQTL studies have
resulted in low replication rates, suggesting the need for a larger overall
sample size to increase the statistical power for identifying eQTLs. We
performed a meta-analysis of four human liver eQTL studies (n= 149, 206,
427, 960), including three previously published, to identify cis-eQTLs (within
100kb of transcription start/stop site). Genotype microarray data from
each study was used for determining sex, genetic ancestry, and common samples
between studies, leading to a total of 1,226 Caucasian liver samples for
analysis. Imputation of genotype data using Impute2 produced a set of
6,684,936 SNPs common among all studies. Probe sequences from each
expression microarray platform were mapped to Entrez genes and a set of
13,669 genes common to all platforms was identified. Using expression data
from each study, probabilistic estimation of expression regions (PEER) was
employed to determine hidden variance components. Matrix eQTL was used
to test for cis- associations, incorporating sex, age, genetic ancestry, and
PEER factors into an additive genetic model. A meta-t-statistic was calcu-
lated using SNP/gene pair t-statistics from each study, and p-values for the
meta-t-statistics were generated using the normal distribution. Testing of
8,236,407 SNP/gene pairs identified significant cis-linked genetic effects on
expression levels (Bonferroni corrected p-value < 6.10−9) for 3,334 genes,
including numerous genes related to drug processing. Ongoing comparative
analysis of the liver eQTL data with the Genotype Tissue Expression (GTEx)
project eQTL data from 9 tissues (whole blood, adipose, muscle, heart,
artery, lung, skin, nerve and thyroid) will allow for a comprehensive identifi-
cation of liver-specific eQTLs, and may help to further elucidate pathogenic
and biological processes in the liver.

668M
Functional mapping of eQTL signals for prostate cancer risk SNPs. L.
Tilmans, S. McDonnell, A. French, Y. Zhang, S. Riska, M. Larson, Z. Fogarty,
Clinic, Rochester, MN.
In order to define the functional role of previously reported risk-SNPs for
prostate cancer (PC), we created a normal prostate tissue eQTL database
using 471 tissue samples from patients with normal prostate using the
Illumina Human Omni 2.5M SNP array and expression was measured by RNA sequencing. Our initial eQTL analysis focused on those risk-SNPs previously identified by multiple PC GWAS studies (n= 132), plus all SNPs (included imputed) that were in linkage disequilibrium (r2>0.5) with each risk-SNP. Additionally, we focused on cis-acting associations
only testing all transcripts within 2Mb of the risk-SNP interval. Of
5116 SNPs identified in the risk-regions, 1002 demonstrated a Bonferroni
significant eQTL signal (p<1.96×10−8) and these were associated with 43
genes. To identify possible functional SNPs for each of the risk-regions
associated with a candidate gene, we utilized a variety of publicly available
databases and bioinformatics tools. These included various GEO datasets
that were the tissue type of prostate tissue or PC cell lines (with and without
drug treatment) and a variety of analysis tools (Regulome, Funcpred,
GWAS3D, etc.) to obtain regulation potential scores for each of the candidate
SNPs. Some of the information retrieved from these databases included
transcription factor binding sites and motifs, previous eQTL data, mRNA
expression, binding, conservation and reported long range interactions. Results
from these analyses were then compiled to determine possible functional regions.
For example, for the risk-SNP rs4962416 on chromosome 10, there were
25 SNPs in LD that had significant eQTL signals (a1<3.71E−12) with the
CTBP2 gene. The strong association of the same SNPs was used to map the
14 most significant to a probable enhancer region as defined by
markers Dnase I, H3K27Ac, H3K4me1 and H3K4me2. The top variant
(rs12769019) as determined by Regulome score and GWAS3D for LinCan
cells is in strong linkage disequilibrium with the risk-SNP and significantly
inhibit disrupting the binding affinities of numerous transcription factors that could be
important for prostate cancer regulation through interactions with the AR
(ie NKX3-1, FOXA1, GR and POU5F1). Utilizing publicly available data
and multiple approaches, we identified a potentially novel functional element for
PC susceptibility loci to help guide subsequent laboratory studies.

669T
Sex-specific genetic architecture of the transcriptome. E.R. Gamaizon2,
N.L. Tigner1, B.E. Stranger2, GTEx Consortium. 1) Section of Genetic
Medicine, Department of Medicine, The University of Chicago, Chicago, IL;
2) Institute for Genomics and Systems Biology, The University of Chicago,
Chicago, IL.
Sexual dimorphism has been observed in a broad spectrum of complex
phenotypes, including autoimmune disorders, psoriasis, and prostate cancers. Here, in support of our efforts to investigate the sex-specific genetic architecture of complex traits, we examine potential functional mechanisms underlying sexual dimorphism. Using RNA sequencing data from the Geno-
type Tissue Expression (GTEx) consortium, we generated a sex-specific gene
expression and mapped sex-biased expression quantitative trait loci (eQTLs)
in a diverse set of human tissues. We identified tissue-specific gene expres-
sion patterns with some genes showing differential expression (p<10-6)
between the sexes in some tissues but not in others. Our approach facilitated
the comprehensive analysis of non-coding RNAs, putatively regulatory, with
substantial sexually dimorphic expression profiles. We identified eQTLs
associated with the expression of a gene in one sex and not the other,
quantified the proportion of shared eQTLs, and highlighted traits with distinct
regulatory variation in the sexes. Notably, our unique dataset allowed us to
evaluate the tissue-specificity and the degree of sharing between tissues of
sex-biased eQTLs. We found in each tissue substantial variation in the
expression of sex-specific eQTLs among different classes of genes, including,
most prominently, antisense transcripts and long intergenic non-coding RNAs.
We identified sex-specific eQTLs among known disease susceptibility
and quantitative trait loci (for example, Parkinson’s disease, Crohn’s disease,
systemic lupus erythematosus, aortic dimension, and folate pathway levels)
as well as associations that were not previously published (115,1450 GTEx
studies). On the basis of the eQTL targets, we propose novel genes, distinct
from the originally reported genes, for some of these phenotypes. Analyses
of WTCCC, the Psychiatric Genomics Consortium, and GIANT GWAS data
reveal that sex-biased eQTLs show tissue-dependent enrichment for trait
associations. Sex-specific eQTL target genes in such tissues as tibial artery
and blood are enriched for antigen processing and presentation as well
as immune-related function whereas those in muscle are in, addition, enriched
for metabolic processes and proteins that localize in the muscle. Our
study shows that genotype-sex interactions have a broad influence on the
human transcriptome and that sex-specific genetic architecture of gene
expression is likely to mediate higher-order traits.

670S
Activating mutations in STIM1 and ORAI1 cause overlapping syn-
dromes of tubular myopathy and congenital miosis. M. Kousi1, V.
Nesin1, G. Wiley1, E.C. Ong1, T. Lehmann2, D.J. Nicholl3, M. Sun4, N.
Shahrokhi5, P.M. Galbraith2, N. Nesin1, K. Wierenga1, L. Tsioka1, N. Katsanis1,
1) Cell Biology, Duke University, Durham, NC; 2) Department of Cell Biology,
University of Oklahoma Health Sciences Center, Oklahoma City, OK; 3) Arthritis
and Clinical Immunology Research Program, Oklahoma Medical
Research Foundation, Oklahoma City, OK; 4) Department of Pathology
and Laboratory Medicine, Faculty of Medicine, University of Malaya, Kuala
Lumpur, Malaysia; 5) Division of Genetics, Department of Pediatrics, University of Oklahoma
Health Sciences Center, Oklahoma City, OK.
Signaling through the store-operated Ca2+ release-activated Ca2+
(CRAC) channel regulates critical cellular functions, including gene expres-
sion, cell growth and differentiation, and Ca2+ homeostasis. Loss-of-function
mutations in the CRAC channel pore-forming protein ORAI1 or the Ca2+
entrance sensing protein stromal interaction molecule 1 (STIM1) result in severe
immune dysfunction and nonprogressive myopathy. Here, we identify gain-
of-function mutations in the cytoplasmic domain of STIM1 (p.R304W) associ-
ated with thrombocytopenia, bleeding diathesis, miosis, and tubular myopa-
thy in patients with Stormorken syndrome, and in ORAI1 (p.P245L), associ-
ated with a Stormorken-like syndrome of congenital miosis and tubular
aggregates in zebrafish embryos, recapitulating key aspects of
immune dysfunction and nonprogressive myopathy. Here, we identify gain-
of-function mutations in the cytoplasmic domain of STIM1 (p.R304W) associ-
ated with thrombocytopenia, bleeding diathesis, miosis, and tubular myopa-
thy in patients with Stormorken syndrome, and in ORAI1 (p.P245L), associ-
ated with a Stormorken-like syndrome of congenital miosis and tubular
aggregates in zebrafish embryos, recapitulating key aspects of
immune dysfunction and nonprogressive myopathy. Here, we identify gain-
of-function mutations in the cytoplasmic domain of STIM1 (p.R304W) associ-
ated with thrombocytopenia, bleeding diathesis, miosis, and tubular myopa-
thy in patients with Stormorken syndrome, and in ORAI1 (p.P245L), associ-
ated with a Stormorken-like syndrome of congenital miosis and tubular
aggregates in zebrafish embryos, recapitulating key aspects of
immune dysfunction and nonprogressive myopathy. Here, we identify gain-
of-function mutations in the cytoplasmic domain of STIM1 (p.R304W) associ-
ated with thrombocytopenia, bleeding diathesis, miosis, and tubular myopa-
thy in patients with Stormorken syndrome, and in ORAI1 (p.P245L), associ-
ated with a Stormorken-like syndrome of congenital miosis and tubular
671M
High-resolution personal genome-wide maps of meiotic double-strand breaks in humans. F. Pratto1, K. Brick1, P. Khil1, F. Smagulova1, G. Petukhova1, R.D. Camerini-Otero1, 1) National Institutes of Health, Bethesda, MD; 2) Uniformed Services University of Health Sciences, Bethesda, MD. Meiotic recombination is initiated by the formation of DNA double-strand breaks (DSBs) that are directed to specific genomic loci called hotspots by a meiosis-specific protein called PRDM9. To date, the most comprehensive map of hotspots sites in humans was generated using computational analysis of patterns of linkage disequilibrium (LD), however, this method cannot resolve gender-specific hotspots or hotspots defined by different alleles of PRDM9. In this work, we exploit a sequencing based method we recently developed to perform the first direct high-resolution genome-wide analysis of meiotic recombination initiation hotspots in individual human males. We mapped up to 39,000 DSB hotspots in each of five individuals with different combinations of PRDM9 alleles. Our PRDM9-specific DSB maps demonstrate that the LD-derived recombination maps reflect hotspots defined by different alleles of PRDM9 and we find that the relatively infrequent C-allele of PRDM9 (an allele found predominantly in African populations) contributes significantly to the LD map. About 75% of DSB hotspots are found at an LD-defined hotspot, yet DSB sites without an LD-hotspot still show a significant elevation in the recombination rate. Thus, our PRDM9-specific DSB hotspot maps greatly expand our knowledge of sites where recombination occurs in the human genome. We observed significant inter-individual variation in DSB frequency at hotspots, however most differences could not be explained by sequence variation at PRDM9 binding sites. Additional factors therefore dictate the efficiency of DSB formation. While sequence changes can influence recombination, we also found that recombination can influence the genomic sequence. Characteristic patterns of polymorphisms at hotspots offer compelling evidence for both GC-biased gene conversion and for a mutagenic effect of meiotic recombination. DSB hotspots were also enriched at structural variants that arise via homology-mediated mechanisms and at chromosomal breakpoints associated with many well-documented genomic diseases. Finally, an analysis of the relationship between meiotic DSBs and crossovers suggests that meiotic DSB initiation frequency is a primary determinant of the genetic crossover landscape.

672T
Investigating the maternal age effect on meiotic recombination in multiple cohorts. H.C. Martin1, J. Hussin1, J. O’Connel1,2, S. Gordon2, K. McAloon2, H. Mbarek1, J.J. Hottenga1, J. Marchini1, D. Boomsma1, N.G. Martin1, P. Donnelly1,2, 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) Illuminia, Inc; 3) Queensland Institute for Medical Research, Brisbane, Australia; 4) Department for Biological Psychiatry, Vrije Universiteit, Amsterdam; 5) Department of Statistics, University of Oxford, Oxford, United Kingdom. Multiple studies have reported that recombination rate increases with maternal age, but several have found an effect in the opposite direction. We investigate this question in new cohorts from the Australian and Dutch twin registries. Our rationale was that the use of dizygotic twins would allow measurement of the variability in recombination rate at a single point in a mother’s lifetime. We analyse these twin cohorts in combination with several previously published studies, and compare different methods for calling crossovers. We then fit a Bayesian hierarchical model to the recombination rates, accounting for differences in the mothers’ baseline rates, and test the hypothesis that the maternal age effect is the same in all cohorts.

673S
Mapping of two neurogenetic disorder genomes with a single molecule nanochannel array platform for genome-wide structural variation analysis. Y.Y.Y. Lai1, E.T. Lam1, A.C.Y. Mak2, V. Searles3, C. Chu1, C. Lin1, N. Anderson2, J.M. Sikela3, P.Y. Kwok1, 1) UCSF, San Francisco, CA; 2) BioNano Genomics, San Diego, CA; 3) Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO. Copy number variation of sequence encoding DUF1220 protein domains has been shown to be associated with human brain size and a number of neuropsychiatric disorders. DUF1220 motifs are embedded in the NBPF and PDE4DIP genes that are mostly found in Chr1q21. They are poorly assembled in the human reference genome due to the repetitive property of that region. Different methods, such as quantitative PCR, array comparative genomic hybridization (aCGH), Sanger or next-generation sequencing have been used for identification of these polymorphic loci. However these methods are too labor-intensive and do not provide positional (break-point) information for the structural variants. Genome mapping utilizes highly parallel nanochannel arrays in which hundreds of very long, fluorescently labeled, single DNA molecules are linearized and imaged. This novel approach is automated on the Irys System (BioNano Genomics, La Jolla, CA), which can scan the entire genome rapidly to generate physical maps that provide a more comprehensive view of the genome. Here, we use this genome mapping approach to detect genome-wide structural variation, including the copy number variation of DUF1220 in the PDE4DIP and NBPF genes, in two diploid cell lines from patients with neuropsychiatric disorders (autism and schizophrenia). To date, we generated over 50X coverage data and constructed de novo assembled genome maps that cover about 90% of hg38 reference genome using an automated assembly pipeline. We identified many structural variants that are found in the Chr1q21 region. Overall, this genome mapping approach is simple and provides a list of genome-wide candidate structural variations that are associated with autism and schizophrenia.

674M
Association study of COL11A2 with aspirin exacerbated respiratory disease and its FEV1 decline. J. Kim1, C. Park2, H. Shin1, 1) Research Institute for Basic Science, Sogang University, Seoul, South Korea; 2) Division of Allergy and Respiratory Medicine, Soonchunhyang University Bucheon Hospital, Bucheon, Republic of Korea. Aspirin exacerbated respiratory disease (AEDR) induces bronchoconstriction in asthma patients and is characterized with a clinical condition of severe decline in forced expiratory volume in one second (FEV1) after ingestion of aspirin. Genetic association studies of several candidate genes within human major histocompatibility complex (MHC) region on chromosome 6p21-24 have reported that this MHC genomic region is implicated in asthma and related respiratory diseases. This study investigated the genetic association of collagen, type XI, alpha 2 (COL11A2) within the MHC genomic region with AEDR and its FEV1 decline by genotyping of 19 tagging COL11A2 SNPs in 93 AEDR patients and 96 aspirin-tolerant asthma controls. As a result, polymorphisms of COL11A2 showed potential associations with AEDR (minimum P = 0.02 in rs2269346), along with the increased significances for the FEV1 decline by aspirin provocation (minimum P = 0.002 in rs2855459). In the haplotype analysis, no significant signals were observed. Despite the need for further replications in larger cohorts and functional evaluations, our preliminary findings suggest that COL11A2 may be a predisposing factor for FEV1 decline-related symptoms.
675T

Bovine animal model for spermatogenic and scrotal alterations: additional clues for an X-chromosome component. P. A. S. Fonseca1, M. P. Almeida2, G. S. Moura3, F. C. Santos4, D. J. Santos5, G. C. Oliveira6, V. R. Vale Filho7, M. V. G. B. Silva8, M. R. S. Carvalho1. 1) Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, 31.270-901, Brazil; 2) Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, 31.270-901, Brazil; 3) Embrapra, Empresa Brasileira de Pesquisa Agropecuária, Embrapa Genética de Plantas, Belo Horizonte, 31.270-901, Brazil; 4) Instituto de Ciências Biológicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil; 5) Departamento de Genética, Universidade Federal de Pernambuco, Recife, PE, Brazil; 6) Departamento de Ciências Médicas, Universidade Federal de São Paulo, São Paulo, SP, Brazil; 7) Departamento de Biologia, Universidade Federal de Campina Grande, Campina Grande, PB, Brazil; 8) BioNano Genomics, San Diego, CA.

We confirm previously reported and identify novel structural variants in human spermatozoa with the detection of sperm chromosome territories via molecular cytogenetics. We study the relationship between sperm DNA integrity and the presence of structural variants. The results suggest that the presence of structural variants is associated with decreased sperm DNA integrity, which may explain the relationship between sperm DNA integrity and fertility.

677M

An integrated platform for the collection of biospecimens to support the Genotype-Tissue Expression (GTex) project. J. C. Keen, L. Canters1, A. Undale1, A. Rao1, M. Barcus2, P. Branton1, L. Sabin3, P. Guan4, N. Roche5, C. Shiver6, A. Smith7, K. Valentino8, S. Volpe9, J. Strzewski9, S. Koester10, R. Little11, H. Moore12 on behalf of the GTEx consortium. 1) Biorepositories and Biospecimen Research Branch, National Cancer Institute, Bethesda, MD; 2) Leidos Biomedical Research, Inc., Rockville, MD; 3) National Human Genome Research Institute, Bethesda, MD; 4) National Institute of Mental Health, Bethesda, MD.

The Genotype-Tissue Expression (GTex) project is an NIH Common Fund study that aims to clarify how genetic variation influences gene expression in normal human tissues. As of May 2014, GTEx has collected 25+ tissue types per donor from over 650 deceased tissue or organ donors. To facilitate the collection and storage of tissues and data, the NCI Biorepositories and Biospecimen Research Branch (BBRB) has developed an integrated biospecimen collection platform. This platform includes collection and storage facilities, a data warehouse, a team of board-certified pathologists, and a molecular analysis core facility. Together, this system has allowed for the acquisition of high quality, well-annotated biospecimens that meet established quality criteria for molecular analysis. Clinical data, genotype and gene expression data are available at dbGaP for each donor and tissue type collected. These data will enable scientists to better understand the genomic and molecular variation in normal human tissues and may facilitate the identification of genes involved in disease states.

676S


Despite recent advances in next-generation sequencing technologies, de novo genome sequence assembly and genome-wide structural variation (SV) detection based on ‘short reads’ remain challenging. For efficient de novo sequence assembly and complete SV characterization, we generated genome mapping data with long DNA molecules (>150kb) fluorescently labeled at Nt.BspQI sites (GTCM@UCSF), linearized and imaged in massively parallel nanochannel arrays [1]. We obtained single-molecule data at 60X genome coverage on CEPH CEU trios samples, NA12878, NA12891 and NA12892 that were extensively analyzed by many groups, including the 1000 Genomes Project. Single-molecule maps were assembled de novo into consensus genome maps that cover ~90% of human genome [2,3]. We confirmed previously reported and identified novel structural variants that were consistent with Mendelian inheritance. We also analyzed genome features of interest (e.g. subtelomeric and subcentromeric regions). Our genome-wide identification of individual genome sequence assembly and structural variations that are otherwise hard to discover with short read sequencing data alone. 1. Lam ET et al. Genome Mapping on Nanochannel Arrays for Structural Variation Analysis and Sequence Assembly. Nat Biotechnol (2012) 30(8):771-776.
679S Impact of genetic polymorphisms in FADS1, FADS2, and FADS3 genes on fatty acid metabolic mechanism on methadone therapy in Taiwan. R. Wang1, H.T. Yang2, Y.H. Lane3, C.L. Huang4, Chieh-Liang Huang. 1) Department of Public Health, China Medical Univ, Taichung, Taiwan; 2) Department of Nutrition, China Medical University, Taiwan; 3) Institute of Clinical Medical Science, College of Medicine, China Medical University, Taichung, Taiwan; 4) Department of Psychiatry, China Medical University Hospital, Taichung, Taiwan.

It had been observed that abnormal fatty acid metabolism in patients with psychological disorders, such as major depression, schizophrenia, and Alzheimer’s disease, etc. Illegal drugs is not only associated with the increased risk of the mental health problems, family and social economic burden, but also the elevated risk of HCV and AIDS infection. Drug abuse is an important public health issue. Heroin is one of the most common increasing abuses in Taiwan. Methadone, a synthetic μ-opioid receptor agonist, is extremely effective to treat to reduce illegal drug use, AIDS-caused infectious transmission and mortality in heroin dependent abusers patients. The aim of the study is to explore the association between the 7 tagSNPs (rs174547, rs174550, rs174570, rs174602, rs498793, rs526126, rs174634) of FADS1, FADS2, and FADS3 (fatty acid desaturase 1, 2, and 3) and fatty acid variation before and after methadone maintenance treatment in drug abusers. A total of 89 subjects with heroin dependence or methamphetamine were recruited from the methadone clinic in China Medical University Hospital. Based on the Chinese Han hapmap LD structure, seven tagSNPs were selected from FADS1, FADS2 and FADS3 genes. Genotyping were carried out using the Applied Biosystem Assay on Demand reagents and were implemented using an ABI Prism 7900HT Sequence Detection System. Fatty acid profiles of RBC were analyzed with gas chromatography. This study was approved by the institutional review board of China Medical University Hospital in Taiwan. All the informed consent was obtained from all participants during their initial clinic visit. All data analyses were performed using SAS version 9.1.3 (SAS Inc., NC, USA). The study results showed that all the 7 tagSNPs are significantly associated with the variation of PUFA and AA/EPA ratio (p<0.0001) but not associated with variations of PUFA and AA/EPA ratio (p<0.0001) but not associated with C18:2n6, C18:3n6 before and after methadone treatment. The data pointed out that methadone treatment would interfere fatty acid metabolism, especially PUFA desaturation in subjects. As abnormal fatty acid profiles were readily associated with several psychological disorders, the correlation between fatty acid metabolism and drug abusers with certain tag SNPs. In conclusion, our study may provide further information regarding the 7tagSNPs of the FADS1, FADS2 and FADS3 were associated with fatty acid metabolism for drug abuser with MMT treatment in Taiwan.

680M Drug Metabolizing Enzyme and Transporter Gene Variation, Nicotine Metabolism, and Cigarette Consumption. A.W. Bergen1, M. Michel1, D. Ntola2, M. Krasnow3, K.N. Connelly4, C. George-Schild1, H. Hops4, A.Z.X. Zhu5, J.W. Baurley6, J.B. McClure7, S.M. Hall8, T.B. Baker9, D.V. Conti5, N.L. Benowitz10, C. Lerman12, R.F. Tyndale13, G.E. Swan11. 1) Center for Health Sciences, SRI International, Menlo Park, CA; 2) Department of Human Genetics, Emory University School of Medicine, Atlanta, GA; 3) Department of Psychiatry, Washington University School of Medicine, St Louis, MO; 4) Oregon Research Institute, Eugene, OR; 5) Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON; 6) BioRealm, LLC, Monument, CO; 7) Research Health Institute, Seattle, WA; 8) Department of Psychiatry, University of California, San Francisco, CA; 9) Center for Tobacco Research and Intervention, Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI; 10) Department of Preventive Medicine, University of Southern California, Los Angeles, CA; 11) Departments of Medicine and of Bioengineering & Therapeutic Sciences, University of California San Francisco; 12) Department of Psychiatry, University of Pennsylvania, Philadelphia, PA; 13) Centre for Addiction and Mental Health, and Departments of Psychiatry, and of Pharmacology and Toxicology, University of Toronto, Toronto, ON; 14) Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine.

The Nicotine Metabolite Ratio (NMR), the ratio of trans-3-hydroxyxococine and cotinine, two nicotine metabolites, is significantly associated with CYP2A6 activity, response to smoking cessation treatments, and cigarette consumption. We searched for drug metabolizing enzyme and transporter (DMET) gene variation associated with the NMR and with prospective abstinence outcomes. A total of 3,546 participants (2,171 smokers and 1,375 non-smokers) were included. DMET genotyping were carried out using the Applied Biosystem Assay on Demand reagents and were implemented using an ABI Prism 7900HT Sequence Detection System. Fatty acid profiles of RBC were analyzed with gas chromatography. This study was approved by the institutional review board of China Medical University Hospital in Taiwan. All the informed consent was obtained from all participants during their initial clinic visit. All data analyses were performed using SAS version 9.1.3 (SAS Inc., NC, USA). The study results showed that all the 7 tagSNPs are significantly associated with the variation of PUFA and AA/EPA ratio (p<0.0001) but not associated with variations of PUFA and AA/EPA ratio (p<0.0001) but not associated with C18:2n6, C18:3n6 before and after methadone treatment. The data pointed out that methadone treatment would interfere fatty acid metabolism, especially PUFA desaturation in subjects. As abnormal fatty acid profiles were readily associated with several psychological disorders, the correlation between fatty acid metabolism and drug abusers with certain tag SNPs. In conclusion, our study may provide further information regarding the 7tagSNPs of the FADS1, FADS2 and FADS3 were associated with fatty acid metabolism for drug abuser with MMT treatment in Taiwan.


Background: Selective Serotonin Reuptake Inhibitors (SSRIs) and cognitive-behavioral therapy (CBT) are effective therapies for pediatric anxiety; however, response to treatment is highly variable. Differential treatment response may be explained by genetic factors. Due to the role of glutamate in learning and memory, we examined genetic variation in the glutamate pathway on the effectiveness of sertraline vs. CBT vs. combination therapy for pediatric anxiety in the randomized, placebo-controlled Child/Adolescent Anxiety Multimodal Study (CAMS). Methods: We examined 15 key variants in glutamate signaling candidates for association with treatment outcome: NMDA glutamate receptor GRIN2B and transporters regulating synaptic glutamate levels (the cystine-glutamate exchange-SLC7A11 and neuronal (SLC1A1) and glial (SLC1A2) glutamate transporters). Results: Enrichment of the glutamate system were apparent only in the CBT treatment group. One SLC1A1 variant was significant in omnibus analyses across all groups. Minor allele carriers (C+) at the glutamate transporter (SLC1A1) rs3335321 variant demonstrated minimal symptom reduction over time with CBT monotherapy (p= 0.06 x10^-1), whereas GG homozygotes showed an excellent response. Interestingly, the opposite allele showed a greater symptom drop with placebo. Within the CBT group, SLC7A11 rs11723658 (p=0.008) and GRIN2B rs1806195 (p=0.005) also predicted treatment response. Conclusion: While no effects on medication response were observed, our data support a possible role for genetic variation in glutamate signaling pathways in response to CBT monotherapy. Our results show promise for eventual personalized anxiety treatment algorithms and warrant replication in independent samples.
682M Clinical and pharmacogenomic features of cisplatin-induced ototoxicity in Asian nasopharyngeal carcinoma patients. S.L. Chan1, L.S. Ng2, C.H. Slow2, C.J. Ross3,4, B.C. Carleton2, M.S. M.R. Hayden1,3, B.C. Goh1,6,8,5, K.S. Loh2, L.R. Brunham7,8,10
1) Translational Laboratory in Genomic Medicine, Agency for Science, Technology and Research, Singapore; 2) Department of Otolaryngology-Head & Neck Surgery, National University Health System, Singapore; 3) Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, BC, Canada; 4) Division of Translational Therapeutics, Department of Pediatrics, Child and Family Research Institute, University of British Columbia, Vancouver, BC, Canada; 5) Pharmaceutical Outcomes Programme, BC Children’s Hospital, Vancouver, BC, Canada; 6) Department of Haematology-Oncology, National University Cancer Institute, National University Health System, Singapore; 7) Cancer Science Institute, Singapore; 8) Department of Pharmacology, Yong Loo Lin School of Medicine, National University Health System, Singapore; 9) Haematology Oncology Research Group, National University Cancer Institute, National University Health System, Singapore; 10) Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

Background: Cisplatin-induced ototoxicity (CIO) is a common and debilitating adverse drug reaction (ADR) that affects up to a quarter of adults and more than half of children receiving the drug, and Asian populations may be at increased risk. Inter-individual variability in susceptibility to CIO suggests a genetic component to this ADR that has not been fully defined. In European populations, TPMT rs1201199 and COMT rs9332377 have been found to be associated with CIO. However, these variants are very rare or absent in healthy individuals from the Chinese and Malay populations. We therefore hypothesized that other pharmacogenomic variants influence rare or absent in healthy individuals from the Chinese and Malay populations.

Methods: Nasopharyngeal carcinoma patients were recruited from the National University Hospital of Singapore and genotyped on a custom Illumina Infinium panel containing 7807 SNPs in genes involved in drug absorption, distribution, metabolism and excretion (ADME). Clinical and audiometric data were collected from all patients. Trends of hearing loss were first delineated in terms of threshold change from baseline averaged across both ears over time. Genetic association was then performed only on patients who received cisplatin. CIO cases were defined as >15dB mean threshold shift at 1 year and controls defined as ≤15dB shift.

Results: Out of 115 patients with complete clinical and audiometric data, 24 received radiotherapy (RT) only and 91 received both RT and chemotherapy (chemoRT). There was no significant threshold shift at lower hearing frequencies (0.5–2kHz). However, at 4kHz, there was a substantial threshold increase from as early as 3 months after initiation of therapy in patients who received chemoRT. At 12 months, 32% of patients who received chemoRT compared to 0% who received RT displayed a threshold increase of >15db. After Qc, 82 patients (24 cases and 58 controls) and 4281 SNPs were included in the genetic association study. COMT and TPMT variants reported previously were extremely rare in this population. None of the ADME variants reached statistical significance after correction for multiple testing.

Conclusion: High frequency hearing loss in NPC patients at time points up to one year is due primarily to cisplatin. Established risk variants for CIO are rare in this population, suggesting the presence of novel genetic factors. Given the small sample size of this preliminary study, we did not detect evidence of ADME variants with large effect size on CIO.


Background: The stimulant dexmethylphenidate (d-MPH) is an effective treatment for hyperactive and inattentive symptoms associated with pediatric attention deficit hyperactivity disorder (ADHD). Treatment-induced growth slowing, however, is a common adverse effect associated with morbidity and mortality. Treatment adherence is inversely related to growth and side effects is common and may be explained by genetic factors. This study investigated whether genetic variation in drug targets (SLC6A3/DAT1, SLC18A2/VMA2, TAARA1) and metabolic enzymes and transporters treatment disposition (ABCBC1, NR1I2) could help explain differential outcomes in treatment response to d-MPH. Methods: In 202 children participating in the NIMH TRECC Study, height and weight were measured at baseline and at Weeks 4 and 8. A subset continued treatment for 18 months and growth was measured regularly. Treatment response was determined by scores on the ADHD-rating scale (ADHD-RS) and Clinical Global Impression Improvement (CGI-I) Scale. Variants previously associated with similar phenotypes or shown to impact protein function were genotyped using TaqMan allelic discrimination assays. Results: Homozygotes for the minor allele (GG) of exonic NR1I2 variant rs7643645 demonstrated rare guanfacine-associated growth suppression (p=1.78 x 10^-13). A second NR1I2 variant (rs1523130) showed differential effects across treatment groups (p=0.09x 10^-10). Heterozygotes for the intronic ABCB1 variant rs1128503 (p=1.7 x 10^-10) and 3-USTR TAARA1 variant rs270722 (p=4.27 x 10^-10) and intronic SLC18A2 variant rs3632227 showed a paradoxical growth acceleration when treated with d-MPH and acceleration when treated with guanfacine. Homozygotes for the minor allele (TT) of 3-USTR SLC6A3 variant rs27072 (p=4.27 x 10^-10) and intronic SLC18A2 variant rs3632227 showed a relatively poor response to either guanfacine or stimulant monotherapy. In contrast, SLC18A2 A+ carriers achieved a significantly greater symptom reduction to combination treatment (p<0.01). Conclusion: Genetic variants impacting d-MPH pharmacokinetic and pharmacodynamic effects influenced treatment outcomes and adverse effects to common ADHD treatments. Understanding the pharmacogenomic factors moderating treatment response and adverse effects could lead to strategies to individualize treatment matching and develop improved interventions.

684M Genetic vulnerability to adverse child attachment quality and stress reactivity. E.H. Gai1, J.L. Borelli2, A. Nguyen1, C.M. Laubacher2, G.S. Heilemann1, E.L. Nurmii1, 1) University of California, Los Angeles, Los Angeles, CA; 2) Pomona College, Claremont, CA.

Background: Stressful parent-child interactions are known to have effects on both child attachment quality and physiologic autonomic reactivity. In this pilot study, we examined the relationship between environmental stress and genetic vulnerability at four polymorphisms that have been linked to stress and attachment phenotypes: serotonin transporter promoter polymorphism (HTT-LPR), oxytocin receptor (OXTR) rs53576, mu-opioid receptor (OPRM1) rs1799971, and glucocorticoid receptor chaperone (FKBP5) rs3800373. Methods: The sample consisted of 45 mother-child dyads recruited from the community. The Child Attachment Interview (CAI) and the Parental Stress Index were collected, as well as cortisol saliva levels during a stressful task. Individual genotypes were assayed and combined genetic load was calculated by adding the number of risk alleles carried by each child. Results: The minor allele at FKBP5 rs3800373 was associated with cortisol reactivity (p=0.038) and attachment quality (p=0.015), both with medium effect sizes (f2 = 0.1). Combined genetic load positively correlated with higher scores on CAI (r = 0.32, p = 0.036). The correlation of a higher number of risk alleles with greater attachment security may be due to a reversal of effects in the absence of stress, which has been previously observed. While this preliminary study was underpowered to sufficiently analyze gene-environmental interactions, the level of parental-stress moderated the interaction between risk alleles and both child attachment and cortisol reactivity at the trend level (p<0.1). Conclusions: These data suggest that genetic risk alleles may have opposite effects on child attachment quality and stress reactivity in the context of high versus low parent-child stress. Our findings are consistent with and extend prior observations. A larger study to further explore these potential gene-environment interactions is currently underway. An understanding of how genetic risk and genetic load may help predict the development of child anxiety and inform the subsequent development of targeted interventions.
685S


Background: The atypical antipsychotic drug, risperidone, has been shown to decrease irritability and aggression in children with autism spectrum disorder (ASD); however, weight gain is a common adverse event. Genetic background may explain the substantial variability in this treatment-limiting side effect. As previous data suggests a potential mitochondrial contribution to antipsychotic-induced weight gain (AIWG), we attempted to replicate prior findings at three specific mitochondrial gene variants (SLC25A4, TSPO and NDUF51). Methods: 225 subjects participating in the NIMH Research Units on Pediatric Psychopharmacology (RUPP) Autism Risperidone trials were genotyped and individual BMIs were measured weekly for eight weeks to examine association of AIWG with mitochondrial gene variants. Results: The rs10024068 promoter variant in SLC25A4 strongly moderated AIWG over the course of just 8 weeks of treatment in both the complete sample (p = 3.21×10^{-6}) and the Caucasian subset (p = 2.49×10^{-6}). Individuals homozygous for the minor G-allele gained significantly more weight (Z-Score change GG = 1) than A-allele carriers (Z-Score change AG = 0.58 and AA = 0.61). This single variant explains 6% of the variation in AIWG in our sample. SLC25A4 encodes the Adenine Nucleotide Translocator 1 (ANT1), part of the mitochondrial permeability pore. There was no evidence for association of weight gain with genetic variants in NDUF51 or TSPO. As seen in prior studies, however, a significant interaction between SLC25A4 and TSPO was observed (p = 0.018), which is especially relevant given the biological association of ANT1 and translocator protein TSPO in forming the mitochondrial permeability pore. Conclusion: Genomewide significant association at SLC25A4 supports the involvement of mitochondrial energy pathways in AIWG and suggests the examination of additional mitochondrial candidates as well as replication in independent samples. These data may help uncover pathways underlying AIWG and may eventually be used to optimize treatment matching.

686M

The application of personalized medicine for understanding clinical variability in rasagiline response in early Parkinson’s Disease. J. Knight, S. Collins, N. Freeman, M. Tampakeras, J. Levy, A. Tchelet, E. Eyal, J. Levy, A. Tiwari, J.L. Kennedy, M. Masellis, A.E. Lang, ADAGIO investigators. 1) Campbell Family Mental Health Research Institute, Centre of Addiction and Mental Health, Toronto, Ontario, Canada; 2) Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada; 3) Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada; 4) Teva Pharmaceutical Industries, Israel; 5) Department of Medicine, Division of Neurology, University of Toronto, Toronto, Canada; 6) Movement Disorders Clinic, Sunnybrook Health Sciences Centre, Toronto, Canada; 7) Neurogenetics Section, Centre for Addiction and Mental Health, Toronto, Canada; 8) Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, University Health Network, Toronto, Canada; 9) The Edmond J. Safra Program in Parkinson’s Disease, Toronto Western Hospital, University Health Network, Toronto, Canada.

Objective: To identify candidate gene polymorphisms associated with peak motor benefit to rasagiline at 12 weeks and associated with sustained benefit over a 36-week evaluation period. Background: Rasagiline, a selective irreversible monoamine oxidase-B (MAO-B) inhibitor, has been shown to be safe and effective for the treatment of Parkinson’s disease (PD). As with all drugs, there is variability among individuals in their clinical benefit to rasagiline, this study is designed to identify genetic factors that may influence this variability. Methods: We performed a retrospective genetic association study using clinical data from the ADAGIO trial. Candidate genes encoding proteins involved in catecholamine synthesis and metabolism, rasagiline metabolism, as well as those reported in Genome Wide Association Studies (GWAS) of PD susceptibility were selected. We examined association between genetic polymorphisms and peak change in unified Parkinson Disease Rating Scale (UPDRS) score from baseline to 12 weeks, using a linear model. We also examined association between genetic polymorphisms and change in UPDRS score from baseline over 12, 24, and 36 weeks, using a mixed effects linear model. Both models controlled for placebo response. Results: 204 Single Nucleotide Polymorphisms (SNPs) and 5 Variable Number Tandem Repeat (VNTR) from 28 candidate genes were successfully genotyped in 694 samples Analysis 1: Two SNPs in strong linkage disequilibrium within the dopamine D2 receptor gene (DRD2) were found to be significantly associated with peak change in UPDRS scores at 12 weeks (rs1076560 and rs2283265, False Discovery Rate [FDR] p=0.045 for each) as well as one in Norepinephrine Transporter (SLC6A2) (rs36023, FDR-corrected p=0.045). Analysis 2: No allelic associations were identified in the model assessing longitudinal data. Conclusions: To our knowledge, this is the largest pharmacogenetic study of an anti-Parkinsonian drug conducted to date. The D2 receptor is a major target of dopamine and its stimulation relates to motor benefits in PD. Prior literature indicates that the two associated SNPs alter transcriptional processing of DRD2. Together with the dopamine transporter, the norepinephrine transporter is involved in the pre-synaptic reuptake of catecholamines, including dopamine. Further investigation of these genes is clearly warranted.
Posters: Pharmacogenetics


Antiplatlet therapy with clopidogrel is standard of care in improving outcome in secondary prevention for cardiovascular patients. While generally effective, substantial inter-individual variation in response to clopidogrel exists, and high on-treatment platelet reactivity to adenosine diphosphate (ADP) is associated with recurrent cardiovascular events. To better assess the genetic determinants of clopidogrel response, we evaluated 28 previously reported single nucleotide polymorphisms (SNPs) on platelet reactivity in 2054 coronary artery disease patients of the ICPC. Consistent with previous investigations, we detected strong association between the loss-of-function CYP2C19*17 variant and increased on-clopidogrel ADP-induced platelet aggregation (P=8.93 × 10^{-20}, Beta = 0.48). Conversely, on-treatment ADP-induced platelet aggregation was significantly decreased in subjects with the gain-of-function CYP2C19*17 variant both before and after adjustment for CYP2C19*2 (P=2.33 × 10^{-7}, Beta = -0.23 and P=3.54 × 10^{-7}, Beta = -0.12, respectively), as well as a functional SNP in CES1 (G143E [rs71647871]), which metabolizes clopidogrel to an inactive form (P=7.76 × 10^{-6}, Beta = -0.62). A variant in P2RY12 (rs1472122), which encodes the platelet ADP receptor was also significantly associated with on-clopidogrel platelet reactivity (P=3.82 × 10^{-4}, Beta = 0.13). We observed no evidence of association between platelet aggregation and genetic variants in other candidate genes including ABCB1 C3435T [rs1045642] and C1236T [rs1128503], FON1 Q192R [rs109551] and L55M [rs854562] [ICBG] (rs9518), and other CYP enzymes (e.g. CYP2SB, CYP2C9, CYP1A2). Our data suggest that, in addition to CYP2C19, genetic variation in genes responsible for generation of the inactive clopidogrel metabolite formation and regulating the P2RY12 ADP-receptor impact on-treatment platelet reactivity and may influence overall clopidogrel efficacy.


Background: While social functioning is a key deficit in autism spectrum disorders (ASD), few current treatments specifically target these impairments. In the NIH Research Units on Pediatric Psychopharmacology (RUPP) study, risperidone reduced social withdrawal ratings in children with ASD; however, variability was substantial. We examined genetic variability in the oxytocin and vasopressin signaling pathways, which are known to play a role in social behavior, for hypothesized association with treatment response. Methods: Risperidone response in 225 children and adolescents with ASD in the NIH RUPP Risperidone trials was measured by weekly aberrant behavior checklist (ABC) social withdrawal subscale (II) ratings over 8 weeks of treatment. Complete common variation at the oxytocin (OXT), arginine vasopressin (AVP), and arginine vasopressin receptor of OXT (AVPR1 and 2) loci was captured with 16 SNPs. Additionally, we included 10 oxytocin receptor (OXT-R) SNPs with evidence for prior association with various social phenotypes. A repeated measures general linear mixed model was used to examine genetic association with symptom improvement. Analyses were restricted to the Caucasian subset to limit effects of ethnicity. Results: Three independent variants (rs2740204, rs2770378 and rs4813625) tagging the genomic locus containing the adjacent AVP and OXT genes were predicted improvement in ABC Social Withdrawal (p=0.002). Individuals homozygous for the C-allele at rs2770378 (p=0.014) and the A-allele at rs2770378 (p=0.0019) showed greater improvement in social withdrawal than G-carriers. The T-allele at rs2740204 demonstrates an allele dose-dependent effect on symptom improvement (p=0.0017). No effects at either OXTR or AVPR receptors was observed. Conclusions: Our results suggest that variants in the OXT/AVP systems influence social withdrawal treatment response to risperidone in children with ASD. If replicated in independent samples, these results may help guide future clinical treatment algorithms.


Background: Growth effects of ADHD pharmacotherapies are important, treatment-limiting side effects. Previously we reported genetic influences on dextemphylenidate (d-MPH) and guanfacine treatment of children with ADHD in the UCLA TRECC sample, including variants associated with treatment response and dopamine and norepinephrine transporter expression. This study examined genetic variation in these signaling pathways for association with growth effects. Methods: 202 subjects between 7-14 years of age were recruited for the acute phase of a randomized, double-blind, placebo-controlled trial of d-MPH and guanfacine for pediatric ADHD. Three treatment groups included guanfacine monotherapy, d-MPH monotherapy, and combination guanfacine plus d-MPH. Medication responders continued in the trial for approximately 14 months (n=99) and height and weight were tracked regularly in addition to ADHD symptoms. We tested association of genetic variation in monoamine and energy balance candidate systems with height and BMI changes during medication exposure. Due to the large number of markers tested (n=96), we only report results meeting genomewide significance.

Results: In our sample, guanfacine monotherapy was associated with height and BMI increases over predictions from CDC growth charts (Z-score increase of 0.14 for height and 0.41 for BMI). Both d-MPH monotherapy Z-score decrease of -0.19 for height and -0.83 for BMI were associated with growth slowing. Variation in treatment group trajectories for height (p=6.47 × 10^{-10}) and BMI (p=5.85 × 10^{-10}) were highly significant. Minor allele (A) carriers at DRD4 rs6277290 showed allele dose-dependent clinically-significant weight loss (2.6 Z-Score loss) on d-MPH (p=1.53 × 10^{-10}). BMI change was moderated by ADRA2A rs1800544, with profound weight loss on d-MPH in C-allele carriers and combination treatment in C-allele carriers (p=2.73 × 10^{-10}). Effects were similar and maintained Bonferroni significance in the Caucasian subset, ruling out population effects. Conclusions: The results presented here suggest that genetic factors contribute significantly to stimulant-mediated weight loss pediatric subjects with ADHD and appear to be able to identify patients more likely to be affected.


Background: Antipsychotic-induced weight gain (AIWG) is a common, treatment-limiting adverse effect of this widely prescribed drug class. Prior examination of the melanocortin system in adults with AIWG has demonstrated a consistent contribution of the melanocortin 4 receptor (MC4R), which is central to energy balance, growth and body weight regulation. We also examined genetic variation in these signaling pathways for association with weight change in children and adolescents with autism spectrum disorder (ASD) treated with risperidone for aggression and irritability. Methods: 225 youth in the NIMH Autism Risperidone studies were genotyped using Life Technologies’s TaqMan platform. Complete common variation in the key melanocortin signaling receptors was captured with 15 SNPs spanning the MC3R, MC4R and neuropeptide Y (NPY) loci. Variants were tested for association with BMI and height change (measured weekly across 8 weeks of treatment) using independent samples and multiple linear mixed model. Results: In contrast to previous literature, MC4R was not associated with AIWG in our sample. Common allele homozygotes (GG) at MC4R promoter SNPs rs8087522 (p=7.65 × 10^{-3}) and rs8093615 (p=6.70 × 10^{-3}) however, showed greater change in height over 8 weeks of treatment in both the complete dataset and the Caucasian subset. On the other hand, we found strong ethnicity-specific association of AIWG with two MC3R gene variants: 3′ variant rs6024733 (p=1.22 × 10^{-3}) and BMI (p=5.58 × 10^{-3}) were highly significant. Minor allele (G) carriers at DRD4 rs5918, CYP1A2 rs1128503, CYP2C19 rs9518, and other CYP enzymes (e.g. CYP2SB, CYP2C9, CYP1A2). Our data suggest that, in addition to CYP2C19, genetic variation in genes responsible for generation of the inactive clopidogrel metabolite formation and regulating the P2RY12 ADP-receptor impact on-treatment platelet reactivity and may influence overall clopidogrel efficacy.
**691S**

**Metformin functional pharmacogenomics: STUB1 functions as an E3 ligase for cyclin A and affects metformin sensitivity.**

N. Niu, T. Liu, M. Deng, R. Ly, L. Wang, M.PET, Mayo Clinic, Rochester, MN.

Metformin, a widely used anti-diabetic drug, is being considered as a highly promising chemopreventive and therapeutic agent. Although the exact mechanism of action of metformin is unknown, it is thought that metformin could activate AMPK and cause G1 cell cycle arrest. In order to identify biomarkers contributing to metformin response, we took advantage of our Human Variation Panel, a 300 EBV-transformed lymphoblastoid cell line (LCL) model with extensive omic data, and conducted a metformin pharmacogenomic analysis. The expression of STUB1, a protein ligase, was found to be associated with metformin cytotoxicity (IC50 value) with p-value <10^-4. Functional validation study using siRNA showed that knockdown of STUB1 significantly altered metformin response in two basal-like breast cancer cell lines, MDA-MB-231 and Hs578T. To further investigate the mechanism by which STUB1 influences metformin response, we performed mechanistic studies using MDA-MB-231 breast cancer cell line as well as STUB1 knock-out mouse embryonic fibroblast (MEF) cells. Knockdown of STUB1 in MDA-MB-231 cell showed an increased proportion of cells in the G1 phase by flow cytometry and a significant increase in cyclin A protein level by Western Blot. Cyclin A is known to be involved in G1/S transition. In our study, we found that STUB1 interacted with cyclin A. In MEF cells, knockout of STUB1 caused an increase in cyclin A protein level as well as a decrease in the ubiquitin level of cyclin A. In addition, cyclin A protein was found to be degraded in a proteasome-dependent manner. Thus, we propose that the E3 ubiquitin ligase STUB1 influences metformin response by facilitating proteasome-mediated degradation of cyclin A. In conclusion, we demonstrated a novel mechanism involving a genomic data-enriched LCL model system, together with functional validation and mechanistic study using cancer cell line, could help us to identify novel genetic biomarkers involved in metformin response.

**692M**

**Association study of polycystic ovary syndrome with two single nucleotide polymorphism of follicle stimulating hormone receptor in Iranian patients: rs1394205 and rs6166.**

B. Dehghan Banadaki1, F. Alsharian2, M. Shiva3, SH. Zarei Motorki1. 1) Faculty of basic sciences and advanced technologies in biology, University of science and culture, Tehran, Iran; 2) Department of Genetics at Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran; 3) Department of Endocrinology and Female Infertility at Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran.

Background: Polycystic ovary syndrome (PCOS) is the most common form of WHO II anovulatory infertility, affecting 5-10% of women in reproductive age. It is characterized by clinical or biochemical hyperandrogenism, menstrual dysfunction with anovulation and presence of polycystic ovaries are the features for PCOS diagnosis. As we know, clinical or biochemical hyperandrogenism, menstrual dysfunction with anovulation and presence of polycystic ovaries are the features for PCOS diagnosis. Hence, many genetic factors might be involved in this disease. In our current study, we evaluated the associations of two SNPs (rs1394205 and rs6166) in this gene with clinical and biochemical traits in Iranian PCOS patients.

Methods: The Genomic DNA was extracted from peripheral blood lymphocytes of 102 PCOS patients (54 drug control, 100 fertile controls). Both PCOS and fertile control groups were undergone same IUI protocols and the clinical, biochemical and ultrasonography traits of all subjects were observed and number of follicles ≥ 15, as a feature of drug response, were evaluated. Final data were analyzed by chi-square test in SPSS statistical software. Results: PCOS patients in comparison to other two control groups, showed significantly association with N680S (p=0.008). Genotype distribution of SS, SN and NN in PCO group was 37%, 42% and 15% respectively. According to drug response (exogenous FSH), 30% of PCOs bearing SS genotype were more sensitive to FSH in comparison to 53% in the controls with the same genotype, while in drug resistant group, 69% of PCOs with SS genotype were more sensitive to exogenous FSH in comparison to 96% in controls with the same genotype. This difference in drug resistance was observed in SN genotype too. Furthermore, 42% in PCOs vs. 20% in drug controls. In patients with SN genotype, 70% of both groups were resistant to drug in comparison to 5% of drug controls with the same genotype. This difference in drug resistance was observed in SN genotype too.

Conclusion: Our observation showed great correlation between rs6166 polymorphism and ovulation induction response in PCOS.

1) University of Illinois, Chicago, Illinois, USA; 2) University of Chicago, Chicago, Illinois, USA.

Personalized medicine proposes the customization of healthcare and optimal therapies based on the context of an individual patient’s genetic makeup and/or other clinical information. Recent progress in pharmacogenetic and pharmacogenomics studies has indicated that both genetic and epigenetic factors may influence drug response to therapeutic treatments. Our objective was to identify histone modification markers associated with drug toxicities using the Lymphoblastoid cell lines (LCL) model, on which we have access to a variety of genetic, epigenetic, and phenotypic data, including the genetic variants associated with chemotherapeutic agents. Specifically, the ChIP-seq data for four post-translational modifications of histone H3 (H3K4me1, H3K4me3, H3K27ac, and H3K27me3) in ten Yoruba LCLs were collected from a publicly available dataset (GSE47791). Drug response association results for five drugs (carboplatin, cisplatin, etoposide, daunorubicin, and ara-C) in the same cell lines were collected from our previously developed pharmacogenetics-cell line database: PACdb (http://www.pacdb.org/). Association studies revealed a substantial number of genes whose local histone modification associated with cytotoxicities based on Akaika information (AIC) and p-value mixed criteria. For example, local H3K4me1 modification levels in 72 genes were observed to be associated with the IC50 of carboplatin. Group comparisons showed that different histone markers exhibited different strength of association. Overall, H3K4me4, a histone mark known to be associated with active gene expression, outperformed other histone markers as a variable for cytotoxicity. Our previous studies had identified a list of potential genes whose expression levels were associated with these drug response phenotypes. We showed that a subset of those associations could be explained by variations of its local histone modification levels. For example, the expression of ARL4A (encoding ADP-ribosylation factor-like 4A) had been identified to be highly associated with the IC50 of cisplatin (p < 0.001). Our results further showed this association could be related to the modification in its local H3K4me3. In summary, our whole genome approach provided a global picture of the contributions of histone modification and other genes to drug-induced cytotoxicity phenotypes. The mapping of histone modification markers provided novel insights into our previous pharmacogenomic findings.


Icahn School of Medicine at Mount Sinai, New York, NY.

Pharmacogenomic studies aim to elucidate the genetic or genomic features contributing to the efficacy of medications. The analysis of these data are challenging due to the large number of genomic features that must be investigated. Shrinkage or regularization techniques can help identify the most important features that can predict sensitivity to drug. This work reviews the properties of several regularization methods, such as ridge regression, lasso, and elastic net, using gene expression data from the Cancer Cell Line Encyclopedia project for predicting the sensitivity of 74 lung cancer cell lines to the drug nilotinib. An empirical illustration of the predictive performance of the regularization methods are provided based on three approaches - the first that directly uses gene expressions, and two methods that use principal components of the gene expressions. In this study, ridge regression models fitted slightly better than lasso and elastic net models.


Children’s Hospital Oakland Research Institute (CHORI), Oakland, CA.

Statins are widely prescribed to reduce cardiovascular disease (CVD) risk. Statins primarily lower plasma low-density lipoprotein cholesterol (LDL-C) levels; but they also cause modest reductions in plasma triglycerides (TG), an independent CVD risk factor, in most people. There is extensive inter-individual variability in statin response, but only a small proportion of this variance has been explained by phenotypic characteristics or by genetic variants identified from association studies. As an alternative approach, we studied statin response variability using RNA-seq data from control and statin-treated lymphoblastoid cell lines (LCLs) derived from 100 Caucasian and 50 African American participants of the Cholesterol and Pharmacogenetics (CAP) simvastatin clinical trial (40 mg/day for 6 weeks). 38% of human genes were statin responsive, and 64% of Epstein-Barr virus genes were upregulated with statin treatment (p<0.0001, N=150). To identify candidate statin efficacy genes, statin-induced changes in gene expression were correlated to plasma LDL-C response and plasma TG response. Initially, we found 151 genes with expression changes that were significantly correlated with LDL-C response (FDR=5%). Next, we adjusted the gene expression changes for potential confounders using probabilistic estimation of expression residuals (PEER) and found that the first two PEER hidden factors were significantly correlated with LDL-C response (p<0.05) but not TG response. We found 23 genes were significantly correlated with TG response (FDR=15%). Insulin-induced gene 1 (INSIG1 [MIM 602055]), a well-known regulator of cholesterol homeostasis, had expression changes strongly correlated with LDL-C response (Spearman’s ρ=0.32, q=0.11, N=150), and this correlation was driven by men (sex interaction p=0.0055). In Caucasians, we identified a SNP that was associated with INSIG1 expression changes (N=99 LCLs, p=5.4x10^-5) as well as with TG response in the combined CAP and PRINCE statin trial populations (p=0.0048, N=1890), predominantly in men. Statin-induced changes in INSIG1 alternative splicing were also correlated with TG response in men, and a combined model including INSIG1 expression level and splicing for the SNP successfully explained the variance (p=5x10^-6, N=86). In summary, our results strongly suggest that variation in INSIG1 contributes to statin-induced changes in plasma TG in a sex-specific manner.


23andMe Inc., Mountain View, CA.

Ninety percent of drugs that are deemed successful in animal models fail human clinical trials, because of adverse drug reactions or lack of efficacy. Human genetics can be a powerful tool in the efforts to increase success of drug development through validation of target genes and prioritization of research programs. 23andMe phenotype-wide association studies (PhеШАS) leverage extremely large sample sizes and phenotypic breadth to yield valuable insights with direct relevance to potential effects in humans. By mining publicly available drug databases for associations between drug target genes and genes harboring variants significantly associated with one or more 23andMe phenotypes, we found over 2,800 drug-phenotype links. Both positive validation of likely successful drug targets and prediction of adverse drug reactions are seen in the 23andMe data, including a replication of association between rare loss-of-function variants in PCSK9 and LDL cholesterol level. A retrospective look at both successful and unsuccessful drug trials confirms our findings, with substantial increase in specificity of predicting successful trials as well as predicting trial failure based on lack of association. Our results support using PheWAS to inform ongoing target validation and identify new drug targets.
699S Association of C677T polymorphism of the MTHFR gene with toxicity in breast cancer patients treated with FEC chemotherapy. A. Ramos-Silva 1, Y. R. Ramírez 1, 2, I. A. Guatrería 1, 3, OM. Soto 1, 2, DL. Carrillo 1, 3, AM. Puebla 1, AR. Rincón 2, 2, MP. Gallegos 1. 1) Lab de Genética Molecular. Div Med Mol, CIBO, IMSS, Guadalajara, Jal, México; 2) Doctorado en Farmacología. CUCS, U de G; 3) Doctorado en Genética Humana. CUCS, U de G; 4) Laboratorio de Inmunofarmacología. Departamento de Farmacología. CUCEI, U de G; 5) Departamento de Biología Molecular y Genómica. CUCS, U de G.

Background: The influence of C677T polymorphism in the MTHFR gene involved in metabolism of chemotherapeutic agents has been studied in different cancers. MTHFR plays a major role in folate metabolism and consequently could be an important factor for the efficacy of a treatment with FEC (fluorouracil-epirubicin-cyclophosphamide) in breast cancer patients. Our aim was to evaluate the association of C677T polymorphism with toxicity effects in breast cancer patient’s neoadjuvantly treated with FEC chemotherapy. Methods: DNA genomic samples from 525 patients (UMAE gynecology and obstetrician Hospital, CMNO, IMSS), that received FEC neoadjuvant chemotherapy were included in the study. Protocol was support by FIS/IMSS/PROY/G1/12/13. The C677T polymorphism was determine by polycrylamide gels electrophoresis, previously PCR and HincII restriction enzyme analysis. The association was determine by odds ratio. Results: The genotype 677TT was associated with gastrointestinal toxicity (diarrhea grade III) in breast cancer patient response at FEC chemotherapy [1.9 (IC95% 1.06–3.53), p=0.028] and genotype 677CT was associated with hematological toxicity (neutropenia) in breast cancer patients non responder at FEC chemotherapy [1.8 (IC95% 1.03–3.41), p=0.046]. Conclusion: The polymorphism C677T could be a good marker of toxicity in breast cancer patients treated with FEC chemotherapy in the analyzed sample.

700M Integration of Genetics into Clinical Development. D. Waterworth 1, L. LF. L. Warren 1, Y. Yeo 1, J. Aponte 2, M. Nelson 1, S. Chissoe 1, Genetics, PCPS, GlaxoSmithKline, King of Prussia, PA; 2) Statistical Genetics, PCPS, GlaxoSmithKline, RTP, NC; 3) Genetics, PCPS, GlaxoSmithKline, Stevanage UK; 4) Genetics, PCPS, GlaxoSmithKline, RTP, NC.

Pharmacogenomic analysis is commonly initiated after a clinical trial completes and questions arise around interindividual variation in safety or efficacy. This approach can limit the number of pharmacogenetic studies conducted, but it can also limit the opportunity to identify important genetic predictors and delay their potential impact on drug development. We sought to circumvent these issues in the STABILITY study that tested darapladib, an Lp-PLA2 inhibitor, against standard of care for stable coronary heart disease (CHD). We genotyped 13,557 patients with the Illumina OmniexpressExome chip prior to study completion. The trial failed to demonstrate efficacy for the primary endpoint of major adverse coronary events (MACE). Despite the large size of this study, no genetic variants reached genome-wide significance for MACE or secondary endpoints. However, we demon- strated that the failure of the primary trial reflected a lack of statistical power and not the absence of genetic effects. The GWAS was significant at the genome-wide threshold of p = 5×10−8. The top hit of the GWAS was rs12924003 (p=1.06×10−5) located downstream of the SAL1 gene. The second hit, SNP rs4771655 (p=9.51×10−5) is ~39kb upstream of IRSG2 gene. The third hit, rs4751427 (p=4.4×10−5) is ~59kb upstream of the Neuropeptide S gene. Our analyses did not detect an association when considering the commonly applied genome-wide correction for multiple testing. Two of our top hits, IRSG2 and NPS, were previously shown to be involved in regulation of insulin sensitivity and food intake in other populations.

702M Genome-wide association study of resistant hypertension in INVEST. N. El Roybu 1, CW. McDonough 1, Y. Gong 1, C.J. Pepine 2, A. Takahashi 2, T. Tanaka 2, M. Kubo 3, RM Cooper-Delhoff 2, JA. Johnson 1, 2, 1) Pharmacotherapy and Translational Research, University of Florida, Gainesville, FL; 2) Division of Cardiovascular Medicine, Department of Medicine, University of Florida, Gainesville, FL; 3) RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

Resistant hypertension (RHTN) is defined as a blood pressure (BP) that remains ≥140/90 mm Hg despite the use of ≥ 3 antihypertensive drugs or BP < 140/90 mm Hg requiring ≥ 4 antihypertensive drugs. RHTN is associated with increased risk for adverse cardiovascular outcomes, especially stroke; therefore, we aimed to identify the association between genetic variant and RHTN using Population-based Association (PCAS) to further characterize those at risk for RHTN. We genotyped 1,199 patient samples from the International Vampamili-SR trial study. Genetic Substudy (INVEST-GENES) on the Illumina OmniExpressExome chip. Non-study drugs were used for BP control except for atenolol in CAS and verapamil in NCAS. In this analysis, a case was defined as a BP ≥ 140/90 despite use of ≥ 3 drugs, or use of ≥ 4 drugs regardless of BP. Overall, BP control in both treatment strategies was similar, which allowed for pooling participants. Logistic regression analysis was conducted sepa- rately for treatment groups on the Illumina OmniExpressExome chip. Results: From both race groups were combined. There was a significant genetic model, adjusting for age, gender, body mass index, diabetes, smoking, left ventricular hypertrophy, congestive heart failure, stroke, percutane- ous coronary intervention, principle components, and treatment assignment. Results from both race groups were consistent. In general, there was an additive effect often leading to patient non-compliance and increased morbidity and mortality due to metabolic syndrome and cardiovascular events. Genetic markers would be extremely helpful to identify prior to antihypertensive exposure and first genetic tests are now becoming commer- cially available. Sample (n=505) consisted of 251 individuals of European and 107 of African-American ancestry with well-documented information on changes. Exclusion criteria included a) related individuals, b) available genotype rate of less than 95%, c) mismatch between genetic markers and assigned sex status and d) heterozygosity of more than four standard deviations from the mean. SNPs with MAF < 5%, and genotypes which deviated significantly from HWE were removed. A series of mixed models was used including all potentially relevant covariances and change in BMI over a period of three months as the outcome. In addition, interactions between time and all other predictors were considered. The final model included a random intercept and slope associated with time, study medica- tion group and baseline BMI. Standard quality control workflow was applied to the genotype data and we analyzed 328,733 SNPs. To avoid population stratification, we plotted the MDS components and selected patients within the cluster corresponding to European ancestry as the largest cohort. The GWAS was conducted on individuals with multiple association statistic- ion analysis was carried out using the R package: ‘nlim’. None of the SNPs was significant at the genome-wide threshold of p = 5×10−8. The top hit of the GWAS was rs12924003 (p=1.06×10−5) located downstream of the SAL-1 gene. The second hit, SNP rs4771655 (p=9.51×10−5) is ~39kb upstream of IRSG2 gene. The third hit, rs4751427 (p=4.4×10−5) is ~59kb upstream of the Neuropeptide S gene. Our analyses did not detect an association when considering the commonly applied genome-wide correction for multiple testing. Two of our top hits, IRSG2 and NPS, were previously shown to be involved in regulation of insulin sensitivity and food intake in other populations.
**703S**


Asthma is a chronic inflammatory condition of the lungs, characterized by excessive responsiveness of the lungs to stimuli, in the forms of infections, allergens, and environmental irritants. Currently, 22.9 million Americans suffer from asthma, and the prevalence has increased dramatically since 1980. Asthma is the leading chronic illness in U.S. children, with 6.8 million affected in 2006. Inhaled corticosteroids (ICS) are a common, effective treatment for asthma. ICS act locally in the lungs to reduce the chronic inflammation that characterizes asthma; this reduction in inflammation reduces the frequency and severity of asthma flares that occur in response to acute triggers. Most studies examining the genetic components of the response of patients to ICS treatment focuses on changes in pulmonary function test results that occur following weeks or months of treatment. In the current study, we wished to determine the possibility of identifying responders and nonresponders to ICS therapy through the use of abstracted electronic medical records (EMRs). We developed an approach to score a patient’s response to ICS by evaluating changes in the frequencies of hospitalization and emergency room visits, along with changes in ICS dose and the need for additional prescription of longer therapy in the form of oral steroids. Through the use of this response score as a quantitative trait in an African American cohort in a genome wide association study revealed a significant single locus, represented by four significant SNPs located within the MSRA gene on chromosome 8. Replication efforts are currently underway.

**705S**

A genome-wide association study identified variants in KCNIP4 associated with ACE inhibitor-induced cough. J.D. Mosley, C.A. McCarty, D.R. Crosslin, C. Yan, A. Haines, J. Wu, P.L. Teixeira, L. Bastarache, D.C. Crawford, J.A. Pacheco, T.A. Manolio, E.P. Bottinger, C.A. McCarthy, J. Linneman, W. Thompson, R.L. Chisholm, G.P. Jarvik, D.R. Crosslin, D.S. Carroll, E.B. Langston, K. Jannins, L.J. Kim, W. Wu, B. Lopatka, J.C. Dennis, D.M. Roden, eMERGE II Network. 1) Department of Medicine, Vanderbilt University, Nashville, TN; 2) Department of Pediatrics, Vanderbilt University, Nashville, TN; 3) Center for Human Genomics Research, Vanderbilt University, Nashville, TN; 4) Biomedical Informatics, Vanderbilt University, Nashville, TN; 5) National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 6) Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY; 7) Mayo Clinic, Rochester, MN; 8) Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; 9) Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington, Seattle, WA; 10) Group Health Research Institute, Seattle, WA; 11) Division of Cardiovascular Diseases, Mayo Clinic, Rochester MN; 12) The Sighetii and Janet Weis Center for Research, Geisinger Health System, Danville, PA; 13) Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA, USA.

The most common side effect of angiotensin converting enzyme inhibitors (ACEIs) is a persistent cough, which effects up to 8% of whites, and often requires a switch to other drug classes. The etiological mechanism of cough is unknown. We conducted a genome wide association study (GWAS) to look for common SNPs that are associated with ACEI-Induced cough in an eMERGE discovery cohort comprised 6,007 whites of European ancestry (EAs) from six sites in the electronic MedEd Records & Genomics (eMERGE) network. The phenotype was defined as a clinically diagnosed cough attributed to an ACEI and recorded in the drug sensitivity/allergy section of the electronic medical record (EMR). Controls were individuals with at least 6 months of ACEI use and no history of an ACEI-associated cough. We performed single SNP tests of association using logistic regression with 3,194,795 SNPs, association with a positive genetic effect, and adjusting for principal components. The most significant SNPs were evaluated in a replication set of 926 eMERGE EA subjects. The GWAS in the discovery set of 1,346 cases and 4,661 controls identified significant associations located in a single region of chromosome 4 within an intron of the gene KCNIP4. The strongest association was in rs145489027 (MAF=0.33, OR=1.3 [95%CI: 1.2–1.4], p=2.4×10−8). Four of the six significantly associated SNPs in this region with p<1×10−6 were available for analysis in a replication set of 157 cases and 769 controls. The strongest association among these 4 SNPs was in rs7675300, also in KCNIP4, (MAF=0.33, OR=1.3 [95%CI: 1.2−1.4], p=0.04).

KCNIP4 is a member of the KCHIP family of EF hand-containing calcium-binding proteins and is primarily expressed in neuronal structures where it regulates Kv potassium channels. These data suggest that ACEI-induced cough may be mediated, in part, through a neurally-mediated mechanism involving KCNIP4.

---

**704M**

A genome-wide association study of hepatic toxicity of methotrexate therapy for rheumatoid arthritis. H. Mo, N. Braggs, L. Bastarache, R.J. Carroll, A. Shah, D. Roden, S. Raychaudhuri, E.W. Karlson, J.C. Denny. 1) Biomedical Informatics, Vanderbilt University, Nashville, TN; 2) Department of Medicine, Vanderbilt University, Nashville, TN; 3) Brigham and Women’s Hospital, Boston, MA; 4) Harvard Medical School, Harvard University, Boston, MA; 5) Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia.

Low dose methotrexate (MTX) is the first-line disease modifying anti-rheumatic drug (DMARDs) for rheumatoid arthritis (RA) and other inflammatory arthritis. While it is generally well tolerated, hepatic toxicity is among one of the most concerning side effects with a cumulative incidence of 31%. We studied individuals identified from BioVU, a DNA biobank linked to deidentified electronic medical records (EMR) at Vanderbilt, to identify genetic susceptibility of hepatic toxicity of MTX therapy. Methods: We identified and genotyped individuals with RA in BioVU using a previously validated algorithm. We evaluated RA individuals for presence of MTX exposure, and then used a combination of algorithms and manual review to ascertained 66 cases and 309 controls with Illumina OmniExpress and HumanExome platforms. Only individuals involving KCNIP4.

---

**706M**


Clopidogrel (Plavix) is a commonly used antplatelet medication in coronary artery disease patients to prevent recurrent thrombotic events. It is a pro-drug that requires activation by CYP2C19, and common loss of function variants (e.g., CYP2C19*2) are associated with decreased antplatelet responsiveness phenotype but explained only a marginal part of this trait. Heritability estimates are high and suggest that there are additional genetic determinants of clopidogrel response. To identify these variants, the International Clopidogrel Pharmacogenomics Consortium (ICPC) amassed clopidogrel antiplatelet responsiveness phenotype data on patients with the on-treatment ADP-induced platelet reactivity and cardiovascular outcomes) and DNA from more than 6,000 clopidogrel-treated patients. A GWAS (Omni Express with exome coverage) was conducted in 2,721 Caucasian patients. We observed strong association between on-treatment platelet reactivity phenotypes and the CYP2C19 locus on chromosome 10 (best SNP rs1926711; r2=0.963 with rs1926711), there remained residual association just upstream of CYP2C19 (rs1998591; p=1.66×10−53). After adjustment for the known CYP2C19*2 loss of function variant (rs4244285; r2=0.963 with rs1926711), there remained residual association just upstream of CYP2C19 (e.g., rs1985951; p=1.66×10−56) suggesting other variants in this region as determinants of clopidogrel response. Other loci nominally associated with platelet response to clopidogrel (p<0.05) included rs12651351, intergenic between FHDC1 and TRIM2; rs10506836 in PLEKHAS; and rs2295306 in DDB1C1.

While we conclude that CYP2C19*2 remains the major genetic determinant of clopidogrel response, other loci are likely to exist, but play a lesser role. Replication studies, now in progress, will be required to affirm potentially novel loci for clopidogrel response.
707S

Class II HLA variants are associated with differential antibody response to toxoid vaccinations in a cohort of Ugandan infants. A. J. Mentzer1,2, A. Muthawi3, T. Carstensen4, G. Smits5, A. Rautanen1,2, D. Gurdasani2, T. Mills1,2, C. Pomilla1, H. Akurut1, D. Kizito1, S. Lule3, K. Mohammed1, K. Elliott1,2, F. van der Klis1, P. Kaleebu1, A. Elliott1, M. Sandhurst1, AVS, HILL1,2.


It is estimated that up to 70% of the variability in response to some vaccines administered in childhood is inherited but the identification of the genetic factors responsible for this variation has not yet been prioritised. We have conducted a genome-wide association study to identify the genetic variants associated with multiple vaccine responses using 2,181,930 autosomal markers genotyped in 1391 Ugandan children enrolled in the Entebbe Mother and Baby cohort study. Vaccine response was defined as quantitative antibody levels measured against 7 antigens targeting 5 separate pathogens which demonstrated little to moderate correlation within individuals (r=0.02 to 0.58). Strong Class II HLA associations were observed with response to diphtheria toxin (most associated SNP rs2647060 P=1.70×10−26, odds ratio (OR) = 10.8, 95% confidence interval (CI) = 5.9–19.6, P = 3.64 × 10−15). To detect additional genetic marker for CBZ-induced cADRs, we performed genome-wide-impulse using the GWAS data and genotype data from 1000 Genomes Project as reference panels by MaCH-Admix software. For a conditional logistic regression analysis of the imputed genotype data, we selected rs1633021 as the covariate with the highest association observed in the analyzed GWAS peak of association. After the conditional analysis, there were 22 loci that reached GWAS-level significant association (P < 5 × 10−8). SNPs in 17 loci were validated and 2 SNPs rs4488702 and rs16899783 were replicated (Combined population: P-logistic = 4.10 × 10−3 and 2.08 × 10−1). When 2 SNPs rs4488702 and rs16899783 were included as the risk marker for CBZ-induced cADRs in the additive model, the percentage was increased from 57% (A‘31:01) to 65% (A‘31:01 and 2 SNPs).

708S


1. Laboratory for Pharmacogenomics, RIKEN Cen. Med. Sci., Yokohama, Kanagawa, Japan; 2. Laboratory for Personalized Medicine, RIKEN Cen. Med. Sci., Yokohama, Kanagawa, Japan; 3. Laboratory for Genotyping Development, RIKEN Cen. Med. Sci., Yokohama, Kanagawa, Japan; 4. Department of Dermatology, Ehime University Graduate School of Medicine, Ehime, Japan; 5. Department of Dermatology, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 6. Department of Dermatology, Showa University School of Medicine, Tokyo, Japan; 7. Department of Dermatology, Kyorin University School of Medicine, Tokyo, Japan.

An anticonvulsant, carbamazepine (CBZ), is known to show incidences of cutaneous adverse drug reactions (cADRs) including Stevens-Johnson syndrome (SJS), toxic epidermal necrolysis (TEN) and drug-induced hypersensitivity syndrome (DIHS). Previously, to identify a susceptible gene(s) for CBZ-induced cADRs, we conducted a genome-wide association study (GWAS) and subsequent HLA typing in 61 cases and 376 patients who showed no cADRs by administration of CBZ (CBZ-tolerant controls), and found that HLA-A*31:01 was present in 60.7% (37/61) of the patients with CBZ-induced cADRs, but only 12.5% (47/376) of the CBZ-tolerant controls (odds ratio (OR) = 10.8, 95% confidence interval (CI) = 5.9–19.6, P = 3.64 × 10−15). To detect additional genetic marker for CBZ-induced cADRs, we performed genome-wide-impulse using the GWAS data and genotype data from 1000 Genomes Project as reference panels by MaCH-Admix software. After conditional logistic regression analysis of the imputed genotype data, we selected rs1633021 as the covariate with the highest association observed in the analyzed GWAS peak of association. After the conditional analysis, there were 22 loci that reached GWAS-level significant association (P < 5 × 10−8). SNPs in 17 loci were validated and 2 SNPs rs4488702 and rs16899783 were replicated (Combined population: P-logistic = 4.10 × 10−3 and 2.08 × 10−1). When 2 SNPs rs4488702 and rs16899783 were included as the risk marker for CBZ-induced cADRs in the additive model, the percentage was increased from 57% (A‘31:01) to 65% (A‘31:01 and 2 SNPs).

710M

A Phenome-Wide Association Study of Numerous Laboratory Phenotypes in AIDS Clinical Trials Group (ACTG) Protocols. S. Pendergrass1, A. Verma1, E. Daar2, R. Gulick3, R. Haubrich4, G. Robbins2, D. Hass2, M. Ritchie1. 1) Center for Systems Genomics, Department Biochemistry & Molecular Biology, Pennsylvania State University, University Park, PA; 2) Departments of Medicine and Microbiology, University of California San Diego, San Diego, CA; 3) Department of Dermatology, Massachusetts General Hospital, Harvard Medical School, Boston, MA; 4) Vanderbilt University, Nashville, TN; 5) Weill Medical College, Cornell University, New York, NY.

Phenome-Wide Association Studies (PheWAS) have the potential to efficiently discover novel genetic associations across multiple phenotypes. Prospective clinical trials data offer a unique opportunity to apply PheWAS to pharmacogenomics. Here we describe the first PheWAS to explore associations between genotypic data and clinical trial data, both pre-treatment and following initiation of antiretroviral therapy. A "pre-treatment" PheWAS considered 27 laboratory variables from 2807 subjects who had participated in 4 ACTG protocols (ACTG384, A5142, A5095 and A5202), and analyzed ~5M imputed SNPs. Lowest p-values were for pre-treatment bilirubin, neutrophil counts, and HDL cholesterol levels. These and multiple other laboratory variables matched associations in the NHGRI GWAS Catalog. An "on-treatment" PheWAS considered data from 1181 subjects from A5202. We considered 838 phenotypes and sub-phenotypes derived from 6 variables: CD4 counts, HIV control, fasting LDL cholesterol levels, fasting triglyceride levels, efavirenz pharmacokinetics (PK), and aztnavir PK. We considered 237 annotated drug-related SNPs from PharmGKB. Of 23 associations tested, the p-values varied from 1.10 x 10^-7 to 3.10 x 10^-1. The variables with matching biological plausibility: LDL cholesterol levels with LPL and APOE; triglycerides with LPL; CD4 counts with innate immune response gene TNF, HIV control with innate immune response gene HLA-DQA1, efavirenz with CHRNA7 and HLA-DQA1, and efavirenz PK with the gene ANG.

This analysis highlights the potential utility of PheWAS to evaluate clinical trials datasets for genetic associations.
714M Selection of cancer patients based on tumor AKT1 or PIK3CA mutation status. J. Fox1, P. Kirk2, J. Whiteley3, H. Ambrose2.

1) Personalized Healthcare, AstraZeneca Pharmaceuticals, Macclesfield, Cheshire, United Kingdom; 2) Innovative Medicines, AstraZeneca Pharmaceuticals, Macclesfield, Cheshire, United Kingdom.

The identification of patients who may experience differential benefit from new cancer medicines as a consequence of the presence of a genetic alteration e.g. mutation in their tumour, requires the timely and accurate identification of the appropriate patient selection methods. Both sourcing and validation of assays for use during drug development bring particular challenges and different solutions are required depending on the clinical situation in which assays will be deployed. We have employed a number of approaches instigated by Suitelab, Silexa, and SNPolisher to assess the performance of the assay when used as a method for patient selection suitable for implementation during the initial, exploratory, stages of drug development in man. For logistical reasons a PCR based test was developed in house for AKT1 whereas for PIK3CA a commercially available (also PCR based) Research Use Only (RUO) assay was used. Both assays underwent preliminary internal validation prior to transfer to external partners for formal validation before their application for patient testing. Identifying acquired mutations, not present in normal patient tissue, in tumour derived material brings additional considerations to DNA testing beyond that necessary for diagnosis of common inherited disorders. First, tests need to have sufficient analytical specificity and sensitivity to allow accurate assignment of mutation status when the mutant sequence exists in only a minority of the DNA sample. Second, the distribution of several variations appears to be unique for our study compared to the SWEGENE database. This means that the specific patient population may have different mutation frequencies compared to the general population.

The SWEGENE project has been performed in a large number of laboratories instigated by Suitelab, Silexa, and SNPolisher to assess the performance of the assay when used as a method for patient selection suitable for implementation during the initial, exploratory, stages of drug development in man. For logistical reasons a PCR based test was developed in house for AKT1 whereas for PIK3CA a commercially available (also PCR based) Research Use Only (RUO) assay was used. Both assays underwent preliminary internal validation prior to transfer to external partners for formal validation before their application for patient testing. Identifying acquired mutations, not present in normal patient tissue, in tumour derived material brings additional considerations to DNA testing beyond that necessary for diagnosis of common inherited disorders. First, tests need to have sufficient analytical specificity and sensitivity to allow accurate assignment of mutation status when the mutant sequence exists in only a minority of the DNA sample. Second, the distribution of several variations appears to be unique for our study compared to the SWEGENE database. This means that the specific patient population may have different mutation frequencies compared to the general population.
Impact of regular physical activity on weekly warfarin dose requirement. P. Shahabi1,2, E. Rouleau-Mailoux1,3, S. Dumas1,4, Y. Feroz Zada1, S. Provost4, J. Hu1, J. Nguyen1, N. Bouchama1, J. Mongrain1, M. Talajic1,2, J.C. Tardif1,2, S. Perreault1,2, M.P. Dubé1,2.

1) Beaulieu-Saucier Pharmacogenomics Centre, Montreal Heart Institute, Université de Montréal, Montreal, QC, Canada; 2) Department of Medicine and Social and Preventive Medicine, Université de Montréal, Montreal, Quebec, Canada; 3) Department of Pharmacology, Faculty of Medicine, Université de Montréal, Montréal, Quebec, Canada; 4) Faculty of Pharmacy, Université de Montréal, Montreal, Quebec, Canada.

Background and aim: Warfarin is the most widely prescribed oral anticoagulant worldwide for the treatment and the prevention of thromboembolic disorders. However, there is a marked inter- and intra-individual variability in the warfarin dose requirement. Genetic factors, including the common variants of vitamin K epoxide reductase complex subunit 1 (VKORC1*2), andcytochrome P-450 2C9 (CYP2C9*2 and *3), can predict up to 40 percent of the variability in warfarin dose and the known non-genetic factors, including weight, height, age and drug interaction, are responsible for another 30%. Hence, the sources of 30% of warfarin dose variability are still unknown. We first investigated the association of regular physical activity (RPA) with warfarin dose in a large cohort of warfarin new users, the Quebec Warfarin Cohort (QWC). We then replicated our findings in an independent on-warfarin hospital-based population, the Montreal Heart Institute (MHI) Biobank.

Methods: The QWC is an observational, community-based, prospective cohort of new warfarin users who were recruited at 18 anticoagulation clinics in the Quebec province of Canada. The level of RPA was assessed in 1,064 patients of the QWC using the Stanford Brief Activity Survey. A regression model was used to evaluate the association between the baseline level of RPA and the warfarin dose requirement at 3-month time point after the initiation of treatment. The model was adjusted for height, weight, age as well asCYP2C9*2, *3 and VKORC1*2 variants. For replication, warfarin dose of prevalent users was modeled in 681 patients of the MHI Biobank in whom the level of RPA was assessed with the Global Physical Activity Questionnaire. The model was adjusted for height, weight, age and CYP2C9 variants in the replication population. Results: Higher levels of RPA were dently associated with warfarin dose in both the QWC (P value = 0.001) and the MHI Biobank (P value = 0.001) in the multiple regression models. Parent diplotypes were modeled using splash allele frequency < 5% coding variants are discovered. Patient diplotypes are reported using standard nomenclature (i.e. CYP2C19*1/*17) along with metabolizer class (Ultrarapid, Extensive, Intermediate, Poor), based on haplotype activity scores from the CPIC guidelines. If non-CPIC, rare (allele frequency < 5%) coding variants are discovered and annotated to sequencing technologies, and can be used with data generated by targeted PCR-based methods (e.g., Sequenom), whole-exome sequencing (WES) and whole genome sequencing (WGS). The tool can be used with WES or WGS data regardless of why the data was generated. In the event of missing data, our tool will display possible haplotypes. The Pharmacogenomics app is currently being piloted as part of a clinical diagnostic exam at The Hospital for Sick Children in Toronto for the evaluation of PGx in paediatric clinical practice.

Background: Pharmacogenomics (PGx) is increasingly gaining acceptance as a guide to prescribing. However, PGx guidelines are complex and human cognition alone is insufficient to implement them in busy clinical workflows. Methods: To implement PGx in the clinical setting we developed clinical decision support (CDS) using synchronous and asynchronous rules integrated in two commercially available electronic medical record (EMR) systems. Results: We were able to retrospectively assess the impact of this information for personalized drug therapy. EMRs for deriving pharmacogenetic information, as well as the potential prescribing patterns. Our study demonstrates the utility of biobank-linked metabolizer status phenotypes for several drugs, including clopidogrel, warfarin, and compared metabolizer status frequencies across different cohorts, ethnicities, and published frequencies associated with certain drugs. By linking genotype data for 103,006 individuals, we phased and imputed a panel of 41 pharmacogenes of clinical importance that also use the canonical star allele definitions, the ability to distinguish alleles defined by a single variant from those defined by multiple variants and the ability to handle rare variants. Potential phased errors, and partial or mixed haplotypes. We believe that assessing why and how often these errors will occur is important in discussing the use of nomenclature to support the clinical implementation of pharmacogenomics in the era of high-throughput sequencing.


Background: Pharmacogenomics (PGx) is increasingly gaining acceptance as a guide to prescribing. However, PGx guidelines are complex and human cognition alone is insufficient to implement them in busy clinical workflows. Methods: To implement PGx in the clinical setting we developed clinical decision support (CDS) using synchronous and asynchronous rules integrated in two commercially available electronic medical record (EMR) systems. Results: We were able to retrospectively assess the impact of this information for personalized drug therapy. EMRs for deriving pharmacogenetic information, as well as the potential prescribing patterns. Our study demonstrates the utility of biobank-linked metabolizer status phenotypes for several drugs, including clopidogrel, warfarin, and compared metabolizer status frequencies across different cohorts, ethnicities, and published frequencies associated with certain drugs. By linking genotype data for 103,006 individuals, we phased and imputed a panel of 41 pharmacogenes of clinical importance that also use the canonical star allele definitions, the ability to distinguish alleles defined by a single variant from those defined by multiple variants and the ability to handle rare variants. Potential phased errors, and partial or mixed haplotypes. We believe that assessing why and how often these errors will occur is important in discussing the use of nomenclature to support the clinical implementation of pharmacogenomics in the era of high-throughput sequencing.


Background: Pharmacogenomics (PGx) is increasingly gaining acceptance as a guide to prescribing. However, PGx guidelines are complex and human cognition alone is insufficient to implement them in busy clinical workflows. Methods: To implement PGx in the clinical setting we developed clinical decision support (CDS) using synchronous and asynchronous rules integrated in two commercially available electronic medical record (EMR) systems. Results: We were able to retrospectively assess the impact of this information for personalized drug therapy. EMRs for deriving pharmacogenetic information, as well as the potential prescribing patterns. Our study demonstrates the utility of biobank-linked metabolizer status phenotypes for several drugs, including clopidogrel, warfarin, and compared metabolizer status frequencies across different cohorts, ethnicities, and published frequencies associated with certain drugs. By linking genotype data for 103,006 individuals, we phased and imputed a panel of 41 pharmacogenes of clinical importance that also use the canonical star allele definitions, the ability to distinguish alleles defined by a single variant from those defined by multiple variants and the ability to handle rare variants. Potential phased errors, and partial or mixed haplotypes. We believe that assessing why and how often these errors will occur is important in discussing the use of nomenclature to support the clinical implementation of pharmacogenomics in the era of high-throughput sequencing.


Background: Pharmacogenomics (PGx) is increasingly gaining acceptance as a guide to prescribing. However, PGx guidelines are complex and human cognition alone is insufficient to implement them in busy clinical workflows. Methods: To implement PGx in the clinical setting we developed clinical decision support (CDS) using synchronous and asynchronous rules integrated in two commercially available electronic medical record (EMR) systems. Results: We were able to retrospectively assess the impact of this information for personalized drug therapy. EMRs for deriving pharmacogenetic information, as well as the potential prescribing patterns. Our study demonstrates the utility of biobank-linked metabolizer status phenotypes for several drugs, including clopidogrel, warfarin, and compared metabolizer status frequencies across different cohorts, ethnicities, and published frequencies associated with certain drugs. By linking genotype data for 103,006 individuals, we phased and imputed a panel of 41 pharmacogenes of clinical importance that also use the canonical star allele definitions, the ability to distinguish alleles defined by a single variant from those defined by multiple variants and the ability to handle rare variants. Potential phased errors, and partial or mixed haplotypes. We believe that assessing why and how often these errors will occur is important in discussing the use of nomenclature to support the clinical implementation of pharmacogenomics in the era of high-throughput sequencing.


Background: Pharmacogenomics (PGx) is increasingly gaining acceptance as a guide to prescribing. However, PGx guidelines are complex and human cognition alone is insufficient to implement them in busy clinical workflows. Methods: To implement PGx in the clinical setting we developed clinical decision support (CDS) using synchronous and asynchronous rules integrated in two commercially available electronic medical record (EMR) systems. Results: We were able to retrospectively assess the impact of this information for personalized drug therapy. EMRs for deriving pharmacogenetic information, as well as the potential prescribing patterns. Our study demonstrates the utility of biobank-linked metabolizer status phenotypes for several drugs, including clopidogrel, warfarin, and compared metabolizer status frequencies across different cohorts, ethnicities, and published frequencies associated with certain drugs. By linking genotype data for 103,006 individuals, we phased and imputed a panel of 41 pharmacogenes of clinical importance that also use the canonical star allele definitions, the ability to distinguish alleles defined by a single variant from those defined by multiple variants and the ability to handle rare variants. Potential phased errors, and partial or mixed haplotypes. We believe that assessing why and how often these errors will occur is important in discussing the use of nomenclature to support the clinical implementation of pharmacogenomics in the era of high-throughput sequencing.
722M

Clopidogrel is an antiplatelet agent used concomitantly with aspirin or as monotherapy in secondary prevention of cardiovascular complications. However, between 5 and 40% of patients develop resistance identified by persistent high platelet reactivity (1). Clopidogrel is metabolized by the hepatic cytochrome P450 (CYP 2C19 enzyme). CYP 2C19 exhibits genetic polymorphisms responsible for the presence of poor metabolizers, intermediate metabolizers and extensive metabolizers. In recent years, research has focused on the CYP2C19 enzyme encoded on chromosome 10. Its polymorphisms may reduce the formation of the active metabolite of clopidogrel and reduce platelet effect. The *2 allelic carriers have increased risk of cardiovascular death, myocardial infarction and stent thrombosis compared to non-carriers (2). The defective mutations of the enzyme and their frequencies change between different ethnic groups; however, the polymorphisms of the CYP2C19 gene has not been studied in colombian population. The objective of this study is to determine the prevalence of the polymorphism CYP2C19*2 in a population treated with clopidogrel in a tertiary clinic. Specifics: 1. To evaluate the genotype and phenotype status of CYP2C19 in colombian population, in order to contribute to the use of appropriate strategies of drug therapy for this population. 2. To determine the frequency of the polymorphisms associated to variability in response to clopidogrel. 3. To identify the phenotype of patients according to the polymorphisms found. 4. Determining the population at high risk of thrombotic events and cardiovascular death according to genotype/phenotype. 5. To calculate the frequency of demographic variables and non-genetic factors which have been associated with variability in response to clopidogrel in the study population.

722S
Clinician Views about Implementation of Pharmacogenomics into Practice. J.F. Peterson1, J.R. Field2, Y. Shi1, J.G. Schildcrout2, J.C. Denny3, T. McGregor4, S.L. VanDriest5, J.M. Pulley1, I. Lubin6, M. Laposata7, D.M. Roden8, E.W. Clayton6,4. 1) Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN; 2) Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, TN; 3) Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN; 4) Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN; 5) Genetics, Centers for Disease Control and Prevention, Atlanta, GA; 6) Department of Medicine, Vanderbilt University Medical Center, Nashville, TN; 7) Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN; 8) Center for Biomedical Ethics and Society, Vanderbilt University Medical Center, Nashville, TN.

Attitudes of clinicians participating in pharmacogenomics implementation projects are not well-described but are likely key to the success of translational programs. Using a 54-item online instrument, we surveyed physicians and nurse practitioners in 2013 taking part in a large pharmacogenomics program within a tertiary care academic medical center. Domains surveyed included attitudes towards ordering pharmacogenomics testing, prior pharmacogenomics education, major factors influencing use of the results, expectations of efficacy of genotype-tailored therapy, and responsibility for applying significant results from a multiplexed genetic panel to clinical practice. Of 121 clinicians solicited, 80 (66%) physicians and nurse practitioners gave a complete survey response. Virtually all respondents (99%) agreed that pharmacogenomic variants influence patients’ response to drug therapy, and a majority agreed with statements regarding the clinical utility of CYP2C19 variants to tailor antiplatelet therapy following percutaneous coronary interventions (80%) and VKORC1 and CYP2C9 variants to tailor warfarin dosing (86%). Strength of evidence for drug-genome interactions (DGIs) was cited by 88% of respondents as the most important factor influencing the ordering of pharmacogenomic testing; systematic reviews and national standards-setting organizations were the most frequently trusted sources for evidence-based guidance. The majority (92%) of respondents were in favor of immediate notification when a clinically significant drug-genome interaction was detected by pharmacogenomics testing. Of 69 respondents who felt most comfortable with responsibility for acting on a result when a prescription change was indicated (51% agreement). When presented with a scenario where the prescription change was affected by a genetic testing performed 6 months earlier, a majority of respondents (67%) favored assigning responsibility to the implementation program to find and contact the patient’s most current providers. Among clinicians practicing within a pharmacogenomics implementation program, genotype results were valued for tailoring prescriptions, but respondents were divided on who is clinically responsible for pharmacogenomic results from a multiplexed panel.
724M Role of APOE4 genotypes in predicting cardiometabolic outcomes in individuals with metformin and metformin-sulfonfonylurea combination therapy. G. Priamvada, B.R. Sapkota, A. Subramanian, P.R. Blackett, D.K. Sanghera. Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK.

Genetic polymorphisms in apolipoprotein E (APOE) locus are known to have a significant impact on various inflammatory and metabolic diseases. APOE genotyping is a well-known marker associated with increased risk of coronary artery disease (CAD) and late-onset of Alzheimer's disease. Recent evidence suggests improved clinical outcomes with metformin in patients with type 2 diabetes (T2D) and heart failure. However, the role of anti-diabetic medications in response to APOE genotypes is less understood. The objectives of this investigation were: 1) to evaluate the distribution of APOE polymorphisms in a large diabetic case-control cohort of the Asian Indian Diabetic Heart Study, 2) to evaluate the impact of APOE polymorphisms on quantitative risk factors of T2D and CAD, and 3) to examine the role of APOE genotypes in response to anti-diabetic therapy. We assessed the role of APOE genotypes with disease and treatment outcome in a total of 3,569 individuals (2,289 T2D cases and 1,280 controls). No significant difference in the distribution of APOE4 was observed among T2D cases and controls. However, APOE4 carriers had higher fasting glucose (p=0.021), higher diastolic blood pressure, higher LDL cholesterol and lower HDL cholesterol (p=0.01) compared to non-APOE4 carriers. Further stratification of data from diabetic patients by APOE genotypes and anti-diabetic treatments revealed a significant decrease in fasting glucose (p<0.001), 2 hours glucose (p<0.001), systolic blood pressure (p=0.007), diastolic blood pressure, and LDL cholesterol (p=0.0001) among the APOE4 carriers compared to non-APOE4 carriers of the same anti-diabetic therapy. Similar improved clinical outcomes were observed in patients with metformin-sulfonfonylurea combination therapy (n=618). Our study suggests APOE4 to be a potential risk factor for cardiometabolic susceptibility in patients with diabetes. Our findings might therefore provide clinical guidance to prescribers and help physicians make appropriate decisions.

725S Genome Liberty: Direct-to-Consumer Pharmacogenetics. J.A. Rosenfeld*, C.E. Mason*. 1) Department of Medicine, Rutgers Medical School, Newark, NJ; 2) American Museum of Natural History, New York, NY; 3) Department of Department of Physiology and Biophysics, Cornell University, New York, NY.

For many medications, there are genetic markers whose genotype will aid in proper dosing. Certain genotypes will rule out the use of a particular medication, while other genotypes will indicate the use of either a higher or a lower dose. There are a wide range of medications that can be tested in this way including Plavix, codeine, Coumadin and estrogen. While the links between the cytochrome p450 genes and proper dosing of medications are widely known, their clinical use has been extremely limited. The vast majority of doctors refrain from performing this or other genetic tests due to their limited knowledge of genetics and their reticence to order a test they do not completely understand. This extreme restraint by doctors has prevented the widespread adoption of pharmacogenetics testing. We have therefore developed a rapid, and low-cost direct-to-consumer test that will empower individuals to help their physicians make proper prescribing decisions. Our test will look at all of the clinically actionable pharmacogenetic markers and produce both a physician-friendly and a patient-friendly report that will allow individuals to know which medications they should avoid, and which they should use at a dose different than that recommended for the general public.

726M CHRNA4 rs1044396 is associated with smoking cessation with varenicline therapy. P.C.J.L. Santos¹, J.R. Santos¹, P.R.X. Tomaz¹, J.S. Issa², T.O. Abreu, J.E. Krieger³, A.C. Pereira¹. 1) Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo Medical School, Sao Paulo, Brazil; 2) Smoker Assistance Program (PAF), Heart Institute, Sao Paulo, SP.

Background: Smoking is one of the biggest risk factors for developing cardiovascular disease, and is considered the most preventable cause of premature death worldwide. It is estimated that there are currently over 1.3 billion smokers globally and although there are several therapies for smoking cessation, the most smokers find quitting difficult. Aim: to evaluate the association of CHRNA4 polymorphisms with the success rate of smoking cessation therapy and with the Fagerstrom Test for Nicotine Dependence (FTND) in patients treated with varenicline. Methods: Four hundred eighty-three smoker patients were evaluated by FTND questionnaire and treated with varenicline and/or bupropion. They were followed for 12 weeks. CHRNA4 rs1044396 and rs2236196 polymorphisms were genotyped by melting curve analysis. Results: We observe that patients with rs1044396 CC genotype had lower frequency of success rate (29.5%) compared with patients with rs1044396 CT or TT genotypes (50.9%). During varenicline therapy (p=0.007; n=167). We also observed that CHRNA4 rs1044396 was associated with smoking cessation success, even in multivariate model (OR: 2.28; 95% CI: 1.03–5.06). However, we did not observe association of CHRNA4 rs1044396 and rs2236196 genotypes with FTND and with bupropion therapy. Conclusion: Our findings indicate that CHRNA4 rs1044396 might be a pharmacogenetic marker for smoking cessation with varenicline therapy since this gene encodes a subunit of the nicotinic acetylcholine receptors more abundant and higher activity.

727S Early drug responses that are followed by an acquired drug resistance in non-small cell lung cancer cells exposed to gefitinib. M. Takahashi¹, M. Fukukawa¹, H. Hajiyo¹, 1) Department of Molecular Pharmacology, National Institute of Neuroscience, NCNP, Kodaira, Tokyo, Japan; 2) Division of RNA Medical Science, The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo, Japan.

Acquired drug resistance is a major problem in medication, particularly in cancer treatment. Early drug responses followed by a drug resistance may provide us with a clue for understanding the mechanism of acquired drug resistance and for developing a therapeutic strategy to overcome such a resistance. We attempted to examine early drug responses of cancer cells that were treated with gefitinib [an epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor]. Non-small cell lung cancer (NSCLC) PC-9 cells that bear an oncogenic mutant EGFR and have the potential for acquiring resistance to gefitinib were used in this study, and we established gefitinib-resistant PC-9 cells (PC-9/GR) by a long-term exposure of a low-dose gefitinib. Comprehensive gene expression analyses against naive PC-9 cells and PC-9/GR cells were carried out, and the results indicated that fibroblast growth factor 2 (FGF2) and its receptor 1 (FGFR1) were significantly increased in PC-9/GR cells. In addition, while the FGF2 and FGFR1 protein levels were increased in PC-9/GR cells, oncogenic EGFR dimer appeared to be decreased. The findings suggested the possibility that a survival signaling system might have been changed from an EGFR dependent-signaling to a FGFR dependent-signaling in PC-9/GR cells. To see early responses of naive PC-9 cells against gefitinib further, we investigated FGFR1 and FGF-2 in naive PC-9 cells exposed to gefitinib. The analysis revealed that FGFR1 was increased in an exposure time- and gefitinib dose-dependent manner. As for FGF-2, although its apparent gene expression level appeared to remain unchanged, the FGF-2 protein in culture medium was markedly increased by gefitinib treatment in a dose-dependent manner: this may mean FGF-2 that was released from dead cells. Additionally, naive PC-9 cells that were treated with exogenous FGF-2 appeared to decrease in sensitivity to gefitinib. Taken together, the results suggested that naive PC-9 cells exposed to gefitinib might be survivable by an up-regulated FGFR1 and by an increase in FGF-2 that was released from neighboring dead cells, and that PC-9/GR cells, after acquisition of gefitinib resistance, would establish a FGF autocrine signaling system. In addition to the above findings, we would like to further discuss the possibility that thyroid hormone-related factors might contribute to up-regulation of FGFR1 in gefitinib-treated cells.
728M
Modeling the pharmacological response to advance the research in pharmacogenetics. J. Bertrand1, M. de lorio2, D.J. Balding1. 1) Genetics Institute, University College London, LONDON, --., United Kingdom; 2) Statistical Science Dept, University College London, LONDON, --., United Kingdom.

Context: Pharmacogenetics studies the genetic part of interindividual variability in drug response. Its main challenges are: i. the phenotype results from physiological processes (e.g. drug elimination or distribution) not directly observed; ii. genetic variation markers lead to uneven and small sample size subgroups and iii. one has to face the dimensionality curse when dealing with high throughput genetic screening. Nonlinear mixed effect models (NLME) handle challenge i. and ii. by capturing the dynamical nature of the phenotype and fitting patients with sparser profiles while borrowing information from more informative ones. Objective: In this work we propose to address challenge iii. with an integrated approach, that combines the strengths of NLME with penalised regression, and compare its performance to a classical stepwise procedure through a simulation study. Methods: We simulated pharmacokinetic profiles in association with genotypes for the 1200 single nucleotide polymorphisms (SNP) of the DMET Chip (Daly, 2007). Several scenarios were studied in terms of design: a typical phase II study or a combined design of such a study with additional through concentrations collected in clinical routine and in terms of genetic association: causal variants affecting one or multiple model parameters. Results: Both approaches showed low power estimates on the phase II study design. Adding patients with only a through concentration led to much higher power estimates for both approaches, with the stepwise procedure detecting more false positives. Further, the integrated approach was more powerful at detecting SNP effects on multiple model parameters. Conclusion: These findings were confirmed by qPCR, and the IgM observation was attributed to higher EFZ and NVP plasma levels, suggesting that it might favor HAART failure in EFZ or NVP containing regimens. The study detected the three previously reported non-synonymous variations, 480 G>C (L160F); 1022 C>T (P341L); 1222 A>G (M408V) and one synonymous variation 156 T>C (S52S) at a minor allele frequency of 0.63, 0.20, 0.43 and 0.27 respectively. This study (M408V) and one synonymous variations 156 T>C (S52S) at a minor allele frequency of 0.63, 0.20, 0.43 and 0.27 respectively. This study

729S
“A Tale Of Genetic Variation In The Human SLC22A1 Gene Encoding OCT 1 Among Type 2 Diabetes Mellitus Population Groups Of West Bengal, India”. D. Sur C.U., Kolkata, Westbengal, India.

The organic cation transporter 1, OCT 1 (also called SLC22A1-Solute Carrier Family 22 member 1), appears to play a role in the efficacy and safety of some drugs, and genetic polymorphisms in the drug transporter have been increasingly recognized as a possible source of variation in drug disposition and response. Genetic variants in OCT1 have been identified largely in European, Asian (Japanese, Chinese and Korean) populations. Interestingly, eight genetic variations were found in the human SLC22A1 gene, which encodes OCT 1, from 50 type 2 diabetes mellitus individuals (T2DM), in West Bengal population. The purpose of this study was to investigate genetic variants of OCT1 in West Bengal populations. The study detected the three previously reported non-synonymous variations, 480 G>C (L160F); 1022 C>T (P341L); 1222 A>G (M408V) and one synonymous variation 156 T>C (S52S) at a minor allele frequencies (MAF) of 0.63, 0.20, 0.43 and 0.27 respectively. This study also found four previously reported intronic variations: IVS1+43(T>G), IVS2 -99(C>T), IVSS -81(G>A), IVS9 +43(C>T) with minor allele frequencies of 0.20, 0.17, 0.18, and 0.37 respectively. This is the first report of SLC22A1 variations among Indian, especially West Bengal’s type 2 diabetes mellitus patients. The present results would be useful for haplotype analysis and pharmacogenetic studies on OCT1.

730M
A role for B cells in Progressive Multifocal Leukoencephalopathy revealed by comprehensive genomic analysis. J. Carulli1, L. Fugger1, A. Day-Williams1, A. Haghi2, C. Sun3, N. Affaire1, H. Li1, C. Dendrou1, T. Plavina1, H. McLaughlin1, P. Cullen1, M. Liu1, A. Thai1, R. Gold4, R. Martin1, I. Jelic3, T. Compton1, T. Harris1. 1) Translational Medicine, Bioengineering, Cambridge, MA; 2) The Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DS; 3) Department of Neurology, St Josef-Hospital, Ruhr University Bochum, Bochum, Germany; 4) Department of Neurology University Hospital Zurich Frauenkindeskasse 26 CH-8091 Zurich.

Multifocal progressive leukoencephalopathy (PML) is a rare and potentially lethal infectious disease caused by JC virus in individuals with compromised immune systems or who take specific immune modulating drugs. JC virus (JCV) has a seroprevalence of 60% in adult populations worldwide, but PML is very rare. Prior to the development of anti-HIV drugs, roughly 5% of AIDS patients developed PML. More recently, PML has been observed at a rate of approximately 1/1000 or lower in patients taking immune-modulating drugs, including natalizumab (anti-VEGFR4) for multiple sclerosis. To identify novel biomarkers of PML risk, we performed whole exome sequencing, whole genome sequencing, and genome wide transcription profiling of multiple sclerosis patients who developed PML while taking natalizumab with appropriately matched MS patients who did not develop PML while taking natalizumab for two years or more. We detected whole genome sequencing of two subjects who developed PML while not taking any immune modulating drugs. Single SNP association testing and gene-based collapsing tests of nearly 100 natalizumab PML cases and 100 natalizumab controls did not reveal any significant associations with PML. One of the two “spontaneous” PML subjects had a complete deletion of the Dock8 gene on chromosome 9p. Dock8 deficiency is a primary immune deficiency with pleiotropic consequences, including B cell migration defects. Genome-wide transcript profiling of 18 natalizumab PML subjects and 150 natalizumab controls revealed that several transcripts, including IgHM, FcRLa, FcRL3, CD72 and CD22, exhibited lower expression levels in PML compared to non-PML patients. These findings were confirmed by qPCR, and the IgM observation was confirmed at the protein level. Pathway analysis of the transcript data identified B cell pathways significantly deregulated in subjects who got PML. Taken together, the genetic data and the transcript data suggest that variation in maturation and/or migration of B cells plays a role in the development of PML and may provide biomarkers for PML risk.

731S
Association between CYP2B6 +516 G>T polymorphism and response to first-line therapy in Brazilian HIV-1+ individuals. T.B. Almeida, M.B. Arruda, R.M. Brindeiro, A. Tanuri, C.C. Cardoso. Laboratorio de Virologia Molecular, Instituto de Biociências, Universidade Federal do Rio de Janeiro, Brazil.

According to World Health Organization, 35 million individuals are infected by HIV-1 in the world and about 9.5 million are currently undergoing highly active antiretroviral therapy (HAART). However, 10–20% of these individuals do not reach therapeutic success, mainly as a consequence of low adherence to the treatment and emergence of drug resistant viral strains. HAART effectiveness is also influenced by host factors that affect drug absorption, activation and metabolism. Single nucleotide polymorphisms (SNP) in the CYP2B6 gene (involved in CYP2B6 gene and HAART failure in EFZ or NVP containing regimens) are considered as important genetic factors. The aim of this study was to investigate the association between 14 candidate SNPs in CYP2B6 gene and HAART failure in EFZ or NVP containing first line regimens. A total of 111 HIV-1+ individuals under HAART selected from the Brazilian cities of Porto Alegre and Curitiba were included in the study. From these, 73 individuals reached therapeutic success, defined by the reduction of viral loads (VL) to undetectable levels (<50 copies/mL). The remaining 38 individuals exhibited HAART failure, with detectable VL after 6 months of treatment. SNP genotyping was performed by SNApshot® or TaqMan® assays (Life Technologies, USA). Nine SNPs were excluded from the study due to deviations from Hardy-Weinberg Equilibrium or because they were monomorphic in our study population. The frequencies of the remaining 16 SNPs were compared using logistic regression models adjusted for sex and HIV-1 subtype. Our results showed an increased risk of HAART failure (OR=2.21; 95%CI= 0.96–5.04; p=0.048) among carriers of the allele +516T (rs3745274). The same allele has been previously associated to higher EFZ and NVP plasma levels, which might favor therapy effectiveness. However, the effect of this SNP may be influenced by additional variations in CYP2B6 or genes coding for drug transporters. It suggests that the genetic background of each population, including linkage disequilibrium patterns, may influence data interpretation, reinforcing the importance of replicating genetic association results in different populations. To our knowledge, this was the first study to describe the association of CYP2B6 gene to HAART response in Brazilian HIV-1+ individuals. The use of pharmacogenetic information in routine clinical practice may be helpful to define personalized regimens, which may be more effective and better tolerated by the patients.
732M

Identifying Differentially Expressed Genes Associated with Extreme Blood Pressure Response to Hydrochlorothiazide monotherapy.

A.C. Costa Sa1,2, A. Webb3, Y. Gong1,2, T. Langas1,2, S. Turner1, A. Chapman4, J. Gums1,2,5, S. Scherer6, R.M. Cooper-DeHoff7,9, W. Sadee4, J.A. Johnson1,3,9, 1) Pharmacotherapy and Translational Research, University of Florida, Gainesville, FL; 2) Center for Pharmacogenomics, University of Florida, Gainesville, FL; 3) Graduate Program in Genetics and Genomics, University of Florida, Gainesville, FL; 4) Department of Pharmacology, College of Medicine, Ohio State University, Columbus, OH; 5) Division of Nephrology and Hypertension, Mayo Clinic, Rochester, MN; 6) The Renal Division, Department of Medicine, Emory University, Atlanta, GA; 7) Department of Community Health and Family Medicine, University of Florida College of Medicine, Gainesville, FL; 8) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 9) Division of Cardiovascular Medicine, Department of Medicine, University of Florida, Gainesville, FL.

Hypertension (HTN) is the most common chronic disease in the world, estimated to affect nearly 1 billion individuals worldwide. Hydrochlorothiazide (HCTZ) is one of the most commonly prescribed antihypertensive (anti-HTN) medications in the US with over 100 million prescriptions annually. Even though HCTZ is one of the preferred options for initial HTN therapy, patient’s response varies widely, and less than 50% of them achieve blood pressure (BP) control. We hypothesize that interpreting the functional elements of the genome will help us in identifying candidate biomarkers associated with variability in the efficacy of HCTZ therapy. The present work aims to identify genes with differential expression levels between respondents and non-responders after HCTZ monotherapy in Caucasian hypertensive PEAR (Pharmacogenomic Evaluation of Antihypertensive Responses) participants. Uncomplicated hypertensive patients were recruited and randomized to either HCTZ or atenolol monotherapy. BP and adverse metabolic responses were assessed after 9 weeks of monotherapy. On 50 European American participants, classified as responders and non-responders to HCTZ, total RNA was isolated from whole blood and used for transcriptomic RNA sequencing with Illumina HiSeq2000®. The generated reads were aligned to the reference genome (Homo Sapiens Hg19) with TopHat2. Cufflinks/Cuffdiff pipeline was used to calculate gene expression levels, reported as fragments per kilobase per million reads (FPKM). Paired t-test was performed to evaluate the significance of the change in FPKM. Differential expression analysis revealed that 2 genes may play an important role in BP response to HCTZ. RHOB and CDC42EP2 (FDR adjusted p-value < 0.05). RHOB is a small GTP binding protein, member of the Rho protein superfamily. It activates serine/threonine kinase MRRK, promoting myosin phosphorylation. Previous studies highlight the convergence of Rho and CDC42 signaling in VSMCs, corroborating our RNA-Seq findings. These results suggest that Rho/Rho-kinase and CDC42/CDC42EP2 pathways may be involved in vascular resistance hypertension and possible reflect in BP response variability associated with thiazide diuretics anti-HTN treatment.

733S

Contribution of rare protein-coding variants to anti-TNF treatment response in rheumatoid arthritis patients. A. Dlugo1,2, J. Cul1, R.S. Futtner3,17, J. Greenberg1,2, D.A. Rappas1,2,4, J.M. Kremers6, A. Barton10, M.J.H. Coenen11, B. Franke11, L.A. Kienemeyer12, X. Mariette13,14, C. Richard-Miceli13,14, H. Canhao16,16, J.E. Fontsea16,16, N. de Vries17, P.P. Tak17, J.B.A. Crusius18, M.T. Nurmmohamed19, F. Kureemaan6, T.W.J. Huizinga20, Y. Okada1,3, E.A. Stank1,3, L. Kaireskog21, L. Padyukov21, E.R. Mardis22, P.K. Gregersen6,7, R.M. Plenge1,3, S. Raychaudhuri1,10,1, Di- vision of Rheumatology, Immunology, and Allergy, Brigham and Women’s, Hospital, Harvard Medical School, Boston, MA; 2) Division of Genetics, Brigham and Women’s, Hospital, Harvard Medical School, Boston, MA; 3) Medical and Population Genetics Program, Broad Institute, Cambridge, MA; 4) Partners HealthCare Center for Personalized Genetic Medicine, Boston, MA; 5) The Genome Institute, Washington University School of Medicine, St. Louis, MO; 6) New York University Hospital for Joint Diseases, New York, NY; 7) CORR, LLC, Southborough, MA; 8) Columbia University, College of Physicians and Surgeons, New York, NY; 9) The Albany Medical College and The Center for Rheumatology, Albany, NY; 10) Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester Academic Health Sciences Centre, Manchester, UK; 11) Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands; 12) Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands; 13) Université Paris-Sud, Orsay, France; 14) APHP-Hôpital Bicêtre, INSERM U1012, Le Kremlin Bicêtre, Paris, France; 15) Rheumatology Research Unit, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal; 16) Rheumatology Department, Santa Maria Hospital-CHLN, Lisbon, Portugal; 17) Amsterdam Rheumatology and Immunology Center, Department of Clinical Immunology & Rheumatology, Academic Medical Center / University of Amsterdam, The Netherlands; 18) Laboratory of Immunogenetics, Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands; 19) Amsterdam Rheumatology and Immunology Center, Department of Rheumatology, Reade, Amsterdam, The Netherlands; 20) Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands; 21) Rheumatology Unit, Department of Rheumatology and Immunology, VU University Medical Centre, Amsterdam, The Netherlands; 22) The Feinstein Institute for Medical Research, North Shore-Long Island Jewish Health System, Manhasset, NY, USA.

There are several effective drugs that treat systemic inflammation in rheumatoid arthritis (RA), including the widely used anti-tumor necrosis factor alpha (anti-TNF) biologic therapies. However, a significant number of RA patients fail to enter remission with the drugs currently available. Despite multiple efforts, no confirmed genetic predictor of anti-TNF treatment response in RA patients has been identified from association studies of common variants. Here, we targeted 750 genes for exon-sequencing and investigated the aggregate contribution of rare protein-coding variants to anti-TNF treatment response (improvement in disease activity score) in 1,131 RA patients of European ancestry. Gene-based association tests resulted in a study-wide significant association at TNF, driven by two rare missense mutations in three patients predicting good treatment response (Beta = +3.04, P=4×10⁻⁷). Interestingly, we observed several additional genes involved in the TNF signaling pathway among the genes with the strongest signals of association (P<0.05). We performed a gene set enrichment analysis based on association Pvalues ranking, and observed a nominal enrichment of association at genes involved in the TNF pathway (Penrichment=0.03, based on 1,000 phenotype permutations). Our results, if validated in independent collections, would suggest that rare protein-coding variants in genes from the TNF signaling pathway contribute to treatment response in RA patients.
734M
Assessing the clinical utility of massively parallel sequencing for pharmacogenomics research in the ClinSeq® study. D. Ng1, L.N. Singh1, C.S. Hong1, K.L. Lewis1, J.C. Mulliken1, L.G. Biesecker1,2, NIH Intramural Sequencing Center, National Human Genome Research Institute, NIH. 1) Medical Genomics and Metabolic Genetics Branch, MIMGB/NIHGR/NHGRI, Bethesda, MD; 2) NIH Intramural Sequencing Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 3) Comparative Genomics and Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD.

ClinSeq® is a medical sequencing project designed to study the clinical application of massively parallel sequencing (MPS) in a cohort of mostly healthy volunteers. One of the goals of personalized medicine is to optimize the efficacy of pharmacologic treatment and to minimize adverse drug reactions. Pilot projects studying the utility of pharmacogenetic (PGx) guided treatment have used Sanger sequencing or genotyping chips to identify a finite number of important PGx variants. To study the utility of MPS for identifying clinically important PGx variants, we compared the accuracy and coverage of whole exome (WES) and genome sequencing (WGS) at 1,934 variant positions interrogated by the Affymetrix Drug Metabolism and Transport (DMET) Plus chip. ClinSeq® cohort is comprised of 973 volunteers age 45–65 from the Washington DC area. Participants received baseline clinical evaluations, WES, consent for return of genetic results and iterative phenotyping to study incidental variants. Five participants had both WES and WGS. Sequence reads were aligned to NCBI assembly GRCh37 with Novoalign v3.02.00 (Novocraft technologies). Genotypes were determined with GATK v3.0-0. Coverage of 1,934 variant positions in 231 DMET genes were analyzed in 973 exomes. High quality reads were defined as GATK score ≥99 and read depth (RD) ≥10. A variant position had high coverage (HC) if ≥80% (n=779–973) of participants had high quality reads, low coverage (LC) if <80% (1–778) and no coverage (NC) if no one had a high quality read at that variant position. WES had the poorest coverage for intergenic (0/31 HC), promoter (3/91 HC) and 5’UTR variants (10/39 HC). Unsurprisingly, WES coverage was the highest in exonic variants (1,239/1,332 HC, 92/1,332 LC, 1/1,332 NC). WES RD can be used to identify copy number variants (CNVs) affecting drug metabolism (i.e., CYP2D6) with XHMM software. The accuracy and precision of MPS was validated by sequencing of each target variants. The accuracy of the genotyping was also validated by sequencing each target variants. The accuracy and precision of the kit ranged from 81.6% to 100% and from 94.8% to 100% on common SNPs. The percentage of correct calls of Korean-specific alleles ranged from 80.2% to 96.6% and from 98.3% to 100%, respectively. The overall accuracy and precision of the kit turned out as 85.8% and 99.1% respectively. The results suggested that the newly developed kit was capable of genotyping both common and Korean-specific CYP2C19 alleles with effectiveness. We are expecting that this kit provides a cost-effect mean for genotyping pharmacogenetically important CYP2C19 alleles in Korean population. Acknowledgement: This research was supported by a grant (152011NFD705) from the Ministry of Food and Drug Safety in 2013. And The Korean DNA was supplied by Center for Genome Science National Institute of Health, KCDC. Korean DNA was originated by Ansung-Ansan cohort data.

735S

Most of drugs are xenobiotics for human body and are metabolized mainly in the liver. Drug metabolizing enzymes in the liver and highly dependant on the activity of enzymes. Cytochrome P450s (CYPs) are known to play essential roles in phase I drug metabolism in the liver. Among the CYPs, CYP2D6, CYP2C9 and CYP2C19 are known to play important roles in pharmacogenomics. Among the list of pharmacogenomic biomarkers in drug labeling, CYP2C19 is the most common valid biomarker involved in metabolisms of 14 different drugs and a number of genotyping kits are already available in the market after US FDA’s approval. National Institute of Food and Drug Safety Evaluation (NIFDS) has put its effort to conduct pharmacogenomic studies since early 2000s for harnessing the new paradigms of drug utilization such as personalized medicine in Korea. As a part of the effort, we developed a genotyping method for CYP2C19 variants using a single base extension technique (SNaPShot) containing Korean specific minor alleles. The kit is capable of genotyping 15 variants of CYP2C19 including common variants, like *2 [A], *3 [A], *4 [G], *5A [T], *8 [C], *17 [T] and also a couples of variants specific to Korean population such as rs11568732, rs3758580, rs3758581, rs4417205 and rs4986894. The validity of the kit was evaluated by triple repetition test on 96 Korean DNA samples. The accuracy of the genotyping was also validated by sequencing of each target variants. The accuracy and precision of the kit ranged from 81.6% to 100% and from 94.8% to 100% on common SNPs. The percentage of correct calls of Korean-specific alleles ranged from 80.2% to 96.6% and from 98.3% to 100%, respectively. The overall accuracy and precision of the kit turned out as 85.8% and 99.1% respectively. The results suggested that the newly developed kit was capable of genotyping both common and Korean-specific CYP2C19 alleles with effectiveness. We are expecting that this kit provides a cost-effect mean for genotyping pharmacogenetically important CYP2C19 alleles in Korean population. Acknowledgement: This research was supported by a grant (152011NFD705) from the Ministry of Food and Drug Safety in 2013. And The Korean DNA was supplied by Center for Genome Science National Institute of Health, KCDC. Korean DNA was originated by Ansung-Ansan cohort data.

736M
NAT2 polymorphisms in a Brazilian indigenous group. V.M. Zembruski1, R.L.F. Teixeira1, P.H. Cabellon1,2, 1) UNIGRANRIO, Duque de Caxias, RJ, Brazil; 2) IOC-Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, RJ, Brazil.

Tuberculosis is an important disease to public health in Brazil and around the world, especially among indigenous people from Amazonia and Central Brazil. Despite treatment to be effective, adverse drug reactions such as hepatotoxicity can occur, decreasing adherence to medication therapy and may lead to drug resistance. Isoniazid (INH) is the major drug associated with drug-induced hepatotoxicity. INH is metabolized via hepatic N-acetyltransferase 2 (NAT2). In subjects with low NAT2 activity there is a higher risk of developing hepatic disorder. On other hand, the high activity shows a correlation between low plasma concentrations of INH and therapeutic failure. The NAT2 protein is encoded by NAT2 gene, and the acetylator phenotypes are segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others. Based in a trimodal classification, individuals homozygous for rapid NAT2 acetylator alleles are deduced as rapid acetylators, individuals heterozygous for slow acetylator NAT2 alleles are deduced as slow acetylators, and individuals possessing combinations of SNPs segregated into clusters possessing a signature SNP either alone or in combination with others.
Pharmacogenomic Analysis of the Ashkenazi Jewish Population by Whole-Genome Sequencing. S.A. Scott1, K.L. Ayers2, Y. Yang3, S. Cheng4, K.Y. Kao5, D. Ben-Avraham6, N. Barzilai7, A. Darvasi8, K. Olfert9, S. Bressmann9, L.J. Ozelius10, J.H. Cho11, H. Ostrer12, G. Atzmon13, L.N. Clark14, T. Lencz15,12, I. Pe'er15,12, B. Reva7,1, I. Peter1. 1) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029; 3) Department of Computer Science, Columbia University, New York, NY 10027; 4) Department of Internal Medicine, Genetics & Pediatrics, Yale School of Medicine, New Haven, CT 06519; 5) Department of Genetics and Department of Medicine, Albert Einstein College of Medicine, Bronx, NY 10461; 6) Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel, 91904; 7) Cancer Biology and Genetics Program and Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065; 8) Department of Neurology, Beth Israel Medical Center, New York, NY 10033; 9) Department of Genetics and Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461; 10) Department of Pathology and Cell Biology and Taub Institute for Research of Alzheimer’s Disease and the Aging Brain, Columbia University Medical Center, New York, NY 10032; 11) Center for Psychiatric Neuroscience, The Feinstein Institute for Medical Research, North Shore-Long Island Jewish Health System, Manhasset, NY 11030; 12) Department of Psychiatry, Division of Research, The Zucker Hillside Hospital Division of the North Shore-Long Island Jewish Health System, Glen Oaks, NY 11040; 13) Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10032.

The Ashkenazi Jewish (AJ) population has a history of bottlenecks, expansions and geographical isolation, which has resulted in a unique genetic architecture that includes a high prevalence of founder mutations for autosomal recessive diseases and higher frequencies of some common disease risk alleles. Although recent candidate gene studies and exome sequencing have identified some unique pharmacogenetic variants in the AJ, no pharmacogenomic studies have been reported in this population. To identify novel variants potentially implicated in drug response variability and to characterize the pharmacogenetic landscape of the AJ, 128 individuals (100X) from the 1000 Genomes Project populations, UGT2B11 p.S195L, FMO2 p.C156R. In addition, after randomly sampling 128 individuals from 1000 Genomes Project populations, ~16.9 Mb of the human genome. Of the 73,683 total variants detected, 373 (30.4%) were novel, with 371 (99.5%) unreported with 5,547 (33.7%) of these observed in two or more individuals. Of the non-synonymous variants, 373 (30.4%) were novel, with 371 (99.5%) having minor allele frequencies (MAF)<5% and 158 (42.4%) predicted in silico to be functional. AJ individuals had a significantly lower burden of predicted functional variants (mean=23) compared to the European (EUR; 28), African (AFR; 38), and Asian (ASN; 29) populations in the 1000 Genomes Project (p<2e-22), consistent with an isolated population. Over 6% of the AJ and EUR predicted functional variants had significant differences in MAF (p<0.0001), including several predicted to have high functional impact (e.g., FMO2 p.S195L, UGT2B11 p.C156R). In addition, after randomly sampling 128 individuals (100X) from the 1000 Genomes Project populations, 149 (normal singleton) non-synonymous variants were unique to the AJ, compared to the EUR (63), AFR (457), and ASN (202) populations. Analysis of clinically actionable pharmacogenetic variants defined by PharmGKB indicated that the AJ harbor most of these variants with comparable frequencies to the EUR; however, ~16% of AJ harbor unique (non-singleton) variants in these clinically relevant pharmacogenes compared to only ~3% of the EUR, the majority predicted to be functional. These data indicate that in addition to common pharmacogenomic variants found in other populations, the AJ harbor many pharmacogenomic variants that are uncommon or absent in other populations and which may play an important role in drug response variability in this unique population.

Pharmacogenetics


The major drug metabolizing enzyme, CYP2D6, is encoded by a highly polymorphic gene. Over 100 star allele haplotypes are known, which can contain SNP, InDel, and copy number variants (CNVs) and which fall into 3 main functional categories (reduced, low, normal). The haplotype with the highest star allele content is predictive of CYP2D6 drug metabolizer phenotype (ultrarapid, extensive, intermediate, or poor). We previously described a workflow whereby sample SNP genotype and CNV results, generated using TaqMan® SNP assays and TaqMan® Copy Number assays, respectively, can be translated to star allele diplotypes using AlleleTyper™ software. However, for samples that carry CYP2D6 duplications and are heterozygous for key SNPs, the specific allele that is duplicated cannot always be identified. A phenotype can be assigned if all 3 alleles are of the same functional category, but if alleles are from different functional categories there may be 2 possible phenotypes. To address this issue, we developed a method to detect allele-specific copy number by digital PCR using the QuantiStudio™ 3D Digital PCR System. Sample DNA is loaded onto nanofluidic chips at concentrations to give one or no copies of target per dPCR. A count of reactions with and without amplification can be used for target quantification purposes. For the allele-specific dPCR application, TaqMan® SNP assays to CYP2D6 variants that are associated with specific duplicated alleles were run in dPCRs on samples of known SNP genotype and CNV status. Samples were first digested with a restriction enzyme to separate tandem duplicated CYP2D6 alleles and enable their partitioning into distinct dPCRs. Sample input amounts and thermal cycling conditions were optimized to best amplify and resolve each allele in cluster plot analysis. Reactions positive for each allele, detected by allele-specific VIC® or FAM™ dye-labeled probes, were counted and the allele ratios determined. For samples heterozygous for target SNPs, 2-copy samples gave close to 1:1 allele ratios, whereas 3-copy samples had close to 1:2 ratios. Tandem duplicated alleles could thus be identified. We have thus shown that allele-specific copy number analysis using dPCR and TaqMan® SNP assays is a simple and effective method for identifying specific duplicated alleles in heterogeneous samples. This method facilitates both CYP2D6 allele genotyping and better prediction of drug metabolizer phenotype.

Pharmacogenetics

Analysis of CYP1A2 gene non-coding region polymorphisms in Roma and Hungarian population samples. B. Meleghi1,2, P. Matyas1,2, B.I. Melégh1, L. Magyari7,2, J. Bene1, B. Duga7,2, Zs. Banfalvi1, A. Szabo7, R. Szalai1,2. 1) Department of Medical Genetics, University of Pecs, Pecs; Hungary; 2) Szentagothai Research Centre, University of Pecs, Pecs, Hungary.

Purpose: CYP1A2 enzyme contributes to biotransformation of wide variety of therapeutically important drugs, including caffeine, clopidogrel, clozapine, warfarin, procarcinogens, and some endogenous substrates. The aim of this study was to determine the pharmacogenetic profile and interethnic differences of variants of CYP1A2 gene in Roma (Gypsy) and Hungarian population.

Methods: A total of 404 Roma and 396 Hungarian healthy subject’s biobanked DNA were genotyped for two non-coding variants of CYP1A2: -163C>A (*1F) and -3860G>A (*1C). Polymerase chain reaction - restriction fragment length polymorphism technique was applied.

Results: The minor allele frequency for CYP1A2*1C variant was 2.02% in Hungarians, while it was undetectable in the Roma samples; the AA homozygous genotype was also not detectable. For CYP1A2*1F polymorphism we found a marked differences in AA genotype in Roma population compared with the Hungarians (31.9% vs. 49.5%, p<0.001) and in minor allele frequency (56.9% vs. 68.6%, p=0.025). The following CYP1A2 genotypes were identified in Roma and Hungarian samples, respectively: *1A/*1A (18.1% vs. 12.4%), *1A/*1F (50% vs. 36.9%), *1F/*1F (31.9% vs. 46.7%). In Hungarian population we found the *1C/*1F genotype (4.04%), but it was not present in Roma subjects.

Conclusions: Analysis of distribution of CYP1A2 gene variants revealed further pharmacogenetic differences between Roma and Hungarian population samples. Theoretically, as a further consequence, the Hungarians have higher chance for rapid metabolism of CYP1A2 substrates, intensified procarcinogen activation, and perhaps thereby elevated risk for cancers.
740M
Evaluation of DNA extracted from up to 16 years old post-mortem blood FTA cards using Quantifier Human Plus Quantification Kit. AL. Rahikainen, JL. Palo, A. Sajantila. Department of Forensic Medicine, Hjelt Institute, P.O.Box 40, FI-00014 University of Helsinki, Helsinki, Finland.

Introduction: The field of pharmacogenetics may significantly benefit from post-mortem (PM) studies which often harvest the extreme cases from the population. When combining genetic information to other autopsy data, especially from toxicological analyses, genetic outliers may be revealed. A potential source of DNA is FTA™ Gene Cards (Whatman™, GE Healthcare). In our department, blood samples from all autopsies have been preserved routinely since June 1998 on FTA cards. However, due to the PM changes it is expected that the starting condition of the samples vary prior to sampling and therefore methods used for PM DNA analysis must be compatible with degraded or inhibited DNA samples. The aim of this study was to determine the quality and quantity of DNA on the FTA cards sampled during 1998–2013 and to evaluate their usability in pharmacogenetic studies. Materials and methods: Four random samples from eight time points covering sampling years 1998 to 2013 resulting 32 DNA samples in total were extracted in triplicates. Four punches (2.0 mm Harris Micro-Punch™) were used in each extraction. DNA extraction was performed with Automate Express™ System (Applied Biosystems (AB)) according to manufacturer’s protocol using PrepFiler Express™ Forensic DNA Extraction Kit (AB). Extracted samples were quantified and qualified by using Quantifier HP Quantification Kit (AB) and analyzed with 7500 Real-Time PCR system (AB). Results: According to the quantification results all samples contained amplifiable DNA, but both quality and quantity were expectedly higher in the more recent samples. There was variation in the DNA quantity between the samples, despite equal amount of input material. Also internal variation was seen among the triplicates. All samples were free of inhibition. Conclusions: Blood samples on FTA cards are rather stable over 16 years although the older samples showed more degradation and lower yields compared to the recent ones. This study showed that also in the oldest samples from 1998 the mean DNA yield was 4.9 ng/µl (small target), and 3.4 ng/µl (large target) making them usable in pharmacogenetic studies, but the higher degradation level must be taken account. In Finland the full toxicological analyses are conducted from approximately 70% of the forensic autopsy cases. Therefore DNA preserved on FTA cards combined with the results of toxicological analyses offers unique opportunity for conducting pharmacogenetic studies.

741S

In pharmacogenetic studies, association tests are routinely performed to identify novel genetic markers predictive of treatment response. Traditionally, three approaches can be used to identify genetic markers in clinical trials that are designed to investigate multiple drug doses: (1) to only use data from the placebo arm and the most efficacious arm and test for genotype by treatment interaction in regression models; (2) to use data from all arms and perform an omnibus test for interactions between genotype and treatment with multiple degrees of freedom in regression models; (3) to use data from all arms and test for interaction between genotype and administered dose in regression models. However, all these approaches have drawbacks: ignoring other treatment arms in approach (1) might lead to decreased power due to reduced sample size and loss of information; approach (2) might not be powerful due to the extra degrees of freedom applied to the omnibus test; administered dose in approach (3) might not be the dose at the target site and therefore less correlated with clinical efficacy. In this study we propose an approach that utilizes data from all available treatment arms and incorporates biological mechanisms into genetic association tests in clinical trials. Drug-target binding or dose-response is estimated using pharmacometric modeling and interaction between genotype and estimated drug-target biding or dose-response is tested in regression models. Genetic and treatment-response data are simulated to mimic dose-ranging Phase II trials. Scenarios with a wide range of genetic effect sizes, mutation frequencies, rate constants for drug-target complex formation, and correlations between drug-target complex and efficacy are considered. Power of the proposed approach and conventional approaches to identify genetic markers for treatment response is compared using simulated data. The proposed approach has great potential to increase the power and probability of success of genetic studies to identify biomarkers at early clinical trial stages.
Exome sequencing of multiplex oral cleft phenotypes detects recurrent shared rare variants in 9 genes. E.R. Holzinger1, Q. Li1, M. Parker2, J.B. Hetemanski3, M.L. Marazita4, E. Mangold1, M.M. Nothen1, J.C. Murray1, A. Scott1, T.H. Beauty1, J.E. Bailey-Wilson1. 1) Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes Health, Baltimore, MD; Select a Country; 2) Department of Epimiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA; 3) Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine; Department of Human Genetics, Graduate School of Public Health; and Clinical and Translational Science Institute, University of Pittsburgh; 4) Institute of Human Genetics, Life and Brain Center, University of Bonn, Bonn, Germany; 5) Departments of Biology, Epidemiology, and Pediatrics, University of Iowa, Iowa City.

Non-syndromic oral cleft phenotypes, including cleft lip with or without cleft palate (CL/P), are complex disorders with some known genetic effects. Previous GWAS have identified multiple loci with small effects on risk of CL/P and recently we have identified a novel, potentially damaging variant in CDH1 in one multiplex CL/P family. For this analysis, we used whole exome sequence (WES) data in families with 2 or 3 distantly related (2nd or 3rd degree relationship) affected individuals to identify genes containing within-family shared rare variants. Fifty-five families containing 114 individuals, 4 duplicate subject controls and 2 unrelated CEPH HAPMAP controls were sequenced on the Illumina HiSeq 2500 and processed through the GATK pipeline. The families were from several different populations including 12 of Indian, 11 of Filipino, 19 of German, 10 of Syrian, 1 of European-American and 2 of Asian descent. Ingenuity Variant Analysis was used to identify variants of uncertain significance (VUS) that were predicted to be damaging. Definitively uncertain VUS were filtered out. We found 92 rare variants in 144 affected individuals in a family and where such sharing was observed in the same gene in at least two separate multiplex families and thus potentially associated with oral clefts. After filtering based on variant quality and frequency (MAF < 0.05 in HapMap), we identified one gene with variants that are homozygous in all affected individuals in at least two families. The variants were sometimes different across families, allowing for allelic heterogeneity. These rare variants are not present in either of the HapMap controls sequenced here. The genes are: ARHGEF14, FBXO32, RREB1, SNRPA3, STAR-D9, ZDHHC11, and ZNF835. Some of these genes have known biological functions which may be related to oral cleft. For example, variants in HSD3B7 have been associated with Hardikur syndrome, which sometimes presents with cleft lip and/or palate. Follow-up will include validation of these genes using Sanger sequencing and genotyping of these variants in other individuals in these same families.

Candidate gene analysis of non-syndromic tooth agenesis in Japenese. Y. Nishiyama1, K. Yasui2, T. Tatematsu3, H. Miyachi4, K. Shimozato5, M. Takenaka6, T. Nishiyama1, Y. Tokita1, 1) Oral and Maxillofacial Surgery, Toyota Memorial Hospital, Toyota City, Aichi, Japan; 2) Prognostic Research, School of Me; 4) Institute of Human Genetics, Life and Brain Center, University of Bonn, Bonn, Germany; 6) Perinatology, Institute for Developmental Research, Aichi Human Service Center, Kasugai, Japan. Tooth agenesis is one of the most common congenital anomalies in humans and it is characterized by developmental absence of teeth. According to the number of missing permanent tooth excluding the third molars, tooth agenesis is classified into two categories as follows. Hypodontia is generally defined as agenesis with absence of less than six teeth, and oligodontia is a condition in which six or more teeth are missing. It is well known that both genes including ethnic background and choosed the common ethnic groups in Japan in common with common defects. For example, oral cleft is ubiquitous disorder world wide, although its incidence varies widely from highest in Asian populations to lowest in those of African descent. Previously we reported the prevalence of permanent tooth agenesis in Japanese was 7.0%, which is similar to that in other population-based studies. Recent studies revealed that mutations in WNT10A gene are presented in 28 to 56% of tooth agenesis in Caucasian population. In this study, we investigated mutation analysis to determine the contribution of WNT10A and other candidate genes such as MSX1 and PAX9 variants. In total, 36 patients (range: 5 to 20 teeth were missing. The mean age of these patients was 19.4 years old. In 17 patients (44.7%), there was a positive family history (3rd degree relative) of tooth agenesis. WNT10A mutations were identified in 16.7% of cases with hypodontia. MSX1 and PAX9 mutations were presented in 21.2%, and 36.4% of cases, respectively. Our findings on the intra oral distribution of agenesis of permanent teeth in children may help us better understand the etiology of agenesis.
Chronic obstructive pulmonary disease (COPD) is characterized by an irreversible airflow limitation in response to inhalation of noxious stimuli, such as cigarette smoke. However, only 15-20% smokers manifest COPD, suggesting a role of genetic predisposition. Although genome-wide association studies have identified common genetic variants that are associated with the susceptibility to COPD, the odds ratios of these variants are much lower than those of familial form of emphysema, such as alpha 1 antitrypsin deficiency. We thus hypothesize that rare genetic variants contribute to the susceptibility to COPD. To test this hypothesis, we performed whole exome sequencing on 62 susceptible smokers with COPD (GOLD stage 3 or 4) and 30 resistant smokers with normal spirometries and without significant comorbidities. Overall we found a significantly higher mutation load in the genomic regions of the susceptible smokers than the resistant smokers. However, the difference was mostly accounted for by non-coding regions and the number of loss-of-function mutations is not significantly different between the two groups. We then used the Variant Annotation and Analysis Selection Tool (VAAST) to prioritize candidate disease-causing genes in susceptible smokers. We identified several hundred candidate genes that show higher prevalence of deleterious mutations in susceptible smokers. Among these candidates, we selected ~100 candidate genes with relatively high expression in human airway epithelial cells for function analysis using siRNA knockdown experiment. We identified several candidate genes that augment cigarette smoke extract-induced cytotoxicity in vitro. These potentially deleterious mutations of candidate genes may contribute to cigarette smoke-induced cytotoxicity, and potentially COPD.

**Posters: Complex Traits and Polygenic Disorders**

**746M Whole Exome Sequencing in Severe Chronic Obstructive Pulmonary Disease.** J. Xing, S. Bruseco, M. Moreau, Y. Bromberg, J. Yang, N. Wang, M. Piccirillo, J. Kliesney-Taiti, J. Zabner, J. Mac, S. Belinsky, T. Nyunoya. 1) Department of Genetics, Rutgers University, Piscataway, NJ; 2) COPD Program, Lovelace Respiratory Research Institute, Albuquerque, NM; 3) Department of Biochemistry and Microbiology, Rutgers University, Piscataway, NJ; 4) Department of Internal Medicine, University of New Mexico and New Mexico VA Health Care System, Albuquerque, NM; 5) Department of Medicine, University of Iowa, Iowa City, IA.

Chronic obstructive pulmonary disease (COPD) is characterized by an irreversible airflow limitation in response to inhalation of noxious stimuli, such as cigarette smoke. However, only 15-20% smokers manifest COPD, suggesting a role of genetic predisposition. Although genome-wide association studies have identified common genetic variants that are associated with the susceptibility to COPD, the odds ratios of these variants are much lower than those of familial form of emphysema, such as alpha 1 antitrypsin deficiency. We thus hypothesize that rare genetic variants contribute to the susceptibility to COPD. To test this hypothesis, we performed whole exome sequencing on 62 susceptible smokers with COPD (GOLD stage 3 or 4) and 30 resistant smokers with normal spirometries and without significant comorbidities. Overall we found a significantly higher mutation load in the genomic regions of the susceptible smokers than the resistant smokers. However, the difference was mostly accounted for by non-coding regions and the number of loss-of-function mutations is not significantly different between the two groups. We then used the Variant Annotation and Analysis Selection Tool (VAAST) to prioritize candidate disease-causing genes in susceptible smokers. We identified several hundred candidate genes that show higher prevalence of deleterious mutations in susceptible smokers. Among these candidates, we selected ~100 candidate genes with relatively high expression in human airway epithelial cells for function analysis using siRNA knockdown experiment. We identified several candidate genes that augment cigarette smoke extract-induced cytotoxicity in vitro. These potentially deleterious mutations of candidate genes may contribute to cigarette smoke-induced cytotoxicity, and potentially COPD.

**747T Autism spectrum disorders and dystrophinopathy in three non-identical twins.** D.P. Moreira, M. Lazar, K.M. Rocha, M. Aguena, G. Yamamoto, G.S. Kobayashi, M.S. Naslavsky, M.R. Passos-Bueno. 1) HUG-CELL, Dept of Genetics and Evolutionary Biology, Institute of Biosciences - University of Sao Paulo, Sao Paulo, SP, Brazil; 2) School of Public Health - University of Sao Paulo, Sao Paulo, SP, Brazil.

Autism spectrum disorders (ASD) are a genetically complex group of neurodevelopmental disorders that have been reported either as an isolated condition or in association with some syndromes. Among them, it has been shown that about 4% of patients with Duchenne/Becker muscular dystrophy (DMD/BMD) have ASD. Here we report three ASD-affected non-identical twins, of which two were also diagnosed with DMD. The ASD diagnosis was made according the DSM-IV-TR criteria. The two ASD-DMD patients harbor a 22-base-pair deletion involving exon 2 (c.64_69del [p.Ser22-Asp25]) in the dystrophin gene, which was maternally inherited. Exome sequencing was performed in the two ASD-DMD patients in order to verify if they harbor other pathogenic mutations that could contribute to the ASD phenotype. To filter for the possible pathogenic variants, we considered only stop codon and frameshift mutations in heterozygosis or homozygosis, with a minor allele frequency (MAF) ≤0.01 in the Exome Variant Server (http://evs.gs.washington.edu/EVS) and 1000 Genomes (http://www.1000genomes.org) databases, and present in ≤5% of a cohort of 136 individuals whose exomes were sequenced at the same laboratory. This filtering led to 10 potential pathogenic mutations. Out of these variants, we highlight two mutations which, in addition to the dystrophin mutation, might contribute to the ASD phenotype: a heterozygous stop codon mutation in OPALIN and a frameshift mutation in DPYSL4, both of which have not been found in 600 Brazilian controls. OPALIN codes for a transmembrane glycoprotein present in oligodendrocytes and is related to myelination. DPYSL4 plays regulatory roles in neuronal differentiation and death and in neurite outgrowth. Dystrophin is a cytoskeleton protein essential for neuronal survival, and has already been associated with neurodevelopmental disorders. However, the precise relationship between ASD and PC remains unknown. Only a small subset of individuals with CDKN2A mutations develops PC, thus, next-generation sequencing (NGS) may present a critical strategy for uncovering genes related to PC development in these families. Since PC patients with CDKN2A mutations have not been systematically assessed for mutations in other known high-risk PC genes, we conducted NGS to examine the known high-risk PC and/or hereditary pancreatic susceptibility genes (ATM, BRCA1, BRCA2, CDKN2A, PMS2, PALB2, PRSS1, STK11, EPCAM, TP53, MLH1, MSH2, MSH6, PM2, APC, VHL, XRC2C2, MEN1, PALLD, FANCC, CTRC, CPA1, CFTR, CASR, SPINK1). DNA was available for NGS on 29 PC patients from 21 American, Dutch, and Italian families with CDKN2A mutations and 12 Italian PC patients without CDKN2A mutations who had a personal or family history of melanoma. We interrogated exome data for rare deleterious (co-segregating) mutations in the 24 PC genes. Deleterious variants were defined as loss-of-function (frameshift, nonsense, in-frame deletions) or substitutions predicted to cause a deleterious amino acid change. Rare was defined as frequency <0.01 in publicly available exome databases. As quality control, we confirmed all CDKN2A mutations in the 29 CDKN2A+ PC patients plus a known BRCA2 frameshift mutation in a CDKN2A- PC patient. Rare, non-pathogenic variants were found in ATM (n=2), BRCA1 (n=1), STK11 (n=1), EPCAM (n=1), TP53 (n=1), MLH1 (n=2), MSH2, MSH6, PM2, APC, VHL, XRC2C2, MEN1, PALLD, FANCC, CTRC, CPA1, CFTR, CASR, SPINK1). CDKN2A variants were found in three American and two Dutch patients (p=0.028). All ATM variants included a novel frameshift mutation in CDKN2A- PC patients. The variants in CDKN2A- PC patients included a novel PMS2 frameshift variant, a MSH6 substitution (frequency=0.0001) somatically mutated in colon cancer, and a CFTR substitution (frequency=0.0003). The findings suggest that rare variants in known high-risk PC-related genes may contribute to PC risk in a subset of patients from melanoma-prone families with and without CDKN2A mutations.
749M

Whole Exome Sequencing of 75 Hereditary Prostate Cancer Families. E. Ostrander1, D. Karyadi1, E. Karlins1, B. Decker1, L. McIntosh2, S. McDonnell1, S. Middha1, D. Schaid1, S. Thibodeau1, J. Stanford2. 1) Cancer Genetics Branch, NHGRI, NIH, Bethesda, MD; 2) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 3) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 4) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN.

Rare inherited mutations are predicted to account for 5% to 10% of all prostate cancer (PCa) cases. However, few causative mutations have been identified. Whole-exome sequencing (WES) of large hereditary PCa (HPC) families offers a new strategy to address genetic heterogeneity in PCa in order to identify rare PCa susceptibility mutations with moderate to high penetrance. Only a small number of families segregating the same rare mutation are needed to highlight genes of potential importance in PCa susceptibility. We performed WES in 75 HPC families from the PROGRESS study. In total, 160 affected men were sequenced with one to six affected men per family. Median read depth was 64.5. GATK UnifiedGenotyper was utilized for SNV and indel detection. All samples were coordinately genotyped and genotype quality was set to 99.9% concordance with the SNP chip. After quality filters, 453,977 variants were identified and genotyped in the 75 HPC families. We searched for rare variants that segregate with prostate cancer in multiple families and are predicted to damage the encoded protein. Rare variants were defined as those with frequencies ≤ 2% in all populations from multiple databases: NHLBI ESP European American and African American, and the five 1000 Genome populations. Protein damaging variants were defined as either high-impact variants (stop gain/loss of frameshift or splice site acceptor and donor) or missense variants that are SIFT deleterious or PolyPhen probably/probably damaging. After population frequency and protein impact filters, 22,139 variants including 2,002 high impact variants remained. In order to determine the segregation with PCa, we linked the exome sequencing data to haplotype data from the Omni Express 700K SNP Chip. This approach allowed us to determine how many affected men in each family potentially carried the identified variants and was especially valuable in the 44 families with only one affected sequenced. After comparing the frequency in our Chinese cohort to the population frequencies, 861 variants were enriched in our set and present in three or more families, 110 of those variants were in six or more families and 89 were predicted to segregate in at least 50% of the cases on average. We have sequenced over 350 of the 861 variants for follow up validation and analysis in our population-based, case-control study of approximately 1,300 men and women. We selected over 350 of the 861 variants for follow up validation and analysis in our population-based, case-control study of approximately 1,300 cases and 1,150 controls.

750T

IDENTIFICATION OF GENETIC VARIANTS IN A CONSANGUINEOUS FAMILY WITH PSYCHOSIS USING AUTOZYGOSITY MAPPING AND WHOLE-EXOME NEXT GENERATION SEQUENCING. A. Al Amri1, J. Ivorra1, M. Ali2, C. Logan2, A. Cardno2, T. Mahmood2, S. Khan3, D. Wu4, M. Prochazkova5, T. McNicholl2, S. Craig2, C. Inglehearn3. 1) School of Biological Sciences, St. James’s University Hospital, Leeds, West Yorkshire, United Kingdom; 2) Leeds Institute of Biomedical & Clinical Sciences, University of Leeds (UK); 3) Leeds Institute of Health Sciences, University of Leeds (UK); 4) College of Medicine, Swansea University (UK); 5) Leeds and York Partnership NHS Foundation Trust (UK).

Psychosis is a condition in which an individual loses contact with reality. It covers profound forms of psychiatric disorders including schizophrenia, bipolar and schizoaffective disorder. These are complex disorders but have been known to run in families. The biological changes taking place in psychosis are not yet well understood but there is clear evidence of the involvement of both environmental and genetic factors. Here we describe the genetic analysis of a first-cousin consanguineous family of eight offspring, with two psychosis-affected cases, that was recruited from the West Yorkshire Pakistani population. Autozygosity mapping was combined with next generation sequencing to identify putative causative mutations. By using Affymetrix Genome-Wide SNP 6.0 arrays, two homozygous regions shared by the two affected siblings were identified on chromosomes 5q14.3-5q14.5 and 9q22.33-9q33.3. Whole-exome capture was carried out on one affected sibling using the SureSelect All Exon V4 reagent and sequenced on an Illumina HiSeq 2500 platform. The sequence reads were aligned against the human reference genome (hg18) and processed in SAM/BAM format using Picard and GATK. Annotar was used to annotate the variants before filtering according to depth of coverage, minor allele frequency, zygosity, mapping information and pathogenicity profiles. Four variants, located in the genes DFNB31, MUSK1, OR1J1 and NUP188, passed the filtering criteria and were identified as the best candidates in this family. Work to assess the implications of these mutations in psychosis is currently ongoing. Keywords: Psychosis, Autozygosity mapping, Schizophrenia.
Examining Associations Between Multiple Sclerosis Cognitive Impairment and Genes Previously Associated with Cognitive Decline in Other Disease, C. Holingue1, M. George1, C. Schaefer2, A. Bernstein3, L. Whitley1, L. Barcellos1,2. 1) Division of Epidemiology, Genetic Epidemiology and Genomics Laboratory, School of Public Health, University of California, Berkeley, CA, USA; 2) Kaiser Permanente Division of Research, Oakland, CA, USA; 3) Palm Drive Hospital, Sebastopol, CA, USA.

Cognitive impairment due to multiple sclerosis (MS) affects over half of individuals with the disease. Cognitive deficits can range from impairments in memory to diminished visual/spatial processing. These symptoms can affect work, educational function, employment status, and quality of life. Yet, there is great variability in the severity and type of impairment that individuals with MS experience. The current evidence regarding the risk factors for cognitive impairment is limited. A recent finding shows evidence for association between the ε4 variant of the APOE gene and exacerbated cognitive decline in MS cases (Shi et al. 2011), which is also implicated in Alzheimer’s disease. This suggests there may be associations between MS cognitive decline and polymorphisms in genes previously associated with neurodegeneration. The current study investigated the association between cognitive score in MS cases and polymorphisms in genes related to cognitive impairment. Cases were identified from Kaiser Permanente, Northern California Region (KPNC). A cognitive score was calculated using the TICS-M (Modified Telephone Interview for Cognitive Status) at study entry. All participants provided biospecimens for DNA extraction. DNA samples were genotyped using Illumina’s Human 660K BeadChip with imputation based on the 1000 Genome Reference and utilizing the IMPUTE2 software. To choose candidate genes, the NCBI gene database was used to search for genes associated with cognitive impairment with a variant associated with Alzheimer Disease, Parkinson’s Disease, Depression, Learning and Memory. Results were extracted from the “Genes to Cognition Online” website (www.g2conline.org). This resulted in a total of 51 candidate genes. A sample of 954 individuals was analyzed. Top-performing variants were prioritized with each SNP in the one of the above genes as the independent variable and cognitive score as the dependent variable. The top two principal components, derived from EIGENSTRAT SMARTPCA were included as covariates in the model, in order to account for possible confounding by ancestry. No evidence for association was observed for SNPs in these genes. This is the first study to investigate cognitive impairment in MS in relation to a broad set of genes previously associated with neurodegeneration or depression.

Genetic contribution to cerebral palsy, G. McMichael1, M.N. Bainbridge2, E. Haan3, M. Corbett4, A. Gardner1,4, S. Thompson5, B.W.M. van Bon6, C.L. van Eyk7, J. Broadbent8, C. Reynolds1, M.E. O’Callaghan1, L.G. Best5,7, L.S. Nguyen4, D.L. Adelson5, R. Russo6, S. Jhangiani1,4, M. Corbett1,4, E. Muzny2, R.A. Gibbs2, A.H. MacLennan1, J. Gecc1,4,5,6,7. 1) The Robinson Institute, University of Adelaide, Adelaide, Australia; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, USA; 3) South Australian Clinical Genetics Service, SA Pathology (and Children’s Hospital), North Adelaide, Australia; 4) School of Pediatrics and Reproductive Health, The University of Adelaide, Adelaide, Australia; 5) Department of Pediatric Neurology, Women’s and Children’s Hospital, North Adelaide, Australia; 6) Department of Human Genetics, School of Medical Science, The University of Adelaide, Adelaide, Australia; 7) School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia; 8) Department of Pediatric Rehabilitation, Women’s and Children’s Hospital, North Adelaide, Australia.

Cerebral palsy (CP) is the most frequent cause of physical disability in childhood with a prevalence of 2.2-5.5 per 100 live births. It is largely sporadic and causally heterogeneous. Evidence of intrapartum fetal compromise is found at birth in <10% of cases. Several known epidemiological risk factors have been identified for CP including preterm delivery, intrauterine growth restriction and intrauterine infection. The contribution of genetic factors to CP is frequently overlooked. We performed whole exome sequencing of 183 cases with CP and their parents when available (88 trio and 67 duo cases). We used multiple prioritization criteria to identify potentially causative variants including the type of mutation (e.g. protein truncating), involvement of a known disease gene, haploinsufficiency index, brain expression, Residual Variation Intolerance Score and Combined Annotation-Dependent Depletion. Based on these strict prioritization criteria, 10 de novo mutations were predicted to be causative for CP; three in known disease genes; TUBA1A (n=2), SCN8A (n=1) and KDMSC (n=1) and six in novel candidate CP genes: AGAP1, JHDM1D, MAST1, NAA35, RFX2 and WIP1. Additionally, we found four inherited from an unaffected mother to affected son X-chromosome variants, in two known disease genes L1CAM and PAK3 and two novel candidate CP genes CD99L2 and ODZ1 respectively. In summary, 14% of CP cases have a predicted pathogenic disease variant with 8 occurring in novel candidates for CP. The genetic heterogeneity of CP revealed by these results reflects the complex and clinically variable nature of this disorder. These results highlight a previously unappreciated role for the contribution of genetics to CP causation, which will influence future management, especially genetic counselling.
755M Testing the effect of compound heterozygosity on anthropometric traits in the general population. S. Lessard1, P.L. Auer2, A. Grif2, C. Schumacher2, J. Knekt3, A. Sarna4, H. Yaghootkar5, B. Hagberg6, T. Edvardsson7, T. Frayling1, R.J.F. Loos1, G. Leitge1, GIANT ExomeChip Consortium. 1) Montreal Heart Institute and Université de Montréal, Montreal, Quebec, Canada; 2) University of Wisconsin, Milwaukee, Wisconsin, USA; 3) Vanderbilt University, Nashville, USA; 4) Icahn School of Medicine at Mount Sinai Hospital, New York, New York, USA; 5) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; 6) Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge 02142, MA, USA; 7) Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK.

Background: In Mendelian genetics, a recessive disease can result when two different recessive alleles in the same gene are inherited in trans, that is on each copy of the gene. This genetic mechanism - termed compound heterozygosity - is observed for several monogenic disorders, such as pterygium, rickets and sickle cell disease. However, the role of compound heterozygosity in human complex diseases and traits has not been explored extensively. Method: We analyzed height, body mass index (BMI) and waist-to-hip ratio (WHR) in up to 88,000 individuals from 9 studies included in the GIANT ExomeChip Consortium. We applied inverse normal transformation on phenotypes after adjusting for sex and age. All participants were genotyped using the Illumina ExomeChip. We defined compound heterozygotes as individuals that carry at least two loss-of-function variants in the same gene, assuming that such variants are rare and not inherited in cis. We compared the phenotypic mean in compound heterozygotes to the mean in the remaining individuals, and assess statistical significance using phenotype-permutations. We combined results across studies using a weighted Z-score meta-analysis method, where the weight is the sample size of the corresponding study. Results: By considering 5,624 nonsynonymous and 10,578 splice site variants genotyped on the ExomeChip, we identified at least one compound heterozygote for 2,911 genes for height, 2,664 genes for BMI and 1,795 genes for WHR. No genes reached a Bonferroni-corrected threshold for statistical significance (P<6x10^-8). However, there was a strong signal for the cyclin D1 gene with height (CCND1 P=3.9x10^-3) and the parathyroid hormone receptor gene (PTH1R, P=5.9x10^-3). In both cases, however, the signal was due to a common variant and reflects a "true" recessive signal. The strongest compound heterozygosity signal observed is for BMI with the taste receptor gene TAS1R1 (P=5.1x10^-4), five nonsense variants suggested. This is a very low frequency variant and compound heterozygotes across 40,210 patients. Conclusion: The role of compound heterozygosity has rarely been tested in the context of complex human diseases and traits. Here, we successfully developed a method of analysing this model in large meta-analysis data. As with other studies of predominantly low-frequency and rare variants, our data suggest that very large sample sizes will be needed to detect robust associations.

756T Locating new genes which may be involved in the development of Primary Congenital Glaucoma. D. Bercovich1, A. Wolf2, O. Geyer2, 1) Human Molec Gen & Pharm, Tel Hai College, Galilee Eylon, Israel; 2) Department of Ophthalmology, Carmel Medical Center, Israel.

Since there are Primary Congenital Glaucoma patients (PCG) who do not contain mutations in genes known to be involved in this disease (like the CYP1B1, MYOC & FOKC1 genes), prenatal screening and diagnosis in fetuses suspected with this disease is not viable. Furthermore, after birth, it is difficult to accurately identify primary Congenital Glaucoma, in many cases it is miss-diagnosed as another syndrome associated with eye diseases. The main problem is the Jewish population, in which individuals with significant mutations in these genes is rare. In order to locate new loci with point mutations or possible copy number variations, 6 PCG Jewish patients from different ethnic groups (Ashkenazi, Sfaradi, Balkan Jews) and their parents pairs, to combine this data with their patients NGS (a total of 10 DNAs) were analyzed. 87 different genes were found be possible candidates in the recessive model. All DNA alterations were found to be in the same way in the parents as well which exclude them as been the pathogenic DNA alterations. One locus with CMA was found to be varied among Congenital Glaucoma patients and family members. This indicated that copy number around this gene could contribute somehow to the Congenital Glaucoma phenotype and more investigation may be done and this possible connection of this gene.

757S Genome-wide Copy Number Variants in Chronic Obstructive Pulmonary Disease (COPD). F. Begum1, I. Ruczinski5, S. L. Forrest, M. Parker1, J. Hetmanski1, T. Beatty6, E. Silverman7, J. Crapo2, COPDGene Investigators. 1) Department of Epidemiology, Johns Hopkins University, Baltimore, MA, USA; 2) Department of Biostatistics, Johns Hopkins University, Baltimore, MA, USA; 3) Cancer Genomics Research Laboratory (CGR), Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; 4) Channing Division of Network Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, USA; 5) Department of Medicine, National Jewish Health, Denver, USA.

Chronic Obstructive Pulmonary Disease (COPD), a common lung disease, is the third leading cause of death in the United States. This chronic and progressive disease has both environmental (mainly cigarette smoking) and genetic risk factors. Several genome-wide association studies (GWAS) have identified multiple genes influencing risk to COPD including CHRNA3, FAM13A, HHIP, RIN3, MMP12, and TJB2. To properly understand the genetic etiology of COPD, it is also important to explore the role of copy number variants (CNVs) since the presence of CNVs can alter gene expression and may be causal for disease. We delineated CNVs using PennCNV on 9076 COPDGene study subjects using genome-wide marker data generated using Illumina’s Omni-Express array. COPDGene subjects are comprised of one-third African-Americans and two-thirds Non-Hispanic white adult smokers, with or without COPD. After employing rigorous quality control procedures to reduce the false positive CNV calls, we tested for association between CNV components (defined as disjoint intervals of copy number regions within race) and several COPD related phenotypes. We detected hemizygous deletions that achieved genome-wide significance on chromosome 3q35.2, near the gene FAM153B, in tests of association with total lung capacity assessed by chest CT among African-Americans. This region includes multiple reported CNVs, and we are currently following up on this signal.

758M Investigating the role of salivary amylase copy number in obesity using low-pass whole genome sequencing. M.A. Yuke1, A.R. Wood1, L. Harris1, L. Boquete-Vilarino1, M. Nalls2, I.O.T. Hetmanski3, T. Beatty4, E. Silverman7, J. Crapo2, S. Bandinelli3, A. Singleton5, D. Meizer6, L. Ferrucci1, M.N. Weedon1, T.M. Frayling1, 1) Genetics of Complex Traits, University of Exeter Medical School, Exeter, UK; 2) Laboratory of Neurogenetics, National Institute of Aging, Bethesda, Maryland, USA; 3) Department of Molecular Neuroscience and Reta Lila Laboratories, Institute of Neurology, UCL, London, United Kingdom; 4) Tuscany Regional Health Agency, Florence, Italy, I.O.T. and Department of Medical and Surgical Critical Care, University of Florence, Florence, Italy; 5) Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Barrack Road, Exeter, UK; 7) Longitudinal Studies Section, Clinical Research Branch, Gerontology Research Center, National Institute on Aging, Baltimore, Maryland, USA.

Copy number variation at the salivary amylase gene (AMY1) has recently been strongly associated with body mass index (BMI) and obesity (Falchi et al. Nature Genetics, 2014). Each additional copy of the AMY1 gene was associated with a 2-4% increase in BMI. However, the number of obese individuals and the variant was reported as explaining a larger proportion of the variance in BMI than the variant at the FTO/IRX3 locus. Here, we aimed to further replicate and characterise this association in the InCHIANTI study. We performed whole genome sequencing on 657 samples to an average depth of 7x. We aligned the reads to a repeat masked GRCh37 31 reference genome using the msFAST alignment algorithm that can align single reads to multiple positions in the genome and so is optimal for regions of low-pass whole genome sequencing. We identified multiple CNVs that were derived an absolute copy number value for almost the full amylase region using mrcNaVaR. We then regressed the resulting copy number against body mass index corrected for age and sex. We had between 88% to >99.9 power to detect association when assuming the published variance explained of between 1.73% and 7.94%. The preliminary analysis did not provide any evidence of an association between salivary Amylase copy number and BMI (P = 0.457). However, our work is ongoing and we are validating the association using quantitative PCR and expanding the analysis to several thousand individuals. In conclusion, our preliminary results fail to support the association between salivary amylase gene copy number.
further phenotype analysis of individuals with specific variants of interest. One initial analysis focused on identifying novel variants predicted to be damaging in genes previously implicated with type 2 diabetes. These variants are currently being genotyped in additional samples (N=7,355) for association analysis with type 2 diabetes. Using this approach, we identified 6 novel rare variants with an OR of 1.057 (OR=1.057; 95% confidence interval 0.98-1.11) in the Ala1369Ser gene that encodes the sulfonylurea receptor 1 protein (SUR1). SUR1, together with inward-rectifier potassium ion channel encoded by the KCNJ11 gene, regulate insulin secretion by ATP/ADP-sensing. Rare activating mutations in SUR1 cause neonatal diabetes while rare inactivating mutations cause hyperinsulinemia of infancy. The common SNP rs757110 has been associated with type 2 diabetes in some studies, but in our sample had no association (minor allele frequency 0.39; odds ratio (OR)= 0.99 [95% confidence interval 0.88-1.11]; P= 0.9 adjusted for age, sex, birth year, fraction of Pima heritage). Of the 6 novel variants, 4 have been genotyped in follow-up samples to date (Arg1420His, Gly1316Gln, Met801Ile, and Ser165Leu) while 2 are pending genotyping (Asp691Glu and Lys1565Ala). Among variants with follow-up genotypic data, Arg1420His provided the most compelling evidence for a role in type 2 diabetes. Individuals with a His allele (243 heterozygotes and 1 homozygote; overall frequency = 0.02 in Pima Indians) were at higher risk for type 2 diabetes (OR= 1.86, 95% CI=1.4 - 2.4; P= 8x10^-10 adjusted as above) despite being 24 years younger on average (Arg1420His n=600, mean age=43, SD= 8 vs. 0.002 adjusted for sex, age, birth year, and fraction of Pima). The one homozygous individual for the risk allele (His/His) had been diagnosed with hyperinsulinemia and hypoglycemia at 4 months of age, and was diagnosed with type 2 diabetes at 2 years of age, suggesting a very early age of onset for this type of variant in homozygosity. Arg1420 is located in the second nucleotide binding domain of SUR1 and therefore a substitution at this position is predicted to have a functional impact. In conclusion, whole genome sequencing identified potential functional variants in ABC8 that appears to contribute to type 2 diabetes in Pima Indians.

Targeted sequencing of genes associated with type 2 diabetes in 6800 individuals. V. Bansal1, 2, J. Gassenhuber1, 5, J. Johnson1, 3, S. Santoricco1, 3, J. Baier1, 4, Y. Muller1, K. Huang1, N. Villarasa1, R. Tisch1, E. J. Topol1, B. O. Boehm3, 4, 5, 1 Scripps Translational Science Institute, La Jolla, CA; 2 Department of Pediatrics, University of California San Diego, La Jolla, CA; 3 Department of Internal Medicine, Division of Endocrinology and Diabetes, University of Michigan, Ann Arbor, MI; 4, 5 University of Ulm, Germany; 4, 5 Nanon-Aventis Germany GmbH, Frankfurt am Main, Germany. Abstract: In high-throughput sequencing studies, it is not rare to encounter unprecedented opportunities for understanding the role of rare variants in complex diseases. For type 2 diabetes, more than 70 common genetic variants have been identified by extensive genome-wide association studies. We utilized a candidate gene approach to search for rare variants associated with type 2 diabetes and focused on genes within loci identified by GWAS for T2D and related traits. In our previous work using the same approach, we observed a significant excess of variants with linkage to the region and three families (n=11 individuals) with no evidence of linkage to the same region. Among the 267 biallelic variants in the linked families, 34 (12.7%) were possibly damaging (PolyPhen>0.80) and 72 had a CADD score >10. In the unlinked families 332 biallelic variants were identified, with 48 variants (14.5%) having a PolyPhen>0.80 and 92 variants with a CADD>10. The minor allele frequency of deleterious variants appears to be higher in our linked families, with 25% having an alternate allele frequency (AF) of 0.127 or below compared to 25% with an AF of 0.078 or below in the unlinked families. Among the linked families 6 variants (2.2%) were private and were not found in the unlinked families or 1000 Genomes. We also describe the distribution of variants in this same linkage region in Mexican, Japanese and African American families sequenced under the same project and compared to 1000 Genomes. Summary: Overall, these findings may partially explain the genetic predisposition to type 2 diabetes in these populations. The current evidence of linkage to type 2 diabetes provides the most compelling evidence for having a role in type 2 diabetes. Individuals with a His allele (243 heterozygotes and 1 homozygote; overall frequency = 0.02 in Pima Indians) were at higher risk for type 2 diabetes (OR= 1.86, 95% CI=1.4 - 2.4; P= 8x10^-10 adjusted as above) despite being 24 years younger on average (Arg1420His n=600, mean age=43, SD= 8 vs. 0.002 adjusted for sex, age, birth year, and fraction of Pima). The one homozygous individual for the risk allele (His/His) had been diagnosed with hyperinsulinemia and hypoglycemia at 4 months of age, and was diagnosed with type 2 diabetes at 2 years of age, suggesting a very early age of onset for this type of variant in homozygosity.

Characterizing variation under linkage peaks in families. K.L. Edwards1, J.Y. Wan1, C. Johnson1, S. Santoricco1, 3, 1 Epidemiology, University of California Irvine, Irvine, CA; 2 University of Colorado, Denver Denver, CO. Abstract: The Metabolic Syndrome (MetS) is a complex condition characterized by a cluster of CVD risk factors, including obesity, lipid abnormalities, hypertension and glucose intolerance. In our previous work using extended pedigree and linkage analysis, we identified three clusters of MetS with evidence for linkage to multivariate traits defined by clusters of MetS risk factors. Methods: Whole exome deep sequencing was performed in a subset of linked families of European, Mexican, African and Japanese-American descent. Results: We provide a descriptive summary of rare and common variation found under one of the linkage peaks on chromosome 2. Results are based on sequencing 5 European American families consisting of two families (n=11 individuals) with linkage to the region and three families (n=17 individuals) with no evidence of linkage to the same region. Among the 267 biallelic variants in the linked families, 34 (12.7%) were possibly damaging (PolyPhen>0.80) and 72 had a CADD score >10. In the unlinked families 332 biallelic variants were identified, with 48 variants (14.5%) having a PolyPhen>0.80 and 92 variants with a CADD>10. The minor allele frequency of deleterious variants appears to be higher in our linked families, with 25% having an alternate allele frequency (AF) of 0.127 or below compared to 25% with an AF of 0.078 or below in the unlinked families. Among the linked families 6 variants (2.2%) were private and were not found in the unlinked families or 1000 Genomes. We also describe the distribution of variants in this same linkage region in Mexican, Japanese and African American families sequenced under the same project and compared to 1000 Genomes. Summary: Overall, these findings may partially explain the genetic predisposition to type 2 diabetes in these populations.

An exome-wide sequencing study for type 2 diabetes-associated kidney disease in African Americans. M. Guan1, 2, P. Mudgar1, J.G. Wilson3, B. Friedman4, D.W. Bowden1, 3, 5, M.C. Y. Wang1, 3, 6, 5, 6 End-stage kidney disease (ESKD) is a significant worldwide public health concern; diabetes accounts for 44% of incident ESKD cases in the U.S. Previous GWAS from our group have identified several strongly associated SNPs for type 2 diabetic (T2D)-ESKD. However, the common SNPs identified in GWAS have modest effects and cumulatively explain only a small proportion of disease variance. Therefore, we evaluated the contribution of low frequency coding variants in T2D patients on dialysis or with high risk for ESKD. Methods: Whole exome sequencing studies were performed on 537 T2D-ESKD cases and 1,028 non-T2D, non-ESKD controls from the T2D-GENES Consortium exome sequencing study. Single variant association analysis and gene-based association analysis of rare variants were performed using the Sequence Kernel Association Test (SKAT) was used for gene-based analyses. Results: One initial analysis focused on identifying novel mutations that affect risk for type 2 diabetes in American Indians. In this study, we developed computational methods for variant calling, assessing population stratification, and rare variant association analysis. Single variant association analysis and gene-based association analysis of rare variants using a number of statistical methods (CAST, Calpha, and SKAT) did not identify statistically significant genes or variants. To improve power, we sequenced an additional 2100 type 2 diabetes cases and 1050 controls using the same approach. Joint analysis of rare variants in the two cohorts identified 10 variants (minor allele frequency in the 0.1-0.5% range) that were associated with type 2 diabetes at 3 years of age, suggesting a very early age of onset for this type of variant in homozygosity. Arg1420 is located in the second nucleotide binding domain of SUR1 and therefore a substitution at this position is predicted to have a functional impact. In conclusion, whole genome sequencing identified potential functional variants in ABC8 that appears to contribute to type 2 diabetes in Pima Indians.
Exome chip meta-analysis identifies novel loci and low-frequency variants contributing to central body fat distribution. A.E. Justice, H.M. Highland, K.L. Young, M. Graf1, T. Karadzic1, N.L. Heard-Costa2, D. Pasko3, V. Turcot1, Y. Liu1, L. Southam5, L.A. Cupples10, C.T. Liu10, C.S. Fox3, T.W. Winkler11, N. Graup2, R.A. Scott12, M.M. McCarthy14, K. Mohlke15, R.J.F. Loos6, I. Borecki16, K.E. North1, C.M. Lindgren14 For the BBMRI, the GDT20, the CHARGE, and the GIANT Consortia. 1) Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA; 2) Human Genetics Center, The University of Texas Graduate School of Biomedical Sciences at Houston, The University of Texas Health Science Center at Houston, School of Public Health, Houston, TX, USA; 3) Welcome Trust Centre for Human Genetics, University of Oxford, UK; 4) National Heart, Lung, and Blood Institute, the Framingham Heart Study, Framingham, MA, USA; 5) Department of Neurology, Boston University School of Medicine, Boston, MA, USA; 6) Genetics of Complex Traits, University of Exeter, UK; 7) Montreal Heart Institute, University of Montreal, Canada; 8) The Genetics of Obesity and Related Metabolic Traits Program, The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 9) The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 10) Boston University School of Public Health, Boston, MA; 11) Department of Genetic Epidemiology, Institute of Epidemiology and Preventive Medicine, University of Cambridge, UK; 12) The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; 13) MRC Epidemiology Unit, University of Cambridge, UK; 14) Oxford Centre for Diabetes Endocrinology and Metabolism, University of Oxford, UK; 15) Department of Genetics, Uniform of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 16) Department of Medical Genetics, Sapienza University of Rome, Rome, Italy; 17) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA.

Increased central fat is a leading predictor of greater cardiometabolic risk. Recent work shows that genetic factors contribute to the distribution of central fat, measured here as waist to hip ratio adjusted for BMI (WHRR), which is correlated with visceral adiposity (r=0.56). Also, among the loci with known association with WHRR+ half-exomes that are enriched for WHRR, more than half show marked sexual dimorphism. The genetic underpinnings of WHRR are complex traits and whole blood cells. Methylation profiles also revealed evidence for methylation quantitative trait locus effects for our novel variant in an established locus (P=4.38×10-7) known to be associated with substantially raised FT4 levels. We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335). We then undertook analysis of rare variants (MAF <1%) using sequence kernel association testing (SKAT) in 40 candidate genes and performed genome-wide complex trait analyses (GCTA) to explore the extent that common SNPs (MAF >1%) explained the variance in TSH and FT4 (N=16,335).
767M
Combining linkage analysis and whole-exome sequencing for the identification of novel ADHD-related variants in multi-generation pedigrees.

I. Corominas-Galbany1, M. Klein1, M. Oinnink1,2, S. Gross-Lesch3, C. Jacob3, A. Reif1, M. Romanos1, B. Franke1,2, K.P. Lesch3,1, Department of Human Genetics, Radboud university medical center, Nijmegen, The Netherlands; 2) Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud university medical center, Nijmegen, The Netherlands; 3) Department of Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany; 4) Department of Child Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany.

Attention-Deficit/Hyperactivity Disorder (ADHD) is a highly heritable and multifactorial disorder characterized by inattention, and/or motor hyperactivity and impulsivity. Susceptibility loci for ADHD have been investigated in the recent years and valuable candidate genes have been suggested. Nevertheless, the understanding of the genetic basis of ADHD is still a challenge. Large multi-generation pedigrees with affected individuals allowed the identification of haplotype blocks shared among all ADHD-affected family members. The characterization of these region is essential for the identification of genes and variants that may be related to the ADHD phenotype. Hence, the main goal of this study is to combine linkage analysis with whole-exome sequencing (WES) in order to identify novel genetic ADHD risk factors. This study was performed with nine ADHD families of German descent, for which WES data was obtained from 3 or more affected individuals to detect rare and common genetic variants shared among all sequenced members of a family. WES was carried out using Agilent SureSelect All Exon 50Mb Target Enrichment kit and single-end sequencing on the 550x1 platform, and variants were genotyped on GeneChip Human Mapping 50K Array Hind240 and SNP data was used to fine-map the ADHD-associated loci previously described. Rare and common candidate variants to be associated with ADHD were validated using Sanger sequencing. ADHD phenotype and haplotype co-segregation in large pedigrees points towards a dominant inheritance pattern of the disease. An initial genome-wide approach focused on one family allowed the identification of 7 rare variants, but none of them segregated with phenotype through the family. This result further supports the inclusion of linkage analysis data, because segregation of ADHD does not have to be due to rare variants only. In fact the combination of several common and rare variants, each with small or intermediate effect, may underlie co-segregating haplotype blocks. Linkage analyses performed in this study identified an haplotype co-segregating with ADHD-affected individuals at 9q31-9q32 that lead to the detection of several common and rare variants of interest in distinct genes. Validation of these variants and segregation analysis through the family are ongoing. This study uses a complementary strategy in addition to GWAS approaches to detect variants that may shed new light on our understanding of ADHD’s genetic basis.

767T
Gene-variants associated with familial mesial temporal lobe epilepsy identified by whole exome sequencing.

R. Secolini1, P.A.O. Ribeiro1, F.R. Torres1, M.G. Borges1, A.C. Coan1, M.E. Morita1, C.V. Soler1, M.L. Santos1, C.V. Maurer-Morell1, B.S. Carvalho1, F. Cendes1, I. Lopes-Cendes1, Department of Medical Genetics, University of Campinas - UNICAMP, Campinas, SP, Brazil; 2) Department of Neurology, University of Campinas - UNICAMP, Campinas, SP, Brazil.

Background: Epilepsy is a chronic common neurological disorder that affects approximately 1% of the population worldwide. Familial mesial temporal lobe epilepsy (MTLE) is a clinically well characterized syndrome with an autosomal dominant inheritance. Objective: To identify genes associated with MTLE in two large families. Methods: Whole exome sequencing was performed in eight individuals from two MTLE families (F-10 and F-26), including six patients and two unaffected parents. Exome was targeted with Nextera Rapid Capture Expanded Exome kit ( Illumina® ) and sequenced in a high-performance Hiseq Illumina 2500 sequencing machine ( Illumina® ) to obtain more than 50X average coverage per sample. A bioinformatics analysis was performed using the GATK software package. Sequences were aligned using BWA algorithm. Variant calling and functional prediction was performed using VariantAnnotator and SnpEff tools. We prioritized non-synonymous, frameshift, splicing, and indel variants according to novelty, quality score, and putative pathogenicity. Results: We found a total 1,955,506 and 2,134,863 variants in F-10 and F-26 families, respectively. After bioinformatics processing, we identified 184 functional variants in F-10 and 193 in F-26, which are shared only by patients and absent in unaffected individuals. Among them, we observed three genes as potential candidates for MTLE in the families studied. Conclusions: Putative roles related to formation of axon connections, protein-protein interaction, phosphorylation-dependent ubiquitination pathway may be the three candidate genes identified relevant for MTLE. Supported by: CEPI-D-BRAINN FAPESP, São Paulo, Brazil.

766S
Mutations in Human Capicua Gene Found in Patients with CFD and NTDs.

Y. Lei1, H. Zhu1, M. Parker1, W. Yang2, G. Shaw2, R. Finnell1,2, 1) Dell Pediatric Research Institute, Department of Nutritional Sciences, The University of Texas at Austin, Austin, Texas; 2) Department of Chemistry, College of Natural Sciences, The University of Texas at Austin, Austin, Texas.

Background: Cerebral folate deficiency (CFD) syndrome is characterized by very low concentration of 5-methyltetrahydrofolate (5-MTHF) in cerebrospinal fluid, while folate levels in plasma and red blood cells are normal. Previously, mutations in several folate pathway genes, including hFRα (folate receptor alpha), DHFR (dihydrofolate reductase), and PCFT (proton coupled folate transporter) have been identified in CFD patients. In an effort to identify causal mutations for CFD, we performed whole exome sequencing analysis of DNA samples collected from a CFD patient, her healthy sibling, and her biological parents. A de novo mutation in human Capicua gene (CIC), c.1057C>T (p.R353X), was identified in the patient. The results were confirmed using Sanger sequencing. In addition, a missense mutation predicted to be damaging, c.1738G>GT (p.G580 GC) was identified in another CFD patient. The characterization of these region is essential for the identification of genes and variants that may be related to the CFD phenotype. Nevertheless, the understanding of the genetic basis of CFD is still a challenge. Large multi-generation pedigrees with affected individuals allowed the identification of haplotype blocks shared among all CFD-affected family members. The characterization of these region is essential for the identification of genes and variants that may be related to the CFD phenotype. Hence, the main goal of this study is to combine linkage analysis with whole-exome sequencing (WES) in order to identify novel genetic CFD risk factors. This study was performed with nine CFD families of German descent, for which WES data was obtained from 3 or more affected individuals to detect rare and common genetic variants shared among all sequenced members of a family. WES was carried out using Agilent SureSelect All Exon 50Mb Target Enrichment kit and single-end sequencing on the 550x1 platform, and variants were genotyped on GeneChip Human Mapping 50K Array Hind240 and SNP data was used to fine-map the CFD- associated loci previously described. Rare and common candidate variants to be associated with CFD were validated using Sanger sequencing. CFD phenotype and haplotype co-segregation in large pedigrees points towards a dominant inheritance pattern of the disease. An initial genome-wide approach focused on one family allowed the identification of 7 rare variants, but none of them segregated with phenotype through the family. This result further supports the inclusion of linkage analysis data, because segregation of CFD does not have to be due to rare variants only. In fact the combination of several common and rare variants, each with small or intermediate effect, may underlie co-segregating haplotype blocks. Linkage analyses performed in this study identified an haplotype co-segregating with CFD-affected individuals at 9q31-9q32 that lead to the detection of several common and rare variants of interest in distinct genes. Validation of these variants and segregation analysis through the family are ongoing. This study uses a complementary strategy in addition to GWAS approaches to detect variants that may shed new light on our understanding of CFD's genetic basis.
768T Targeted exome sequencing in extended pedigrees with type 2 diabetes identifies a novel diabetic nephropathy susceptibility gene. M.G. Pozzolotti1,2, A.M. Smiles1, J.C. Mychaleckyj1, R. Magi1, M. Mangino1,2, J.J. Warram3, A.S. Krolevsaki1,2, J.H. Warram1, S.S. Rich3, R. Magi1, M. Mangino1,2, J.J. Warram3, A.S. Krolevsaki1,2, J.H. Warram1, S.S. Rich3, C.M. Lindgren2,6,1), Estonia; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 3) Department of Biotostatistics, University of Liverpool, Liverpool, UK; 4) Department of Obstetrics and Gynaecology, University of Tartu, Tartu, Estonia; 5) Institute of Bio- and Translational Medicine, University of Tartu, Tartu, Estonia; 6) Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA.

Genome-wide linkage analysis in extended pedigrees from the Joslin Study of Genetics of Nephropathy in Type 2 Diabetes Family Collection (Joslin T2DN Family collection) previously identified several genomic regions with evidence of linkage for urinary albumin excretion levels or variation in renal function in individuals with type 2 diabetes (T2D). Several in independent studies have reported evidence of linkage to many of these same regions, however, to date no gene located within or near these loci that contributes to these diabetic nephropathy (DN) phenotypes has been identified.

To advance these efforts, we recently performed targeted exome sequencing of all protein-coding genes across 6 loci (4 genomic regions with evidence for linkage with urinary albumin excretion levels: chromosomes 5q, 7q, 21p, and 22q; and 2 genomic regions linked to variation in renal function: chromosomes 2q and 7p) using a custom target enrichment library followed by Next-Generation sequencing. A total of 662 members from 49 extended pedigrees in the Joslin T2DN Family collection were resequenced for the coding regions of the 361 genes across the 6 linkage regions. Following quality control analyses, multi-sample variant calling was performed for 603 samples that exceeded 35X on-target read depth and with greater than 15-fold sequence coverage across 70% of the target regions. In total, more than 6,000 non-reference variants were identified, including > 1,200 missense and 100 nonsense variants. Using data from the NHLBI Exome Sequencing Project reference panel, rare functional variants were selected from among these variants and segregation analysis was performed in families contributing evidence of linkage to each of the 6 linked genomic regions. Analyses were performed to date have identified 3 rare variants in the prion rich 14-like (PRR14L) gene that segregate in DN cases from 4 families with evidence of linkage on chromosome 22. These data suggest that rare variants in PRR14L account for the linkage peak identified at this locus.

769S Mis-matches between adiposity and metabolic traits: A replicated genomewide association study for metabolic disparity. L.J. Corbin1, K. Burrows1, M. Mangina2, SM. Ring3, NJ. Timpson4. 1) MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, United Kingdom; 2) DTR Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom; 3) School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom.

While increased Body Mass Index (BMI) is correlated with a range of cardiometabolic risk factors, two sub-phenotypes of obesity have been characterised that contradict this. Metabolically obese normal-weight (MONW) individuals present with a normal BMI but have a poor metabolic profile. In contrast, metabolically healthy obese (MHO) individuals have high BMI, but do not show major metabolic disturbances expected. By developing a novel phenotypic outcome acting as an indicator of metabolic disparity, we explore the biological pathways underlying these sub-phenotypes using genomewide association analyses in two independent cohorts. Analyses were performed on a collection from the Avon Longitudinal Study of Parents and Children (ALSPAC, mean age 15.5yrs) and in a replication cohort of adults from the UK Adult Twin Registry (TwinsUK, mean age 50.9yrs). A compound phenotype was generated by subtracting standardised BMI from standardised fasting plasma glucose to yield a continuously distributed variable where positive values were indicative of “healthier” adiposity status relative to glucose, zero was indicative of matching adiposity/glucose status and negative values were indicative of “healthier” glucose status relative to adiposity. Genomewide analysis of this trait in ALSPAC (n=2564) gave evidence of association at the G6PC2 locus where each additional T allele at the well-known variant rs560887 was related to a -0.25(SE 0.04, p=3.9e-10) change in standardised disparity score. This was replicated in TwinsUK (n=2599): -0.11(SE 0.04, p=0.003) and a meta-analysis gave a joint p-value of 6.1e-11. An additional signal was also seen when using DEXA derived fat mass in place of BMI (meta-analysis p=4.4e-10) and, as anticipated, with an inverted beta coefficient when replacing glucose status with HOMA_b. This work illustrates how derived phenotypes can be used as indicators of disparity between cardiometabolic risk factors and glycaemic traits. Using a standardised disparity score allows a powerful analysis of metabolic features avoiding the costly procedure of defining MONW and MHO by threshold. Results are in line with the carriage of additional T alleles at G6PC2(rs560887) being associated with improved glycaemic profile. Work here did not yield signals of variants showing independent contributions to adiposity; this likely the result of genetic architecture and the causal effect that BMI has on many downstream causal effects.

770M Association analysis of exome chip data of Polycystic Ovary Syndrome in Estonian Biobank. R. Magi1, A.P. Morris1,2, T. Karaderi1, T. Laasko1, T. Tammiste1, A. Metspalu1, A. Salumets3,4, C.M. Lindgren2,6, K. Burrows1, A. Metspalu1, A. Salumets3,4, C.M. Lindgren2,6. 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 6) Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA.

Polycystic ovary syndrome (PCOS) is a common multifactorial disease affecting up to 10% of women of reproductive age, therefore being their most frequent endocrine problem. It is one of the leading causes of female subfertility as ~40% of women with PCOS are infertile. Genome-wide association studies (GWAS) have revealed several candidate genes, but most of the heritability of PCOS is unexplained. To investigate the contribution of potentially causal coding variants to PCOS, we have genotyped 167 cases and 711 population controls (363 females) from the Estonian Biobank with the Illumina exome array.

We conducted single variant and burden tests of association using SKAT-O within genes for (i) loss of function (LOF) and (ii) rare non-synonymous (NS) variants with minor allele frequency (MAF) <1%. The association analyses were adjusted for first two principal components to account for the population stratification. In the autosomal analysis, both male and female samples were used in the control group but in the X chromosome analysis, only female samples were used.

Altogether 55,545 polymorphic variants were successfully tested in single variant analysis with 127 NS and 254 missense variant which was showing exome-wide evidence of association (p<5x10^-7). Bonferroni correction for 100,000 variants: exm233350 in the nebulin coding NEB gene (p=4.9 x 10^-10, MAF=0.05%). Mutations in NEB have previously been associated with myopathy, however, none of the associations were statistically significant in the gene-based tests after multiple testing correction for 20,000 genes (p=2.5x10^-6). The strongest associations came from aggregating non-synonymous rare variants within POLK (p=4.3 x 10^-5) and PEL13 (7.3 x 10^-6) genes. A replication and immune response related genes. Our study suggests that rare variants can contribute to the genetic component of PCOS, but cannot explain previously reported association signals in established GWAS loci.
Large-scale exome genotyping reveals novel coding variation associated with endometriosis. A.P. Morris1,2,3, R. Mägi1, N. Rahmioglu1, A. Mahajan1, N. Robertson1, M. Peters1, M. Saare2, A. Salumets2,3, K.T. Zondervan1, UK Exome Chip Consortium. 1 Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom; 2 Estonian Genome Centre, University of Tartu, Tartu, Estonia; 3 Welcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 4 Department of Obstetrics and Gynaecology, University of Tartu, Tartu, Estonia; 5 Institute of Bio- and Translational Medicine, University of Tartu, Tartu, Estonia.

Genome-wide association studies (GWAS) have identified nine loci harbouring common variants associated with endometriosis. However, these variants typically map to non-coding genomic regions, and together explain only ~3% of the estimated 52% heritability of the condition. To investigate the contribution of coding variation to endometriosis pathogenesis, we undertook genotyping with the Illumina Exome Chip of two studies of European ancestry: (i) 910 cases from the Oxford Endometriosis Gene (OXEGENE) study and 13,334 population controls (6,828 females) from the UK Exome Chip Consortium; and (ii) 326 cases and 711 population controls (363 females) from the Estonian Biobank. Within each study, we evaluated the association of endometriosis with: (i) individual coding variants; and (ii) burden/over-dispersion of loss of function (all frequencies) and rare non-synonymous (minor allele frequency [MAF] less than 1%) variants within genes using SKAT-O. Analyses were adjusted for principal components to account for population structure. Association summary statistics were combined across studies by meta-analysis (fixed-effect sample size weighted Z-score method for single variants and Fisher’s method for gene-based tests). We conducted pathway analysis on the basis of single variant meta-analysis summary statistics using the most up to date curated pathway gene-sets from the Molecular signatures database as implemented in MAGENTA. No individual coding variants achieved exome-wide significant evidence of association (p<5x10^-5, Bonferroni correction for 100,000 variants). The strongest signals include deleterious missense variants in TAF1I (p=1.5x10^-5, MAF=0.077%) and BMP3 (67YN, p=3.2x10^-5, MAF=2.7%). We observed exome-wide significant evidence of association (p<2.5x10^-5, Bonferroni correction for 26,000 genes) with burden/over-dispersion of loss of function variants in C10orf69. We observed rare non-zero burden of loss of function variants in C10orf69 (p=1.7x10^-5), ZNF485 (p=1.1x10^-6), and RSDA2 (p=2.1x10^-6). MAGENTA analyses highlighted potential involvement of cell adhesion/structure, immune function and cancer-related pathways in endometriosis. Our study provides evidence that rare coding variation contributes to the genetic component of endometriosis. None of the identified genes from these analyses map to established endometriosis loci, providing no support for the hypothesis that rare coding variation can explain common GWAS association signals.
775S


Necrotizing enterocolitis (NEC) and spontaneous intestinal perforation are gastrointestinal emergencies in premature neonates. NEC is a common in neonates affecting approximately 11% of premature infants with a mortality rate of from 20% to 40%. Recent studies have suggested that genetic polymorphisms, specifically variants in genes affecting inflammation, may contribute to susceptibility for neonatal bowel inflammation and NEC. We utilized whole exome sequencing (WES) to search for variants that may increase the risk of developing NEC. Buccal swabs were collected for DNA extraction from infants ≤ 32 weeks gestation with and without a diagnosis of NEC. Infants with congenital heart disease, congenital anomalies, and inherited blood/metabolic were excluded. We performed exome sequencing using an Illumina TruSeq Exome Enrichment Kit on an Illumina HiScanSQ system for 54 samples (n = 27 without NEC; n = 27 diagnosed with NEC). We aligned the resulting reads to the hg19 reference genome with BWA, applied GATK base quality score recalibration, indel realignment, duplicate removal, and performed SNP and INDEL discovery. Variants were annotated using ANNOVAR. We analyzed 11 genes (IL1A, IL1B, IL17A, IL17D, IL17F, NFKB1, NFKB2, IL1A, IL1B, IL6, PTAFR, and TLR4) for non-synonymous or missense variants generated by WES and selected 26 variants that we tested for an association with development of NEC. None of the variants tested were associated with the development of NEC. However, we uncovered 26 candidate variants in inflammatory genes. None of the discovered variants were associated with the development of NEC. In this limited, candidate gene approach to the analysis of WES data, we did not discover a causal variant for NEC; however, we will be expanding the search for variants to the entire exome.

776M

Whole-exome imputation of sequence variants identified two novel associations with adult body height in African Americans. M. Du1,2, P.L. Auer1,2, S. Jiao1,3, J. Haessler2,4, D. Altshuler2,5, E. Boerwinkle6, C.S. Carlson1, C.L. Carty1, Y.I. Chen2, K. Curtis1, N. Franceschini6, L. Hsu1, R. Jackson1, L.A. Lange7, G. Lettre2, K.L. Monda8, D.A. Nickerson2, A.P. Reiner1,9, S.S. Rich4,10, S.A. Ross1, J.I. Rutter4, C.J. Willer8,11, J.G. Wilson12, K. North9, C. Koop*ber13, N. Heard-Costa14, U. Peters1, National Heart, Lung, and Blood Institute (NHLBI) Go Exome Sequencing Project. 1) Fred Hutchinson Cancer Research Center, Seattle, WA, USA; 2) University of Wisconsin-Milwaukee Joseph Z. Silber School of Public Health, Milwaukee, WI; 3) Broad Institute, Cambridge, MA; 4) The University of Texas Health Science Center at Houston, Houston, TX; 5) LABiomed at Harbor-UCLA Medical Center, Torrance, CA; 6) University of North Carolina Gillings School of Global Public Health, Chapel Hill, NC; 7) The Ohio State University Wexner Medical Center, Columbus, OH; 8) University of North Carolina School of Medicine, Chapel Hill, NC; 9) Montreal Heart Institute and Université de Montréal, Montreal, Quebec; 10) University of Washington, Seattle, WA; 11) University of Virginia School of Medicine, Charlottesville, VA; 12) University of Michigan Medical School, Ann Arbor, MI; 13) University of Mississippi Medical Center, Jackson, MS; 14) Boston University School of Medicine, Boston, MA.

Adult human body height is a quantitative trait for which genome-wide association studies (GWAS) have identified numerous loci, primarily in European populations. These loci, composed of common variants, explain less than 10 percent of the phenotypic variance in height. Here, we searched for novel associations between height and common (minor allele frequency, MAF≥5 percent) or less frequent (0.5 percent<MAF<5 percent) variants across the exome in African Americans. Using a reference panel of 1,692 African Americans and 471 Europeans from the National Heart, Lung, and Blood Institute’s (NHLBI) Exome Sequencing Project (ESP), we imputed whole-exome sequence data into 13,719 African Americans with existing array-based GWAS (discovery set). Variants achieving a height-association threshold of P<5·10^-6 in the imputed dataset were followed up in an independent sample of 1,989 African Americans with whole-exome sequence data (replication set). We used a Bonferroni-corrected P<2.5·10^-7 to define statistically significant associations in meta-analyses combining the discovery and replication sets (N=15,708 participants). We discovered and replicated 3 independent loci for association following Bonferroni-correct: 5p13.3/C5orf22 (rs17410035) and 13q14.2 (SPRYD7/rs114089985) harbored novel height alleles independent of previous GWAS-identified variants in all populations (τ2<0.01 with GWAS loci); in contrast, 17q23.3/GH2/rs2006123 was correlated with GWAS-identified variants in European and African populations. Notably, 13q14.2/rs114089985 is uncommon in African Americans (MAF=0.03) and is monomorphic in European and Asian populations, suggesting it is an African American-specific height allele. In summary, our findings demonstrate that whole-exome imputation of sequence variants can identify low frequency variants as well as discover novel variants in non-European populations.
777T Next steps for whole exome sequenced cases: imputing non-coding regions and incorporating whole genome sequenced controls. A.E. Hendrick1,2*, S.A. McCarthy3, E. Schrader1, S. Furlong1, E. Anand1, J. Fung1, A. Nascimento1, U.K. Obesity Group. 1) Mathematical and Statistical Sciences University of Colorado -- Denver, CO; 2) Wellcome Trust Sanger Institute, Cambridge, CB10 1HH.

The number of whole-exome sequenced (WES) case, or case-control, sets has recently increased substantially. After researchers complete an initial family or case-control exome based analysis, one might ask what else can be done. Two possible ways to proceed are to impute non-segregated regions or to use external data sets to increase the number of samples in the control set and subsequently the power to detect association. Here, we do both.

Others have shown that imputation of non-coding regions and association tests of simulated quantitative phenotypes is possible given the <1 off target reads produced by high-depth WES. Here, we investigate the imputation and association of non-coding regions using real data from the UK10K project (www.uk10k.org). Specifically, we use high depth (50x) WES case (N=926) and control (N=1233) samples as well as low depth (6x) whole-genome sequenced (WGS) control samples (N=3825). The case patients are from the Severe Childhood Onset Obesity Project consisting of UK white patients, with an age of onset below 10, and with a body mass index (BMI) Standard Deviation Score (SDS) > 3, a design that is enriched for rare, highly penetrant causes of obesity.

We focus on case-control analysis, which is particularly susceptible to bias, and add WGS controls to an existing WES control set as well as use the WGS separately. Using IMPUTE2, we impute the WES cases and controls merged with both the 1000Genomes only and the 1000Genomes plus the UK10K WGS as the imputing backbone. As a true positive, we look for the FTO signal, a signal that is not in high LD with variants in the coding region and was thus notably missing when we performed our case-control analysis on the WES coding regions. We also work to show that, after sufficient quality control and filtering, we attain results without inflated test statistics despite imputation into non-coding regions and the use of WGS controls.

We believe this work is vital to getting the most information and benefit out of WES data. Successful research in this area will enable groups to increase power to replicate signals and find new genetic associations using WES case or case-control data regardless of whether the association is inside or outside of a coding region.

778S Whole-exome DNA sequencing to find new variants associated with fetal hemoglobin levels. K.S. Lo1, G. Letter4-6, *1). Montreal Heart Institute, Montreal, Quebec, Canada; 2) Faculté de Médecine, Université de Montréal, Montreal, Quebec, Canada.

The most promising strategy to treat patients with the commonest Mendelian diseases in the World - sickle cell disease (SCD) and beta-thalassemia - is to increase the endogenous production of fetal hemoglobin (HBF). HBF is a highly heritable trait (h2~0.5-0.8) and genetic studies have identified SNPs at the BC11A, HBS1L-MYB and beta-globin loci that are associated with its levels. To find new HBF-associated loci, and test the role of rare coding gene variants that could increase HBF, we imputed successfully (RSQR>0.5) 906, 5,478 and 28,091 variants with MAF<5% and MAF≥5%, respectively. To increase statistical power, we imputed the coding variants identified by WES plus the UK10K WGS as the imputing backbone. As a true positive, we look for the FTO signal, a signal that is not in high LD with variants in the coding region and was thus notably missing when we performed our case-control analysis on the WES coding regions. We also work to show that, after sufficient quality control and filtering, we attain results without inflated test statistics despite imputation into non-coding regions and the use of WGS controls.

We believe this work is vital to getting the most information and benefit out of WES data. Successful research in this area will enable groups to increase power to replicate signals and find new genetic associations using WES case or case-control data regardless of whether the association is inside or outside of a coding region.

779M Whole-Exome Sequencing Identifies Rare, Functional CFH Variants in Families with Macular Degeneration. J. Seddon1, Y. Yu1, M. Triebwasser1, K.A. Davis1, M. Daily2, S. Raychaudhuri3, D. Kavanagh4, *1) Tufts Medical Center, Department of Ophthalmology and Ophthalmic Epidemiology and Genetics Service, 800 Washington St. #450, Boston, MA 02111, 2) University Washington School of Medicine, Seattle, WA, 3) Newcastle University, UK; 4) Mass. General Hospital, Boston, MA; 5) Brigham and Women's Hospital, Boston, MA.

We sequenced the whole exon of 35 cases and 7 controls from 9 age-related macular degeneration (AMD) families in whom known common genetic risk alleles could not explain their high disease burden and/or their early-onset advanced disease. Two families harbored novel rare mutations in CFH (R53C and D90G). R53C segregates perfectly with AMD in 11 cases (heterozygous) and 1 elderly control (reference allele) (LOD= 5.07, P= 6.7x10-7). In an independent cohort, 4 out of 1,676 cases but none of the 745 examined controls or 4300 NHBLI Exome Sequencing Project (ESP) samples carried the R53C mutation (P=0.0039). In another family of 6 siblings, D90G similarly segregated with AMD in 5 cases and 1 control (LOD=1.22, P=0.009). No other sample in our large cohort or the ESP had this mutation. Functional studies demonstrated that R53C decreased the ability of FH to perform decay accelerating activity. D90G exhibited a decrease in cofactor-mediated inactivation. Both of these changes would lead to a loss of regulatory activity, resulting in excessive alternative pathway activation. This study represents an initial application of the whole-exome strategy to families with early-onset AMD. It successfully identified high impact alleles leading to clearer functional insight into AMD etiopathogenesis.

780T Whole genome sequencing of 3,514 individuals from the founder population of Sardinia. C. Sidore1,2, F. Busonero1,2, A. Maschio1,2, M. Zoledziewska1, A. Mulas1,2, E. Porcu1,2, G. Pistis1,3, M. Sten1, F. Daniele1, A. Kwongt1, C.W. Chiang2, R. Lyons3, A. Angus1,3, H.M. Kang1, J. Novembre6, S. Sanna1, D. Schlessinger6, F. Cucca1,3, G. Abecasis2.

1) Istituto di Ricerca Genetica e Biomedica, CNR, Monserato, Cagliari, Italy; 2) Center for Statistical Genetics, Ann Arbor, University of Michigan, MI, USA; 3) Università degli Studi di Sassari, Sassari, Italy; 4) University of Michigan, DNA Sequencing Core, Ann Arbor, MI, USA; 5) Center for Advanced Studies, Research, and Development in Sardinia (CRS4), AGCT Program, Parco Scientifico e tecnologico della Sardegna, Pula,Italy; 6) Department of Human Genetics, University of Chicago, IL, USA; 7) Laboratory of Genetics, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA.

Recently, the combination of genome wide association studies with reference panels based on next generation sequencing has profoundly improved the resolution of genome analysis and the ability to identify causal variants associated with complex traits or diseases. The commonly used sequencing-strategy is based on a collection of populations that do not allow imputation of population-specific variants, particularly those present in isolated populations. Here we substantiate the increased informativeness that is recoverable by further analysis of population-specific variant present in genetically distant populations, and especially in isolated founder populations whose genetic distance from more cosmopolitan groups is most marked. To investigate the Sardinian specific genetic variation we used whole genome sequencing (coverage ~4x) of 3,514 Sardinian individuals, identified as members of AMD families, to detect rare mosaics (SNPs and indels) and generated a reference panel for imputation. Strikingly, for Sardinian individuals, imputation using our Sardinian reference panel was much improved for low frequency variants (MAF 0.5-5%) compared to imputation based on 1000 Genomes Project haplotypes (increasing r2 with directly measured genotypes from 0.59 to 0.90). We will show how relative isolation affects haplotype length, enrichment of deleterious variants, and genetic differentiation with Europeans. Next, we used this reference panel to study the genetics of LDL-cholesterol in the island by first imputing missing variants in 6,502 individuals from an isolated population with >800,000 SNPs. Variants detected by our sequencing explained >90% of the genetic heritability of LDL-cholesterol and point to many interesting association signals. Particularly interesting is the C39X mutation in the HBB gene, which causes sickle cell disease in Europe, not previously identified by any GWAS analysis. The variant results in a decrease of 14.4 mg/dL in blood among carriers (p < 10-21). Our results illustrate the benefits of large-scale sequencing efforts in founder populations and their ability to uncover functionally relevant variants that may be very rare elsewhere.

784S
Quantitative trait loci for plasma proteins in current and former smokers with and without chronic obstructive pulmonary disease (COPD). R.P. Bowler1, W. Kneissl2, K.Krehl3, W. Sun1, G. Jacobson1, T.H. Cheit1, M.B. Drummond1, R.E. Kanner1, P.G. Woodruff2, D. Couper1, G.A. Hawkins2. COPDGene and SPIROMICS. 1) Department of Medicine, National Jewish Health, Denver, CO; 2) Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC; 3) Department of Biostatistics and Informatics, School of Public Health, University of Colorado Denver, CO; 4) Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 5) Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah; 6) Department of Medicine and Cardiovascular Research Institute, University of California San Francisco, CA; 7) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston Salem, NC.

Background: Chronic obstructive pulmonary disease (COPD) occurs most commonly in current and former smokers and is the third leading cause of death in the United States. Genetic association studies have identified multiple genetic loci associated with COPD phenotypes such as airflow obstruction and emphysema. Similarly protein biomarker studies have identified blood biomarkers associated with these COPD phenotypes. The relationship between genetic markers and protein levels has not been explored in at risk subjects. Methods: Using two similar cohorts of older smokers (COPDGene and SPIROMICS), we explored the relationship between genome wide genetic markers and 96 candidate blood biomarkers for COPD in 750 (SPIROMICS) and 602 (COPDGene) non-Hispanic Whites. Genotyping was performed on Illumina platforms. Candidate COPD blood biomarkers were assessed in fresh frozen serum and plasma using 13 Myriad-RBM single and multiplex panels. cis and trans protein quantitative trait loci (pQTL) were identified using an additive regression model adjusted for covariates (age, gender, BMI, smoking status, pack-years), as well as genetic principal components, and for correction for multiple comparisons. The Stouffer-Liptak test was used for the meta-analysis to combine the p-values from the two studies. Results: We identified 539 significant pQTL SNPs (69% cis; 31% trans) in 36 proteins. The top cis-association was for GC (rs7041; Vitamin D binding protein; P = 5 X 10^-29). This SNP leads to an Asp to Glu amino acid change at position 29 and is associated with a more severe COPD. The top trans association was rs507666 (chromosome 9) for SELE (chromosome 1; E - selectin; P = 7 X 10^-102). This SNP is an intrinsic variant in ABO (alpha 1-3-N-acetylgalactosaminyltransferase) and has also been associated with blood group and altered glycosyltransferase activities. Other notable pQTLs include rs2070600, a non-synonymous intron variant in ABO (alpha 1-3-N-acetylgalactosaminyltransferase) and D binding protein; P= 5 X 10^-22). The top 31% of significant cis-associations (P <5 X 10^-5) had a minor allele frequency (MAF) > 5%, small effect sizes, and in aggregate explain a small fraction of the population variation in these reproductive ageing traits. It is predicted that functional genetic variants with lower MAF, but larger effects on phenotype, may contribute substantially to the genetic variance underlying these traits.

Aim: To identify low frequency coding variants contributing to the genetic variation in age at menarche and age at natural menopause.

Methods: Exome chip genotyping data were collected from 17 studies of European populations. Single marker and burden testing meta-analyses were carried out in 67,628 women for age at menarche and 35,605 women for age at natural menopause.

Results: New low frequency signals were identified for both age at menarche and age at natural menopause. For age at menarche, a nonsynonymous variant with MAF 1.1% was identified in ABO51 (2p13) (p=4.95×10^-10), which increased age at menarche by approximately 3 months per allele. Burden testing identified the gene PRKAG1 (12q13.12, containing 4 variants with MAF <3.5%) as associated with age at menarche (p=2.2×10^-8). For age at natural menopause, two low frequency nonsynonymous variants were identified in HELB (12q14.3) (MAF 2.5% p=3.87×10^-23 and MAF 3.5% p= 1.37×10^-21); these variants are 48 bp apart and are in linkage disequilibrium (r^2=0.7), each increasing age at natural menopause by approximately 1 year. Two-directional burden testing identified the gene SLC04A1 (20q13.33, containing 18 variants) as associated with age at menopause (p= 4.9×10^-8).

Conclusions: We report the first large-scale exome chip genotyping project for reproductive ageing. Two genes containing low frequency or rare variants were associated with age at menarche and a further two genes were associated with age at natural menopause. None of these genes have effects on phenotype, may be implicated in genetic associations for reproductive traits in humans and the effect sizes were large compared to GWAS findings.

786T

We ascertained an Algerian consanguineous family in which two sibs presented with psychomotor delay, progressive microcephaly, spasticity, thin corpus callosum, and severe and early onset obesity. Exome sequencing identified two homozygous substitutions cosegregating with the phenotype and locating 170 kb apart on 7q22.1: a c.1313+4G>T splice mutation in AP4M1 previously described in a Moroccan family and a c.595A>T missense variation in AZGP1 which encodes zinc-alpha2-glycoprotein (ZAG). Haplotyping analysis indicated that the AP4M1 mutation was found in two families shared between both families, whereas the AZGP1 mutation occurs secondarily and is unique in the family. Mutations in AP4M1 cause AP4-deficiency syndrome, a condition characterized by severe intellectual disability, progressive microcephaly and spasticity. Notably, none of the 25 previously reported cases with AP4-deficiency syndrome exhibited obesity. On the other hand, ZAG is an adipokine stimulating lipolysis in adipocytes; ZAG likely regulates body weight since administration of human ZAG to ob/ob mice resulted in progressive weight loss. We propose that the phenotype of our patients resulted from the additional effects of the two mutations in AP4M1 and AZGP1 accounting for the neurological signs and the precocious morbid obesity, respectively. The contiguous gene syndrome was proposed in 1986 to explain the association of multiple and unrelated clinical features due to the deletion of multiple adjacent genes: the phenotype results from the combination of the endophenotypes of each contiguous gene sensitive to haploinsufficiency. Today, high-throughput sequencing allows us to enlarge this concept to describe simultaneous transmission of independent mutations that are genetically linked.
Sequence of genes expressed in podocytes uncovers FSGS risk alleles in European-American population. M. Artemov, E. Kopp, A. S. Shew, M. Dakey, 1) Chemistry and Chemical Biology Dept, Harvard University, Cambridge, MA; 2) Broad Institute, Cambridge, MA; 3) Analytic and Translation Genetics Unit, Massachusetts General Hospital, Boston, MA; 4) Department of Pathology and Immunology, HHMI, Washington University School of Medicine, St. Louis, MO; 5) Kidney Disease Section, NIDDK, National Institutes of Health, Bethesda, MD; 6) Molecular Genetic Epidemiology Studies Section, NCI, Frederick, MD.

Focal segmental glomerulosclerosis (FSGS [MIM 607832, 603965, 613937, 613937, 613937]) is a rare disorder characterized by nephrotic syndrome in adults and children. Over the last 10 years, various genetic approaches have identified at least eight genes: ARHGAP24, TRPC6, MYH9, INF2, APRIL, NPHS2, ACTN4, and CD2AP as FSGS susceptibility genes. Since then, genome-wide association studies (GWAS) of FSGS cases, it seems very likely that other genes await discovery. The FSGS phenotype is known to be specific to podocytes, thus we focused only on sequencing 2500 genes that are highly expressed in the podocyte. Here we present our analysis of 214 European FSGS cases aiming to search for novel germline mutations in Europeans predisposing to the disease. We used expression studies in both human and mouse to assemble a list of approximately 2500 genes representing about 7 Mb and sequenced these genes in about 200 European individuals by using 378 full exome sequences of European ancestry controls for the case/control study. Most of the genes so far reported to be associated with FSGS have been found in studies involving individuals of African-American descent. Here we focused on genetic factors driving disease in European individuals. We found promising rare mutations in 14% of the FSGS cases that have previously been found in controls in Africans. We then performed unique to cases mutations analysis and found that top hits are ARHGAP24, which is a known risk gene in Africans, APRIL, which is in a close linkage with known APRIL risk gene and has been previously been found in controls in Togo, MYH9, which directly interacts with MYH9 at the protein level and potentially could be a new risk gene for FSGS. We have used allelic association test to search for associated variants. Our two top hits were the known FSGS allele - G1 APRIL variant and a stop codon mutation in MYH9, which we observed in 11% of the European FSGS cases and 10% of the African controls. These observations suggest that the genetic nature of FSGS in Europeans is similar to Africans and mostly is driven by the same alleles and risk genes. We find several new genes that seem to increase risk of FSGS in Europeans and were not previously reported.


Keratoconus (OMIM 184800) is a complex disease characterised by progressive thinning and conical protrusion of the cornea, resulting in major visual impairment. Recent genome-wide association studies have implicated genetic variation near the VSX1 and MIR184 genes and GWAS for central corneal thickness (CCT) have revealed associations with TSPAN4 and VSX1, which may act as a protective variant when combined with haploinsufficiency of TBX3 on 22q11.21. Overall, this analysis reveals the utility and challenges of using WGS for identifying potential modifier genes (ZFM2 and ADNP2) for this complex disorder.

Whole exome sequencing of Cold Medicine-Related Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis (CM-SJS/TEN) with Severe Mucosal Involvement. Y. Hitomi, S. Khor, M. Ueta, S. Sotozono, S. Kino, T. Horiuchi, S. Kameyama, T. Ogino, T. Kikuta, H. Yu, C. Sotozono, M.J. Daly, K. Tokunaga 1) Genetics, Albert Einstein College of Medicine, Bronx, NY; 2) Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY; 3) Department of Obstetrics & Gynecology and Women's Health, Montefiore Medical Center, Bronx, NY; 4) Childrens Healthcare of Atlanta at EastGueston, Atlanta, GA.

The 22q11.2 Deletion Syndrome (22q11DS) is characterized by highly variable expressivity, ranging from mild craniofacial defects to severe congenital heart disease and of psychiatric illness. Genetic modifiers may account for variable expressivity through a multi-hit model where the 22q11.2 deletion acts as the primary hit and genetic variants act as second hits to modify the phenotype. To identify candidate genetic modifiers for physical and behavioral phenotypes, we used whole genome sequencing (WGS) of two trios with normal parents and probands (C1 and C2) with 22q11DS, who have similar de novo 3 Mb deletions but discordant phenotypes. On the remaining 22q11.2 allele in each case, we found 30 coding variants in each case and 1,038 and 940 noncoding variants in C1 and C2 respectively. For C1, we found 349,978 rare (MAF < 0.05), 57,319 novel, and 48 de novo variants elsewhere in the genome. For C2, we found 371,778 rare, 61,367 novel, and 35 de novo variants. C1, a male, recruited at 14 years old, had a history of Fallot (TOF), cognitive deficits, and attention deficit hyperactivity disorder. He was later diagnosed with schizophrenia (SCZD) at age 28, C2, a female, age 14, had juvenile rheumatoid arthritis that has resolved, no heart aneurysms, and no major cognitive deficits or psychiatric illness. Our first goal was to determine if there was an enrichment of potentially pathogenic variants in candidate genes in TOF and C2, compared to controls (C3).

We identified 62 genes in C1 and 65 genes in C2 with potentially pathogenic novel, rare compound heterozygous, or de novo variants using ANNOVAR. We used TopGene to prioritize these target genes against 1000Genomes controls from our cohort and controls from OMIM. We found enrichment of candidate genes for TOF, but we did not find enrichment of candidate genes for cognitive/SCZD, compared to C2 (Fisher’s p-value = 9.72E-05). Among the interesting variants in C1, were common in C2 (homozygous) KANK1 which is associated with renal cell carcinoma, MPRIP, which directly interacts with MYH9 at the protein level and potentially could be a new risk gene for FSGS. We have used allelic association test to search for associated variants. Our two top hits were the known FSGS allele - G1 APRIL variant and a stop codon mutation in MYH9, which we observed in 11% of the European FSGS cases and 10% of the African controls. These observations suggest that the genetic nature of FSGS in Europeans is similar to Africans and mostly is driven by the same alleles and risk genes. We find several new genes that seem to increase risk of FSGS in Europeans and were not previously reported.
791M Identification of putatively causative variants in three anorexia nervosa multiplex families by whole exome sequencing. D. Li1, R.D. Coney2, B. Lo3, H. Hakonarson4,5, S. Neuhaus1, J. Eskin1, 6) The Center for Applied Genomics, Abramson Research Center, The Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA; 2) Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; 3) Center for Quantitative Sciences, Vanderbilt University, Nashville, TN 37232, USA; 4) Division of Human Genetics Department of Pediatrics, The Children’s Hospital of Philadelphia and The Perelman School of Medicine, Philadelphia, PA 19104, USA; 5) Division of Pulmonary Medicine, The Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA.

Anorexia nervosa (AN) is a perplexing illness characterized by low body weight and persistent fear of weight gain during period of growth, resulting in extreme emaciation. Family studies have consistently demonstrated that AN occurs in families and twin studies have revealed the contribution of additive genetic factors to the observed familial aggregation, but so far just few genetic factors have been found to be specific to AN and no single gene has been shown to be necessary or sufficient to express the phenotype. While GWAS relies on proxy association of genetic variants with unknown disease causality, in which it only focuses on one affected child from multiplex families, the published studies don’t specifically analyze inherited mutations despite the fact that AN is highly heritable. Familial AN with young age of onset, extremely low body weight, extended course of illness or male affected constitutes a unique and extreme form of the disease. Here we report the discovery of 5 ultra rare (MAF=0.000077 in ESP6500SI) or novel variants in CMT1, DRD4, KCNNA3, GRPR and ESRR1, in 3 AN multiplex kindreds as strong biological candidates. Polymorphisms in these genes have been associated with other complex disease including, but not limited to AN, Attention Deficit Hyperactivity Disorder (ADHD), schizophrenia, depression and bipolar disorder. We are currently recruiting additional family members to further validate the cosegregation pattern and conducting follow-up functional assessment to elucidate the roles these genes play in AN and potentially other eating disorders. The resulting variants will be presented together with detailed phenotypic characterization of the families examined.

791S Exome sequencing of multiplex pedigrees for the identification of novel rare susceptibility variants for CD. B.-S. Petersen1, S. Nikolaus2, S. Schreiber-Arnke1, 1) Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Germany; 2) Department of Internal Medicine, University Hospital Schleswig Holstein, Kiel, Germany.

Crohn’s disease (CD) is a complex, chronic inflammatory bowel disease (IBD). A heterogeneous mixture of genetic and environmental factors likely play a role in causing CD. Genome-wide association studies (GWAS) and meta-analyses have so far identified 163 genetic susceptibility loci for IBD. These, however, explain less than 30% of the heritability so far. Apart from the mostly common variants included in classical GWAS, further sources of heritability include rare and novel variants with possibly higher penetrance which can be identified through systematic resequencing studies. We have therefore carried out exome sequencing for 50 affected individuals from 17 pedigrees with at least three individuals affected by CD for finding rare high penetrance variants and novel candidate genes involved in disease etiology. The results were analyzed pedigree-wise and filtered for rare and novel coding variants shared by the patients. In several cases this led to the identification of candidates likely to play an important role for disease development in the underlying pedigree. A total of 300 variants were selected and are currently being genotyped in a large patient control cohort to test for recurrence and for proving that the identified variants are not common in healthy individuals. Our preliminary results have revealed novel rare candidate variants in known and novel CD genes shared by the affected individuals of a pedigree which may have an important impact on the development of the disease, also in other patients.

792T Whole exome sequencing implicates novel rare genetic variants in susceptibility to Legionella pneumonia. A. Ndungu1, KS. Elliot2, TC. Mills3, A. Rautanen4, P. Hutton5, C. Garrard6, AC. Gordon7, CJ. Hinds8, AVS. Hill9, SJ. Chapman4, GAIns Investigators. 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 2) John Radcliffe Hospital, Oxford, UK; 3) Intensive Care Unit, John Radcliffe Hospital, Oxford, UK; 4) Section of Anaesthesiology, Pain Medicine & Intensive Care, Imperial College, London, UK; 5) William Harvey Research Institute, Barts & The London School of Medicine, London, UK; 6) University Psychiatry Medicine, Churchill Hospital Site, Oxford Radcliffe Hospital, Oxford, UK.

Community acquired pneumonia (CAP) is associated with considerable mortality and morbidity worldwide. 2-9 % of all CAP cases are attributable to the flagellated bacteria, Legionella pneumophila. Although contaminated water supplies may be relatively common, yet there is great inter-individual variation in susceptibility to disease. Host genetic variation is increasingly recognized as playing an important role in determining susceptibility to infectious agents, however, the role of the host genome in contributing to Legionella pneumonia has not been fully characterised. Few candidate gene studies have mainly focused on genes in the Toll-like receptor pathway that plays a central role in mammalian cholesterol metabolism suggesting a potential role for these genes in susceptibility to Legionella. We use a whole exome sequencing approach in an attempt to identify novel genes and large effect genetic variants associating with Legionnaires’ disease in humans. We performed whole exome sequencing of 16 Caucasian individuals from 4 multiplex families by whole exome sequencing to identify new genes for NS-ARID in several families. Due to the heterogeneous nature of the disorder, exome sequencing is a powerful tool as almost all known genes can be assessed simultaneously in one study.

In this study, we performed exome-sequencing for several families with non-syndromic autosomal recessive Intellectual Disability Using Exome Sequencing. N. Vasiili1, I. Ahmed2, K. Mittal1, M. Ohadi3, A. Mikhailov4, MA. Rafia5, A. Bhatti6, M. Carter7, DM. Andrade8,9, M. Ayub4, P. John4, JB. Vincent1,2.

1) Molecular Neuropsychiatry and Development Lab, The Campbell Family Brain Research Institute, The Centre for Addiction & Mental Health (CAMH), Toronto, Ontario, Canada; 2) Atta-ur-Rehman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad-Pakistan; 3) Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto ON Canada; 4) Division of Neurology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada; 5) Krembil Neuroscience Centre, Toronto Western Research Institute, Toronto, Canada; 6) Division of Developmental Disabilities, Dept. of Psychiatry, Queen's University, Kingston, ON, Canada; 7) Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada; 8) Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada.

Intellectual disability (ID) is a genetically heterogeneous disorder with more than 50 mutated genes to date. ID is characterized by deficits in memory skills and language development with difficulty in learning and problem solving. It affects ~2% of population with difference in severity based on the Intelligence quotient (IQ) scores from mild forms with IQ range from 50 to 70, to profound forms with IQ below 20-25. For detection of disease-causing mutation in such a heterogeneous disorder, exome sequencing is a powerful tool as almost all known genes can be assessed simultaneously in one study.

In this study, we performed exome-sequencing for several families with non-syndromic autosomal recessive ID (NS-ARID). After genotyping and exome sequencing, several interesting genes were detected. For example, a homozygous missense mutation was detected in LRP2 gene in a Pakistani patient with mild ID from a marriage between second cousins. The variation was absent in 400 ethnically matched healthy control chromosomes and is not listed in SNP databases. The LRP2 mutation identified here is located in one of the LDL-receptor class Adomains which is a cysteine-rich repeat that plays a central role in mammalian cholesterol metabolism suggesting that alteration of cholesterol processing pathway can lead to intellectual disability. In another Pakistani family, a homozygous missense mutation in NUP205 gene was detected in a female patient with mild ID. Sanger sequencing analysis showed complete segregation within the family. Functional studies are ongoing to verify the pathogenicity of variants detected.

Here we describe the use of autozygosity mapping and whole exome sequencing to identify new genes for NS-ARID in several families. Due to the heterogeneous nature of the disorder, exome sequencing is a powerful tool as almost all known genes can be assessed simultaneously in one study.
Low-frequency coding variation in PRF1 and GALC mediate multiple sclerosis risk. C. Cotsapas1,2, M. Mitrović1, International MS Genetics Consortium. 1) Neurology, Yale School of Medicine, New Haven, CT; 2) Broad Institute of MIT and Harvard, Boston MA.

Multiple sclerosis (MS) is an autoimmune disease in which the myelin sheath surrounding brain neurons is destroyed by the immune system, leading to progressive physical and cognitive morbidity, and often death. Over the last decade we have conducted a series of genome-wide association studies showing that the genetic architecture of MS is complex and likely involves hundreds of risk loci. We have identified almost 100 such loci explaining 15% of the heritability and our ongoing common variant mapping efforts in ~36,000 cases and ~44,000 controls will substantially increase this number. To complement these efforts we have interrogated 250,000 low-frequency non-synonymous (NS) coding variants across all exons in the genome using Illumina’s exome chip in a total of ~40,000 cases and ~60,000 controls of European descent. In our interim analysis, we have analyzed 26,231 cases and 24,031 controls in 14 country-level strata following stringent quality control to eliminate technical artefacts and population stratification. We tested >80,000 NS polymorphic variants for association to MS risk and find that two low-frequency variants (MAF < 0.05) show convincing evidence of association (p < 6.25 x 10^-7, Bonferroni-adjusted p < 0.05 for variant tested). These variants are in genes PRF1 (preformer, OR = 1.2) and GALC (galactosylceramidase, OR < 0.77). A further variant in HDAC7, a histone deacetylase, has suggested evidence of association (p = 1.97 x 10^-5, OR = 0.77). We are currently analyzing our remaining samples to perform a joint analysis across approximately 100,000 cases and controls. Recognizing that even in a cohort of this size, individual variant testing is underpowered to detect rare risk alleles forming gene burden tests to identify genes harboring more low-frequency MS risk variants than expected by chance. As not all genes tolerate substitutions equally, we shall also look for genes with strong mutational constraint and association to MS risk. Overall, our results suggest that low-frequency variation identifies a number of MS risk genes in both known risk loci and elsewhere in the genome. These effects account for a small but significant proportion of disease risk heritability and reveal novel risk genes and aspects of MS susceptibility biology.

Exome sequencing of 487 Community Acquired Pneumonia patients. S. Elliott1, J. Fellay, C. McLaren1, L. Schlabschbach, J. Fellay. 1) School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; 2) Paediatric Critical Care Research Group, Mater Research, University of Queensland, Brisbane, Australia.

Background: Respiratory viruses are the most common pathogens leading to non-elective admission to Pediatric Intensive Care Unit. Human genetic variation plays a key role in determining individual responses after exposure to infectious agents. Current knowledge on the genetic basis of susceptibility to common respiratory infections is limited. Stosiek and colleagues identified a significant proportion of the inter-individual differences in response to the virus leading to the finding of rare susceptibility variants of relatively recent origin. In order to identify such variants an exome sequencing study was undertaken. DNA samples from 487 adult UK individuals admitted to an intensive care unit with severe community-acquired pneumonia (CAP) were collected as part of a study of genetic predictors of death from sepsis in critically ill patients (Genomic Advances in Sepsis [GAIN Se]). Analysis of sepsis susceptibility was performed on a discovery cohort of 270 CAP samples compared to the UK10K ALSPAC control dataset. After stringent QC criteria were applied, 135,392 variants were identified. Of these, 43 reached ExWAS significant threshold for association (p < 3.6 x 10^-7) and an additional 217 variants were suggestive (p < 1 x 10^-6). The exomes from the remaining 217 CAP patients are being analysed as a replication dataset against the UK10K TWINSKU control dataset. The sepsis outcome phenotype was also analysed, measured as 28 day mortality post-ICU admission within the 487 discovery and 412 CAP cohorts (died n=237, survivors n=237, unknown n=13). In single variant analysis of - sepsis outcome, seven variants were identified reaching the ExWAS significance threshold including variants in two related genes known to be involved in thrombosis. Collapsing methods with rare deleterious variants are being performed to detect gene centric associations. Identification of novel, large-effect genetic variants has the potential to significantly expand current understanding of sepsis biology and may have clinical applications.
Genome-wide exome array analyses reveal novel rare variants for refractive error in Asia populations. Q. Fan1, J. Liao2, CC. Khor3, YY. Teo1,2, SM. SAW1,2, YY. Wong1,2,3, CY. Cheng1,2,3, 1) SSSH School of Public Health, National University of Singapore, Singapore, Singapore; 2) Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore; 3) Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore; 4) Department of Statistics and Applied Probability, National University of Singapore, Singapore; 5) Department of Ophthalmology, National University of Singapore, Singapore, Singapore.

As one of the leading causes of visual impairment, myopia poses a significant burden to public health in Chinese population. Genome-wide association studies have identified genetic loci contributing to refractive error, however, a substantial portion of the heritability remains unknown. We conducted exom array analysis for low-frequency (MAF 1-5%) and rare variants (MAF < 1%) on spherical equivalent in 2,253 Chinese and 2,091 Malay adults residing in Singapore. We genotyped the whole genome exome variants using the Illumina HumanExome Beadchip. Among 274,00 variants passing quality control, 64,000 (31%) were polymorphic variants. We conducted the regression-based association test for single-variant (MAF > 0.1%) and burden test (SKAT) for 53,991 nonsynonymous variants (MAF < 5%) with 14,310 genes. One gene harboring 4 rare/low frequency nonsynonymous variants (cumulative MAF = 3.6%) on chromosome 1 was identified to be associated with SE variation after Bonferroni correction for the number of genes (P = 2.25 x 10-7, respectively). Conditional analysis suggested the signal of the top gene was mainly driven by the single low frequency variant within this gene. This protein coding gene is a part of the epidermal growth factor receptor (EGFR) signaling pathway, whereas EGFR signaling has been implicated in the myopia development in animal models. Six genes containing at least two rare nonsynonymous variants also exhibited suggestive association with SE (P < 5.0 x 10-5). Our data suggest the likely role of rare genetic variants in influencing refractive error. The interpretation of the gene-based tests needs carefully considering the effects of single rare variants where the large sample size is required.

Analysis of the rare variant burden in the exomes of candidate HIV-target genes in relation to HIV-acquisition and AIDS-progression. M.C. Turchin1, S. Penugonda2, E-Y. Kim3, K. Kunstman1, M. Stephens1,2, S.M. Wolinsky1, 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) The Feinberg School of Medicine, Northwestern University, Division of Infectious Diseases, Chicago, IL; 3) Department of Genetics, University of Chicago, Chicago, IL.

Multiple genome-wide association studies (GWAS) have been conducted attempting to link common human genetic variation (minor allele frequency, MAF, >5%) to various aspects of HIV and AIDS pathology. Despite using large sample sizes (up to 6,000s) and samples representing ancestries beyond Western Europeans (Fellay et al. 2009, Pelak et al. 2010, McLaren et al. 2011), the majority of associations have only been found within the human leukocyte antigen (HLA) region. Here, we present the results of a more focused gene-exome sequencing study utilizing ~550 genes whose protein-products have previously been implicated as significantly associating with the 18 HIV-1 proteins (Jäger et al. 2012). These ~550 genes were sequenced in over 900 individuals, the majority of which are of Western European descent. Individuals in this sample were classified as either seropositive or sero-negative for associations involving HIV-acquisition, with HIV-positive and HIV-negative groups being further classified as slow/very-slow progressors of rapid/very-rapid progressors (where applicable), thus allowing us to look at associations regarding AIDS-progression as well. Analyses were conducted focusing on both the rare variant burden (MAF <5%) in the exomes of these genes as well as variants that are classified as ‘damaging’ as designated by multiple functional-prediction algorithms. Overall, we find a handful of genes outside the HLA region that appear to contain an excess of rare, ‘damaging’ variants in one class of individuals versus another for both HIV-acquisition and AIDS-progression. These results provide further insight into the biology of HIV-human interactions and proper follow-up and validation, may provide future directions for the development of HIV-drug targets.


Cerebrovascular diseases are the second most important cause of death in Spain, and together with other neurodegenerative diseases, they are the leading cause of disability in adults. Variability in functional outcome after a stroke can be influenced by many factors. Irrespective of clinical factors such as age, stroke etiological subtype, vascular stenoses, location of the injury and the size of the affected area, inter-individual variation in capacity of neuronal recovery is considerable. A number of systems and metabolic pathways are important for response to cerebral ischemic damage, and their activity may be modulated by variation in the genes that encode their various components. We aim to identify genetic variants, genes and pathways influencing the functional recovery process. With this aim, we have selected 81 patients with extreme phenotypes (36 bad vs 45 good outcome), suffering an anterior territorial ischemic stroke, with similar stroke severity, and matched for basal functional level, age and gender, from a cohort of over 4000 stroke cases. The patients underwent exome sequencing (Nimblegen v3 and Illumina sequencing). Downstream analysis was performed using a mixed-model association test, with variants collapsed by gene and weighted by their frequency and condel scores, with the aim of identifying genes with an accumulation of variants in either set of samples. This provided a set of approx. 400 nominally significant genes (although none passed multiple testing correction), which includes some relevant genes for stroke risk and stroke outcome. Further analysis including CNV analysis and integration with GWAS results from this and other cohorts of stroke cases is underway. These findings will be validated in the extended cohort.


The UK10K Consortium (http://www.uk10k.org/) is a collaboration between multiple research centres, mainly in the UK, aiming to uncover genetic variants contributing to disease and health status by sequencing 10,000 people. As part of the Rare Disease Group, the exomes of a 1,000 individuals with rare, extreme conditions have been sequenced to an average read depth of 72x using next generation sequencing technology (NGS), with an aim to detect novel causal variants. Amongst these, we focused our analysis on patients with disorders of insulin action of unknown aetiology. These are a group of highly heterogeneous conditions with a spectrum from insulin resistance (SIR) syndrome (a collection of rare disorders of extreme resistance to the glucose-lowering effects of insulin) and unrestrained metabolic or mitogenic insulin-like activity in the absence of insulin. In the discovery phase, 132 families from across the world (totaling approx. 1,000 individuals) were whole-exome sequenced, with follow-up in a further 250 index cases targeted by sequencing of 76 genes selected either as good candidates for disease from previous NGS or GWAS studies, or already known to be disrupted in individuals with disorders of insulin action (e.g. INSR, LMAN1, PPARG, AGPAT2 and AKT2). After sample quality control, 117 discovery and 248 follow-up samples were carried through for further analysis. To prioritise putative causative variants after removing those with low sequencing depth and quality, we filtered using UCSC Genome Browser (http://www.1000genomes.org/), NHLBI Exome Sequencing Project (ESP; http://evs.gs.washington.edu/EVS/) and the remaining UK10K exomes and genomes to exclude common variants and focused on those predicted to alter protein sequence (potentially functional). Family trios were also assessed to detect variants with a high probability of being de novo using DeNovoGear (https://sourceforge.net/projects/denovogear/). Possible candidates were confirmed by Sanger sequencing and co-segregation with available family members was examined wherever possible. Combined analysis of discovery and replication samples has provided possible diagnosis to 5.6% of the index cases to date. These include mutations within genes known to be disrupted in syndromes of SIR (INSR, PPARG, AGPAT2, ALCMS1) and in strong candidates either with a role in insulin action (MTOR, RICTOR, SNAP23, TBC1D4) or those within pathways relevant to pheno-type features (POCI1A).
Admixture mapping of exome genotyping data implicates region 15q21.2-22.3 with keloid risk in African Americans. K.S. Tsosie1,2, D.R. Velez Edwards1,3, S.M. Williams4, T.L. Edwards1,2,3,4, S.B. Russell1,7. 1) Center for Human Genetics, Vanderbilt University, Nashville, TN; 2) Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, TN; 3) Vanderbilt Epidemiology Center; 4) Institute for Medicine and Public Health; 5) Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, TN; 6) Department of Genetics, Geisel School of Medicine, Dartmouth University, Hanover, NH; 7) Division of Dermatology, Department of Medicine, Vanderbilt University, Nashville, TN.

Keloids (MIM 148100) are benign dermal fibrotic tumors with no effective clinical remedy that affect people of recent African ancestry approximately 20 times more than individuals of Caucasian descent. Possible related fibroproliferative diseases with increased prevalence in African populations include hypertension, nephrosclerosis, allergic disease, and uterine fibroma. Familial aggregation and ancestral differences in risk among geographic subpopulations strongly suggests a genetic association between African ancestry, keloids and fibroproliferative disease risk. There are no published genome-wide studies of keloid risk in African ancestry subjects. We conducted admixture mapping (AM) and whole exome association in 478 African Americans (AAs: 122 cases, 356 controls) with exome arrays to identify regions of local ancestry and SNP genotypes under AM peaks associated with keloid risk. Results: The most significant association with keloids discovered by AM was observed on chr15q21.2-22.3. This 5Mb region includes NEDD4, which was previously implicated in keloid formation by GWAS in Japanese and later validated in Chinese. Though our study nominally replicated this finding by AM and genotype association, the most significant SNP genotype association under the AM peak was observed at MYO1E (rs747722, odds ratio [OR]=4.41, 95% confidence interval [CI]=2.29-8.50, p=9.07×10^{-6}). A scan of all common genotype associations also identified associations at MYO7A (rs35641839, OR=4.71, 95% CI=2.38-9.32, p=8.34×10^{-6}) at chr11q13.5. GWAS have linked the chr15q21.2-22.3 region with hypertension in AAs, asthma in Europeans, and atherosclerosis in a Finnish cohort, providing evidence for common genetic elements. Examining expression microarray data of fibroblasts from keloids and normal scars that included some subjects from this study, markedly upregulated expression of MYO1E was observed. Conclusions: This study is the first to use AM and exome array association analysis to explore the genetics of keloids in AAs. Our findings, strengthened by support from expression data, further elucidate a potential region on chr15q21.2-22.3 for a role in risk of keloids in AAs, Japanese, and Chinese populations.
1 Department of Epidemiology, Case Western Reserve University, Cleveland, Ohio.

Speech sound disorders (SSD) are communication disorders occurring in 16% of 3 year olds and cost an estimated $30 - $154 billion annually in lost productivity, special education, and medical care. SSD exhibits high heritability, and for childhood apraxia of speech (CAS), a severe subtype of SSD, there are a few candidate genes such as FOXP2. However, apart from this knowledge, a comprehensive understanding of SSD genetics remains to be developed. We hypothesize that SSD is similar to neurodevelopmental disorders such as autism and intellectual disability and occurs not due to a single variant, but rather as the result of a high burden of multiple variants. The present study to test this hypothesis is a 26-year longitudinal cohort study based in Cleveland, OH, involving 135 cases with SSD, 34 (35%) of whom are affected with CAS, and 213 unaffected parental controls.

To examine the genetic component of SSD, we used both exome sequencing and HumanOmni2.5Exome genotyping. Analysis of FOXP2 in exome sequenced (n=11 CAS) and genotyped (n=135 SSD and CAS) samples did not identify any single variant or CNVs that could be considered causative mutations. We expanded our analysis to examine overall variant burden using a single measure of deleteriousness, Combined Annotation-Dependent Depletion scores (C-scores), which allows for comparison of the distribution of deleterious variants between groups. Preliminary results on a subset of 7,000 variants indicate that while there is no significant difference between the total number of rare (<1%) variants per individual (mean # case=3.43 vs control=3.26, p=0.5), there is a significantly higher burden of deleterious rare variants (p=0.08). Additionally, compared to cases without ASD (mean RawC=4.16), apraxic individuals have a significantly higher burden of deleterious variants (p=0.05). Finally, a missense variant in BTBD8, a gene expressed prominently in the fetal, but not the adult brain, is best examined in an integrative, whole genome manner and imply that human speech evolved through changes in many genes and interactions between them rather than as a single event. Acknowledgements: DC00528, DCO12380, T32-HL070567.

Functional follow-up, fine mapping and haplotype meta-analysis improve insight in findings from exome chip analyses and reveal potential genetic variants causing glucokinase deficiency, S.M. Willems
A.Y. Chu, R.A. Scott, H. Yaghootkar, J.I. Rotter, J.E. Huffman, M.O. Goodarzi, S.K. Iyengar, C. M. Stein, W.S. Bush, B. Truitt, L. Freebarin, A. Z. H. Goh, D. Siscovick, J.R. Mead, R.A. Scott, S. Wang, S. M. Goodarzi, CHARGE GLYCEMIA-T2D EXOME CHIP WORKING GROUP. 1) MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK; 2) Genetic Epidemiology Unit, UMC Utrecht, Department of Biometry, Erasmus University Medical Center, Rotterdam, The Netherlands; 3) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; 4) Fairbanks School of Public Health, Department of Epidemiology, University of Alaska Fairbanks, Fairbanks, AK, USA; 5) Institute of Medical Genetics, University of Abertay, Dundee, UK; 6) Division of Preventive Medicine, Brigham and Women's Hospital, Boston MA, USA; 7) Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK; 8) Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI, USA; 9) Department of Neurology, Wayne State University School of Medicine, Detroit, MI, USA; 10) Harvard Pilgrim Health Care Institute, Department of Population Medicine, Harvard Medical School, Boston, MA, USA; 11) Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, AL; 12) Department of Epidemiology, University of Washington, Seattle, Washington, USA; 13) Division of Statistical Genomics and Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA; 14) The Children's Monarch Institute, Monarch Institute, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; 15) The Icahn School of Medicine at Mount Sinai, New York, NY, USA; 16) Department of Genetics, University of Michigan, Ann Arbor, MI, USA; 17) Department of Biology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA; 18) MRC Human Genetics Unit, MRC IGM, University of Edinburgh, Edinburgh, Scotland, UK; 19) Massachusetts General Hospital, General Medicine Division, Boston, MA, USA; 20) Department of Medicine, Harvard Medical School, Boston, MA, USA; 21) Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy; 22) Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL; 23) The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York, NY, USA; 24) New York Academy of Medicine, New York, New York, USA; 25) Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA, USA; 26) Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; 27) National Heart, Lung, and Blood Institute (NHLBI) Framingham Heart Study, Framingham, MA, USA; 28) Division of Endocrinology, Diabetes and Metabolism, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

To explore the role of coding single nucleotide variants (SNVs) on fasting glucose (FG) levels, we performed exome-wide association scans (GWAS) in up to 30,564 non-diabetic individuals of European (84%) and African (16%) ancestry from 23 studies with Illumina Exome chip data. We performed single variant analyses for all coding and non-coding SNVs with MAF>0.02%. Gene-based tests included rare SNVs predicted to be protein-altering. Single variant analyses revealed two novel FG-associated loci: GLP1R (SNV A316T) and ABO (four non-coding SNVs, lead rs651007). Follow-up of GLP1R in exome sequencing data from up to 14,118 individuals, performed to investigate sequence variation not captured by the exome chip, identified an intronic SNP 100 bp from A316T (rs761386, MAF=3%, [5.09, p=7.6x10^-6]) that was more strongly associated with FG than A316T (p=0.01). rs761386 was close to a splice site, suggesting a possible effect on GLP1R pre-mRNA splicing. Follow-up of the four ABO SNVs using GTex and ENCODE Consortium resources identified that rs507666 resides near the transcription start site of a long non-coding RNA that is antisense to exon 1 of ABO and expressed in pancreatic islets. rs607666 is also an eQTL for the glucose transporter SLC2A4 (p=1.6x10^-6). Gene-based analyses identified a set of 15 SNVs associated with FG in the previously established G6PC2 locus. Using a similar approach, we performed haplotype meta-analysis showing significant association (p=1.1x10^-2) for haplotypes with the 15 SNVs. Haplotypes carrying the single rare allele at R283X (MAF=0.26%, p=2.8x10^-6), P32455 (MAF=0.19%, p=1.4x10^-5) or S207Y (MAF=0.59%, p=1.5x10^-6) were most strongly associated with FG compared to the common haplotype. The association at the three SNVs was not explained by the known common intronic variant in G6PC2 rs66887, which is close to the splice acceptor of intron 3 and may be implicated in G6PC2 pre-mRNA splicing. rs560887 is also near the transcription start site of expressed sequence tag DB031864, a potential cryptic minor isoform of G6PC2 mRNA. Together, these findings suggest that both coding and non-coding SNVs in G6PC2 affect glucose homeostasis. In conclusion, further follow-up on exome chip results with exome sequencing, functional annotation and haplotype meta-analysis offers new insight into the functionality of observed associations and reveals novel variants implicated in regulation of FG levels.
811S Exome Array Analysis of Quantitative Traits related to Glaucoma. A.I. Iglesias Gonzalez1, H. Springlek1,², S. van der Lee1, N. Amin1, C.C. Klei1,˙2, C.M. van Duijn1,1. 1) Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands; 2) Department of Ophthalmology, Erasmus MC, Rotterdam, The Netherlands.

**Purpose:** Intraocular Pressure (IOP) and Vertical Cup to Disc Ratio (VCDR) are directly linked to the development of glaucoma. However, there has been major progress identifying common variants with small effect, there has been limited progress identifying rare variants. To evaluate the association of exonic variants to IOP and VCDR in Caucasians, we analysed exome-sequencing data from genome-wide expression data in both a population-based and a family-based study.

**Methods:** The IlluminaHumanExome BeadChip containing ~250,000 variants was used to genotype 3,163 samples from the Rotterdam Study I (RS-I), a population-based cohort, and 1,512 subjects from the Erasmus Rucphen Family (ERF) Study, a family-based cohort. In addition, overlapping variants of ~1,879 individuals from RS-I (~600) and ERF (~1,279) with exome-sequencing data were included in the analysis. In total, 5,781 individuals with IOP data and 3,758 individuals with VCDR data were analysed. Rare variants were analysed for an association with IOP and VCDR using both single variant analysis and burden test implemented in SexMeta; an additive genetic model was assumed adjusting for age, sex, and the first five principal components (RS-I) or family structure (ERF). Variants with p-values <2×10⁻⁶ were selected for replication in RS-II and RS-III imputed with the Genome of the Netherlands (GoNL).

**Results:** Single variant analysis for IOP identified a novel common missense variant rs1154353 (p-value = 2.33E⁻¹¹) located in the NAPAS2 gene, involved in a circadian mechanism known as contrast sensitivity which is regulated by retinal ganglion cells. Our findings using classical method supported the association. Gene-based analysis showed significant association (p-value <1.85E⁻⁹) between IOP and three genes: RSNAHE2B, SPOPL and IVL when using a minor allele frequency (MAF) upper bound of 5%. Gene-based analysis for VCDR identified one novel significant association (p-value <1.85E⁻⁹) with CDK13 which belongs to the family of cyclin-dependent kinases and interact with CDKN2B, a key gene involved in glaucoma.

**Conclusions:** We have identified multiple genes associated with IOP. NAPAS2 is involved in the circadian clock mechanism and changes in clinical practice. The role of RSNAHE2B and SPOPL belongs to the ubiquitin ligase complex, and IVL is important for keratinocyte differentiation. In addition, in the VCDR analysis we found a new gene involved in the cell cycle pathway.

812M Relationship between neutrophil count, white blood count and TCIRG1 variation. E.A. Rosenthal1, V. Makaryam2, D.C. Daley1, D.R. Crosslin3, 1,4, 5, D.A. Nickerson3, A.P. Reiner3, 4, 5, G.P. James2, 1) School of Human Sciences, Univ of Wisconsin, Madison, WI; 2) Department of Medicine, LBL Gen, Univ of Washington School of Medicine, Seattle, WA; 3) Dept Genome Sciences, Univ of Washington, Seattle, WA; 4) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA; 5) Dept of Epidemiology, Univ of Washington School of Public Health, Seattle, WA.

Cognitive neutropenia is a hematological condition characterized by low neutrophil counts and recurrent bacterial infections. Recently, we reported linkage and association between congenital neutropenia and a novel TCIRG1 missense mutation (NG00778.1:c.2206C>A) in a five generation pedigree (PMID:24753205). We investigated the role of TCIRG1 SNVs in neutrophil count (NPC) and total white blood count (WBC), of which neutrophils are the major component, using data from the exome sequencing project (N=3560, WBC range 2x10⁵ to 11x10⁹ cells/Liter). After quality control, only 934 individuals (367 European ancestry (EA), 563 African ancestry (AA), 4 other ancestry) had measured NPC (NPC range 10-87% of total WBC, correlation with WBC = 0.40), reducing power to detect an association. Gene-based analysis showed significant association (p-value <1.85E⁻⁹) between NPC and three genes: RSNAHE2B, SPOPL and IVL when using a minor allele frequency (MAF) upper bound of 5%. Gene-based analysis for VCDR identified one novel significant association (p-value <1.85E⁻⁹) with CDK13 which belongs to the family of cyclin-dependent kinases and interact with CDKN2B, a key gene involved in glaucoma.

**Conclusions:** We have identified multiple genes associated with IOP. TCIRG1 is important for keratinocyte differentiation and in the VCDR analysis we found a new gene involved in the cell cycle pathway.

813T Identification of COPD causal variants by combining GWAS associated SNPs, lung eQTLs, and pathogenicity prediction tools. M. Lamontagne, 1 C. Labruyere, 1 P. Deschepper, 1 K. Hazout, 1 D. Nickle, 2 D. Sin, 2 PD. Pare4, 5 M. Laviolette, 1 Y. Bosse 1,2,3.

1) Centre de recherche de l’Institut universitaire de cardiologie et de pneumologie de Québec, Québec, Canada; 2) University of Goningen, University Medical Center Goningen, GRIAC research institute, Goningen, The Netherlands; 3) Merck & Co. Inc., Rahway, New Jersey, USA; 4) University of British Columbia James Hogg Research Center, Center for Heart and Lung Health, St. Paul’s Hospital, Vancouver, BC, Canada; 5) Department of Molecular Medicine, Laval University, Quebec City, QC, Canada.

**Introduction:** Chronic obstructive pulmonary disease (COPD) is the fourth most common cause of death worldwide and is predicted to be the third leading cause of mortality by the year 2030. COPD is a complex disease whose pathogenesis is poorly understood and is strongly influenced by genetic factors. Genome-wide association studies (GWAS) have identified polymorphisms associated with COPD and lung function. However, most of these SNPs are located in non-coding regions (intron and intergenic) and their mechanisms of action are unknown. The aim of the present study is to identify GWAS-associated SNPs that are also associated with gene expression levels in the lung and to verify whether these SNPs are in LD with possible pathogenic variants.

**Methods:** Genome-wide association studies on COPD and lung function were reviewed to identify significant SNPs. GWAS-associated SNPs were analyzed to identify expression Quantitative Trait Loci (eQTLs) in a large-scale lung eQTL mapping study including 1,111 patients that underwent lung surgery at three participating sites. eQTL analyses were performed independently in the three datasets and meta-analysed using PLINK. We used the Combined Annotation-Dependent Depletion (CADD) method to determine the potential pathogenicity of selected SNPs. Rare variants were analysed for an association with COPD and lung function from previous GWAS were available in the three cohorts. Thirty-three GWAS-associated SNPs were significantly regulating the expression levels of 30 genes (43 probe sets) after meta-analysis, with 11 genes associated with the pathogenesis of obesity and an increased risk of developing type 2 diabetes (T2D) in adulthood. Objective was to determine whether the IRS2 G1057D polymorphism gene is associated with obesity in young people.

**Results:** A public health problem in the country and is associated with insulin resistance (IR). The substrate insulin receptor 2 (IR2) in the signaling pathway of insulin is considered essential in the development and/or survival of the pathway of insulin. The G1057D polymorphism is involved in the development and/or survival of the pathway of insulin is considered essential in the development and/or survival of type 2 diabetes (T2D) in adulthood. Objective was to determine whether the IRS2 G1057D polymorphism gene is associated with obesity in young people.

**Methods:** A descriptive study was conducted on students of the Universidad Juarez Autonomia de Tabasco, Villahermosa, Tabasco, Mexico; 2) Hospital Civil de Guadalajara, Guadalajara, Mexico.

**Conclusion:** The G1057D polymorphism IRS2 gene is associated with obesity in young people.
815M Revealing the detailed MHC implication in seven common diseases from the WTCCC by HLA imputation. N. Vince1,2, A. Bashirova1,2, G. Wilson1, M. Carrington1,2. 1) Cancer and Inflammation Program, Laboratory of Experimental Immunology, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD; 2) Ragon Institute of MGH, MIT and Harvard, Cambridge, MA.

The Wellcome Trust Control Consortium (WTCCC) comprise a study of 7 complex human diseases each of which comprise approximately 2000 individuals: bipolar disorder (BD), coronary artery disease (CAD), Crohn's disease (CD), hypertension (HT), rheumatoid arthritis (RA), type 1 diabetes (T1D), type 2 diabetes (T2D). The genome wide association study (GWAS) published in 2007 showed various associations across the genome, and particularly strong associations within the MHC for RA and T1D. Most HLA alleles are not efficiently tagged by any GWAS SNPs, and therefore, HLA alleles effects can be missed in SNP analyses. To test for allelic effects, imputation approaches have been recently developed. We used the SNP2HLA software to impute the HLA class I and class II alleles in the 7 disease cohorts. Imputed HLA class I data was further applied to predict HLA-C level of expression. Each of the 7 disease cohorts were compared to 3000 controls from the WTCCC, all with UK ancestry. Logistic regression with stepwise selection as a statistical model was used to ascertain the independency of each association. As expected RA and T1D show the strongest association with HLA: HLA-DOA*10501 (P=7.1E-15, OR=0.43) and HLA-DRB1*1001 (P=7.2E-26, OR=0.18), respectively. In addition, HLA-DRB1*1003 shows a strong association with CD (P=9.0E-13, OR=3.51). The other 4 studied diseases have less significant associations with HLA, which need to be replicated elsewhere for confirmation: HT, HLA-DQB1*0602, OR=1.55; BD, HLA-DBP1*0101 (P=0.002, OR=0.72); T2D, HLA-B*3701 (P=0.003, OR=0.54); CAD, HLA-A*0205 (P=0.005, OR=0.44). Apart from the HLA alleles, the HLA-C expression level associates with CD: P=4.3E-6, OR=1.39. Thus, available GWAS data can be used for imputing HLA alleles and exploring the HLA allelic effects on human diseases.

816T A polymorphism in the peptidyl arginine deiminase type IV gene (PADI4) associated with radiographic joint destruction in patients with rheumatoid arthritis who are negative for anti-citrullinated peptide antibody (ACPA). K. Ikari, S. Yoshida, K. Yano, A. Taniguchi, H. Yamanaka, S. Momohara. Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan.

Rheumatoid arthritis (RA) is a complex polygenic disease of unknown etiology and is characterized by progressive joint damage. Anti-citullinated peptide antibodies (ACPA) are the most specific autoantibody for RA. Genetic polymorphisms in the PADI4 gene, encoding the peptidyl arginine deiminase 4, have been associated with susceptibility to RA in several populations. PADI4 polymorphisms have also been reported to be associated with joint damage in patients with RA. In the present study, we investigated whether a polymorphism in PADI4 is associated with radiographic joint destruction in ACPA-negative patients with RA.

This study used DNA samples from 122 Japanese ACPA-negative patients with RA who satisfied the American College of Rheumatology 1987 revised criteria for RA. Most of the patients were female (81.1%), 51.6% were rheumatoid factor (RF) positive, and the mean age was 55 years. We used the Sharp/van der Heijde score (SHS) of the hands at a 5-year duration to measure joint damage. Single nucleotide polymorphism (SNP) rs2240340 was selected for the study because it has the best evidence of association with RA in the Japanese population. Genotyping was performed by using a TaqMan assay. The genetic risk of joint damage associated with rs2240340 was assessed by multiple regression analysis adjusted for the possible genetic risk associated with the shared epitope and with the disease duration. The DR beta 1 major histocompatibility complex gene (HLA-DRB1). These alleles are thought to be associated with joint damage in RA patients. The PADI4 SNP was significantly associated with radiographic joint destruction in the Japanese population (P = 0.0035). The result remains significant when the analyses was adjusted by RF status (cutoff = 15 IU/ml) with HLA-DRB1 SE alleles (P = 0.0287). A SNP in PADI4 may contribute to joint destruction in ACPA-negative patients with RA. The PADI4 gene is likely to play a role in the disease progression of RA in addition to its role in the shared epitope. Further studies will be required to be able to determine the impact of PADI4 polymorphisms on radiographic joint destruction.

817S Risk for nonsyndromic cleft lip and palate from rare coding variants. K. Asrani1, W. Yang2, J. Rine1, E. Lammer2, G. Shaw2, N. Marini1. 1) California Institute for Quantitative Biosciences, University of California, Berkeley, CA; 2) Department of Pediatrics, Stanford University Medical Center, Stanford, CA; 3) Children's Hospital Oakland Research Institute, Oakland, CA.

Cleft lip and/or cleft palate are common craniofacial malformations with complex and heterogeneous etiologies, reflecting both genetic and environmental factors. Both linkage and association studies have shown multiple genes influence risk for clefts. Murine models of clefts as well as studies in human suggest that developmental genes involved in craniofacial development and craniofacial malformations are most likely to play a role in the disease progression of RA in addition to its role in the shared epitope. However, the list of causal genes is largely unknown. The list of newly identified genes is likely to include genes that have not yet to be identified. Association studies, whether case-control or family-based, have little power to detect unknown rare variants that may be causal. Thus, our goal was to identify rare, potentially causal variants among a list of randomly chosen candidate genes. To this end, we sequenced the exons of 51 candidate genes in 322 ethnically diverse cases and 150 non-malformed controls. 31 target genes are involved in folate metabolism and comprise nearly all cellular enzymes that utilize a folate cofactor. 20 additional genes have been implicated in disease risk by murine models of clefting (e.g. Wnt5b), studies of normal craniofacial development (e.g. Bmpr1b), or by multiple association studies (e.g. Irf6). Our findings indicate that rare, protein-altering variants (missense, nonsense, frameshift) confer risk for nonsyndromic clefts. First, the aggregate burden of rare, coding alleles was considerably higher in cases than controls. Overall, cases were 33% more likely to harbor a rare, protein-altering change in 51 genes (case mean = 1.52 mutations/individual; control mean = 1.15 mutations/individual). Second, mutations that are likely to result in functional impairment were significantly enriched in cases. 15 of 322 cases (5%) harbored one such change within this set of 51 genes, whereas no such changes were found in controls (0 of 150; P-value = 0.01). Third, in gene-level analyses, the burden of rare alleles showed a strong case bias for several genes/genome regions previously implicated in cleft risk. For example, within the Bmih7/Dmgd1 gene cluster on chromosome 5, 41 rare alleles were found in the case population versus 9 in controls (OR=2.3 95% CI = 1.1-4.9). Other loci with stronger allele burdens in cases were e.g. Bmpr4 and Bmpr1b. Thus, we conclude that rare coding variants may confer risk for isolated clefts. Additional rare and common variant association analyses on this dataset will be discussed.


The epithelial cadherin is a calcium-dependent cell adhesion molecule, essential to the formation of adherens junction. Mutations in its encoding gene, CDH1, have been largely associated with different types of cancer, including breast, lung, ovarian, renal, and thyroid cancer [1-3]. Recent studies have suggested a potential involvement of mutations in this gene with cleft lip with or without palate (CL/P) in individuals without known history of cancer in the family. Nonsyndromic CL/P is a complex trait, and, although many associations with common variants have been reported, they confer small effect and do not explain the whole heritability of the disease. To investigate the role of rare (<1%) CDH1 variants in susceptibility to nonsyndromic CL/P in the Brazilian population, we analyzed the coding regions and 3' exon-intron boundary of 31 target genes (sequenced by Sanger method) and 609 controls (from our in-house control exome database - sequenced with HiScan SQ - Illumina Inc; mean coverage of 100x). We found a total of 15 non-described variants in our affected sample and 14 in our control sample. We observed a 5, 41 rare alleles were found in the case population versus 9 in controls (OR=2.3 95% CI = 1.1-4.9). Other loci with stronger allele burdens in cases were e.g. Bmpr4 and Bmpr1b. Thus, we conclude that rare coding variants may confer risk for isolated clefts. Additional rare and common variant association analyses on this dataset will be discussed.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
819T

YJ. Lee -5

SNP rs657555 demonstrated association with RA (p = 6.13×10

EIRA and NARAC (AP = 0.208 [0.071-0.345] and AP = 0.209 [0.019-0.399],

DRB1 SE alleles in autoantibody positive RA in the two independent cohorts

We found

PTPN2

calculating the attributable proportion due to interaction (AP) between 11

(3151 patients with RA and 2247 matched healthy controls). Interactions were investigated by

(3151 patients with RA and 2247 matched healthy controls) and NARAC

two independent cohorts were used in this study, the Swedish EIRA study

Genotypes from

HLA-DRB1 SE alleles in development of RA.

Methods:

through the analysis of statistical interactions between SNPs in

PTPN2

and

PTPN2

We therefore aimed to identify additional susceptibility SNPs in

region remain undefined.

consequences of genetic variations in the

PTPN2

gene is associated with CALs in KD. Allele 8 conferred protection against

of Kawasaki disease. However, the functional

CTLA4

gene encodes a T-cell receptor, cytotoxic T-lymphocyte-associated protein 4 (CTLA4), which involved in controlling the proliferation and apoptosis of T lymphocytes. A microsatellite (AT)n repeat in the 3’ untranslated region has been reported to be associated with the organ-specific autoimmune disorders in several racial groups. CTLA4 expression is increased during the acute stage of KD and significantly higher than at the convalescent stage. This suggests that CTLA4 may be important in pathogenesis of KD or CALs. We investigated whether the CTLA4 gene was associated with CALs in Han Chinese children with KD. Material and methods The patients were 576 unrelated children (339 boys, 237 girls) with KD. Their age at diagnosis was 1.9 ± 1.7 years (range 0.1 - 10.2 years). We typed 3’UTR (AT)n microsatellite using fluorescence-based methods. Polymerase chain reaction (PCR) primers were 5’-GCCAGTGATGCTAAGGGTGG-3’ (forward) and 5’-ACACAAAAACATCGTGGCTC-3’ (reverse). The forward primer was labeled with fluorescent dyes. PCR products were gel-purified on an agarose gel and sequenced. Genotyping analysis was performed with GENESCAN 4.0. We designate the alleles according to the number of repeat units. Statistical analysis Statistical difference in allele distribution between patients with CALs and those without CALs were assessed by the chi-square test. Odds ratios and 95% confidence intervals were also calculated. Only those alleles of >2.0% in frequency in either patients or control were compared. The Bonferroni correction, Pc = 1 - (1/Pn), was used for multiple comparisons where P is the corrected P value, P the uncorrected value and n is the number of comparisons. In this study, n is 8 for each allele. A Pc value of less than 0.05 was considered statistically significant. Results The PCR products were 95-141 base pair (bp) in length corresponding to alleles 8-31. Alleles 16, 8, 15, and 17 were major alleles with a frequency of >10%. Allele 8 was 0.40 (19.8% in patients with CALs and significantly less frequent compared to that in patients without CALs which was 0.2752 [27.9%]; OR= 0.64 [0.47-0.85], P = 0.0023, Pc = 0.018. Conclusions The microsatellite (AT)n in the 3’UTR of the CTLA4 gene is associated with CALs in KD. Allele 8 conferred protection against the formation of CALs in patients with KD.

821M

Genetic variants of SMADs in the TGF-β/SMAD signal pathway are related specifically to susceptibility to ulcerative colitis in Japanese patients. T. Inamine1, S. Suzuki1, A. Yamashita1, S. Fukuda1, S. Kondo1, H. Isomoto2, K. Tsukamoto1. 1) Dept Pharmacotheapeutics, Nagasaki Univ Grad Sch, Nagasaki, Japan; 2) Dept Gastroenterol and Hepatol, Nagasaki Univ Grad Sch, Nagasaki, Japan.

PURPOSE Inflammatory bowel diseases (IBD), comprised of Crohn's disease (CD) and ulcerative colitis (UC), is attributed to inappropriate inflammatory response in the intestinal epithelia of patients. TGF-β/SMAD signals play a key role in differentiation of naive CD4+ T cells to Th17 cells or regulatory T (Treg) cells. In order to identify genetic determinants of IBD, we investigated an association between susceptibility to IBD and SMADs polymorphisms in the Japanese population. METHODS The study subjects consisted of 108 patients with UC, 81 patients with CD, and 199 unrelated healthy control subjects. A total of 21 tag single nucleotide polymorphisms (SNPs) in four genes (SMAD2, SMAD3, SMAD4, and SMAD7), which are involved in the TGF-β/SMAD signal pathway, were genotyped by PCR-restriction fragment length polymorphism, -direct DNA sequencing, or -high resolution melting curve analysis. The frequencies of alleles and genotypes were compared between control subjects and UC patients or CD patients by chi-square test or Fisher's exact test in three inheritance models: the allele, the minor allele dominant, and the minor allele recessive models. RESULTS Five SNPs (rs13381619, rs9955626, rs1792658, rs1792684, and rs1792671) of SMAD2, rs4147358 of SMAD3, two SNPs (rs7229678 and rs9304407) of SMAD4, and rs12956924 of SMAD7 showed the significant association with susceptibility only to UC. In each gene, rs13381619 of SMAD2, rs4147358 of SMAD3, rs3004407 of SMAD4, and rs12956924 of SMAD7, respectively, were associated with susceptibility to CD (P = 0.0409 in the minor allele recessive model). CONCLUSION Genetic variants of SMAD signaling molecules in the TGF-β/SMAD pathway may disturb the signal transduction and alter the balance of differentiation to Th17 and Treg, leading to dysregulation of immune response and eventually resulting in the development of IBD, especially UC. However, a genome-wide association study in European ancestry has shown the association of SMAD3 with susceptibility to CD. Although the TGF-β/SMAD signal pathway is crucial in the etiology of IBD, genetic variants of the key SMAD molecules in this pathway may be different between Caucasian and Japanese patients.

820S

Interactions between PTPN2 and HLA-DRB1 SE alleles in rheumatoid arthritis. M. Houtman, K. Shchelnytsky, L. Hadza, Hv. Rheumatology Unit, Department of Medicine Solna, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden.

Background: One of the most common chronic autoimmune disorders worldwide is rheumatoid arthritis (RA). It has a strong genetic component, and over 100 risk loci have been discovered in genome-wide association studies (GWAS). The major risk factor for RA is HLA-DRB1 shared epitope (SE) alleles. Outside of this region, a recently identified candidate gene for RA is protein tyrosine phosphatase non-receptor type 2 (PTPN2), a suggested phosphatase in cytokine receptor signaling. However, the functional consequences of genetic variations in the PTPN2 region remain undefined. We therefore aimed to identify additional susceptibility SNPs in PTPN2 through the analysis of statistical interactions between SNPs in PTPN2 and HLA-DRB1 SE alleles in development of RA. Methods: Genotypes from two independent cohorts were used in this study, the Swedish EIRA study (3151 patients with RA and 2247 matched healthy controls) and NARAC (673 patients with RA and 1196 controls). Interactions were investigated by calculating the attributable proportion due to interaction (AP) between 11 SNPS in the PTPN2 region and HLA-DRB1 SE alleles. Results: We found PTPN2 SNPs rs867555 and rs11080606 in statistical interaction with HLA-DRB1 SE alleles in autoantibody positive RA in the two independent cohorts EIRA and NARAC (AP = 0.208 [0.071-0.345] and AP = 0.209 [0.019-0.399], respectively). When individuals from the EIRA study were analyzed only PTPN2 SNP rs867555 demonstrated association with RA (p = 6.13×10

β/SMAD signal pathway, were genotyped by PCR- gel-purified on an agarose gel and sequenced. Genotyping analysis was performed with GENESCAN 4.0. We designate the alleles according to the number of repeat units. Statistical analysis Statistical difference in allele distribution between patients with CALs and those without CALs were assessed by the chi-square test. Odds ratios and 95% confidence intervals were also calculated. Only those alleles of >2.0% in frequency in either patients or control were compared. The Bonferroni correction, Pc = 1 - (1/Pn), was used for multiple comparisons where P is the corrected P value, P the uncorrected value and n is the number of comparisons. In this study, n is 8 for each allele. A Pc value of less than 0.05 was considered statistically significant. Results The PCR products were 95-141 base pair (bp) in length corresponding to alleles 8-31. Alleles 16, 8, 15, and 17 were major alleles with a frequency of >10%. Allele 8 was 0.40 (19.8% in patients with CALs and significantly less frequent compared to that in patients without CALs which was 0.2752 [27.9%]; OR= 0.64 [0.47-0.85], P = 0.0023, Pc = 0.018. Conclusions The microsatellite (AT)n in the 3’UTR of the CTLA4 gene is associated with CALs in KD. Allele 8 conferred protection against the formation of CALs in patients with KD.
822T

An estimated 50% of the variability in acquiring human immunodeficiency virus type 1 (HIV-1) upon exposure is attributable to host genetic factors. The only genetic polymorphism conclusively associated with risk of acquiring HIV-1, a low frequency 32-base pair deletion in the chemokine (C-C motif) receptor 5 (CCR5) gene, accounts for little of the population variability. Identification of other genetic polymorphisms underlying HIV-1 susceptibility is needed to better understand disease pathogenesis. The BST2 gene encodes a host restriction factor that suppresses the release of HIV-1 particles by tethering them to the cell surface based on in vitro studies. However, little is known about the influence of BST2 polymorphisms on HIV-1 acquisition and disease progression in vivo. In this study, we tested a dense set of polymorphisms spanning BST2 for association with HIV acquisition and viral load using 955 HIV antibody negative cases and 2,181 HIV antibody positive controls from the Urban Health Study of European American and African American people who inject drugs. We tested 470 single nucleotide polymorphisms (SNPs) and insertions/deletions (indels), genotyped or imputed from 1000 Genomes across BST2 and its flanking regions. The novel SNP rs113189798 was associated with HIV acquisition, exceeding the corrected P value threshold for multiple testing. The rs113189798-G allele, which occurred at frequencies of 16% in AAs and 4% in EAs, was associated with reduced risk of HIV-1 acquisition across both ancestry groups (meta-analysis P=1.43×10^-4); odds ratio (95% confidence interval) of 1.22 (1.01-1.49) in AAs and 2.17 (1.43-3.33) in EAs. This SNP is located 17.8 kb downstream of BST2. The SNP rs12609479, which was previously reported for affecting BST2 polymorphisms on HIV-1 acquisition and disease progression in vivo. In this study, we tested a dense set of polymorphisms spanning BST2 for association with HIV acquisition and viral load using 955 HIV antibody negative cases and 2,181 HIV antibody positive controls from the Urban Health Study of European American and African American people who inject drugs. We tested 470 single nucleotide polymorphisms (SNPs) and insertions/deletions (indels), genotyped or imputed from 1000 Genomes across BST2 and its flanking regions. The novel SNP rs113189798 was associated with HIV acquisition, exceeding the corrected P value threshold for multiple testing. The rs113189798-G allele, which occurred at frequencies of 16% in AAs and 4% in EAs, was associated with reduced risk of HIV-1 acquisition across both ancestry groups (meta-analysis P=1.43×10^-4); odds ratio (95% confidence interval) of 1.22 (1.01-1.49) in AAs and 2.17 (1.43-3.33) in EAs. This SNP is located 17.8 kb downstream of BST2. The SNP rs12609479, which was previously reported for affecting BST2 expression and subsequently decrease risk of acquiring HIV-1. Rs113189798 and rs12609479 are only weakly correlated (r^2=0.2-0.4), and represent distinct association signals. No significant associations were found for HIV viral load. Our findings provide support to BST2 as a genetic susceptibility factor for HIV-1 acquisition: identifying a novel SNP association for rs13189798 and linking the previously reported regulatory SNP rs12609479 to risk of acquiring HIV-1.

823S
Assessing genetic association between RASGRP3 and SLE susceptibility, X. Kim-Howard1, C. Sun1, A. Adler2, H. Zhang3, L.H. Lian2, K.H. Chua4, S.-C. Bae5, S. Natch1. 1) Oklahoma Med Res Foundation, Oklahoma City, OK; 2) Peking University First Hospital, Beijing, China; 3) Department of Molecular Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia; 4) Department of Biomedical Science, University of Malaya, 50603 Kuala Lumpur, Malaysia; 5) Hyundai University Hospital for Rheumatic Diseases, Seoul 133-791, Korea.

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disease with diverse clinical manifestations. Despite evidence for a distinct genetic architecture of SLE in Asians, little research has investigated the full spectrum of genetic variants that influence Asian SLE. Recent GWAS in Han Chinese identified association between SLE and RASGRP3 (rs13385731). RASGRP3 is responsible for Ras-ERK signaling mediated by B-cell receptor ligation in B cells and is involved in immunoglobulin production. It is not known whether rs13385731 is the primary variant responsible for SLE association with RASGRP3, or if it is in linkage disequilibrium with another causal variant. Objectives of this study were to: (a) perform comprehensive analysis using dense fine-mapping in 3 Asian populations, (b) identity robust, independent SLE-predisposing variants, and (c) test whether these variants are associated in ethnically diverse populations. Genotype data from 125 SNPs within RASGRP3 (chr2: 33,656,416-33,794,798 bp) were extracted from the ImmunoChip array, on 2487 cases and 3958 controls from Korean (KR), Han Chinese (CH), and Malaysian Chinese (MC) cohorts. To increase statistical power we performed ethnicity-specific imputation-based association analysis. For replication in other ancestries the most significant SNPs were confirmed by TaqMan assay in European-Americans (EA), African-Americans (AA), and Hispanics (HS). We also assessed association between significant SNPs and SLE clinical sub-phenotypes. After imputation we analyzed 162 SNPs from RASGRP3. The most significant association was at rs13425999 (P(META)=7.8E-8; EA=9.8E-5; P(CH)=1.0E-3; P(MC)=0.03; P(META)=4.5E-7). Published rs13385731 was significantly associated in EA and AA (P(za)=2.1E-3; P(AA)=2.2E-3) but not HS (P(HS)=0.68); rs13425999 was confirmed (P(KR)=9.8E-5; P(CH)=1.0E-3; P(MC)=0.03). SLE association with RASGRP3 was replicated in other sub-phenotypes. After imputation the association signals were consistent across EA, AA, and HS.

823M
Polymorphism in MEN1 gene is associated with increased risk and earlier age of pituitary adenoma development. J. Klovins1, R. Peculis1, I. Balcerie2, V. Pirags1,2. 1) Genome Centre, Latvia Biomedical Research and Study Centre, Riga, Latvia; 2) University of Latvia.

Clinically significant pituitary adenomas affect one individual out of approximately 1000 to 1300 people in general population. So far mutations in a number of genes responsible for familial cases of PA contributing up to 5% of all tumors have been described. Genetics of cases not clearly attributed to familial adenoma is unclear with multiple genes potentially involved. In this study, we analyzed 96 tag-SNPs from seven genes known for their involvement in familial PA cases and 3 receptor genes (AIP, SSTR2, SSTR5, DRD2, GNAS, MEN1 and PRKar1a) including their neighboring regions in 143 pituitary adenoma patients and 354 age and sex matched controls (1:2.5) in order to investigate potential involvement of these genes in determination of the risk for non-familial pituitary adenomas. We identified one rare SNP in MEN1 locus strongly associated (P=2.1e-4; OR =17.8; CI95[2.18-145.5]) with increased risk of PA development. The same SNP was also associated with younger age at diagnosis of PA (beta=-17.4). We also confirmed previously found association of SNP in SSTR5 gene and two nominal associations of SNPs in DRD2 gene with either risk of PA development or clinical characteristics of the disease. We thus have showed that SNP in MEN1, the gene known to be responsible for familial pituitary tumors also plays role in sporadic cases.
**825T**

Targeted resequencing of CFH-CFHR genes identifies new putative functional common and rare variants conferring susceptibility to Meningococcal disease. V. Kumar1, Z.Y. Phua1, T.W. Kuipers2, F. Marti-non-Torres3, A. Salas5,7, M.L. Hibberd1, E.D. Carroll5, W. Zen1, M. Levin2, S. Davila1. 1) Human Genetics, Genome Institute of Singapore, Singapore; 2) Division of Infectious Diseases, Department of Medicine, Imperial College London; 3) Division of Pediatric Hematology, Immunology and Infectious diseases, Emma Children’s Hospital Academic Medical Center, Amsterdam, The Netherlands; 4) Pediatric Emergency and Critical Care Division, Department of Pediatrics, Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain; 5) Grupo Gallego de Genética, Vacunases e Investigación Pediátrica, Instituto de Investigación Sanitaria de Santiago, Galicia, Spain; 6) Unidade de Xenética, Departamento de Anatomía Patológica e Ciencias Forenses Facultade de Medicina, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain; 7) Instituto de Medicina Legal, Facultade de Medicina, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain; 8) Infectious Diseases, Genome Institute of Singapore, Singapore; 9) Institute of Child Health, University of Liverpool, Alder Hey Children’s National Health Service Foundation Trust, Liverpool, UK; 10) Department of General Pediatrics, Medical University of Graz, Graz, Austria.

Meningococcal disease (MD) is an infection caused by Neisseria meningitidis. In recent years, genome wide association studies (GWAS) have suggested the involvement of a number of genes in determining host susceptibility and progression to disease. For instance, variants within the CFH-CFHR region have been robustly shown to be associated with susceptibility to disease. Yet the putative functional variants within this genomic region remain to be elucidated. Moreover, this reported association accounts for a small fraction of heritability and it is plausible that rare variants with moderate penetrance could contribute to disease susceptibility as well. CFH and CFHR genes share sequences of high homology within chromosome 1 that has made extremely challenging its analysis. Using a Nimblegen capture design we have been able to cover more than 92% of the targeted region (~300 kilobases). We sequence 238 MD cases and 237 controls from Western Europe and perform a case-control analysis. Calling of 3,032 high-quality single nucleotide polymorphisms (SNPs) with less than 10% sample missingness rate was obtained. Within the CFH region we have identified a SNP (in linkage disequilibrium with the top GWAS SNP, rs1065489) exceeding original GWAS P value. In addition we have identified a rare variant with P value < 10^-6, a genetic effect of the opposite direction to that of GWAS SNP. We are currently genotyping these SNPs in 1,300 MD cases and 3,000 controls to validate our results.

**826S**

Candidate-gene association study of sciatica. S. Lemmelä1, S. Solovieva1, R. Shiri1, M. Heliovaa2, J. Vilkan3, O.T. Raitakari1,5, T. Lehtimäki6, E. Vilkan-Jurtuna1, K. Huusafvel-Pursiainen1. 1) Health and Work Ability, Finnish Institute of Occupational Health, Helsinki, Finland; 2) Population Health Unit, National Institute for Health and Welfare, Helsinki, Finland; 3) Department of Medicine, University of Turku, and Division of Medicine, Turku University Hospital, Turku, Finland; 4) Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; 5) Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; 6) Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere School of Medicine, Tampere, Finland.

Sciatica is a common and often disabling low back disorder (LBD) in working-age populations. It presents as pain radiating from the back to the leg, usually caused by compression or irritation of one of the lumbosacral nerve roots. It is a complex disorder with relatively high heritability (35-75%) but poorly understood molecular mechanisms. Candidate gene studies have mostly focused on genes associated with LBD, such as those related to cartilage structure and stability, pain signaling, obesity or inflammation. We investigated 11 SNPs in genes previously associated with LBDs (sciatica, lumbar disc degeneration [LDD] or low back pain), osteoarthritis or nicotine addiction (smoking is a known risk factor for LBD). The study comprised two large Finnish population cohorts; Young Finns Study (YFS; 171 sciatica cases, 1777 controls) and Health 2000 (H2000; 294 sciatica cases, 5296 controls). Four candidate gene SNPs were genotyped in both populations; rs1107946 of COL1A1, rs2294995 of COL9A3, rs7775 of FRZB and rs1317286 of CHRNA5. Seven additional SNPs were genotyped in YFS; rs61734651 of COL9A3, rs1799007 of COL1A2, rs1800587 of IL1A, rs1800796 of IL6, rs731236 of VDR, rs288326 of FRZB, rs1669968 of CHRNA5 as well as rs2187689 in HLA region in H2000. Genotyping was done by qPCR using Taqman assays or OpenArray system. Allele and genotype frequency differences between cases and controls were estimated by Pearson’s X2-test with Yates’ continuity correction in R program. Odds ratios with 95% confidence intervals were estimated using R. We found an association with rs2187689 on HLA region (6p21.32) and sciatica in H2000 (p<0.04; OR=1.31, 95%-CI 1.02-1.69). This SNP was selected due to its previous association with LDD in a GWAS of Northern Europeans2. None of the other SNPs were significantly associated with sciatica. Further analysis taking into account work-related and lifestyle factors will be carried out. Current findings, based on relatively large population cohorts, are in concord with earlier studies in suggesting that common candidate genes may have a more limited role alone but a wider range of genes each with a subtle effect as well as gene-environment interaction are likely involved. GWA studies and exome/whole-genome sequencing approaches may aid in identifying novel candidate loci for low back disorders, including sciatica. 1Eskola, Lemmelä et al., PLoSOne 7(11):e49995, 2012. 2Williams et al., Ann Rheum Dis 72(7):1141-8, 2013.
827M
Association study between NOD2 and CCD122-LACC1 genes and leprosy in Brazilians. C.S. Marques1, H. Salomão2, V.M. Fava2, L.E.A. Armez3, E.P. Amaral3, C.C. Cardoso1,4, I.M.F. Dias-Batista3, W.L. Silva3, P. Medeiros3, M.C.L. Virmont3, F.C.F. Lana3, A.G. Pacheco3, M.O. Moraes3, M.T. Mira1, A.C. Pereira1, 1) Leprosy Laboratory, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Rio de Janeiro, Brazil; 2) Core for Advanced Molecular Investigation, Graduate Program in Health Sciences, School of Medicine, PUCPR, Curitiba, Brazil; 3) Departamento de Enfermagem, Materno-Infantil e Saúde Pública Federal de Minas Gerais, UFMG, Brazil; 4) Laboratório de Virologia Molecular, Departamento de Genética, UFPR, Uruçuca, Brazil; 5) Instituto Laboratório E. D. C. do Rio de Janeiro, Brazil;

Leprosy is a complex disease with phenotypes strongly influenced by genetic variations. Previously, a Chinese genome-wide association study (GWAS) identified non-synonymous single nucleotide polymorphisms (SNPs) at the leprosy susceptibility locus, which was only partially replicated by independent studies in different ethnicities. The aim of our study was to perform a validation and replication study of the Chinese GWAS in Brazilians, using a stepwise strategy that involved two family-based and three independent case-control samples, totaling 3,614 individuals enrolled. We selected 36 tag SNPs at five candidate genes from the Chinese study (CCDC122-LACC1, NOD2, TNFSF15 and RIPK2) which were tested in the discovery sample from Prata Village, a former leprosy colony. The DNA samples were genotyped using allelic discrimination Real Time PCR (TaqMan assay, Applied Biosystems).

The family-based association analysis was performed using Transmission Disequilibrium Test (TDT) in the FBAT software, version 2.0.2. In case-control studies, comparative analyses for allelic, genotype and carrier frequency were conducted. Gene-centric chip analyses were done to identify an under and over represented regions in R environment. Also, overall analysis combining the case-control samples were performed controlling for geographic region, gender and ethnicity. Linkage disequilibrium was estimated by Haploview software, version 4.2. In the discovery sample we observed an association between leprosy and two SNPs at NOD2 (rs8057431-A) and CCDC122-LACC1 (rs4942254-C) alleles, both under-transmitted to affected offspring, indicating protection to leprosy. After, these associations were consistently detected in all replication populations from Brazil. The combined analysis showed the following odds ratios towards protection: rs8057431-AA (OR=0.49, P=1.93e-06) and rs4942254-CC (OR=0.72, P=0.003). These results indicate an association between NOD2 and CCD122-LACC1 genes with leprosy protection in Brazilians, and suggest them as important markers to this disease across diverse populations.

828T
The role of SIRT2 in human longevity; converging evidence from gene expression, epigenetics and genetic variation. D.R. Mazzotti1, E.P. Amaral2, C.C. Cardoso2,4, I.M.F. Dias-Batista3, W.L. Silva3, P. Medeiros3, M.C.L. Virmont3, F.C.F. Lana3, A.G. Pacheco3, M.O. Moraes3, M.T. Mira1, A.C. Pereira1, 1) Leprosy Laboratory, Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Rio de Janeiro, Brazil; 2) Core for Advanced Molecular Investigation, Graduate Program in Health Sciences, School of Medicine, PUCPR, Curitiba, Brazil; 3) Departamento de Enfermagem, Materno-Infantil e Saúde Pública Federal de Minas Gerais, UFMG, Brazil; 4) Laboratório de Virologia Molecular, Departamento de Genética, UFPR, Uruçuca, Brazil; 5) Instituto Laboratório E. D. C. do Rio de Janeiro, Brazil; 6) Center for Applied Genomics - The Children’s Hospital of Philadelphia, Philadelphia, United States; 7) Center for Sleep and Circadian Neurobiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, United States.

The increase in life expectancy indicates that aging at the population level is unavoidable. The identification of factors that may help the development of interventions to promote healthy aging is fundamento. Using an in vitro approach, we aimed to identify whole blood oxidative stress gene expression changes and associated epigenetic mechanisms as well as genetic variants associated with human longevity. For the gene expression and epigenetic study, oldest old individuals (N=10, 85 to 105 years old), older adults (N=13, 60-70 years old) and young adults (N=15, 20-50 years old) had oxidative stress and antioxidant defense related gene expression and microRNA expression evaluated by qRT-PCR, and DNA methylation by bisulfite sequencing. For the association study, 2055 individuals (20-105 years) were genotyped at 1,727 single nucleotide polymorphisms (SNPs) and with 830M, SVEP1 c.2080A>C (p. Gin581His) gene is associated with altered mortality of septic shock. T. Nakada1,2, J. Russell1, J. Boyd1, S. Thair1, E. Nakada2, K. Walley2, 1) Department of Emergency and Critical Care Medicine, Chiba University Graduate School of Medicine, Chiba, Japan; 2) University of British Columbia, Critical Care Research Laboratories, Heart + Lung Institute, St. Paul’s Hospital, Vancouver, BC, Canada.

Background: Septic shock is a leading cause of death in intensive care units. Genetic factors are associated with altered mortality of septic shock. However the key genetic variations have not been fully elucidated. To identify the key genetic variations, we first identified non-synonymous single nucleotide polymorphisms (SNPs) in conserved genomic regions that are predicted to have significant effects on protein function. We then test the hypothesis that these variants across genome after clinical outcome of septic shock. Methods: Septic shock patients (n=520, European ancestry) were genotyped for 843 non-synonymous SNPs, which we identified in conserved regions of the genome and are predicted to have damaging effects from the protein sequence using PolyPhen2 and phastCons. The primary outcome is 28-day mortality. Productions of anoderm molecules including IL-8, GRO-alpha, MCP-1 and MCP-3 were measured in human umbilical vein endothelial cells (HUVECs) after SVEP1 gene silencing by RNA interference. We measured IL-8 levels in the plasma obtained on admission from the septic shock patients. Results: Of 843 non-synonymous SNPs, SVEP1 c.2080A>C (p. Gin581His, rs10817033) was significantly associated with altered 28-day mortality (Armitage trend test, uncorrected P=3.8 x 10^{-5}, Bonferroni corrected P=0.032). Patients with septic shock having the C allele of SVEP1 c.2080A>C had a significant increase in the hazard of death over the 28-day (hazard ratio 1.72, 95% confidence interval, 1.31 - 2.26, P=9.7 x 10^{-4}) and increased organ dysfunction, and needed more organ support (P<0.05). SVEP1 gene silencing increased IL-8, GRO-alpha, MCP-1 and MCP-3 in HUVECs under LPS stimulation (P<0.01). Non-survivors at day 28 had increased IL-8 levels (P=0.0005). There was a trend toward increased plasma IL-8 with increasing C allele of SVEP1 c.2080A>C, most notably among homogygotes carriers, though this was not statistically significant (CC vs. AC vs. AA, P=0.22). OR with SVEP1 C allele carriers was 4.52 and region surrounding SVEP1 was associated with sepsis susceptibility in protein sequence and is predicted to be damaging. Conclusions: C allele of SVEP1 c.2080A>C (p. Gin581His), a non-synonymous SNP in conserved regions and predicted to have damaging effects on protein structure, was associated with increased 28-day mortality of septic shock. SVEP1 may involve the leukocyte adhesion pathway.
831T
Generalization and fine-mapping of CDKN2B-AS1 for primary open-angle glaucoma in African Americans from the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study. N. Restrepo, R. Goodloe, E. Farber-Eger, D. Crawford. 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN.

Primary open-angle glaucoma (POAG) is the second leading cause of permanent vision loss and blindness in the U.S. Genome-wide and candidate gene association studies have identified several loci associated with POAG risk in populations of European descent. African Americans are ~15 times as likely to develop permanent vision impairment from glaucoma vs. European Americans, yet few studies have been performed in this population. To begin to fill these gaps in knowledge, we have performed a generalization and fine-mapping study for one POAG-associated region, CDKN2B-AS1, in African Americans from the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study as part of the larger Population Architecture using Genomics and Epidemiology study. A total of 242 CDKN2B-AS1 SNPs were targeted for genotyping using the Metabochip. African American POAG cases (n=135) and controls (n=1,739) were identified in BioVU, the Vanderbilt University Medical Center DNA repository. Cases and controls were identified via a combination of International Classification of Disease diagnostic codes, Current Procedural Terminology billing codes, and manual review of clinical records. We performed single SNP tests of association for common variants (MAF>0.05) using logistic regression assuming an additive genetic model adjusted for age and sex. We did not replicate associations previously described in European Americans for rs523096, rs564398, rs2157719, rs1412829, rs1063192 at a corrected threshold of 2.06x10^-4. These SNPs have minor allele frequencies (MAF) ranging from 0.37-0.45 in HapMap CEU populations compared to 7.9% among EAGLE BioVU African Americans. In addition to striking MAF differences, we observed differences in linkage disequilibrium (LD) for this gene between HapMap CEU and ASW data, and mostly low levels of LD in EAGLE BioVU African Americans when plotted using LocusZoom. Although none of the associations tested survived multiple correction testing, several were nominally associated with POAG including six variants in a 37kbp region (Chr 9: 22,105,026-22,068,305) that were associated at p<0.01 with ORs ranging from 1.38-3.21, uncorrected P=0.0006) and IHPS. Significantly reduced associations were observed between both MLNR SNPs and IHPS. Statistically significant increased associations were observed between both MLN rs1547668 (homozygous minor: OR=2.78; 95% CI=1.51-5.12, uncorrected P<0.001), MLNR genes and IHPS. Given the predominance of IHPS among non-Hispanic Whites, we used the population-based New York State (NYS) Congenital Malformations Registry to identify non-Hispanic White cases with IHPS born 1998-2005. A sample of sex- and racial/ethnic-matched controls born during the same time period was randomly selected from unaffected NYS live births. Case and control data were linked with NYS Newborn Screening Program data to obtain residual newborn blood spots. DNA was purified from de-identified blood spots for cases (n=656) and controls (n=656) and genotyped for single nucleotide polymorphisms (SNPs), identified by haplotype tagging, in the MLN (12 SNPs) and MLNR (1 SNP) genes. Using unconditional logistic regression, we estimated odds ratios (ORs) and 95% confidence intervals (CIs), corrected for multiple comparisons (P<0.0038), between each SNP genotype and IHPS. Significantly increased associations were observed between both MLN rs1547668 (homozygous minor: OR=2.78; 95% CI=1.51-5.12, uncorrected P<0.001), MLNR genes and IHPS. Significant decreased associations were observed for the MLNR SNP and IHPS. Using this large population-based sample of European ancestry, we observed significantly increased associations in two MLN SNPs and significantly decreased associations in another MLN/SNP and IHPS. The hormone encoded by the MLN gene is secreted by cells of the small intestine to regulate gastrointestinal contractions and motility. The associations identified for MLN SNPs provide insights into potential genetic susceptibilities for IHPS and support the observed associations between maternal and infant erythromycin use and this defect.
833M Intracranial Aneurysm Genetics: A south India perspective. S. Sathyans, L. Koshy1, H.V. Easwer2, S. Premkumar3, J. Alapatt4, S. Nair5, R.N. Bhattacharyya6, M. Banerjee7. 1) Human Molecular Genetics, Rajiv Gandhi centre for technology, Thiruvananthapuram, KERALA, India; 2) Department of Neurosurgery, SCTIMST, Thiruvananthapuram; 3) Department of Neurosurgery, Calicut Medical College, Calicut, India. Intracranial aneurysm (IA) is a fairly common condition that is often asymptomatic until the time of rupture. Rupture of cerebral aneurysm is the foremost cause for spontaneous Subarachnoid hemorrhage (SAH). In general population 2-3% of the individuals are likely to harbour intracranial aneurysm. Autopsy studies from India have shown the prevalence of IA ranges from 0.2% to 10.3% with a mean prevalence of 5.3%. Aneurysmal SAH is associated with mortality rate as high as 40% to 50%. Genetic and environmental factors are reported to influence development of IA. The genes involved in vascular remodelling, endothelial dysfunction and immune response converge in addressing various hypothesis that has been put forward for pathogenesis of IA. Till date five genome wide association studies have been carried pointing out involvement of 4q31.23, 8q21.3, 9p21.3 and 13q.13.1 9p21 genomic loci for IA. The objective of this study is to screen these hypothesis driven candidate genes and validate the positional regions in the GWAS studies in the pathogenesis of intracranial aneurysm in South Indian population. The study populations consisted of 225 radiologically confirmed aneurismatic cases and 236 ethnically and age matched controls from Dravidian Malayalam speaking population of Kerala. Genotyping was carried out based on allelic discrimination and Sequencing chemistry. In silico validation of functional effect was carried out. While evaluating the environmental factors our study show that history of hypertension (P-value<0.001, OR<2.98) and cigarette smoking (P-value<0.001, OR=3.58) where associated with intracranial aneurysm while diabetes showed an protective effect on intracranial aneurysm (P-value=0.009, OR=0.34 (0.15-0.76). While evaluating the pathway focused genetic associations for possible endophenotyping we observe that MMP2, COL1A2 and VCAN genes of extracellular matrix remodeling and proinflammatory cytokines of TNFalpha and IFN gamma of immune response are associated with IA in South Indian population. Their functional implication and interaction in developing a precise ECM/proinflammatory endophenotype would be interesting. Among positional GWAS hits located in 4q31.23, 8q21.3, 9p21.3 and 13q.13.1 9p21 harbouring EDNRA, SOX17, CDKN2BAS and STARD13 respectively we could not replicate any of these regions with exception to 8p21.3 to be associated for IA in South Indian population. This is the first study on genetics of IA in any Indian population.

834T Association between IRF6 polymorphisms and 8q24 region in non-syndromic cleft lip with or without cleft palate in Brazilian population. L.T. Souza1,2, T.K. Kowalski2, J. Ferrari2, I.L. Monleò2, E.M. Ribeiro3, J. Souza3, G.F. Leal3, A.C. Fett-Conte3, V.G. Silva-Lopes2, A.C.K. Ribeiro-dos-Santos4, S.E.B. Santos2, T.M. Felix1,2,3,5. 1) Programa de pós-graduação em Saúde da Criança e do Adolescente-Universidade federal do Rio Grande do Sul, UFRGS, Brazil; 2) Laboratório de Medicina Genômica, Hospital de Clínicas de Porto Alegre, HCPA, Brazil; 3) Projeto Crânio-Face Brazil, Brazil; 4) Laboratório de Genética Humana e Médica, Universidade Federal do Pará, UFPA, Brazil; 5) Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, HCPA, Brazil. Cleft lip and palate (CL/P) has a multifactorial inheritance. Identification of genetic risk factors has been the subject of intensive research. Three polymorphisms, rs2235371 and rs642961 in IRF6 and rs987525 in 8q24, have been associated with CL/P risk in several studies. Variants in IRF6 gene are responsible for 12% of clefts, the first variant associated was rs2235371 (V274I) and it is in linkage disequilibrium with rs642961. The rs642961 was suggested to cause disruption of the binding site of transcription factor AP-2c. The function of 8q24 is still unknown. A variant rs987525 (C>A) have involved in the pathogenesis of CL/P in Caucasians and admixed populations. The aim of this study was to evaluate the association between three polymorphisms and CL/P in the Brazilian admixture population. In this study were included subjects with non-syndromic CL/P and their parents that from three different geographical regions of Brazil. This research was approved by the Research Ethics Committee of Hospital de Clínicas de Porto Alegre and all the individuals signed an Informed Consent. For ancestry analysis, a panel of 48 INDELS was selected to measure the proportions of three different ancestries in the probands with CL/P. All selected markers were analyzed by multiplex. The SNP were analyzed using Taqman assay (Applied Biosystem). Statistical analysis was performed with FBAT and Haplin softwares. Acrety was analyzed in 228 probands. We select triads whose probands had European ancestry >=0.6 to further genetic analysis of the polymorphisms. We analyzed 151 nuclear families totaling 390 individuals. MAF for rs2235371 A allele was 0.08, 0.19 for rs642961 A allele and 0.41 for rs987525 A allele. TDT analysis showed an overtransmission of G allele in probands with CL/P. Haplotype analysis in IRF6 showed association for the children haplotype (rs2235371G and rs642961A) in single and double dose and mother haplotype (rs2235371G and rs642961A) for single dose. This data confirms association of rs2235371 G allele in IRF6 and rs987525A allele in 8q24 region with non-syndromic CL/P, similar to the previous studies in Caucasian and Mixed populations. The association between these polymorphisms and CL/P is due to European ancestry. We found association with the haplotype (rs2235371 G and rs642961 A) for non-syndromic cleft lip and/or palate. This is the first study that was able to show this association in Brazil.
Background: The prevalence of childhood obesity is increasing worldwide and has become a major health problem. Identification of susceptibility genes in early life could provide the foundations for interventions in lifestyle to avoid obese children to become obese adults. Genome-wide association studies (GWAS) have led to the identification of several loci in the human genome containing genetic variants conferring an increased risk of developing overweight and obesity. Furthermore, association studies in children and adolescents remain useful in the investigation of candidate genes for obesity, in which rare variants might be identified and responsible for unexplained phenotypic variation. In this study we evaluated the influence of genetic variants related to obesity identified by GWAS (MC4R, TMEM18, SH2B1, SEC16B) and association studies (ADIPOQ, PPARG, LEPR) on anthropometric phenotypes and food intake in a cohort of 424 children followed-up since birth until 3 years old. Methods: The polymorphisms were genotyped using real-time polymerase chain reaction and the dependent variables were compared among genotypes at the ages of 1 and 3 years old by t-tests and analysis of variance. Multilevel mixed models were also used to analyze genotype effects on the patterns of individual change in anthropometric and dietary variables over time. Results: In this sample 52.6% were boys. At 3 years, overweight prevalence was 43.1%. TMEM18 rs8548238 was associated with BMI Z-score at 3 years and over the years (P=0.029, P=0.011; respectively). We also found associations with intake of lipid dense foods at 1 year (P=0.045) and total energy intake over the years (P=0.017). MC4R rs17782313 was associated with intake of sugar dense foods at 1 year and intake of lipid dense foods over the years (P=0.023, P=0.020; respectively). SH2B1 rs7488665 was associated with intake of sugar dense foods at 3 years (P=0.029). SEC16B rs10913469 was associated with sum of skinfolds at 1 year in girls (P=0.028). Association between ADIPOQ rs17300539 and BMI-Z score was observed at 1 year (P=0.022) and ADIPOQ rs6548238 was associated with sum of skinfolds at the years (P=0.027). Conclusions: This study provides indications that genetic variants in TMEM18, MC4R, SH2B1, SEC16B and ADIPOQ genes might be associated with anthropometric phenotypes and food intake in young children.

835S 836M The QT-interval prolonging variant p.D85N of KCN1 associates with reduced levels of insulin after an oral glucose load. A. Jonsson¹, M.N. Harder¹, A. Stancakova², A.P. Gjesing¹, T.S. Ahluwalia³, ⁴, N. Granup⁵, S.S. Torekov¹, ⁵, C. Graff¹, ⁶, A. Linneberg⁷, T. Jørgensen⁷, ⁸, T.I.A. Sørensen¹, ⁹, J.J. Holst¹, ¹, M. Laakso², J.K. Kanters³, ⁴, O. Pedersen¹, T. Hansen¹, ¹. ¹ The Novo Nordisk Foundation Center for Basic Metabolic Research; Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; ² Department of Medicine, University of Eastern Finland, Kuopio, Finland; ³ Copenhagen Prospective Studies on Asthma in Childhood, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; ⁴ The Danish Pediatric Asthma Center, Gentofte Hospital, the Capital Region, Copenhagen, Denmark; ⁵ Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark; ⁶ Department of Health Science and Technology, Aalborg University, Aalborg, Denmark; ⁷ Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark; ⁸ Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; ⁹ Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, the Capital Region, Copenhagen, Denmark; ¹⁰ Gentofte, Aalborg and Herlev University Hospitals, Hellerup, Denmark; ¹¹ Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.

Aim/hypothesis: Mutations that cause functional changes in an ion channel may result in different diseases, such as cardiac arrhythmias and beta cell dysfunction, depending on which tissue the channel is expressed in. KCN1 is expressed in both cardiomyocytes and in pancreatic islets and a low frequent functional variant in this gene, KCN1 p.D85N (CEU MAF 0.02), has been shown to affect the QT-interval. We aimed to test for associations between the QT-interval-associated allele of KCN1 p.D85N and insulin release from pancreatic beta cells. Methods: Associations of the KCN1 p.D85N variant with QT-interval and insulin levels were studied in 5,738 Danish individuals from the Inter99 study without known diabetes. An association with 2-h insulin was taken forward for replication in a total of 1,662 Danish and Finnish non-diabetic individuals and the results were combined in a meta-analysis. The beta cell response to an intravenous load of glucose and tolbutamide was examined in 303 individuals from the Danish Family study. Results: The QT-interval-prolonging allele of KCN1 p.D85N associated with increased QT-interval (β=0.023, p=1.4×10⁻⁵⁷) and decreased 2-h insulin (β=−0.149, p=0.013) in Inter99. The association with decreased 2-h insulin stayed significant in the combined meta-analysis (β=−0.174, p=7.6×10⁻⁷). This variant was also associated with decreased intravenous tolbutamide-induced insulin secretion (incremental AUC (mean±SD) 1,574±1,744 vs. 2,398±1,625, p=0.036) in the Danish Family study. Conclusion/interpretation: We found that heterozygote carriers of the low frequent QT-interval-prolonging allele of KCN1 p.D85N have lower insulin levels after an oral glucose tolerance test and intravenous tolbutamide injection suggesting that these carriers have both impaired glucose and tolbutamide-stimulated insulin secretion.
837T
The role of selected ion channel genes in dental caries. D. Lewis1, J. Shaffer2, E. Feingold2,3, M. Cooper2,4, M. Vanyukov3,5,6, B. Maher7, S. Reis8,9, W. McNeel10, R. Roux11, R. Weyant12, S. Levy13,14, A. Vieira5,4, M. Marazita1,3,4,6,7. 1) Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA; 2) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3) Center for Craniofacial & Dental Genetics, Department of Oral Biology, School of Dental Medicine, Pittsburgh, PA; 4) Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 5) Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA; 6) Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA; 7) Department of Mental Health, John Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 8) Clinical and Translational Science Institute, School of Medicine, University of Pittsburgh, Pittsburgh, PA; 9) Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA; 10) Dental Practice and Rural Health, West Virginia University, Morgantown, WV; 11) Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV; 12) Department of Dental Public Health and Information Management, School of Dental Medicine, University of Pittsburgh, Pittsburgh PA; 13) Department of Preventive and Community Dentistry, University of Iowa College of Dentistry, Iowa City, IA; 14) Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA.

Ion channels play an important role in regulating and maintaining the calcium and pH homeostasis that is critical for tooth development. Mutations in genes encoding for ion channels can be related to a heterogeneous group of diseases called "channelopathies." One common channelopathy is Timothy Syndrome, a Mendelian disease that affects many parts of the body including the heart, fingers and toes, the nervous system and causes small, misplaced teeth and frequent cavities in children. As part of a larger candidate gene study we investigated 480 single-nucleotide polymorphisms (SNPs) in several ion channel genes, including CACNA1C, CACNA2D1, CACNB2, CACNG2, KCNH1, KCNK5, and KCNK17, several of which are known to play a role in Timothy syndrome or other channelopathies. We tested association of these genes with dental caries in 13 race and age-stratified cohorts from six independent studies of Caucasians and African Americans. We performed analyses independently for each cohort and synthesized results by meta-analysis. After gene-wise adjustment for multiple testing, two SNPs for CACNA2D1 were significantly associated with dental caries via meta-analysis across the five childhood cohorts and in one individual childhood cohort (p-values<0.0007). In adults, genetic association was observed in three individual cohorts for potassium channel genes KCNH1 and KCNK5 (p-values<0.001), but no single SNP was significant via meta-analysis across all eight adult cohorts. These findings strengthen the hypothesis that ion channel genes, particularly those involved in channelopathies, may affect the risk of dental caries. Grants:R01-DE014899, U01-DE018903.

838S
Analysis of Haptoglobin Duplication with Type 2 Diabetes and Diabetic End Stage Kidney Disease. JN. Adams1,2,3, LM. Raffield1,2,3, AJ. Cox2,4, ET. Barton1, CD. Langefeld2, BI. Freedman3,5, MCY. Ng2,3, ND. Palmer2,3, DW. Bowden2,3,4. 1) Molecular Genetics and Genomics, Wake Forest School of Medicine, Winston Salem, NC; 2) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC; 4) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC; 5) Division of Public Health Sciences, Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC; 6) Department of Internal Medicine - Nephrology, Wake Forest School of Medicine, Winston-Salem, NC, 27157.

Haptoglobin (HP) is an acute phase protein that binds to freely circulating hemoglobin. HP assists in the removal of hemoglobin from the circulatory system, thus preventing its accumulation in the kidney. This accumulation could lead to kidney damage and oxidative damage caused by the heme iron. The HP protein exists as two distinct forms, HP1 and HP2. The longer HP2 form arose from a duplication of exons 3 and 4 and has been associated with cardiovascular disease (CVD) events and mortality in individuals with type 2 diabetes (T2D). HP has also been reported to be associated with T2D in multiple ethnicities, and with end stage kidney disease (ESRD) in individuals with type 1 diabetes. This novel duplication is not tagged or genotyped on conventional genotyping arrays and thus requires direct genotyping of the locus or Western blot. In this study we investigated the association of the HP duplication with T2D and diabetic end stage kidney disease (DMESKD) in African-Americans. HP genotyping was performed by PCR amplification using two sets of allele-specific primers, followed by resolving the products by agarose gel electrophoresis. Genotyping was performed in 4560 African American individuals: 2022 with T2D and DMESKD, 891 with T2D only, and 1647 controls with neither disease. The HP duplication was then analyzed for association with T2D and DMESKD. The HP duplication was not associated with T2D in this population (p=0.23). However, analyses revealed that the HP2 allele was associated with increased DMESKD risk with an odds ratio (OR) of 1.18 (95% confidence interval (CI) 1.06-1.28, p=0.0012) when comparing individuals with DMESKD to control individuals. When comparing individuals with DMESKD to individuals with T2D, the HP2 allele remained associated with DMESKD with an OR of 1.15 (95% confidence interval (CI) 1.05-1.25, p=0.00183). This study suggests an association between the HP2 allele and susceptibility to DMESKD in the African-American population. HP is a strong biological candidate for kidney disease. HP has been shown to decrease oxidative damage of the kidney. However, the HP2-2 protein has been shown to have decreased antioxidant properties compared to HP1-1. In addition the HP2-2-hemoglobin (Hb) complex is cleared much slower than the HP1-1-Hb complex. This leads to an accumulation of iron in the proximal tubule cells of the kidney. This accumulation leads to increased oxidative stress and damage to the tubule, and increased kidney disease.
Type 2 diabetes (T2D) is a metabolic disease characterized by both insulin resistance and impaired insulin secretion. African Americans (AAs) have higher disease prevalence of T2D (12.6%) compared to European Americans (7.1%). Although GWAS of primarily European populations have identified ~70 loci associated with T2D, these loci are usually represented by a common index variant spanning a large region of linkage disequilibrium (LD). GWAS in a single population may not have sufficient resolution to identify causal SNPs, address the impact of rare variants, or assess genetic architecture underlying ethnic disparities in disease prevalence. With these considerations, we examined SNP- and locus-wide association of 44 T2D loci reported in Europeans (P<5x10^-8, odds ratio [OR] 1.07-1.39) for association in an AA cohort consisting of 1617 T2D cases and 676 healthy controls. Considering the pathophysiology of T2D, we also examined 29 reported European glucose homeostasis (GH) loci (P<5x10^-8). Individuals were genotyped using an Affymetrix BioBank array customized for fine-mapping T2D and GH candidate genes/loci. Our results show that 16 T2D index SNPs, or their proxies (r^2>0.5) at the THADA, GRB14, IRS1, WFS1, ANKRDR55, JAZF1, CDKN2A/B, TCF7L2, KCNQ1, MTNR1B, KLDHC5, HMGAG, TSPAN8/LGR5, FTO, MCHR1, and CILP2 loci were significantly associated with T2D (P<0.05, OR 1.17-1.55) in AAs. Additionally, 6 GH index SNPs or proxies at the DNA/TMEM185, GLIS3, ADRA2A, MADD, FADS1, and VPS13C/CDC44A/B loci were significantly associated with T2D (P<0.05, OR 1.19-1.35). The strongest association was TCF7L2 rs7903146 (P=0.005). This finding is consistent with current literature. Of the 22 associated loci, the most strongly associated SNP in AAs was identical to or in LD (r^2>0.5) with the European index SNP at 10 loci: GRB14, IRS1, ANKRDR55, JAZF1, KCNQ1, FTO, TCF7L2, MADD, FADS1, and VPS13C/CDC44A/B. Gene-based analyses including both common and rare coding variants in SKAT revealed significant association at the IRS1 (P=0.025) and TSPAN8 (P=0.020) loci, but these associations were lost when only SNPs with predicted functional consequences were included in the model. These findings suggest that loci discovered in European GWAS influence T2D risk in AAs with similar effect sizes, that the lesser degree of LD in AAs may facilitate lower disease prevalence of T2D (12.6%) compared to European Americans.

Candidate Genes for Non-syndromic Orofacial Clefts Identified by GWAS Were Assessed in Two African Populations. A. Butali1,2,3, P. A Mossey4,5,6, W.L. Adeyemo3,4, E. Mekonnen4,6, L. Gaines1, T. Busch1, R.O. Braimah6, S.B. Aregbesola1, W.L. Adeyemo3,4, E. Abate1, T. Hailu1, M. Ibrahim1, P. Gravenm1, M. Deribew1, M. Gessesse3, A. Adeyemo4,5, M.L. Marazita1, J.C. Murray1. 1) University of Iowa, Iowa City, IA, U.S.A; 2) University of Dundee, UK, 3) University of Lagos, Nigeria; 4) Addis Ababa University, Ethiopia; 5) Obafemi Awolowo University, Nigeria; 6) National Institutes for Health, Bethesda, MD, U.S.A; 7) University of Pittsburgh, Pittsburgh Pennsylvania, U.S.A. Genome-wide association studies (GWAS) for non-syndromic cleft lip with or without cleft palate (NSCL/P) have identified 14 new loci. These loci have been replicated in several studies confirming the role of common variants in increasing risk to NSCL/P. Rare variants in these candidate genes have also been reported in resequencing studies. Here, we present analysis of common variants reported in the GWAS studies and rare variants identified following genotyping and Sanger sequencing in two sub-Saharan African populations from Nigeria and Ethiopia. Genotyping of 24 common variants in these 14 loci was done using Fluidigm technology which allows multiplexing of samples. Genotyping of 228 pedigrees was completed with a 95% genotyping call rate. Transmission disequilibrium tests and parent of origin analyses were conducted using PLINK. We sequenced the coding regions of the following genes: MAPF, PAX7, VAX1, ARHGAP29, and IRF6 in 220 probands from Africa [140 Nigerians and 80 Ethiopians] in order to identify rare functional variants. The TDT results for the common variants were not significant (p>0.05). However, we observed significant paternal effects for SPRRT2 rs9574565 (X2=4.5; p=0.03). We also found new rare variants — p.His165Asn in the NAFB gene, p.Asp268Gly in the PAX7, a splice-site variant that creates a new donor splice-site in PAX7 following sequencing. All these variants segregate in parents. It is important to conduct an unbiased genome-wide association and sequencing studies using samples from this understudied African population in order to understand the genetic background of blood lipids may be largely shared with that of glycaemic and anthropometric traits. Recent GWAS have identified ~100 SNPs associated with lipid levels, but knowledge of their relevance to other metabolic traits is incomplete. We sought to assess the association of lipid SNPs with metabolic traits in 4,202 participants of the Fenland cohort study. We genotyped or imputed a total of 137 SNPs identified in a recent large-scale GWAS. Using linear regression models, we studied the association of individual SNPs and genetic risk scores of lipid traits (i.e. triglycerides, total-, HDL- and LDL-cholesterol) with detailed glycaemic and anthropometric phenotypes. These included fasting and post-challenge plasma glucose, insulin, HbA1c as well as anthropometric measurements from ultrasound and dual-energy X-ray absorptiometry (DEXA). Quantile-quantile plots of the association of SNPs with metabolic traits were systematically inflated, highlighting considerable pleiotropy. A 63-SNP HDL-increasing genetic score was associated with multiple measures of reduced glycaemia, insulin resistance and adiposity. Individual SNPs with evidence of pleiotropic effects included those at known metabolic-trait or type 2 diabetes (T2D) loci (e.g. FTO, MC4R, FADS1-2-3, FAM13A and GCKR). The SNPs with the strongest effect across multiple phenotypes were rs1121980 at FTO and rs629301 at SORT1 - a known LDL and coronary artery disease (CAD) gene encoding an insulin-sensitive intracellular sorting protein. The LDL-reducing allele of rs629301 was associated with increased fasting insulin (beta [SE] in SD per allele, 0.07 [0.03]), HOMA-insulin resistance index (0.07 [0.03]) and body fat mass (0.06 [0.03]). DEXA-scan analyses also revealed increased fat mass in each body compartment - in particular the trunk (0.07 [0.03]) - while bone or lean masses were unaffected. In the publicly-available CARDiogram and DIAGRAMv3 datasets, the SNP was associated with CAD (OR [95% CI], 0.90 [0.87-0.93]) and T2D (1.04 [1.00-1.08]) in opposite directions, consistent with the association observed in the corresponding reference T2D traits - LDL-cholesterol (beta [SE], -0.13 [0.03]) and fasting glucose (0.05 [0.03]). Our study highlights opposite effects of SORT1 on cholesterol levels and glycaemic traits. This finding is relevant to discussions on the development of lipid-lowering drugs targeting sorting.
Interaction of immune-related genetic polymorphisms and breastfeeding duration with Helicobacter pylori prevalence: the Pasitos Cohort Study. M.L. Grove1, K.A. Volck1, E.A. Brown1 1) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX; 2) Department of Biochemistry and Molecular Biology, University of Texas Medical School at Houston, Houston, TX; 3) Center for Infectious Disease, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX.

Helicobacter pylori (H. pylori) is a bacteria that colonizes in the mucosal layers of the gastric and sometimes duodenal or esophageal epithelium in up to 50% of the human population. In 3% of individuals, H. pylori contributes to the progression of gastric cancer which is the second leading cause of cancer mortality worldwide. Numerous risk factors have been associated with H. pylori infection risk, including genetic variation and environmental contributors such as breastfeeding. Therefore, we investigated the associations of immunological candidate gene loci with H. pylori prevalence within the context of breastfeeding (BF) status. We examined 87 SNPs in 27 immune-related genes in 239 Hispanic infants from the Pasitos Cohort Study which were ascertained from two populations residing on the United States (US) and Mexico border at El Paso, Texas and Juarez, respectively. H. pylori infection prevalence was ascertained using serum measures, breastfeeding status was categorized as <6 months or ≥6 months duration, and genetic variation was captured with the Sequenom MassARRAY® system using established genotyping methods. Additive logistic regression models were used to test each single nucleotide polymorphism (SNP) by country of origin which included gender, BF status, and a SNP×BF multiplicative interaction term. Models with statistically significant interactions were further stratified by BF categories and the genotypes assessed for additive effects including gender as a covariate. We observed significant SNP×BF interactions for three variants in the US population (CCRF P=0.003, IL15 P=0.004, and SFTPD P=0.042), and IL6R (P=0.037) in the Mexican population. After stratification of these loci by BF status, variants in CCRF and SFTPD were shown to be associated with 25% and 26% lower prevalence in infants BF ≥6 months (rs2022001, odds ratio [OR]=0.25, 95% confidence interval [CI]=1.00-0.70, P=0.006; and rs721917, OR=0.26, 95% CI=0.08-0.80, P=0.019), respectively. Homozygote infants BF <6 months with the risk-raising allele in the US population and IL6R in the Mexican population were 2.6 and 2.1 times more likely to be infected with H. pylori (rs10519613, OR=2.66, 95% CI=1.49-4.77, P=0.001; and rs6427641, OR=2.12, 95% CI=1.09-4.14, P=0.028), respectively. To our knowledge, we are the first to report variations in immune-related genes may interact with BF status to influence H. pylori prevalence in Hispanic infants.

Why do some athletes with sickle cell trait suffer from heat illness? A.C. Stone1, R. Grieger1, J. Lund1, M. Ciambella1, C. Flansburg2, L. Madrigal1 1) Human Evolution & Social Chg, Arizona State Univ, Tempe, AZ; 2) Department of Anthropology, University of South Florida, Tampa, FL.

The NCAA recently declared sickle cell trait (SCT) to be a risk factor for sudden illness and death among student athletes. Although its effect on SCT is not fully understood, fetal hemoglobin (HbF) concentration in adults is negatively correlated with disease severity in sickle cell anemia, and the concentration of HbF in adults is found to be highly variable across populations. In order to test whether lower HbF levels in some SCT athletes increases their risk for illness during exercise, we conducted preliminary analysis of five single nucleotide polymorphisms (SNPs) located in the human beta globin gene cluster (rs7482144 and rs10128556), the HBS1L-MYB intergenic interval (rs9402686), and the BCL11a gene (rs10189857 and rs4671393). These SNPs contribute to the heritable variation in HbF levels and are associated with increased HbF concentrations in adults. A sample (n=22) of NCAA football student athletes was genotyped at these SNPs, and their allele frequencies were compared to those of other populations. For rs7482144 and rs10128556 in the beta globin gene cluster, the minor allele frequencies of both were 0.091 in the sample population. The minor allele frequencies for rs9402686, rs10189857, and rs4671393 were 0.07, 0.34, and 0.41, respectively. These results compared closely with other populations of recent African heritage. The results of this study will be combined with data from medical records to assess whether these polymorphisms can be used to predict susceptibility to exercise related illness in NCAA student athletes with SCT.
845M IDENTIFICATION OF SIGNIFICANT ASSOCIATION AND GENE-GENE INTERACTIONS OF POLYMORPHISMS IN THREE INFLAMMATORY GENES CRP, TNF-α, AND LTA FOR LOWER EXTREMITY PERFORMANCE IN COMMUNITY-DWELLING ELDERS IN TAIWAN-TAICHUNG COMMUNITY HEALTH STUDY FOR ELDERS (TCHS-E). T.C. Li1,2, C.C. Lin3,4, C.I. Li5,6, N.H. Meng4,5, W.Y. Lin6,4, C.S. Liu4,4, C.H. Lin6,4, Y. Wang7, C.K. Chang6, F.Y. Wu4, L.N. Liao4,1) Graduate Institute of Biostatistics, China Medical University, Taichung, Taiwan; 2) Department of Healthcare Administration, College of Health Science, Asia University, Taichung, Taiwan; 3) Department of Family Medicine, China Medical University Hospital, Taichung, Taiwan; 4) School of Medicine, College of Medicine, China Medical University, Taichung, Taiwan; 5) Department of Medical Research, China Medical University Hospital, Taichung, Taiwan; 6) Department of Physical Medicine and Rehabilitation, China Medical University Hospital, Taichung, Taiwan; 7) Ph.D. Program for Aging, College of Medicine, China Medical University, Taichung, Taiwan; 8) Department of Public Health, College of Public Health, China Medical University, Taichung, Taiwan.

Purpose: We conducted a cross-sectional study to assess the main effects and possible gene-gene interaction of three inflammatory cytokine gene polymorphisms tumor necrosis factor α (TNF-α), lymphotoxin α (LTA) and interleukin-6 (IL-6) on markers of appendicular skeletal muscle mass by investigating nine single-nucleotide polymorphisms (SNPs) of these three genes in community-dwelling elders in Taiwan. Methods: Among leg lean muscle mass (LMS, kg) and appendicular skeletal muscle mass (ASM, kg), were measured by Dual-energy X-ray absorptiometry. Height-adjusted skeletal muscle index (hSMI; kg/m²) was defined as ASM divided by height squared. Weight-adjusted skeletal muscle index (wSMI; %) was tallied as ASM divided by weight. Height-adjusted lower extremity index (LtEI; %) was calculated as total lower extremity muscle mass (LMM; kg) divided by height squared. Significant genotypic association was found for SNP rs2239704 (in LTA; P <0.05) on leg LMS, ASM, hSMI, and wSMI in men and for SNP rs1880243 (in IL-6; P <0.05) on leg LMM in women. We also identified significant two-locus gene-gene effects involving rs909253 in LTA gene and rs1799964 in TNF-α gene for arm and leg LMS, ASM, hSMI, and wSMI in men; and for SNP rs1880243 (in IL-6; P <0.05) on leg LMM in women. We also identified significant two-locus gene-gene effects involving rs909253 in LTA gene and rs1799964 in TNF-α gene for arm and leg LMS, ASM, hSMI, and wSMI in women. Conclusions: These results support the hypothesis that inflammatory genes are involved in lean muscle mass markers, most likely via complex gene-gene interactions. In addition, there exists gender difference in the association between inflammatory genes and lean muscle mass markers.
847S
Genetic risk of rheumatoid arthritis conferred by HLA-DRB1 in African Americans stratified by local ancestry. R.J. Reynolds1, N.M. Rajewszki2, S. Raychaudhuri3,4, R.M. Plenge3, S.L. Bridges1. 1) University of Alabama at Birmingham, Birmingham, AL; 2) Wake Forest University, Winston Salem, NC; 3) Brigham and Women’s Hospital, Boston, MA; 4) Broad Institute, Cambridge, MA; 5) Merck & Co.

HLA-DRB1 Val11 is the residue with highest risk for rheumatoid arthritis in Caucasians. We sought to validate the rheumatoid arthritis (RA) risk of HLA-DRB1 exon 2 amino acid positions and residue substitutions in African Americans. We inferred local ancestry (copies of ancestral alleles) flanking HLA-DRB1 to test the hypothesis that RA risk from residues at amino acid position 11 is independent of European ancestry at this locus. 579 seropositive RA patients and 893 healthy controls were four digit HLA-DRB1 genotyped and all samples were genotyped on the Immunochip. Ancestral population reference panels of Africans (Zambia) and Europeans (1200 haplotypes each) were genotyped on the Immunochip, and for all samples, the extended major histocompatibility complex of chromosome six was phased using BEAGLE. Local ancestry was estimated with HAPMIX. Association analysis was performed using PLINK. Binomial regression models were fit in R to test the hypothesis that the Val11 and Asp11 residues had equivalent risk with local ancestry. Local European ancestry (P-value = 9.02E-04), but not global European admixture (P-value = 0.07) was associated with increased RA risk. Amino acid position 11 was the most strongly associated position (P-value <1E-05), but positions 71 and 74 were not significantly associated. Within position 11 the valine, OR(CI) = 3.0 (2.3, 4.0), and aspartic acid, OR(CI) = 1.9 (1.3, 2.7), residues conferred RA risk. The aspartic acid residue is specific for HLA-DRB1 *09:01. There was no joint effect of local European ancestry and Val11 on RA risk (P-value = 0.69). In contrast, there was an interaction effect of local ancestry and Asp11 on RA risk (P-value = 8.59E-03). Stratified by two copies of African ancestry, the observed Asp11 risk increased from OR 1.9 to 2.6 (1.7, 4.1). Asp11 was not a significant risk residue for individuals with either 1 or 2 copies of local European ancestry. These results demonstrate that the major source of RA risk comes from HLA-DRB1 amino acid position 11 and the valine residue, which does not vary by African or European ancestry. However we also found evidence of heterogeneity in RA risk between Caucasians and African Americans for both HLA-DRB1 amino acid positions 71 and 74 and the Asp11 residue. Importantly, we document that local ancestry of class II HLA genes may mediate the individual genetic risk for immune system related diseases.

848M
THE MAIN EFFECTS AND GENE-GENE INTERACTIONS AMONG CRP, TNF-α AND LTA IS ASSOCIATED WITH HANDGRIP STRENGTH IN COMMUNITY-DWELLING ELDERS IN TAIWAN-TAICHUNG COMMUNITY HEALTH STUDY FOR ELDERS (TCHS-EC). T.C. Li1, W.Y. Lin1, L.N. Liao1, F.Y. Wu1, S.C. Lin2,3, C.C. Lin2,3, C.H. Lin2,3, W.Y. Lin2,3, C.I. Li4,5, C.W. Yang6,7, N.H. Meng6,8, C.K. Chang6,8, T.C. Li4,5, L.N. Liao1. 1) Department of Public Health, China Medical University Hospital, Taichung, Taiwan; 2) Department of Family Medicine, China Medical University Hospital, Taichung, Taiwan; 3) School of Medicine, College of Medicine, China Medical University Hospital, Taichung, Taiwan; 4) Ph.D. Program for Aging, College of Medicine, China Medical University, Taichung, Taiwan; 5) Department of Medical Research, China Medical University Hospital, Taichung, Taiwan; 6) Department of Physical Medicine and Rehabilitation, China Medical University Hospital, Taichung, Taiwan; 7) Graduate Institute of Biostatistics, China Medical University, Taichung, Taiwan; 8) Department of Healthcare Administration, College of Health Science, Asia University, Taichung, Taiwan.

Purpose: We undertook a cross-sectional study to assess the main effects and possible gene-gene interactions of inflammatory gene polymorphisms C-reactive protein (CRP), tumor necrosis factor α (TNF-α) and lymphotixin α (LTA) on handgrip strength, by investigating eleven single-nucleotide polymorphisms (SNPs) of three genes in community-dwelling elders in Taiwan. Methods: Eleven SNPs (rs2794520, rs1205, rs1130864, rs1800947, and rs3093059 in CRP gene; rs1799964, rs1800629, and rs3093662 in TNF-α gene; rs2239704, rs909253, and rs1041981 in LTA gene) were utilized to genotype 472 unrelated elderly subjects (221 women and 251 men; mean age, 73.8±6.1 years). Handgrip strength was measured by handgrip dynamometer (TM-110D, TTM Co. Japan). We systematically utilized Hardy-Weinberg equilibrium (P >0.05). For the comparison of handgrip strength among the genotype, dominant and recessive models, the linear regression models were adjusted for age, BMI, smoking and physical activity. In the dominant models, there were significant differences in handgrip strength in the G allele of CRP rs3093059 in women (P <0.05). In women, CRP rs2794520 and LTA rs2239704 were significant markers for handgrip strength. In the genotype and recessive models, men carrying the AA genotype of TNF-α rs1800629 had greater handgrip strength; women carrying the GG genotype of CRP rs1205 had lower handgrip strength. We identified significant gene-gene interactions of CRP rs1205 & LTA rs2239704 (P <0.05). In women, as CRP rs3093059 & LTA rs2239704 in men, were significant interactions on handgrip strength. Conclusion: Our study show that an observed main effect and interaction between CRP, TNF-α and LTA polymorphisms on handgrip strength in elders.
A gene-gene interaction in a shared Alzheimer disease/age-related macular degeneration pathway. M.W. Logue1,2,3, M. Schu1, J. Farell4, K.L. Lunetta5,6, G. Jun1,2,4, C.T. Baldwin1, M.M. DeAngelis5, L.A. Farrell1,2,4,6,7, 1) Biomedical Genetics, Boston University School of Medicine, Boston, MA; 2) Biostatistics, Boston University School of Public Health, Boston, MA; 3) Research Service, VA Boston Healthcare System, Boston, MA; 4) Ophthalmology, Boston University School of Medicine, Boston, MA; 5) John A. Moran Eye Center, University of Utah, Salt Lake City, UT; 6) Neurology, Boston University School of Medicine, Boston, MA; 7) Epidemiology, Boston University School of Public Health, Boston, MA.

There is a growing body of biochemical and epidemiological evidence that Alzheimer disease (AD) and age-related macular degeneration (AMD) share pathogenic mechanisms. Recently, we found that risk genes for both disorders are enriched in clathrin-mediated endocytosis (CME), LXR/RXR activation, and atherosclerosis signaling pathways (Logue et al., 2014). Examining these pathways in summary GWAS data, we identified several novel AMD loci including HGS and TNF. Here, we explored the possibility that AD risk variants modify the effect of AMD risk variants within these three pathways by testing gene-gene interaction models in HapMap 2 imputed GWAS data for 1,336 AMD cases and 1,121 controls of European ancestry from the MMAP Study cohort (Chen et al., 2010; Fritsche et al., 2013). We analyzed the peak AD variants (or LD proxies) in PICALM, CD2AP, APOE, and CLU for interaction with the peak AMD risk loci in APOE, CSNK2B, HGS, TNF, CETP, C3, C4A, and COL10A1, restricting AD and AMD risk variant interactions to genes within one of the 3 pathways (22 independent tests; significance determined using robust SE estimates). Using a logistic-GEE model of AMD risk which included an AMD risk variant and AD risk variant, a QxG interaction, and age as a covariate, we found significant interaction (p = 0.00019; exceeding a Bonferroni-corrected threshold) within the CME pathway between the known AMD risk variant rs10948363 in CSNK2B and the known AD risk variant rs10948363 G allele (MAF=46% in 1000 Genomes EUR) and even reverses (rare homozygous state) the risk of AMD conferred by the rs10948363 G allele. The rs10948363 G allele, which involves genes in the CME pathway (CD2AP, CETP, C3, C4A, and COL10A1), was nominally significant (p<0.01). These findings indicate that AD and AMD share genetic mechanisms and that the clathrin-mediated endocytosis pathway may also play an important role in AMD pathogenesis.

Targeted regulome sequencing reveals common, rare, and private regulatory variants associated with fetal adiposity. C. Guo1, I. McDowell2, A.A. Pab3, OM. Scholleners2, GE. Crawford1,2, BE. Engelhardt1,2, MG. Hayes3, W.L. Lowe4, TE. Reddy1,4. 1) Institute for Genome Science and Policy, Duke University, Durham, NC; 2) Department of Biology, Massachusetts Institute of Technology, Cambridge, MA; 3) Department of Preventive Medicine, Division of Biostatistics, Northwestern University Feinberg School of Medicine, Chicago, IL; 4) Department of Pediatrics, Division of Medical Genetics, Duke University, Durham, NC; 5) Department of Biostatistics & Bioinformatics, Duke University Medical School, Durham, NC; 6) Division of Endocrinology, Metabolism and Molecular Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL.

Fetal growth has immediate and long-term effects for the health of the newborn. In particular, small and large for gestational age babies are at risk for later life obesity and type-2 diabetes. In our previous genome-wide association study of 4281 newborns, we demonstrated association of variation between CCNL1 and LEKR1 in 3q25.31 with fetal adiposity. As the associated variants were not in protein coding exons, we hypothesized that they alter the expression of nearby genes. To test this hypothesis, we sequenced regulatory elements within 250kb of the lead GWAS SNP, encompassing all variants in strong LD (r2 > 0.8) with rs990400, across 760 babies in the upper and lower deciles of fetal adiposity in the Hyperglycemia and Adverse Pregnancy Outcome Study. The babies represented four ancestry groups: European, Afro Caribbean, Mexican-American, and Thai. To define regulatory elements for sequencing, we used genome-wide measures of open chromatin across ~50 different cell types including pre-adipocytes, liver, and iLE-β cells. Open chromatin regions are a general indicator of transcription factor binding and gene regulation, and several studies have demonstrated that a substantial fraction of complex phenotypes involve genetic variation in those regions. In total, we sequenced ~100 regulatory elements at a median coverage of 1100x for each sample. We found 995 variants (921 SNPs and 74 indels), 560 of which are not found in dbSNP or the 1,000 Genomes Project. To assess the regulatory function of the identified variants, we mapped eQTLs for protein coding genes and long-noncoding RNAs in four different cell types using data from the GTEx project and in lymphoblastoid cell lines from the gEUVADIS dataset. Several variants in our sequencing are eQTLs for nearby genes, including two long-noncoding RNAs. Together, by combining targeted regulome sequencing and genome-wide eQTL analyses, these results suggest a potential role for novel genes including long-noncoding RNAs in fetal adiposity.

Genome-wide Sequencing to Identify Novel Variants for Obesity in Pima Indians. K. Huang, P. Piaggi, S. Kobes, R. Hanson, C. Bogardus, L. Baier. Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ.

To identify genetic variation that increases risk for type 2 diabetes (T2D) or obesity, we obtained whole genome sequence data on 335 Pima Indians (51%; male, age: 25.1±5.8 years, BMI: 33.5±7.3 kg/m2). Sequencing was performed by Illumina (N=301) and Complete Genomics, Inc (N=34). ~13 million variants were found, including ~11 million SNPs, ~1.6 million indels and 255,802 substitutions. Among all SNPs, 2.7 million were novel. Individuals who were sequenced were part of a study of health among Pima Indians living in the Gila River Indian Community, and were informative for longitudinal measures of BMI and type 2 diabetes status. As a preliminary screen for selecting variants for follow-up genotyping, we performed association analyses between all novel common SNPs (N=71,021 with a MAF>0.05 and a genotype calling rate>0.85) and maximum BMI childhood z-score (age 5-20 years; N=287) and maximum BMI in adulthood (after age 15 years from an exam where the individual was non-diabetic, N=289). The most significant preliminary association with childhood z-score was a novel exonic SNP in ZNF595 (chr4:86248, A/C, MAF=0.15; p=3.2×10^-12 after adjusted for birth year, β=1.2 units per copy of the risk allele). In addition, variants in CHRNA4 and OR4A15/OR4C15 showed genome-wide significant association with childhood z-score (smallest p=1.7×10^-10 after adjusted for birth year, β=1.1 units per copy of the risk allele) and adulthood BMI chr11:55226415, G/A, MAF=0.15; p=2.0×10^-9 after adjusted for age, sex and birth year, β=1.0 kg/m2 per copy of the risk allele), respectively. These novel variants and others selected for having the lowest p values or being positioned within the excellent biologic candidates will be genotyped in 5,880 American Indians for validation of the preliminary associations with z-score during childhood and BMI during adulthood.
852T Fine-mapping eGFR susceptibility loci through trans-ethnic meta-analysis. A. Mahajan1, J. Haessler2, N. Franceschini3, A. Morris1,4,5. 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 2) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; 3) University of North Carolina, Chapel Hill, NC, USA; 4) Department of Biostatistics, University of Liverpool, Liverpool, UK; 5) Estonian Genome Center, University of Tartu, Tartu, Estonia.

Chronic kidney disease (CKD) is a major public health problem and affects nearly 10% of the global population. Reduced estimated glomerular filtration rate (eGFR), a measure of kidney function used to define CKD, is associated with cardiovascular disease morbidity and mortality, acute kidney injury, and progression to end stage renal disease. Genome-wide association studies (GWAS) have been successful in identifying loci for eGFR. However, these loci are typically characterised by common lead SNPs with association signals extending over large genomic intervals containing multiple transcripts. As a result, limited progress has been made in identifying causal variants for eGFR and understanding the downstream pathogenesis of CKD.

To address these drawbacks, we performed trans-ethnic meta-analysis to fine-map known eGFR loci by leveraging differences in distribution of linkage disequilibrium between diverse populations. We considered six GWAS comprising of 23,568 individuals of European, African American, and Hispanic ancestry, each supplemented by imputation up to the 1000 Genomes Project reference panel (March 2012 release). Within each study, association with eGFR (MDRD equation) was tested under an additive model. We then combined association summary statistics across studies with MANTRA, 500kb up and down of the lead SNP at known eGFR loci, and constructed “credible sets” of SNPs that encompass 99% of the posterior probability of being causal. We resolved fine-mapping of potential causal variants to less than 500kb up and down of the lead SNP at known eGFR loci, and constructed fine-mapping panels of sequence variants for eGFR and understanding the downstream pathogenesis of CKD.

To address these drawbacks, we performed trans-ethnic meta-analysis to fine-map known eGFR loci by leveraging differences in distribution of linkage disequilibrium between diverse populations. We considered six GWAS comprising of 23,568 individuals of European, African American, and Hispanic ancestry, each supplemented by imputation up to the 1000 Genomes Project reference panel (March 2012 release). Within each study, association with eGFR (MDRD equation) was tested under an additive model. We then combined association summary statistics across studies with MANTRA, 500kb up and down of the lead SNP at known eGFR loci, and constructed “credible sets” of SNPs that encompass 99% of the posterior probability of being causal. We resolved fine-mapping of potential causal variants to less than 500kb up and down of the lead SNP at known eGFR loci, and constructed fine-mapping panels of sequence variants for eGFR and understanding the downstream pathogenesis of CKD.

To address these drawbacks, we performed trans-ethnic meta-analysis to fine-map known eGFR loci by leveraging differences in distribution of linkage disequilibrium between diverse populations. We considered six GWAS comprising of 23,568 individuals of European, African American, and Hispanic ancestry, each supplemented by imputation up to the 1000 Genomes Project reference panel (March 2012 release). Within each study, association with eGFR (MDRD equation) was tested under an additive model. We then combined association summary statistics across studies with MANTRA, 500kb up and down of the lead SNP at known eGFR loci, and constructed “credible sets” of SNPs that encompass 99% of the posterior probability of being causal. We resolved fine-mapping of potential causal variants to less than 500kb up and down of the lead SNP at known eGFR loci, and constructed fine-mapping panels of sequence variants for eGFR and understanding the downstream pathogenesis of CKD.

To address these drawbacks, we performed trans-ethnic meta-analysis to fine-map known eGFR loci by leveraging differences in distribution of linkage disequilibrium between diverse populations. We considered six GWAS comprising of 23,568 individuals of European, African American, and Hispanic ancestry, each supplemented by imputation up to the 1000 Genomes Project reference panel (March 2012 release). Within each study, association with eGFR (MDRD equation) was tested under an additive model. We then combined association summary statistics across studies with MANTRA, 500kb up and down of the lead SNP at known eGFR loci, and constructed “credible sets” of SNPs that encompass 99% of the posterior probability of being causal. We resolved fine-mapping of potential causal variants to less than 500kb up and down of the lead SNP at known eGFR loci, and constructed fine-mapping panels of sequence variants for eGFR and understanding the downstream pathogenesis of CKD.

To address these drawbacks, we performed trans-ethnic meta-analysis to fine-map known eGFR loci by leveraging differences in distribution of linkage disequilibrium between diverse populations. We considered six GWAS comprising of 23,568 individuals of European, African American, and Hispanic ancestry, each supplemented by imputation up to the 1000 Genomes Project reference panel (March 2012 release). Within each study, association with eGFR (MDRD equation) was tested under an additive model. We then combined association summary statistics across studies with MANTRA, 500kb up and down of the lead SNP at known eGFR loci, and constructed “credible sets” of SNPs that encompass 99% of the posterior probability of being causal. We resolved fine-mapping of potential causal variants to less than 500kb up and down of the lead SNP at known eGFR loci, and constructed fine-mapping panels of sequence variants for eGFR and understanding the downstream pathogenesis of CKD.
854M
Meta-analysis on the 22q11.21 region identifies an autoimmune disease risk allele as associated with systemic lupus erythematosus. Y. Zhang1, Y. Wang2, J. Yang1, N. Hirankarn2, X.J. Zhang2, Y.L. Lau1, W. Yang1,2. 1) Paediatrics, The University of Hong Kong, Hong Kong, Hong Kong; 2) Lupus Research Unit, Department of Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; 3) State Key Laboratory Incubation Base of Dermatology, Key Laboratory of Dermatology, Anhui Medical University, Ministry of Education, China, Heifei, Anhui, 230032, China; 4) Centre for Genomic Sciences, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong.

Systemic lupus erythematosus (SLE) is an autoimmune disease with unclear etiology. It usually presents with a diverse spectrum of clinical manifestations. Genetic predisposition is responsible for 50 to 60 percent of the likelihood of developing the disease. Genome-wide association studies (GWASs) on SLE have identified more than 40 loci with genome-wide significance. The 22q11.21 genomic region is a shared locus for several autoimmune diseases including SLE, systemic sclerosis (SSc), Crohn’s disease (CD), celiac Disease (CeD), Rheumatoid Arthritis (RA), Psoriasis (PS) and Inflammatory bowel disease (IBD). Here we examined the association for SLE in this region, and emphasized on the relationship of all the reported variants with various autoimmune diseases. Through meta-analysis of two existing GWASs on Chinese Han populations with a total of 1,659 cases and 3,398 controls matched geographically, we have identified SNP rs2298428 as the most significant SNP for SLE in the 22q11.21 region (P = 2.17e-09). The association of rs2298428, a missense mutation in YDJC gene, was further replicated in three cohorts from Hong Kong, Anhui and Thailand, independent of the GWAS samples used (P = 3.1e-11, OR = 1.23).

This SNP also showed consistent correlation with the expression of UBE2L3 in multiple cell types, highlighting the functional mechanism of the abnormal expression of UBE2L3 to SLE.

855T
Involvement of GTF2IRD1 in the complex hearing phenotype of Williams-Beuren Syndrome. C.P. Canales1, A.C.Y. Wong2, G.D. Housley3, E.C. Hardeman1, S.J. Palmer1. 1) Cellular and Genetic Medicine Unit, School of Medical Sciences (SOMS), UNSW Australia, Sydney, Australia; 2) Translational Neuroscience Faculty & Department of Physiology, School of Medical Sciences (SOMS), UNSW Australia, Sydney, Australia.

Amongst the spectrum of physical and neurological defects in Williams-Beuren Syndrome (WBS), it is common to find a distinctive response to sound stimuli that includes extreme adverse reactions to loud or sudden sounds (usually called ‘hyperacusis’), and a fascination with certain sounds that may manifest as strengths in musical ability. However, hearing tests have indicated that sensorineural hearing loss (SNHL) is frequently found in WBS patients and the genetic origins of this unusual auditory phenotype are currently unknown. However, the use of the term ‘hyperacusis’ is inaccurate since hyperacusis implies increased functional sensitivity leading to detectably lower hearing thresholds, and auditory allodynia, meaning averse to, or fear of certain sounds that are usually acceptable to others, is a more accurate description of this condition. Here, we investigated the involvement of GTF2IRD1, a gene located within the WBS deletion that has been implicated as a contributor to the WBS neurocognitive profile and craniofacial anomalies, in WBS. Using mouse models of knockout mice, we analysed the expression of the gene in the inner ear and examined hearing capacity. Rather than using acoustic tests in mice that rely on behavioural responses, which could be confounded by the altered anxiety responses previously reported in Gtf2ird1 knockout mice, we used objective physiological assays under general anaesthesia to evaluate hearing capacity, minimizing the behavioural component. Using auditory brainstem response (ABR), which measures sound-evoked auditory neurotransmission from the cochlear nerve to the auditory midbrain, and distortion product otoacoustic emissions (DPOAE), which measures sound-evoked electromechanical amplification via the outer hair cell ‘cochlear amplifier’, we demonstrated that Gtf2ird1 null mice have hypoacusis (higher hearing thresholds) in both assessments. These data, together with the Gtf2ird1 expression pattern in diverse cell types within the inner ear, indicate that the principal hearing deficit in the mouse can be traced to impairments in cochlear amplifier, suggesting that similar mechanisms may underpin the SNHL experienced by WBS patients. The origin of the auditory allodynia remains unexplained but may form part of a central defect in the processing of sensory input and/or emotional control mechanisms, which GTF2IRD1 may also contribute to.

856S
Genetic variant at ETS1 locus increases lupus risk and affects Stat1 binding. L. Kottyan1,2, X. Lu1,2, E. Zoller1, M. Weirauch1,2, B. Namjou1, K. Greis1, N. Shen1,2, K. Kaufman1,2, J. Harley1,2,3. 1) The International Consortium on the Genetics of Systemic Lupus Erythematosus (SLEGEN); 2) Center for Autoimmune Genomics and Etiology, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH; 3) United States Department of Veterans Affairs Medical Center, Cincinnati, Ohio, USA; 3) Immunology Graduate Program, University of Cincinnati College of Medicine, Cincinnati, OH; 4) Cancer Cell Biology, University of Cincinnati Medical Center, Cincinnati, OH; 5) Institute of Rheumatology, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Over the past 10 years, genetic association studies have identified over 55 systemic lupus erythematosus (SLE)-risk loci. Genetic variants in ETS1 have been associated with SLE in several independent studies in populations of Asian ancestry. Several recent studies have implicated ETS1 as a critical driver of immune cell function; indeed, mice deficient in ETS1 develop an SLE-like autoimmunity. rs1128334, in the 3’ UTR of ETS1, has been associated with the decreased expression of ETS1. We performed a fine-mapping study of over 20,000 subjects of two multi-ancestral cohorts using genotyped and imputed variants spanning the ETS1 locus. By constructing genetic models using frequentist and Bayesian association methods, we identified a set of variants that are most likely to be causal. Of the three variants that we tested, only rs6590330 differentially binds lysate from B cell lines as assessed by electrophoretic mobility shift assays and DNA affinity precipitation assays. We found that the active form of the transcription factor STAT1 binds the risk allele of rs6590330 but not the non-risk allele using mass spectrometry, a finding confirmed by Western blot analysis. rs6590330 is in strong linkage disequilibrium with rs1128334 and is predicted to also be associated with decreased ETS1 expression. We propose a model in which the risk allele of rs6590330 increases SLE risk by binding pSTAT1 and depressing the expression of ETS1.

857M
Rs738409 Polymorphism in PNPLA3 Gene is Associated with Lower Insulin Resistance in Korean Men. JH. Park1, JM. Yun1, HT. Kwon2. 1) Family medicine, Seoul National University Hospital, Seoul, South Korea, MD; 2) Family Medicine, Healthcare Research Institute, Seoul National University Hospital Healthcare System Gangnam Center, Seoul, South Korea, MD.

Background & Aims: The rs738409 polymorphism in adiponutrin is a well-known genetic risk factor for nonalcoholic fatty liver disease (NAFLD) development. When we consider that NAFLD is closely associated with insulin resistance, the association between rs738409 variant and insulin resistance is highly suspected. Methods: We enrolled 1,189 Korean men who visited the Family Medicine Center of Seoul National University in Seoul, South Korea from December 2009 to June 2012 and 1,189. Hepatic steatosis was evaluated by abdominal ultrasound and subjects with secondary causes of NAFLD or alcohol consumption of more than 30g/day were excluded from analysis. The rs738409 genotyping was performed using TaqMan assay on a ViATM 7 Real-Time PCR System. Results: The participants were predominantly middle-aged men (49.1 ± 7.0 years; range, 30-60 years), and the frequencies of NAFLD and insulin resistance were 43.7% and 17.5%, respectively. The minor allele frequency of the rs738409 G allele was 0.43. The rs738409 G allele was not associated with insulin resistance status (Fisher’s exact P = 0.653) and HOMA-IR index level (one-way ANOVA P = 0.068) in the analysis without adjustment for NAFLD status. When we performed a multivariable regression analysis after adjustment for NAFLD status, the rs738409 G allele was associated with lower insulin resistance (HOMA-IR) index compared to the CC genotype (P = 0.031 and < 0.001, respectively). Then, we checked the effect of the rs738409 G allele on HOMA-IR index levels in each NAFLD and non-NAFLD subgroup stratified by NAFLD status using one-way ANOVA. In the non-NAFLD group, the GG and GT genotypes compared to the CC genotype showed significantly lower mean levels of HOMA-IR index in an additive manner (Sidak-adjusted P = 0.018 and < 0.001, respectively). However, in the NAFLD group, those inverse associations were not shown. Conclusion: The adiponutrin polymorphism rs738409 known as a NAFLD genetic risk factor was also associated with decreased levels of insulin resistance assessed by HOMA-IR index, and the inverse association was prominent in subjects without NAFLD.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
The present study revealed significant associations between and haplotypes of controls regions of the HLA-G with TD1 susceptibility.

The 5' regulatory region and 3'UTR fragments were amplified using the polymerase chain reaction (PCR) and were directly sequenced in an ABI310 Genetic Analyzer. The presence of significant associations between the SNPs detected of the HLA-G gene was evaluated by means of a likelihood ratio test of linkage disequilibrium (LD), using the ARLEQUIN. The PHASE method and expectation-maximization (EM) algorithm were used for haplotype inference. Regarding 3'UTR haplotypes, we observed two significant haplotypes, a conferring susceptibility to TD1 (UTR17), and another suggesting a protective role to TD1 (UTR3). On 5' URR, we found four significant haplotypes, a conferring susceptibility to TD1 (UTR17), and another suggesting a protective role to TD1 (UTR3). Thus, with relation extended haplotypes, suggest a balance between these regions, which were according to the previous papers, but we found two new haplotypes (PG01049 + UTR-13, G0103f + UTR-17), on the total two conferring protection (G010101c + UTR-4, G0104a + UTR-3), and one of the new (G0103f + UTR-17) conferring susceptibility. The present study revealed significant associations between and haplotypes of controls regions of the HLA-G with TD1 susceptibility.
860M  
Association of variants in GALNT10 and related pathway genes with body mass index in African Americans. M. Stromberg1,2, P. Mudgal2, B.I. Freedman2, D.W. Bowden3,4, M.C.Y. Ng2,3,5  
1) Department of Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 2) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC; 4) Department of Internal Medicine - Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, NC; 5) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC; 6) Internal Medicine-Endocrinology, Wake Forest School of Medicine, Winston-Salem, NC.  

Obesity is a chronic health problem which has a higher prevalence in African Americans (AAs) (35.7%) compared to non-Hispanic whites (23.7%) in the United States. Genetic contributors to obesity in AAs remain poorly understood, although recent genome-wide association studies (GWAS) identified several regions including the GALNT10 locus as being associated with body mass index (BMI) in AAs. GALNT10 is a member of the N-acetylgalactosaminyltransferase family which is located in the Golgi apparatus and is expressed in the hypothalamus of the brain, which regulates hunger. There are 20 members in the GALNT family, each with a unique expression pattern. Previous studies have not identified the causal variant(s) at GALNT10 and have not examined biological candidate genes from the relevant pathways in detail. As such, this study assessed 71 candidate genes for association with BMI in 2,293 African American subjects. Genotyping was performed using an Affymetrix Biobank array customized for fine-mapping of GWAS and candidate loci related to metabolic traits. Single SNP association was performed in two cohorts using linear regression under an additive genetic model. For variant with independent support for association, a sliding window meta-analysis was performed following meta-analysis. Nominal significance (P<0.0005) was observed at ATF4 (p=0.000337), GALNT18 (p=7.91E-06), STGAL1 (p=0.000191), WBSCR17 (p=0.000215), with the strongest association at GALNT16. GALNT18, or GALNTL4, has been previously associated with Type 2 Diabetes in AAs. The most significant association observed at the GALNT10 locus is rs1346482 (p=0.001015). Previously associated SNPs in this locus (rs815611, rs2033195, rs7708584) are shown to have pairwise r2 values ranging from 0.00-0.84 (0.84, 0.84) and are highly associated in the African American population. These findings suggest that several members of the GALNT family and related pathway members may play roles in modulating adiposity in AAs, although further replication is needed.  

861T  
Targeted Sequencing of an Admixture Mapping Peak in Latinos Implicates Rare Non-coding Variation in Asthma Susceptibility. D.G. Torger-son1, M. Pino-Yanes1, C.R. Gignoux2, C. Eng1, E.G. Burchard1,4, 1) Department of Medicine, University of California San Francisco, San Francisco, CA; 2) Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA.  

The prevalence and morbidity of asthma varies widely among global populations and racial/ethnic groups within the U.S. We previously identified a genome-wide significant admixture mapping peak at 18q21.1, whereby local Native American ancestry was associated with an increased risk of asthma in a meta-analysis of predominantly Hispanic/Latino individuals. The ancestry association was replicated in a large independent study of 3,774 Latinos; however, the variation behind the ancestry association remains to be identified. Towards this goal, we sequenced a 342 Kb contiguous region across the top of the admixture mapping peak including all coding sequences including SMAD2 and ZBTB7C and complete non-coding/intergenic sequences in 1,978 Mexicans and Puerto Ricans with and without asthma from the Genomics and Admixture in Latino Americans study (GALA II Study). No new individual variant was found to be associated with asthma in the Mexican or Puerto Rican populations, suggesting that the admixture mapping peak affecting asthma may be present at lower frequency and/or contribute to disease susceptibility in a unique population.  

862S  
Refinement of mapping and fine-mapping of trait-stratified genome-wide association study identifies novel genetic associations with cytokine phenotypes in systemic lupus erythematosus. T.B. Niewold1, M. Imru1, Y. Ghodke-Purank1, J.M. Dorshcer1, J.A. Kelly2, M. Marion2, J.M. Guthbridge2, C.D. Langefeld3, J.B. Harley4, J.A. James5, K.L. Sivils2, 1) Rheumatology and Immunology, Mayo Clinic, Rochester, MN; 2) Arthritis & Clinical Immunology Program, Oklahoma Medical Research Foundation, Oklahoma City, OK; 3) Department of Biostatistics, Wake Forest University, Winston-Salem, NC; 4) Cincinnati Children’s Hospital Medical Center and Cincinnati VA Medical Center, Cincinnati, OH.  

Introduction: Systemic lupus erythematosus (SLE) is a highly heterogeneous disorder, characterized by differences in serum cytokines and clinical manifestations. High serum interferon alpha (IFN-α) is an important heritable phenotype in SLE which is involved in primary disease pathogenesis, and this phenotype accounts for some of the biological heterogeneity between SLE patients. The aim of the study was to replicate and fine-map previously detected genetic associations with serum IFN-α in SLE. Methods: We previously undertook a case–case genome-wide association study of SLE patients stratified by ancestry and extremes of phenotype in serum IFN-α. Single nucleotide polymorphisms (SNPs) in seven loci identified in this screen were selected for follow up in a large independent cohort of more than 1300 SLE patients (672 European-ancestry, 420 African ancestry, and 231 Hispanic/ American-ancestry). Each ancestral background was analyzed separately, and a panel of 334 ancestry-informative markers was used to control for ancestry and admixture. Results: SNPs in the PPM1H and EFNAS loci demonstrated strong association with serum IFN-α in eQTL analyses (p<10-4 for each association) in European ancestry. In African ancestry subjects, SNPs at YST2, GHR, ST3GAL1, or GALNT10 were associated with serum IFN-α, and one of the SNPs was common between European- and African-ancestry subjects. In Hispanic/American ancestry patients, LRR2C20 demonstrated nominal evidence for association with serum IFN-α. Conclusions: This study demonstrates the power of using a serologic subphenotype to elucidate genetic factors involved in complex autoimmune disease. The distinct associations observed in different ancestral backgrounds emphasize the heterogeneity of molecular pathogenesis in SLE, and the need for stratification by subphenotypes. We hypothesize that these genetic variants play a role in disease manifestations and severity in SLE.  

863M  
Refinement of association signals and residual heritability in host control of HIV viral load. P.J. Ma2, P.J. Fellay2, 1) the International Collaboration for the Genomics of HIV. 1) School of Life Science, EPFL, Lausanne, Vaud, Switzerland; 2) Swiss Institute of bioinformatics, Lausanne, Vaud, Switzerland.  

Background: Genome-wide association studies of HIV outcome consistently identify the MHC region as the major genetic influence on disease progression. Through establishment of the International Collaboration for the Genomics of HIV we sought to bring together all existing GWAS data in patients to uncover further genetic signals.  

Methods: Genomic-wide SNP data were collected from 25 clinical centers. Plasma viral load measurements obtained during the chronic phase of untreated infection were available for 6,315 individuals of European ancestry, patients to uncover further genetic signals.  

Results: Consistent with previous studies, the top association signal was amino acid position 97 in HIV-B (p=4e-143), with independent associations at positions 67 (p=4e-112) and 45 (p=8e-49). Controlling for these positions fully accounted for classical HLA-B allele associations. Residual association at rs9264942, a marker of HLA-C expression level, was retained after controlling for amino acid effects (conditional p=1e-11). Outside of the MHC, a second peak of association in the CCR5 region (top SNP rs4317138 p=8e-16) was also observed. Conditioning for the known effects of CCR5Δ32 and CCR2R64V we found additional association with other variants in the HLA region, especially in the HLA and CCR5 regions, we observed a reduced, yet still significant h2 estimate of 16% (p=5e-8).  

Conclusions: By combining available GWAS data in HIV infected individuals we refined the known association signals in HLA-B and CCR5. Controlling for these effects uncovered additional associations in the CCR5 region and yet another signal in the CCR2 region. Heritability analysis suggested that variation outside known regions, captured through common SNPs, also contribute to HIV control.
864T

Study of genetic risk factors for susceptibility to leprosy - the chromosomal region 6q25-q27 revisited. G.B. RAMOS,1 H. SÁLROMAO,1 A.S. FRANCIO,2 C.C. CARDOSO,2 M.O. MORAES,2 A.C. PEREIRA,2 R.I. WERNECK,2 M.T. MIRA3,4 1 School of Medicine, Pontifícia Universidade Católica do Paraná, Curitiba, Paraná, Brazil; 2 Health and Biosciences School, Pontifícia Universidade Católica do Paraná, Curitiba, Paraná, Brazil; 3 Instituto Lauro de Souza Lima, Bauru, São Paulo, Brazil; 4) Laboratório de Hanseníase, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Rio de Janeiro; 5) Laboratório de Virologia Molecular, Departamento de Genética, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

Leprosy, a chronic infectious disease, affects approximately 220,000 new individuals worldwide every year. Genetic studies have identified several genes and genomic regions contributing to the control of host susceptibility to leprosy. In 2004, a study identified variants of the PAR2 gene associated with the disease, but the association signal observed does not seem to entirely explain the genetic effect observed for this locus, suggesting the presence of additional candidate genes. Here, we test variants of SOD2, located near PAR2/PACRG at the 6q25-q27 region, for association with leprosy in four Brazilian population samples of distinct design and ethnic background. The primary population included 186 volunteers distributed in 71 trios (55 independent) composed by one affected individual and their parents, recruited at the Prata Village, located at the Amazonic state of Pará; the replication study involved three case-control samples from the southern, southeastern and central-western regions of Brazil: Curitiba, Paraná (198 cases and 183 controls); Rio de Janeiro, Rio de Janeiro (507 cases and 236 controls) and Rondônia, Mato Grosso (411 cases and 424 controls), respectively. SNP markers covering the entire SOD2 locus were genotyped by fluorescence-based TaqMan and/or Sequenon technology. Family-based association analysis was performed as implemented in the software FBAT, version 2.0.4. Logistic regression was applied to the case-control samples, as implemented in the software SAS, version 9.1.

Family-based analysis revealed borderline evidence of association, after further analysis of LD patterns and association using a combined population sample is currently ongoing.

865S

Next-generation sequencing and targeted fine linkage disequilibrium mapping reveal FREM1 mutations associated with HIV acquisition in a sub-Saharan African cohort of female sex workers. J.F. Tuff1,2, D. La1,3, N. Klaponski1, V. Ly1,2, A. Yuen1, B. Liang2, S. Tyler2, F.A. Plummer3,4, M. Luo,1,3 1) JC Wilt Infectious Diseases Research Centre, Winnipeg, MB, Canada; 2) National Microbiology Laboratory, Winnipeg, MB, Canada; 3) University of Manitoba, Department of Medical Microbiology, Winnipeg, MB, Canada.

There is considerable heterogeneity in susceptibility to HIV-1 (HIV) infection among individuals. This variability in risk can be attributed to a combination of environmental, viral, and host genetics. A group of women in the Pumwani Sexworker (ML) Cohort, established in Nairobi, Kenya, exhibit a HIV-exposed seronegative (HESN) phenotype, despite repeated exposure through high-risk sex work. We previously described the identification of an intronic single nucleotide polymorphism (SNP) (rs1552896: NM_144966.5:c.1881+58G>C) in FRAS1 related extracellular matrix 1 (FREM1 [MIM: 608944]) as the highest-ranking correlate of this protective phenotype in a low-resolution genome-wide association study. To further investigate this association, we herein sought to identify FREM1 mutations in linkage disequilibrium (LD) with rs1552896, followed by fine mapping of polymorphic loci deemed most likely to be functionally causal as determined by SIFT/PolyPhen-2 or ENCODE via RegulomeDB. We conducted pyrosequencing of FREM1 (n=69) with Roche 454 GS FLX high throughput sequencing technology, and identified 1576 polymorphisms, 966 of which were newly discovered. Four of the 69 polymorphisms showing putative LD with rs1552896 (LOD ≥2) were chosen for fine mapping via PCR and Sanger sequencing: (i) rs2779500 (NM_144966.5:c.1315G>C, NP_659403.4:p.Val439Leu (n=541)), (ii) rs1353223 (NM_144966.5:c.1495A>G, NP_659403.4:p.Ile499Val (n=633)), (iii) rs10810271 (NM_144966.5:c.1262-439Leu (n=541)), (ii) rs1353223 (NM_144966.5:c.1495A>G, NP_659403.4:p.Ile499Val (n=633)), (iii) rs10810271 (NM_144966.5:c.1262-439Leu (n=541)), (iv) a novel microsatellite allele (c.1261+1207CCCT[8]TCCT[7]; Mi8*7, n=1956) found within a newly characterized complex tandem tetranucleotide repeat (CTTR). Fine mapping confirmed significant LD between rs1552896 and each of rs2779500 (LOD 6.94, OR=2.084), rs1353223 (LOD 3.92, OR=2.17), rs10810271 (LOD 5.29, OR=2.18), and Mi8*7 (LOD 85.37, OR=2.22), 95% CI 1.21-3.88). Analysis of samples successfully genotyped at both rs1552896 and CTTR loci (n=1046) revealed that Mi8*7 is more strongly associated with HESN (p=7.50 x 10^-3, OR=2.15, 95% CI 1.21-3.88) than is rs1552896 (p=7.67 x 10^-3, OR=2.18, 95% CI 1.21-3.90) than is rs1552896 (p=7.67 x 10^-3, OR=2.18, 95% CI 1.21-3.90) than is rs1552896 (p=7.67 x 10^-3, OR=2.18, 95% CI 1.21-3.90), suggesting that Mi8*7 is a better marker for HESN in FREM1. We further show that Mi8*7 belongs to a subgroup of FREM1 CTTR ((CCCT)₉₋₁₀(TCCT)₇₋₁₀) that, when homozygous, is also correlated with the HESN phenotype (p=1.99 x 10^-2, OR=2.22, 95% CI 1.45-3.40).

West Africa is experiencing some of the fastest growth in Type 2 Diabetes (T2D) prevalence in the world. It is believed that much of this increase is secondary to ongoing changes in body composition (BC). In this study, we investigated the association of BC indices (BMI, Fat Mass [FM], Fat Free Mass [FFM], and % Fat Mass [PFM]) with a locus (5q22-q31) previously identified in a T2D linkage analysis in West Africans (WA). Previously we performed fine-mapping of this region using 1405 SNPs genotyped in 931 WA. Seven SNPs were associated with at least one BC index, with the strongest association for rs2306617 (ALDH7A1) and BMI (β 1.3, p=1.0 x 10^{-4}). Based on these prior results, we conducted targeted resequencing of ALDH7A1 in 142 WA (48 obese, T2D; 48 non-obese, T2D; 48 obese, non-T2D). Identified SNPs were genotyped in 1726 WA, and the 37 SNPs that met quality control standards were evaluated in linear regression models. Three variants were associated with the BC indices, with the most striking association for rs3738174 with FFM (β 2.1 kg, p=9.9 x 10^{-11}) and BMI (β 0.9, p=2.8 x 10^{-9}). Other associated variants include rs1138005 with FFM (β -2.2 kg, p=2.6 x 10^{-4}) and rs7447380 with BMI (β 2.7, p=0.001). Of note, these variants have been annotated for promoter and enhancer marks, with rs3738174, in particular, influencing gene expression in cell types of relevance to BC (chondrocytes, adipocytes, and mesenchymal stem cells). The stronger results with FFM, and the previous identification of ALDH7A1 in a GWAS of osteoporosis, led us to investigate the association of this gene on bone phenotypes. We evaluated phenotype data for Aldh7a1 knockout mice (Jackson Laboratory, Project NIH-0076). Among the 16 mice on which DEXA scans had been performed, the null genotype was associated with a higher ratio of bone mineral content to lean body mass (p=0.04). The null genotype was not associated with markers of adiposity. The associations of this locus with FFM in WA and bone mineral content in mice suggest that BMI associations with ALDH7A1 may, in fact, reflect an influence of this locus on body weight through bone mineralization. ALDH7A1 encodes an aldehyde dehydrogenase, and reactive aldehydes have been shown to affect bone remodeling. Taken together, this work suggests that ALDH7A1 influences body composition, which may result from perturbations in bone mineralization.
868S
Integrating functional data to prioritize variants in statistical fine-mapping studies. G. Kichaev1, W.Y. Yang2, S. Lindstrom3, P. Homozogiana1, E. Eskin1,2, A.L. Price1,4, P. Kraft3,4, B. Pasaniuc1,2,5, L. Nelson2,5, 1) Bioinf. genet. traits Inter-departmental Program, UCLA, Los Angeles, CA., USA; 2) Dept of Computer Science, UCLA, Los Angeles, CA., USA; 3) Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston MA., USA; 4) Dept of Biostatistics, Harvard School of Public Health, Boston MA., USA; 5) Dept of Human Genetics, David Geffen School of Medicine, UCLA, Los Angeles, CA., USA; 6) Dept of Pathology and Laboratory Medicine, David Geffen School of Medicine, UCLA, Los Angeles, CA., USA.

Standard statistical approaches for prioritization of variants for functional testing in fine-mapping studies either use marginal association statistics or estimate posterior probabilities for variants to be causal under simplifying assumptions. Here we present a probabilistic framework that integrates association strength with genomic functional annotation data (e.g. ENCODE) to improve accuracy in selecting plausible causal variants for functional validation. A key feature of our approach is that it empirically estimates the contribution of each functional annotation to the trait of interest while allowing for multiple causal variants at any risk locus. We devise efficient algorithms that estimate the parameters of our model across all risk loci to further increase performance. In fine-mapping simulations based on 1000 Genomes data our approach reduces the average number of SNPs to be tested in functional assays to identify 90% of all causal SNPs from an average of 12.3 variants per locus when ranking by marginal association statistics (or 25.0 when ranking by posterior probabilities under single causal variant assumption [Maller et al 2012]) to 9.7 variants per locus. We validate our method using published data from a large-scale melanoma case-control study.[6] We find novel functional associations for at least 4 melanoma and 3 skin traits [Teslovich et al 2010] and find that the relative probability for causality is increased for variants in exons and transcription start sites and decreased in repressed genomic regions at the risk loci of these traits. Using these results, we predict trait-specific functional annotations, we estimate causality probabilities across all traits and variants, reducing the size of the 90% confidence set from an average of 17.5 to 13.5 variants per locus in this data.

869M
Analysis of genes involved in carnitine metabolism and functions in autistic patients by targeted sequencing. J. Ge1, X. Wang1, H. Cui1, B. Zhang1, H. Xu,2, P. Fang2, A. Beaudet1, 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Fudan University, Shanghai, China.

A recent study reported the link between the deficiency of the trimethyllysine hydroxylase enzyme (TMLHE), the first enzyme on the carnitine biosynthetic pathway, and the risk of autism, raising the possibility that carnitine might be important in the autism etiology. The well-established function of carnitine is its role in fatty acid metabolism as a molecule transporting the fatty acids and acetate across the mitochondrial membrane. The carnitine homeostasis in humans is maintained through intake from dietary sources, endogenous synthesis, and reabsorption in the kidney. To understand whether carnitine has influence on the risk of autism, we analyzed the genes functioning in carnitine biosynthesis, carnitine transporters, carnitine-acetyl carnitine acyltransferases, carnitine acylcarnitine translocases, as well as several other genes that can influence carnitine metabolism, such as genes that can cause secondary carnitine deficiency. Using the Fluidigm Access Array™ System for high-throughput PCR amplification and Illumina MiSeq next-generation sequencing, we have sequenced the exons of these 23 genes in 384 autistic probands from the Simons Simplex Collection (SSC). Totals of 108 heterozygous and 2 hemizygous protein changing rare mutations were found as follows: ACADL: c.928_929delC; 8ACAT2: p.D122Y; 8ACAT2: p.D266G and ACADM: p.D270G 2 times. The two hemizygous mutations are SLC6A14: p.T4185S and SLC25A43: p.R3230. Interestingly, we found that 2 patients carrying 3 heterozygous mutations in different genes, and 18 patients carrying 2 heterozygous mutations. We are now checking the inheritance of these mutations and ongoing our study to the full SSC cohort.

870T
Genetic contributions to obesity and metabolic risk in Mexican-American children. R.J. Mudgway1, S.L. Spilman2, A.P. Maitya3, J.T. McCracken3, L.J. Nurni1. 1) Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA.; 2) Background: Hispanic populations have among the highest rates of obesity among other racial/ethnic groups. Genetic moderators of weight gain and metabolic factors may help explain vulnerability to obesity in these higher risk populations. Previous studies have identified signaling networks involved in obesity risk, including energy balance, monoaminergic, and growth factor systems. We examined common genetic variants associated with BMI and metabolic factors in a sample of Mexican-American children. Methods: Baseline data from 356 Mexican-American children 6-12 years old from the Los Angeles area were collected from matched overweight and lean groups. Physiological measures were recorded for each subject, including BMI, fat mass, physical measurement, serum chemistry (such as lipid and glucose levels), respiratory function, and resting energy expenditure. Families underwent a month-long nutrition and diet intervention and returned for follow-up measures. Association between metabolic phenotypes and complete common genetic variation in energy balance candidates (MC4R, NPY), monoamine receptors (ADRA2A, HTR2C, DRD1-5), and brain-derived neurotrophic factor (BDNF). Key variants showing association with obesity risk in prior literature (FTO and PRKAR2B) were also included. Results: In agreement with published reports in Caucasians, the minor (C) allele of rs1421075 in FTO, was associated with increased BMI in Mexican-American children (p=0.027) while the minor allele (G) of rs13224682 in PRKAR2B was associated with leanness (p=0.03). We observed novel functional variants associated with baseline BMI (p=0.038) and rs5326 with BMI change over 1 month (p= 0.59). In the X-linked HTR2C gene, rs185147 demonstrated a gender specific effect, where the common allele (C) was associated with higher BMI in girls (p=0.003). Additionally, exploratory analyses revealed a potential role for ADRA2A in resting energy expenditure, MC4R in baseline lipid levels, and both FTO and BDNF in lipid and glucose changes over time. Conclusions: Common genetic variation in energy balance, monoamine, and growth factor candidate pathways were nominally associated with weight and metabolic profile in Mexican-American children. These results warrant replication in large, independent samples and could help target for intervention children at risk for obesity and suggest personalized treatment strategies.

871S
The analysis of MC1R polymorphisms can be used as a tool to predict complex phenotypes, such as skin and hair color in Brazilian population. O. Figueiredo, C. Goncalves, O. F. Goncalves, 1) FMUSP, Legal Medicine, Ethics and Occupational Health, Medical School, University of Sao Paulo, Sao Paulo, Brazil.

Human pigmentation traits, including color variation in skin, eye, and hair, belong to the most visible and differentiating human traits. The genetic basis underlying variation in human pigmentation has been the subject of intensive research by investigators in a variety of life science communities, including forensic purposes, in which the use of these trials can help identifying missing person or guide some police investigations. The two types of melanin in human skin and hair coloration are the protective dark-colored eumelanin and the sulphur containing light red-yellow pheomelanin. Eumelanin together with pheomelanin constitute the two main pigments of the skin and hair. The events leading to the melanogenesis are controlled by different genes, that we can highlight the melanocortin 1 receptor gene (MC1R) which encodes a protein in melanocytes responsible for melanin synthesis regulation. Polymorphisms in MC1R, which result in a loss of function of the receptor, are associated with increased pheomelanin production and skin and hair color in a sample of 401 individuals of admixed population from Brazil, intending to use the data in forensic genetics casework in several situations. No deviation in Hardy-Weinberg equilibrium was observed for all the polymorphisms analyzed. We found a strong association between the SNP rs858479 (G>A) and yellow skin color (OR: 4.15; 95% CI: 1.51-10.64) and between the SNP rs1110400 (T>C; OR: 1.49; 95% CI: 1.05-2.11) and pheomelanin production (G>A) (p=0.04). In the X-linked HTR2C gene, rs6119654 demonstrated a gender specific effect, where the common allele (C) was associated with higher melanin in both genders (p=0.0343). This study demonstrates the potential role of MC1R in forensic genetics casework in several situations. No deviation in Hardy-Weinberg equilibrium was observed for all the polymorphisms analyzed. We found a strong association between the SNP rs858479 (G>A) and yellow skin color (OR: 4.15; 95% CI: 1.51-10.64) and between the SNP rs1110400 (T>C; OR: 1.49; 95% CI: 1.05-2.11) and pheomelanin production (G>A) (p=0.04). In the X-linked HTR2C gene, rs6119654 demonstrated a gender specific effect, where the common allele (C) was associated with higher melanin in both genders (p=0.0343).
872M
Genetic risk variants for body mass index are associated with decreased excessive daytime sleepiness in 9,832 individuals of European ancestry from NHLBI cohorts. J.M. Lane,2,3 R. Saxena,2,3 R. C. Bjonnes1,2,3, B. Cade4, S. Redline5, F.A.J. Scheer6, N. Pujan4, D. Gottdiener7,1. 1) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 2) Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; 3) Division of Sleep Medicine, Brigham and Women’s Hospital and Harvard Medical School Boston, MA; 4) Division of Pulmonary and Critical Medicine, Johns Hopkins University, Baltimore, MD. Near half of US adults report suffering from excessive daytime sleepiness (EDS) at least a few days a month, leading to cognitive, psychological and metabolic issues, such as obesity. We therefore asked if there is a common genetic basis for both mass index (BMI) and EDS. We evaluated the relationship between EDS and genetic risk scores for BMI comprised of genetic variants identified in genome-wide association studies. We used data from NHLBI cohorts with genotype and phenotype data in subjects of European ancestry (self-report dichotomous trait) (n=11,737). Analysis was adjusted for age, gender, and ancestry principal components. A fixed effects, inverse-variance meta-analysis was performed. Risk scores were calculated using 18 SNPs previously associated with BMI. We find increased odds of EDS in overweight (OR 1.020, p=0.041), obese (OR 1.083, p less than 0.0001), and morbidly obese (OR 1.195, p less than 0.0001) individuals of European ancestry versus normal BMI (n=11,737). We find a significant association between EDS and genetic variants in BDNF and FTO. In total 14/18 SNPs demonstrated decreased risk of EDS with BMI risk alleles, even after adjustment for BMI. A weighted BMI risk score is significantly associated with EDS (OR=0.96 [0.95-0.98], p=0.001). There were independent associations with percent weight change at 12 months post-transplant were done using dose dependent and risk allele analyses. ANOVA showed a dose dependent effect for rs6347 (p=0.05) and rs1946816 (p=0.019). Risk allele chi squares p=0.057 at 5 years. Age significantly modulates the relationship between EDS and BMI genetic risk score (Pint=0.017). There was no significant interaction with gender, sleep apnea, mood disorder, or season of measurement. Our association results capture an effect on EDS beyond BMI. Paradoxically, BMI was associated with decreased EDS in individuals associated with a decreased odds of EDS, indicating divergent physiological roles for underlying pathways. Mechanistic understanding of BMI risk variants may provide parallel insights into excessive daytime sleepiness.

873T
Promoter Polymorphism and low Serum Levels of Mannose Binding Lectin as risk factor for Rheumatoid Arthritis in Indian population. A. Sodhi1, J. Singh2, S. Singh3, S. Arora4, M. Kaur5. 1) Department of Human Genetics, Guru Nanak Dev University, Amritsar, India; 2) Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar, India; 3) Rheumatology clinic, Amritsar, India. Mannose Binding Lectin-2 (MBL-2) is a C-type serum lectin synthesized by the liver as a acute phase protein. MBL-2 gene is located on chromosome 10q11.2-q21 and various single nucleotide polymorphisms (SNPs) in this gene has been reported to be associated with infection as well as autoimmune diseases. Promoter polymorphism in MBL-2 gene is reported to be associated with altered serum MBL levels which are associated with increased risk of various disease conditions. Rheumatoid arthritis (RA) is an autoimmune disease characterized by chronic inflammation and destruction of articular cartilage as well as synovial hyperplasia. Prevalence of RA is 0.5% - 1.0% of the global population and is increasing dramatically. The present study aimed to investigate the frequency of the mutations at positions -550 and -221 of promoter region in MBL-2 gene along with serum MBL and hs-CRP serum levels. The study was approved by institutional ethical committee as per declaration of Helsinki and informed consent was obtained from each individual. Blood samples were collected from 202 RA patients and 200 age, gender and ethnicity matched healthy subjects. Information regarding anthropometric variables, Demographic features, disease duration, and disease activity score was also recorded. Single nucleotide polymorphism in the MBL promoter (-550/ -221) were amplified by Amplification refractory mutation system-PCR (ARMS-PCR). Functional serum MBL-2 and hs-CRP levels were analyzed using commercially available kits. Out of two SNPs studied, significant difference was observed in genotypic as well as allelic frequencies of H/L (-550) polymorphism (p<0.0001). L allele of -550 is at risk for RA patients (odd ratio: 1.079 CI 95%: 0.021-2.021; p<0.0001). Combination of two promoter polymorphism YXLL is risk over YXHL (odd ratio: 12.12 CI95%: 3.71-39.19; p<0.00001) and H allele is protective over L allele (Odd ratio: 0.132; CI 95%: 0.04-0.388; p<0.0002). Serum MBL-2 levels were found to be significantly (p<0.001) lower in RA patients as compared to healthy controls and hs-CRP serum levels were found significantly (p<0.01) higher in RA compared to healthy controls. Therefore, this study showed that the polymorphism in the promoter region of the MBL gene may be a genetic marker associated with RA and presence of mutation in the -550 H/L promoter region of the MBL-2 gene correlated to low MBL serum levels.

874S
Carboxypeptidase E and dopamine transporter SNPs are associated with percent weight change in kidney transplant recipients. A.G. Stansfield1,2, A.K. Cashin3, D.K. Hathaway4, Y.P. Conley1. 1) Health Promotion and Development, University of Pittsburgh, Pittsburgh, PA, USA; 2) National Institute of Nursing Research, National Institutes of Health, Bethesda, MD; 3) University of Tennessee Health Science Center, Memphis, TN. Variations in dopamine pathway genes have recently gained attention for being associated with weight gain and obesity. These genes include the dopamine active transporter gene (SLC6A3 or DAT1), which codes for the transporter responsible for dopamine reuptake from the synapse, and the carboxypeptidase E (CPE) gene, which has been shown to have a regulatory effect on that transporter. While the associations are exciting, there are often logistical difficulties in the longitudinal study of the genetics of weight gain. Kidney transplant recipients make an ideal population for these types of genetic association studies, as approximately 30% of recipients gain a significant amount of weight (>10 kg) in the first year after surgery. Although many clinicians attribute this weight gain to immunosuppressant therapies, this assumption has not been supported in the literature. Furthermore, there is no known relationship between dopamine function and renal disease, although gene expression work in this population has shown a significant relationship between weight gain and dopaminergic pathway genes. The purpose of this study was to test variations in SLC6A3/DAT1 and CPE for associations with weight gain, using kidney transplant recipients as a model population. Blood samples were previously collected from 70 kidney transplant recipients (43% female, 57% African American, aged 50.7 ±13.2 years, baseline weight 181.9 ±39.4 pounds) at the time of surgery. Two SNPs in SLC6A3 (rs6347, rs6350) and three SNPs in CPE (rs1533645, rs1946816, rs1541903) were genotyped rs6347 and rs1946816 (p<0.018). Risk allele chi squares p=0.057 at 5 years. Age significantly modulates the relationship between EDS and genetic risk score (Pint=0.017). There was no significant interaction with gender, sleep apnea, mood disorder, or season of measurement. Our association results capture an effect on EDS beyond BMI. Paradoxically, BMI was associated with decreased EDS in individuals associated with a decreased odds of EDS, indicating divergent physiological roles for underlying pathways. Mechanistic understanding of BMI risk variants may provide parallel insights into excessive daytime sleepiness.
875M Candidate gene association study of chronic obstructive pulmonary disease using a targeted high throughput sequencing approach. J. Klar1, H. Matsson1, C. Söderhäll1, H. Backman2, A. Lindberg3, E. Rönmark3, J. Kere4, B. Lundbäck4, N. Dahl4. 1) Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; 2) Department of Biosciences and Nutrition and Center for Innovative Medicine (CIMED), Karolinska Institutet, Huddinge, Sweden; 3) The OLIN studies, Sunderby Hospital of Norrbotten, Luleå, Sweden; 4) Krefting Research Centre, Institute of Medicine, University of Gothenburg, Sweden.

Background: Chronic obstructive pulmonary disease (COPD) is a common disease in Sweden, affecting approximately one in six in age over 40 years, with severe impact on health and quality of life. COPD results from environmental factors, especially cigarette smoking, with contributions from yet unknown genetic background factors and by gene-environment interactions. Not all smokers develop COPD suggesting that genetic factors modulate the life-time risk and conversely, non-smokers can be at elevated risk for COPD. The general objective of this project is to get insight into the mechanisms and interactions of specific genes and genetic pathways that contribute to the development of COPD.

Aim: To investigate candidate genes, with emphasis on genes important for lung development and homeostasis, in defined populations of smokers and non-smokers in search for genetic variations associated with development of COPD.

Methods: We conducted a candidate gene analysis on 200 kb of enriched sequences, including 22 genes implicated in lung development and 71 genes and regions previously associated with COPD. Targeted enrichment (HaloPlex; Agilent) and high throughput sequencing (Illumina) was performed on 96 patients and 96 healthy controls retrieved from the Swedish Obstructive Lung Disease in Norrbotten (QLIN) Studies sample set.

Results: We identified a total of 2,151 SNPs of which 78, distributed in 45 gene regions, have significantly different allele frequencies in COPD cases than controls. All COPD associated SNPs showed a strong effect on the development of COPD, as indicated either by low or high odds ratios (OR). The SNPs are mainly located in genes that cluster in pathways associated with cell proliferation, including genes involved in both development and repair of the lung.

Conclusion: Our preliminary results confirm previous findings and, in addition, high effect sizes of SNP variants associated with COPD compared to other studies. The strong effect sizes are possibly attributed to the genetic background of our study populations. The results support the idea that genetic factors in genes for lung development are important determinants of adult lung function that may ultimately contribute to COPD.

876T SNP variants in MHC are associated with sarcoidosis susceptibility and subgroups - a joint case-control association study in four European populations. A. Wennerström1,2,3, E. Lahelma2, V. Antilla4, J. Grunewald5, C. van Moorese6, M. Petrek7, A. Eklund8, J. Grunewald5, A. Backman1, M. Ronningen9, M. Seppänen10, O. Selroos11, M.-L. Klokki12. 1) National Institute for Health and Welfare (THL) Public Health Genomics Unit, Helsinki, Finland; 2) Transplantation Laboratory, Haartman Institute, University of Helsinki, Finland, Helsinki, Finland; 3) University of Helsinki The Institute for Molecular Medicine Finland (FIMM) Biomedicum, Helsinki, Finland; 4) Analytical and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States; 5) Respiratory Medicine Unit, Department of Medicine Solna and CMM Karolinska Institutet and Karolinska University Hospital, Solna, Sweden; 6) Department of Pulmonology, St Antonius Hospital Nieuwegein, and Heart and Lung Center University Medical Center Utrecht, Utrecht, Netherlands; 7) Laboratory of Immunogenomics and Immunoproteomics, Faculty of Medicine and Dentistry, Palacky University Olo mouc, Olomouc, Czech Republic; 8) Raa sepori Health Care Centre, Raa sepori, Finland; 9) Immunodeficiency Unit, Division of Infectious Diseases, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland; 10) Rheumatology Unit, Department of Medicine, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden Finland; 11) Semeco AB, Vejbystrands, Sweden; 12) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, United States.

Sarcoidosis is a multiorgan inflammatory disorder of unknown aetiology. The most probable pathophysiology of sarcoidosis, the dysregulation of the immune response strongly suggests benefits from a better understanding of the role of the immune mediating genes (e.g. MHC genes) in sarcoidosis susceptibility. We present results from a Finnish case-control discovery sample set as well as three independent replication studies from the Swedish, Dutch and Czech populations. We studied four genes in the MHC Class III region (LTA, TNF, AGER, BTNL2) and HLA-DRA in relation to HLA-DRB1 alleles to detect variants predisposing to sarcoidosis and to identify genetic differences between patient subgroups. Patients with sarcoidosis (n=805) were further subdivided based on the disease activity and the presence of Löfgren syndrome. In a meta-analysis, seven SNPs were associated with non-Löfgren sarcoidosis (NL; the strongest association with rs3177928 in non-Löfgren sarcoidosis (NL; the strongest association with rs3128843 in BTNL2/HLA-DRA region, P=3.4E-12, OR=3.4) when compared with healthy controls (n=870). The high LD between SNPs and an HLA-DRB1 challenged the result interpretation. In addition to these SNPs, population-specific associations for sarcoidosis were observed. In conclusion, there is clear evidence that polymorphisms in the BTNL2 and HLA-DRA have a role in sarcoidosis susceptibility. Most importantly, our study revealed sarcoidosis-related variants that were shared across ethnicities as well as ethnicity-specific genetic markers. Future functional studies are required to reveal the causal variants of these associations and the immunogenetic basis related to sarcoidosis.
Assessment of LGALS3 genetic variants rs4644, rs4652, rs2075601 and galectin-3 levels as risk factor in Rheumatoid Arthritis.

T. Kaur1, M. Kaur1, S. Singh5, S. Arora5, J. Singh5, 1) Department of Human Genetics, Guru Nanak Dev University, Amritsar, India; 2) Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar; 3) Rheumatology Clinic, Amritsar.

Galectin-3 is a pro-inflammatory molecule family encoded by LGALS3 gene on 14q21 chromosome. It can act as a key player in various inflammatory diseases by activation of macrophages and neutrophils. Rheumatoid arthritis (RA) is a chronic, inflammatory, autoimmune disease characterized by destruction of peripheral joints leading to deformity and disability. Worldwide, 1% of population is suffering from RA. Elevated levels of galectin-3 have been implicated in serum and synovial fluid of RA patients. A single study involving LGALS3 genetic variants have indicated its association with RA. The present study has been proposed to evaluate LGALS3 variants and its protein serum levels in RA in Indian population. The present case-control study included 200 RA patients, diagnosed according to 1987 revised criteria of American college of Rheumatology and acohort of 200 age, gender and ethnicity matched controls. The study was approved by institutional ethical committee in accordance with declaration of Helsinki and written informed consent was obtained from each participant. Genetic typing for rs4644 and rs4652 was done by PCR-RFLP method while for rs2075601 by tetra primer PCR method and by Sanger Sequencing. Serum galectin-3 and serum hs-CRP levels were assessed in both patients and controls using commercial available ELISA kits (Abcam, UK and MyBioSource, USA respectively). Suitable statistical analysis was performed using SPSS version 18.0. Significant difference in allelic distribution has been observed in RA patients as compared to controls (p<0.05) and there was suggestive evidence of an association in a Co-dominant model (AA vs AC = AC vs CC; OR= 1.37, 95% CI 1.00-1.88, p = 0.048) for rs4644 SNP. For rs4652 SNP, present study found a dominant mode of association with RA (AC/CC vs AA; OR= 1.37, 95% CI 1.00-1.88, p = 0.048) for rs4644 SNP. For rs4652 SNP, the frequency of TT genotype was risk factor for our RA population. Significantly elevated serum galectin-3 and serum hs-CRP levels have been observed in RA patients than controls was risk factor for our RA population. Significantly elevated serum galectin-3 was found to be prevalent in RA patients (11%) than controls (7%). Different variants were found as a risk factor. For rs2075601 SNP, the frequency of TT genotype was significantly higher in RA patients as compared to controls (p<0.05) and there was suggestive evidence of an association in a Co-dominant model (AA vs AC = AC vs CC; OR= 3.947, 95% CI 0.86 -19.67, p= 0.047) and AA genotype was risk factor. For rs2075601 SNP, present study found a dominant mode of association with RA (AC/CC vs AA; OR= 1.37, 95% CI 1.00-1.88, p = 0.048) for rs4644 SNP. For rs4652 SNP, the frequency of TT genotype was risk factor for our RA population. Significantly elevated serum galectin-3 was found to be prevalent in RA patients (11%) than controls (7%).

Significant association was observed between LGALS3 gene polymorphism (tagSNPs) and risk of RA susceptibility. The association was performed by logistic regression using Plink software. A stepwise regression analysis was performed using the R program. The statistical results were corrected for multiple testing using the Bonferroni correction for each polymorphism evaluated. The sample showed 80.11% of statistical power to detect genetic association. Twenty-four SNPs spanning 16 genes were statistically associated with the etiology of NSCLP in the sample. The associated genes were: TCEB3 rs22355451*T (p = 0.032); MSX1 rs3775261*T (p = 0.009); SPRY1 rs300566*A (p = 0.017); MSH2 rs4868442*T (p = 0.020); PRSS3 rs1700180*C (p = 0.015); nms35585*G (p = 0.001), rs303048*T (p = 0.006); rs1675414*C (p = 0.042); SHH rs1233556*T (p = 0.047); VAX1 rs10787760*G (p = 0.011); rs7086344*T (p = 0.000); rs65585429*T (p = 0.013); TFX10 rs3758938*G (p = 0.017); JAG2 rs1555374*T (p = 0.001); BMP4 rs17563*C (p = 0.016); JAG2 rs11621316*G (p = 0.003); KIF7 rs4932248*C (p = 0.007); rs4932230*T (p = 0.029); AXIN2 rs11655966*T (p = 0.030); DVL2 rs2074222*G (p = 0.002); rs222850*C (p = 0.015). The stepwise regression analysis showed that SNPs associated contribute together to 15.5% of the determinants of etiology of NSCLP in this sample. In conclusion, these data suggest that these genes are associated with NSCLP in the Brazilian population. In addition, this is the first study to suggest association between KIF7 and TCEB3 genes and etiology of NSCLP. Supported by: FAPESP, CNPq.
879T
Exome sequencing and targeted DNA resequencing reveals association of the MYO5B SNP rs183559995 with risk of Familial Nonsyndromic Cleft Lip and Palate. S. Beiraghi1, H.A. Stessman2, A.K. Mirza2, R.J. Schafer2, W. Wang2, C.L. Myers1, B.G. Van Ness1. 1) Division of Pediatric Dentistry, University of Minnesota, Minneapolis, MN; 2) Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN; 3) Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN.

Non-syndromic cleft lip with or without palate (NSCL/P) is one of the most common congenital birth defects with an incidence of 1/500-1/1000. Genetic factors have been shown to play significant role in the development of NSCL/P. Using non-parametric linkage analysis study on a six-generation family (n=27) with probable autosomal dominant, low penetrance inheritance of NSCL/P, we have previously reported that a 5.7-Mb genomic region on 18q21.1 that potentially contains a pathogenic, high-risk variant for NSCL/P. In the current study, we performed exome sequencing on 6 affected individuals, 2 obligate carriers, and 4 unaffected individuals from the NSCL/P family using Illumina HiSeq with TruSeq Exome Enrichment. 100 Western European (CEPH) genomes from the 1000 Genomes Project were utilized as unaffected controls owing to low penetrance for NSCL/P within this family. All sequences were mapped to hg19 human reference genome using Burrows-Wheeler Aligner 0.5.9 (BWA) and variants were called using SAMTools and GATK Unified Genotyper for all sites with greater than 8 reads. High quality variants were used as markers in genome-wide association analysis linked to the affected phenotype (NSCL/P) using PLINK whole genome association analysis toolset. Four (4) candidate SNPs within the same gene, MYO5B, a myosin family member involved in protein trafficking, as well as associated variants (SNVs and INDELs) were identified that may contribute to NSCL/P disease etiology. Subsequent targeted Sanger re-sequencing in 33 family members demonstrated that the MYO5B intronic SNP rs183559995 (GA) was significantly associated with the NSCL/P trait (p=0.001, 95% CI=2.25-1853.3, OR=15). Our results indicate that this SNP is a strong candidate gene for familial NSCL/P. Further functional studies will be required to determine its significance with regard to MYO5B structure and function.

880S
Rare Variants Within 7p Region Associated with Carotid Bifurcation Intima-Media Thickness Among Dominican Republic Families. N.D. Dueker1, A. Beecham1, L. Wang2, S. Blanton3, D. Cabral1, E. Sabala2, T. Rundek3, R.L. Sacco2,3,1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL; 3) Department of Neurology, Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL.

Ischemic stroke (IS) and myocardial infarction (MI) are leading causes of mortality and disability in the US. Though both IS and MI are suggested to be genetically controlled, only a small proportion of the risk for both is explained. One approach to identifying genetic variants associated with these disorders is to identify the genetic determinants of subclinical phenotypes, such as carotid intima-media thickness (cIMT). We have previously shown cIMT measures to be heritable and found evidence for linkage and association of common variants on 7p with carotid bifurcation IMT (BIF). Therefore, we sought to further characterize the 7p region and to identify additional candidate variants, hypothesizing that rare variants in this region are associated with BIF. To test this hypothesis, we sequenced the 1 LOD unit down region of 7p in 33 family members demonstrating evidence for NSCL/P disease etiology. Subsequent targeted Sanger re-sequencing in 33 family members demonstrated that the MYO5B intronic SNP rs183559995 (GA) was significantly associated with the NSCL/P trait (p=0.001, 95% CI=2.25-1853.3, OR=15). Our results indicate that this SNP is a strong candidate gene for familial NSCL/P. Further functional studies will be required to determine its significance with regard to MYO5B structure and function.

881M
No association of PTPN22 and SUMO4 Polymorphisms with predisposition to type 1 diabetes (T1D) in a cohort of south Indian subjects. B.G. Gorjai1, U. Ratnamala2, S.K. Nath3, U. Radhakrishna3. 1) Oncology Department, Krishna Institute of Medical Sciences, Secunderabad, Andhra Pradesh, India; 2) Department of Pharmacology, Creighton University, Omaha, NE, United States; 3) Arthritis and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, OK, United States; 4) Beauumont Research Institute, Royal Oak, MI, USA.

Type 1 diabetes, (T1D) (juvenile or insulin-dependent diabetes mellitus) is an autoimmune disease characterized by the lack of insulin due to an attack on insulin-producing beta cells. The incidence of T1D varies from 8-17/100,000 in Northern Europe and the U.S. The general incidence of T1D in India is 10.6 cases/year/100,000. The exact etiology and pathogenesis of T1D is still unknown, however genetic factors are believed to be a major component for the development of T1D. Additionally, modifying epigenetic factors such as diet, environmental, infections and lifestyle play an important role in disease expression. To date, several genome-wide association and candidate gene studies have identified more than 25 genetic associations with high confidence, including PTPN22, SUMO4, CTLA4 and IL2RA. However no common genetic mutations or pathogenic causative genomic variations have been identified. We have recruited 500 sporadic T1D patients, and an equal number of age-matched controls from southern India. We have recently sequenced PTPN22 and SUMO4 genetic variants in 100 T1D patients and equal controls using a SNP array. Our results did not show any association with PTPN22 (rs2476601, 1858C>T). This SNP was monomorphic (G/G) in both T1D. This is consistent with other Asian populations. The data of SUMO4 variation (rs237025, 183A>G, rs10985658, 1858C>T) was not significantly different between T1D and controls. However, the “G” allele frequency between cases and controls are not statistically significant (40% in cases vs. 35% in controls, OR (CI)= 1.23 (0.82-1.85), p=0.30), most likely due to small sample sizes.

882T
Evaluation of Genetic Polymorphism of MBL2 Gene and Pulmonary Function Test in Chronic Obstructive Pulmonary Disease. A. Sharma1, G. Gandhi1, B. Mathothra2, J. Singh3, S. Singh4, M. Kaur4. 1) Department of Human Genetics, Guru Nanak Dev University, Amritsar, India; 2) Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar, India; 3) Department of Tuberculosis and Chest Diseases, Govt. Medical College, Amritsar, India.

Chronic Obstructive Pulmonary disease (COPD) is characterized by narrowing of airways due to abnormal inflammatory response to noxious particles leading to progressive impairment in the pulmonary function. Worldwide, COPD is the fourth leading cause of morbidity and mortality in India. In India, its prevalence is 4.1%. Several molecules of innate and adaptive immune system are implicated in disease etiology and pathogenesis. Mannose-binding lectin (MBL) is a pattern-recognition protein which binds to carbohydrates of infectious agents or cells resulting in activation of complement system and thus causing stimulation of inflammatory reactions. Objectives of the present study were to evaluate MBL2 gene polymorphisms in COPD patients and its association with pulmonary function test. The study was approved by Institutional Ethical Committee. After informed consent, blood samples were collected from 80 COPD patients and 72 age, gender and ethnicity matched controls. Genomic DNA was isolated from blood samples using inorganic method. Single nucleotide polymorphisms (SNPs) of codon 54 and 57 of the MBL2 gene were studied by PCR-RFLP using restriction enzymes BanI and MboI respectively. SNP of 5UTR region (allele P/Q) was studied by ARMS-PCR. Pulmonary function test was performed using standard spirometric system for the diagnosis and categorization of COPD patients GOLD (2012) guidelines were used. All statistical analyses were performed using SPSS version 16.0. No significant difference was observed in genotypic and allelic distribution of codon 54, 57 and P/Q alleles in comparison with controls. Genetic combination, heterozygosity for codon 54, P/Q and 5UTR region 57 (GAAACT), appears to be a risk factor (OR=4.301; CI=1.06-17.45; p=0.04) for susceptibility to develop COPD with respect to AAAACT. All the pulmonary function test parameters viz. forced vital capacity (FVC), forced expiratory volume in one second FEV1, FEV1/FVC ratio and PEFR rate failed to distinguish COPD cases and controls significantly in patients than controls (p=0.001). On the basis of disease severity, patients were stratified into mild (3.75%), moderate (23.75%), severe (40.00%) and very severe (21.50%) stages. Significant association was found with GAAACT for FEV1 (p<0.001) in case of severe COPD patients and GAAACT for FVC (p<0.001) in case of mild COPD patients. The results of the present study showed that MBL2 polymorphisms may be involved in pathogenesis of COPD.
1) Department of Molecular Genetics, Yakut Scientific Centre of Complex Medical Problems, Siberian Branch of the Russian Academy of Sciences, Yakutsk, Russia; 2) Laboratory of Molecular Biology, Institute of Natural Sciences, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation; 3) Department of Radiology, Republican Hospital #2 - Center of Emergency Medicine, Ministry of Public Health of the Sakha Republic, Yakutsk, Russian Federation; 4) Republican Centre of Professional Pathology, Republican Hospital #2 - Center of Emergency Medicine, Ministry of Public Health of the Sakha Republic, Yakutsk, Russian Federation; 5) Audiology-Logopaedic Center, Republican Hospital #1 - National Medical Centre, Ministry of Public Health of the Sakha Republic, Yakutsk, Russian Federation; 6) Department of Pediatric, Medical Institute, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation; 7) Department of Normal and Abnormal Anatomy, Operative Surgery with Topographic Anatomy and Forensic Medicine, Medical Institute, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation; 8) Institute of Foreign Philology and Regional Studies, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation; 9) Institute of Natural Sciences, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation; 10) Institute of Humanitarian Research and Indigenous Cultures, Bashkir State University, Ufa, Russian Federation; 11) Institute of Radiology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation; 12) Novosibirsk State University, Novosibirsk, Russian Federation; 13) Department of Genetics, Ufa Scientific Centre, Russian Academy of Sciences, Ufa, Russian Federation; 14) Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russian Federation.

Age-related hearing impairment (ARHI) of the Yakut population belongs to the frequent sensory disorders registered in 50% of individuals over 80 years. ARHI is a multifactorial disorder due to environmental and poor-known genetic components. In this study, we present the data on age-related hearing impairment of 48 heterozygous carriers of mutation IVS1+1G>A (GJB2 gene) and 48 homozygous carriers of single mutation IVS1+1G>A in the Yakut population. The age of deafness manifestation in individuals with genotype IVS1+1G>A/wt was estimated to be ~ 40 years (rs=0.504, p=0.003). These findings demonstrate that the single IVS1+1G>A mutation with genotype IVS1+1G>A/wt was estimated to be ~ 40 years (rs=0.504, p=0.003). Age of hearing loss manifestation in individuals with genotype wt/wt (p<0.05). Moreover, the average age of individuals with GJB2 genotype IVS1+1G>A/wt (both sexes) were significantly worse than in individuals with genotype wt/wt (p<0.05). Age of hearing loss manifestation in individuals with genotype IVS1+1G>A/wt was estimated to be ~ 40 years (rs=0.504, p=0.003). These findings demonstrate that the single IVS1+1G>A mutation in the Yakut population is associated with age-related hearing impairment (ARHI) of the IVS1+1G>A carriers in the Yakuts.
886S Functional study of Peptidylarginine deiminase type 4 as genetic risk factor for RA
A. Suzuki, Y. Koyab, F. Shoda
In a Japanese population a case-control association study (Ref 1). PAD4 is a member of the PAD family and converts arginine residue (peptidylarginine) to citrulline residue (peptidylcitrulline). PAD4 is highly expressed in bone marrow, macrophages, neutrophils and monocytes. Peptidylcitrulline is an interesting molecule in RA, because it is an antigen of ACPA and only PAD4 can generate peptidylcitrullines, via modification of protein substrates. To evaluate the importance of PAD4 gene in the progression of RA, we generated Pad4-/- D8A1J mice. We used Pad4-/- mice to show that Pad4 is significantly affected to progress of collagen induced arthritis (CIA), well known as an RA model animal. Expression of various inflammatory cytokines and Pad4 mRNA in immune cells was detected by real-time TaqMan assay. Cytokine concentrations in sera were measured by enzyme-linked immunosorbent assay. We demonstrated that Pad4 expression was induced by CI immunization. In Pad4-/- mice, inflammatory cytokine levels were significantly decreased compared with those in wild-type mice. Interestingly, Pad2 expression was induced in immune cells of Pad4-/- mice in compensation for the defect in Pad4.

887M Characterization of a TERT-CLPTML multi-cancer risk locus on chr5p15.33
Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD; 2) Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD; 3) Department of Health Sciences Research, Mayo Clinic, Rochester, MN.

Genome-wide association studies (GWAS) have mapped multiple independent cancer risk loci (n=6) to a small region on chr5p15.33 for at least ten distinct cancers, including bladder, breast, glioma, lung, melanoma, non-melanoma skin, ovarian, pancreas, prostate, and testicular germ cell cancer. This region harbors two plausible target genes, TERT which encodes the catalytic subunit of telomerase reverse transcriptase which maintains chromosomal ends by adding telomeres repeats, and CLPTML which encodes the exit and regulated transmembrane protease which prevents cancer cell growth, protect cells from apoptosis, and can induce abnormal cytokinesis. The most significant pancreatic cancer GWAS SNP on chr5p15.33 was rs401681 (P=3.7×10^-10; OR=1.19) located in the 3rd intron of the CLPTML gene. CLPTML knockout mice improved this signal by three orders of magnitude (P=1.4×10^-5) of the background. We identified a risk-conferring variant at this locus for multiple cancers. We were consistent with elevated levels of the TERT specific enhancer effects and influences that for pancreatic cancer, the multi-cancer risk locus within CLPTML. Our results indicate TERT late genotypes to expression of CLPTML in inhibition of expression. Furthermore, allele specific siRNA targeting showed stronger expression (average 60%, range 45-85%) but no effect on expression. Furthermore, allele specific siRNA targeting showed stronger expression (average 60%, range 45-85%) but no effect on expression. 2 per cancer type). Promoter siRNA targeting across the regulatory region harboring the putative functional SNP resulted in a strong inhibition of TERT expression (average 60%, range 45-85%) but no effect on CLPTML expression. Furthermore, allele specific siRNA targeting showed stronger inhibition of TERT expression from this risk allele than the protective allele. Current work focuses on identifying the protein(s) that differentially bind this variant through proteomics, and performing eQTL analysis to correlate genotypes to expression of TERT and CLPTML. Our results indicate that for pancreatic cancer (Ref 2), the multi-cancer risk locus within CLPTML (tagged by rs401681) may be explained by a single SNP that confers allele specific enhancer effects and influences TERT gene expression. The effects were consistent with elevated levels of the TERT expression in carriers of the risk allele (C allele) but not with the T (wild-type) allele of rs401681. We observed strong correlation of allele specific expression of CLPTML in the brain region of the transgenic fish with the C (risk) allele, but not with the T (wild-type) allele of rs1884302. In our in vitro results suggest that the C allele at rs1884302 may, in fact, be regulating expression of CLPTML while in our in vivo results indicate that the presence of the C allele at this SNP acts as an enhancer.

888S A non-coding variant near BMP2 associated with sagittal non-syndromic craniosynostosis in zebrafish, C. M. Justice, J. Kim, S. D. Kim, G. Yagnik, B. Carrington, R. Soo, A. F. Wilson, S. A. Boyadjiev
1) Genomics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 2) Department of Pediatrics, University of California Davis Medical Center, Sacramento, CA; 3) Zebrafish Core, Translational and Functional Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD.

Craniosynostosis (CS) is a common congenital malformation in which one or more of the cranial sutures of an infant skull fuse prematurely. Approximately 80% of the CS cases are non-syndromic with unknown etiology and approximately 18% of these involve the sagittal suture. Sagittal non-syndromic craniosynostosis (sNSC) has an estimated prevalence of about 2 per 10,000 live births. Our recent genome-wide association study of sNSC cases in total. Finally we demonstrated in an in vitro model that the "GG" genotype of the functional rs6265 lead to higher BMP2 expression, which could result in a higher risk to the AR phenotype. A common variant of the BDNF gene is associated with increased risk towards moderate-severe AR. The potentially causative "GG" genotype of rs6265 appears to augment expression of BDNF mRNA in whole blood as demonstrated in an in vitro system and thereby lead to disease through activating an inflammatory cascade.

889T Functional BDNF gene variants increase risk to moderate-severe allergic rhinitis (AR) A. Andiappan, J. M. Quiek, C. Schumann, A. National Institutes of Health, Bethesda, MD; 2) Claxton Jones, University of Rochester, Rochester, NY; 3) Department of Allergy and Rheumatology, Graduate School of Community Medicine, University of Tartu, Tartu, Estonia.

Background: BDNF is a secretory protein belonging to the neurotrophin family which regulates the inflammatory cascade leading to allergic disease. Interestingly, patients with BDNF levels have a high likelihood to suffer from allergy phenotypes such as AR, asthma and eczema. No prior genetic study has investigated the relationship between BDNF polymorphisms and moderate-severe AR susceptibility. Objective: Aim of the study was to assess the association of genetic variants of BDNF with moderate-severe allergic rhinitis, and to determine whether this has in any functionally consequences. Methods: TagSNPs spanning the BDNF gene were selected from the human HapMap CHB (Chinese) population. These BDNF tagSNPs were then tested for association with moderate-severe AR in a population of 2216 Shandong Chinese. The association was replicated in another independent Chinese population. TagSNPs with P values ≤0.001 were selected for final model. P values ≤0.05 were considered statistically significant. Results: After correction for multiple testing, 2 SNPs, rs1077664 and rs6265 were significantly associated with BDNF mRNA expression with whole blood. Furthermore, this association was also validated in a large eQTL meta-analysis comprising of 5416 individuals in total. Finally we demonstrated in an in vitro model that the "GG" genotype of the functional rs6265 lead to higher BMP2 expression, which could result in a higher risk to the AR phenotype. A common variant of the BDNF gene is associated with increased risk towards moderate-severe AR. The potentially causative "GG" genotype of rs6265 appears to augment expression of BDNF mRNA in whole blood as demonstrated in an in vitro system and thereby lead to disease through activating an inflammatory cascade.
890M
Functional Investigation of Celiac Susceptibility Gene LPP in T Cells.
B. Molloy, M. Freeley, E. Quinn, R. McGinn, A. Long, R. McManus. Department of Clinical Medicine, Institute of Molecular Medicine, St. James’s Hospital, Dublin 8, Ireland.

Celiac Disease (CeD) is a common, complex and chronic immune-mediated disease affecting the small intestine. CD4 T cells are known to play an important role in celiac disease etiology as they initiate an immune response to gluten displayed by antigen presenting cells. A large case control study using the Immunochip identified the transcription factor/adhesion protein Lipoma-preferred partner (LPP) as the most significantly associated non-HLA risk locus with an associated p value of 10^-49 (Trynka et al. 2011). LPP has known roles in cancer and smooth muscle migration but little is known if it plays a role in T cells or how it may contribute to celiac disease pathophysiology (Grunewald et al. 2009). mRNA sequence data of CD4 T cells from our lab (data not shown) shows that LPP is expressed at higher levels in celiac samples compared to controls. We aimed to investigate the role of LPP in T cells. In this study we wanted to firstly confirm the expression of LPP in CD4 T cells and examine the effect LPP may have on cell migration through siRNA knockdown. In addition, using qPCR we tested a number of potential LPP targets that demonstrated dysregulation in our sequencing study for differences in expression in the presence or absence of LPP. We confirmed LPP expression in peripheral blood T lymphocytes. T cells knocked down for LPP showed defects in transwell migration in response to chemotactic signals. Furthermore, preliminary data shows that when stimulated with the chemokine CXCL12, knockdown of LPP is associated with alterations in the mRNA levels of the potential LPP interactors or transcriptional targets MMP25, TIMP1 and CXCR4 suggesting a possible mechanism by which LPP contributes to disease pathophysiology. Ongoing investigation including the use of flow cytometry aims to further delineate the role of LPP in T cells.

891T
Why do Genetic “Risk Factors” for Major Diseases not Always Negatively Affect Survival? S. Ukrainskaya1, K. Arbeev1, A. Kulminski1, I. Akushevich1, D. Wu1, G. Joshi2, I. Culminekaya3, K. Land4, E. Stallard4, A. Yashin1. 1) Center for Population Health Aging, Duke University, Durham, NC; 2) University at Buffalo, NY.

Common complex diseases, such as cancer, CVD, diabetes and AD, are major contributors to mortality. However, genetic variants that have been associated with increased risks of such diseases are often found in genomes of long-lived people, and do not seem to compromise longevity. Here we discuss several genetic mechanisms that might plausibly explain the seemingly paradoxical situations in which genetic “risk factors” for major diseases may be neutral or even beneficial in relation to survival and longevity of their carriers. Such mechanisms include (but are not limited to): (i) trade-off-like effects of genes on risks of mortalities from different health disorders; (ii) age-specific influence of genes on vulnerability to diseases and death; (iii) gene-gene interaction (epistasis); (iv) gene-environment interaction. We review current evidence in support of this explanation and conclude that being a genetic risk factor for major disease does not necessarily mean being the risk factor for all-cause mortality. The net effect of a disease risk allele on person’s survival may be negative, neutral or positive, depending on a balance of detrimental and beneficial effects of such allele on various health and aging related traits. This balance may change with age, environmental conditions and genetic surrounding. Facilitating research on conditional effects of genes is of critical importance for better understanding the complex relationships between diseases and longevity and for advancing the area of personalized prevention.

892S
Disruption of the CTNN2D gene causes learning problems within the dyslexia spectrum. A. Lindstrand1,2,3, W. Hofmeister1,2, D. Nilsson1,2,4, A. Topa1,2, BM. Andersson1,2,3, F. Darko6, H. Matsson1, I. Tapia Páez1,2, T. Klingberg6, L. Samuelsson5, V. Wirta6, F. Vezzi9, J. Kere1,10, M. Nordenskjöld1,2,3, E. Syk Lundberg1,2,3. 1) Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; 2) Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden; 3) Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden; 4) Science for Life Laboratory, Karolinska Institutet Science Park, Solna, Sweden; 5) Department of Clinical Genetics, Sahlgrenska University Hospital, Gothenburg, Sweden; 6) Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden; 7) Department of Biosciences and Nutrition, and Center for Innovative Medicine, Karolinska Institutet, Huddinge, Sweden; 8) SciLifeLab, School of Biotechnology, KTH Royal Institute of Technology, Stockholm, Sweden; 9) SciLifeLab, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden; 10) Molecular Neurology Research Program, University of Helsinki, and Folkhälsoanst Institute of Genetics, Helsinki, Finland.

Cytogenetically visible chromosomal translocations are highly informative as they can pinpoint strong effect genes even in complex genetic disorders. Here we report a mother and daughter with borderline intelligence and learning problems within the dyslexia spectrum and two apparently balanced reciprocal translocations; t(1;8)(p22;q24) and t(5;18)(p15;q11). By low coverage mate-pair whole genome sequencing we were able to pinpoint the genomic breaks to 2 kb intervals. We then located the chromosome 5p breakpoint to intron 9 of CTNN2D by direct sequencing. An additional case with similar phenotypic presentation and a 163 kb microdeletion exclusively involving CTNN2D was identified with genome-wide array comparative genomic hybridization. This microdeletion at 5p15.2 is also present in mosaic state in the patient’s mother, but absent from the healthy siblings. We investigated the effect of CTNN2D polymorphisms on normal variability and identified a polymorphism (rs2561622) with significant effect on phonological processing. Finally, given the potential role of CTNN2D in neuron motility, we used morpholino knockdown in zebrafish embryos to assess its effects on neuronal migration in vivo. Analysis of the zebrafish forebrain revealed a subpopulation of neurons misplaced between the diencephalon and telencephalon. Taken together, our human genetic and in vivo data suggest that defective migration of subpopulations of neuronal cells due to haploinsufficiency of CTNN2D contribute to the cognitive dysfunction in our patients.
Aim: Decreased beta-cell function is a key player in development of type 2 diabetes. Large hypothesis-free studies aiming to identify new gene variants that explain the heritable component of beta-cell function apply various surrogate measures of insulin release. We examined to what extent common surrogate measures of insulin release have shared genetic causes. Methods: Genetic and phenotypic correlations were calculated in a family-cohort (n=269) in which beta-cell indices were estimated based on fasting and oral glucose-stimulated plasma glucose and serum insulin levels. Furthermore, we genotyped by the Metabochip a large population-based cohort (n=6269) for common genetic variants known to associate with type 2 diabetes, fasting plasma glucose levels or fasting serum insulin levels to examine their association with various indices. Results: The phenotypic and genetic correlations differed noteworthy for the traits compared, emphasizing that the phenotypic correlation is an insufficient measure of the magnitude of shared genetic impact. Also, we found that corrected insulin response, insulinogenic index, and area under the curve for insulin after an oral glucose challenge shared the majority of their genetic background with genetic correlation between 0.81 and 0.99. The BIGTT index for acute insulin release differed genetically from the latter traits with genetic correlations between 0.51 and 0.81 due to less influence by incretin-related genes and more influence by insulin sensitivity-related genes on the BIGTT index. The homeostasis model assessment for beta-cell function was genetically closely related to fasting insulin with a genetic correlation of 0.90. By examining variants known to associate with type 2 diabetes, fasting plasma glucose levels or fasting serum insulin levels in a large population-based study sample, it was evident that traits displaying a high level of genetic correlation tended to share a larger number of associated SNPs. We also identified a few SNPs that associate exclusively with only one of the measures of insulin secretion and the biological effect of these SNPs gave clues to understand how the indexes for insulin secretion could reflect different physiological mechanisms. Conclusion: The level of shared genetic background varies between surrogate measures of insulin release, and this should be considered when designing a genetic association study to best obtain information on different mechanisms of insulin release.

**Potential Transcriptional Mediators for Established Type 2 Diabetes Variants in Southwestern American Indians.**


A number of single nucleotide polymorphisms (SNPs) reproducibly associated with type 2 diabetes mellitus (T2DM) have been identified, but the molecular mechanisms by which these variants influence susceptibility to T2DM remain largely unknown. We analyzed potential mediation between T2DM variants and gene transcription in peripheral blood. Participants included 1416 American Indians (23% with T2DM), from urban Phoenix, Arizona, in whom transcriptomic measurements had been made on the Illumina HumanHT-12 v4 Expression Beadchip. Genotypes were generated for 44 established T2DM-susceptibility SNPs, 42 ancestry informative markers, used to control for admixture, and 42 “random” SNPs, which were used for genomic control. In these samples 7 of the T2DM markers had nominally significant (p<0.05) associations consistent with the established direction, including SNPs in GCK, CDC123, GRB14, SLC16A11, FTO and 2 SNPs in KCN11. We analyzed association of these 7 SNPs with all 15,854 unique transcripts that were significantly expressed and that, according to REMOAT, did not contain SNPs. This resulted in 85 transcripts that had suggestive association (p<0.005) with at least one of the 7 SNPs, and these were tested for potential mediation of the relationship between the SNP and T2DM using the Sobel test. Nominally significant mediation was identified for 4 transcripts, involving 3 SNPs (in SLC16A11, FTO and CDC123). The strongest results were for the SLC16A11 SNP rs75493593 and the RNASEK transcript, located 33 kb away. This SNP was associated with T2DM (odds ratio=1.36 per copy of the risk allele, p=0.0017) and the risk allele was associated with lower expression of RNASEK (by 0.17 SD per copy, p=1.6x10^-5). Lower expression of RNASEK was associated with T2DM (odds ratio=1.22 per SD decrease in expression, p=0.0062, controlled for, age, sex, heritage and the SNP). Mediation analysis was consistent with RNASEK expression as a partial mediator of the SNP effect (p=0.012). Additional potential mediators included rs75493593 and expression of ADAM15 (p=0.021), rs8050136 and expression of GSTF2 (p=0.041), and rs108591 (p=0.011) on CDC123 and expression of PRO0641 (p=0.048). These analyses identify potential transcriptional mediators of the relationship between T2DM-associated variants and the risk of type 2 diabetes, but the results need to be confirmed in additional populations.
A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. I. Molte1, 2, N. Grarup2, M.E. Jørgensen3, P. Bjerregaard2, J.T. Treебæk2, M. Fumagalli3, T.S. Kornelussen3, M.A. Andersen1, T.S. Nielsen3, N.T. Kraarup3, A.P. Gjesing2, J.R. Zierath2, A. Linneberg2, X. Wu2, G. Sun2, X. Jin2, J. Al-Aam2, J. Wang3, K. Borch-Johnsen1, 4, O. Pedersen2, R. Nielsen2, A. Albrechtsen2, T. Hansen1, 4, Department of Human Genetics, University of Chicago, Chicago, IL, USA; 2) The Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen, Denmark; 3) The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 4) Steno Diabetes Center, Gentofte, Denmark; 5) National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark; 6) Department of Integrative Biology, University of California, Berkeley, CA, USA; 7) Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark; 8) Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark; 9) BG-Shenzhen, Shenzhen, China; 10) Holbæk Hospital, Holbæk, Denmark.

The Greenlandic population is a historically small and isolated population, which has experienced a dramatic increase in type 2 diabetes (T2D) prevalence. Motivated by this, we performed genetic association mapping of four T2D-related quantitative traits in up to 2757 Greenlandic individuals without clinically diagnosed T2D. Using a linear mixed model to test for association while controlling for false positives due to admixture and relatedness, we analyzed Illumina MetaboChip data and exome sequencing data. Our analyses led to the discovery of a nonsense variant in the gene TBC1D4 with an allele frequency of 17% in the Greenlandic population. Under a recessive model, homozygous carriers of this variant have markedly lower plasma glucose (β=-0.18 mmol/L, P=1.1×10^-14) and serum insulin (β=165 pmol/L, P=1.5×10^-29) two hours after an oral glucose load compared to individuals with other genotypes. Furthermore, they have marginally lower plasma glucose (β=-0.18 mmol/L, P=1.1×10^-4) and serum insulin (β=-8.3 pmol/L, P=0.0014) at fasting and markedly higher T2D risk (OR 10.3, P=6.1×10^-24). Heterozygous carriers of the variant have a significant, but moderate, increase in 2-hour plasma glucose (β=0.43 mmol/L, P=5.3×10^-10) and a 3.4-fold increase in body mass index compared to individuals from a different Greenlandic cohort. The nonsense variant is located in an exon that is exclusive to a long isoform of TBC1D4, which is mainly expressed in skeletal muscle. Analyses of muscle biopsies showed decreasing mRNA and protein abundance and increasing regulation of copies of the variant. A similar decrease was observed in protein abundance of the glucose transporter GLUT4. Since TBC1D4 is known to be a mediator of insulin-stimulated glucose uptake in cells through regulation of GLUT4 mobilization, these results suggest that the variant, which leads to premature termination of the long TBC1D4 isoform, causes insulin resistance in skeletal muscle and thereby increased risk of developing T2D. The observed effect sizes are several times larger than any previous findings in large-scale genome-wide association studies of these traits and the identified variant accounts for more than 10% of all cases of T2D in Greenland. This finding provides new insights into T2D and constitutes further proof of the value of conducting genetic association studies outside the traditional setting of large homogeneous populations.

The Type 1 Diabetes Susceptibility Gene CLEC16A encodes protein which restrains NK Cells function. R. Pandey1, M. Bakay2, S. Yoeun3, J. Kushner4, H. Hakonarson4, 1 Centre for Applied Genomics, Children’s Hospital Of Philadelphia, Philadelphia, PA; 2) Department of Pediatric Medicine, Endocrine-Metabolism, Texas Children’s Hospital, Houston, TX.

Type 1 Diabetes (T1D) is a multi-factorial childhood disease with a strong genetic component. Several GWAS had showed association of CLEC16A with T1D, which encodes a protein of unknown function. The highest levels of CLEC16A expression were identified in NK cells, which are required for development of T1D. Therefore, defining the role of CLEC16A in NK cells will provide insight into the pathogenesis of T1D. We reported previously that protective alleles of CLEC16A [A/A] are associated with higher levels of mRNA. Thus, we hypothesize that CLEC16A functions in NK cells to restrain secretory functions including cytokine release and cytotoxicity. First we investigated the expression of CLEC16A in human immune cells and non-immune tissues on mRNA level by RT-PCR and on protein level by western blot. To address the role of CLEC16A in NK cells we studied consequences of knockdown in NK cell lines and ex-vivo NK cells and over-expression of this protein in NK cell line. Using a retroviral expression system, we created NK cell lines stably over-expressing the canonical form of CLEC16A with a GFP expression reporter. Expression was validated by Western blot analysis. Cytotoxicity and INF-γ production were decreased in NK cells stably overexpressing CLEC16A. Optimized protocols using CLEC16A siRNA mediated knockdown enabled a 70% reduction in Clec16A protein levels in NK cells and 35% increase in cytotoxicity compared to cells receiving control siRNA. We performed conjugation assay to rule out decreased target cell killing. CLEC16A over-expressing NK cells formed smaller number of conjugates for all time points with no difference in CD107a expression. Subcellular localization studies revealed cytosolic localization. CLEC16A knockdown in mice resulted in increase in NK cell cytotoxicity of the YAC-1 targets in comparison to control. Taken together, our results indicate that CLEC16A serve a role in restraining two major functions of NK cells, cytotoxicity and cytokine release. Studies are in progress to validate the mechanism. Our improved understanding of this novel TD1-linked gene and the protein it codes will likely suggest new therapeutic interventions in T1D.
989S
DIO2 rsSNPs, transcriptional factor binding sites and disease. N. Buroker. Pediatrics, 356320, University of Washington, 1959 Pacific Ave NE, Seattle, WA.

The TIO2 gene transcribes deiodinase type 2 that converts the thyroid prohormone, thyroxine (T4), to the biologically active triiodothyronine (T3). The thyroid hormone T3 plays an important role in the regulation of energy balance (SNP). Regions (SNPs) in the promoter region novel SNP (-2035bp), 5'UTR (rs12885300), intron one (rs225010, 225011 and rs225012), exon two (rs225014) and 3' UTR (rs6745495 and rs225015) of the DIO2 gene are in linkage disequilibrium. These rsSNPs alter the DNA landscape for potential transcriptional factors (TFs) to attach resulting in changes in transcription factor binding sites (TFBS). The alleles of each rsSNP were found to produce unique TFBS resulting in potential changes in TF DIO2 regulation. These regulatory changes are discussed with respect to disease and sickness.

989M
Harnessing genome engineering to characterize the role of STRs in gene regulation. D. Zielinski1, M. Gynmek1, 2, Y. Erlich1. 1) Whitehead Institute for Biomedical Research, Cambridge, MA; 2) Harvard-MIT Division of Health Sciences and Technology, MIT, Cambridge, MA.

Most variants associated with complex traits reside in non-coding DNA, suggesting an important role for cis-regulatory elements. Efforts to discover cis-eQTLs (expression Quantitative Trait Loci) have mainly focused on the contribution of SNPs to gene expression. However, several candidate gene studies in human and model organisms suggest that Short Tandem Repeat (STR) variations can modulate expression levels and splicing of nearby transcripts. After identifying significant associations between STR variations and expression profiles across hundreds of samples from the 1000 Genomes Project, we are experimentally validating the effects of these candidate STRs on expression levels in human lymphoblastoid cell lines. Specifically, we are using the CRISPR/Cas genome engineering system to generate a distribution of STR alleles and measuring the associated gene expression levels. In addition to validating our in silico findings, this technique will allow us to fine map individual eQTL signals and determine whether the STR is the true causal signal at these candidate loci.

900T
Functional genomics of the costimulatory locus in autoimmune disease. L. Petukhova1, 2, T. Yamany2, L. Bian2, Z. Dai2, E.N. Drill2, E.D. Broadbent1, P.L. Nagy2, R. Clynes1, 2, A.M. Christiano1, 2, 3, E.N. Drill1, 2, 3, 5, J.D. Broadbent4, D. Zielinski1, L. Bian2, Z. Dai2, E.N. Drill2, E.D. Broadbent1, P.L. Nagy2, R. Clynes1, 2, A.M. Christiano1, 2, 3, 1) Department of Epidemiology, Columbia University, New York, NY; 2) Department of Dermatology, Columbia University, New York, NY; 3) Department of Genetics and Development, Columbia University, New York, NY; 4) Department of Pathology and Cell Biology, Columbia University, New York, NY; 5) Department of Biostatistics, Columbia University, New York, NY; 6) Department of Medicine, Columbia University, New York, NY.

The costimulatory locus contains three key immunoregulatory genes (CD28, CTLA4, and ICOS), and was one of the first genomic regions associated with autoimmune disease outside of the HLA locus. GWAS have provided robust agnostic evidence for association of this region with type 1 diabetes, rheumatoid arthritis, celiac disease, Graves disease and alopecia areata (AA). Despite the extensive evidence for conferring risk, the mechanisms by which genetic variants contribute to disease have remained elusive. Furthermore, the efficacy of therapies that target the costimulatory pathway demonstrates that it is a critical axis of autoimmunity for some patients. In order to systematically identify all disease risk variants at this locus, we performed targeted re-sequencing of the costimulatory locus in 122 AA patients from our GWAS cohort, targeting the entire 297Kbp region. This experiment identified 1209 variants that passed rigorous QC filters, which we computationally phased, allowing us to assign alleles to chromosomes. Among the 244 sequenced chromosomes, we identified 88 chromosomes that carried GWAS-identified risk haplotypes. We next identified 208 variants that were significantly enriched on the chromosomes carrying GWAS risk haplotypes (p<4.1e-5). Among these enriched SNPs, there is one CTLA4 protein coding variant (rs231775; p. T17A), and eight SNPs annotated with regulatory mutations in CD4 cells in public databases. Importantly, we validated regulatory effects by demonstrating that four of these regulatory variants affect the distribution of CTLA4 isoforms in activated human T-cells and show that this shift inhibits T cell cytokine production, and TCR-mediated MAPK activation in CTLA4 deficient Jurkat cells. The identification of risk variants with biological consequences perturbing the costimulatory axis provides a crucial step forward in understanding how this locus contributes to autoimmune disease.
902M A Genomewide Association Study of Alcohol Dependence in the Irish Affected Sib Pair Study of Alcohol Dependence. B.P. Riley1, 2, 3, 4, A.E. Adkins1, 2, 3, 4, C.M. Hackot1, 2, 3, 4, T.B. Bigdeli1, 2, 3, 4, M.S. Groteveld1, 2, 3, 4, A.G. Davies1, 5, J.C. Bettinger1, 5, C.A. Prescott1, M. Mamdani1, 2, 3, 5, V. William-1, 2, 3, 4, D.M. Dick1, 1, 2, 3, 4, V.I. Vladimirov1, 2, 3, 4, B.T. Webb1, 2, 3, 4, K.S. Kendler1, 2, 3, 4. 1) VCU Alcohol Research Center, Virginia Commonwealth Univ, Richmond, VA; 2) Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth Univ, Richmond, VA; 3) Dept of Psychiatry, Virginia Commonwealth Univ, Richmond, VA; 4) Dept of Human & Molecular Genetics, Virginia Commonwealth Univ, Richmond, VA; 5) Dept of Pharmacology and Toxicology, Virginia Commonwealth Univ, Richmond, VA; 6) Dept of Psychology, University of Southern California, Los Angeles, CA; 7) Dept. of Psychology, Virginia Commonwealth Univ, Richmond, VA.

Background: We report results from a genomewide association study (GWAS) in an ethnically homogeneous Irish sample (N=706 related cases; 1755 population controls) with strong supporting evidence from VCU Alcohol Research Center (VCU ARC) model organism (MO) studies. Methods: GWAS cases from the Irish Affected Sib Pair Study of Alcohol Dependence (IASPASD) were diagnosed using DSM-IV criteria. Affymetrix V6.0 arrays were genotyped at 3 separate core facilities and BeagleCall was used to call genotypes. IMPUTE2 and the 1000 Genomes reference haplotype panel (March 2012 freeze) were used to impute unmeasured genotypes. After QC filtering, imputation, and post-imputation filtering, 710 AD cases, 1755 controls and 8.2 million SNPs remained. Probabilities were converted to dosages with MACH2. Case/control association analysis was run using MQLS to correct for the non-independence of siblings. A sex weighted prevalence estimate of 8.9% was used for controls. We used a significance threshold of p<3.06E-8 based on the number of independent LD blocks in 1000 genomes data. FDR q-values were calculated with QVALUE in R. Results: SNPs in the COL6A3 gene on chromosome 2 and an intergenic region of chromosome 3 near the ECT2 gene were genomewide significant. SNPs in 8 independent genomewide regions had FDR q-values <1% (notably in KLF12, POMT2/TMED8, LOC339975 and RYR3). Knockdown of a COL6A3 ortholog by RNAi in C. elegans leads to significant reductions in the development of acute functional tolerance (AFT), Knockdown of POMT2 by RNAi in D. melanogaster led to increased spasticity as part of the immune system. We previously identified a missense polymorphism, rs1143679 within ITGAM acts resistance to ethanol. Human replication data support LOC339975, a long loss of function allele of unc-68, the C. elegans homolog of RYR3, confers resistance to ethanol. Human replication data support LOC339975, a long non-coding RNA predicted to regulate 12 microRNAs that in turn regulate genes including numerous genes previously implicated in alcohol phenotypes such as SLC6A3, GABRG1, CNR1, ESR1, ACSL4 and multiple alcohol aldehyde dehydrogenase genes. Discussion: Our case-control GWAS detected numerous associated loci supported by functional assays in model organisms or directly linked to implicated genes in humans, a powerful combined approach elucidate genetic factors underlying alcohol dependence.

903T eMERGE Phenome-Wide Association Study (PheWAS) Identifies Clinical Associations and Pleiotropy for Functional Variants. A. Verma1, S. Verma1, S. Pendergrass1, D. Crawford2, D. Crosslin3, H. Kuvianen4, W. Bush2, Y. Bradford1, I. Kullo1, S. Bielinski5, R. Llo3, J. Denny6, P. Peissig6, S. Hebing7, E. Pugh5, M. Andrade6, M. Ritchie7, G. Tromp7. 1) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA; 2) Geisinger Health System, Danville, PA; 3) Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA; 4) Vanderbilt University, Nashville, TN; 6) Marshfield Clinic, Marshfield, WI; 7) Mayo Clinic, Rochester, MN; 8) National Human Genome Research Institute, Bethesda, MD; 9) John Hopkins University, Baltimore, MD.

We performed a phenome-wide association study (PheWAS) exploring the association between stop-gained genetic variants and a comprehensive group of phenotypes to identify novel associations and potential pleiotropy. Using multiple bioinformatics tools we selected 38 functionally relevant stop-gained/missense genetic variants within the genotypic data of 37,972 unrelated patients from seven study sites in the Electronic Medical Records and Genomics (eMERGE) Network. We calculated comprehensive associations between these variants and case-control status for 3,518 ICD-9 diagnosis codes (requiring ≥ 3 visits per individual to identify case status, ≥10 case subjects per ICD-9 code). All analyses were adjusted for sex, site, platform and the first 3 principal components. A total of 418 associations passed a liberal significance threshold of p< 0.01. The most significant association was between GLG1 rs9445 and “chronic non-alcoholic liver disease” (p=1.2×10-5, β=2.60). We identified many potentially pleiotropic associations at p< 0.05, 35 out of 68 SNPs demonstrated associations with more than one phenotype, and 17 SNPs were each associated with > 10 different ICD-9 codes. For example, we found associations for IL34 rs4985556 with 25 diagnoses, such as “lupus erythematous” (p=5.94 ×10-3, β=0.98) and for GBE1 rs2229519 with 33 diagnoses, such as “hypertension” (p=1.2×10-3, β=0.067), “hyperlipidemia” (p=6.66 ×10-3, β=0.058), and “ocular hypertension” (2.49 ×10-3, β=0.21). We will seek replication of these results. In conclusion, our PheWAS shows stop-gained variants may have important functional effects and that PheWAS are a powerful strategy to mine the full potential of the EMR for genome-phenome associations.

904S Lupus associated coding polymorphism rs1143679 within ITGAM acts in both nucleotide and protein level to develop disease phenotypes. A.K. Maitil1, K. Bhattachari1, P. Molghare1, X. Kim-Howard1, J-P. Anaya1, S.K. Nath1. 1) Gen Epidemiology Unit, A & CI, OMRF, Oklahoma City, OK; 2) Universidad del Rosario-Corporación para Investigaciones Biológicas, Bogotá, Colombia.

Integrin alpha M (ITGAM; CD11b) is a component of the macrophage-1 complex, which mediates leukocyte adhesion, migration and phagocytosis as part of the immune system. We previously identified a missense polymorphism, rs1143679 (R77H) at the exon 3, that is strongly associated with systemic lupus erythematosus (SLE). We explored the molecular mechanism by which the rs1143679 risk allele in ITGAM contribute to the development of SLE disease manifestations. We show that rs1143679 carrying DNA sequence acts as strong transcriptional enhancer and the risk allele affects the expression of ITGAM gene in vivo. This SNP binds with NFκB1, EBF1 and Ku70/80 protein in vitro and in vivo in the monocytes of SLE patients and healthy individuals but with reduced efficiency in the presence of risk allele. The observed reduced ITGAM expression in risk SLE allele carrying individual's monocytes is attributed to reduced interaction of these proteins with RNA pol II at the ITGAM promoter implicating reduced transcription of this gene resulting in less mRNA and protein production. Ku70/80 is a lupus auto-antigen and apart from its DNA repair function, also acts as transcription factor and binds with several enhancer SNPs that are associated with lupus. Using NextGen ChIP-seq, we mapped Ku70/80 binding genomic regions that could indicate lupus associated genomic regions. Our genome-wide analysis revealed the discovery of lupus associated genes that increased transcription levels due to reduced transcription of ITGAM gene resulting in reduced protein production as well as mutant protein production that could explain the strong association of this SNP with lupus.
**905M**

Host genetic variation and Kaposi's Sarcoma-associated herpesvirus infection. N. Sallat, C. Franklin, A.-Palser, P. Kellam, I. Barroso. Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom.

Kaposi's sarcoma-associated herpesvirus (KSHV) is an oncogenic gammaherpesvirus associated with a variety of lymphoproliferative diseases particularly Kaposi's sarcoma (KS). Primary effusion lymphoma (PEL) and Multicentric Castlemans disease (MCD). The virus establishes a lifelong infection with a lytic phase in which active virus replication facilitates spread and pathogenesis; and an immunologically silent latent phase that promotes persistence in B-cells. Despite high seroprevalence, only a small proportion of infected people develop tumours, in addition, striking geographic distribution and familial clustering of disease is suggestive of a possible genetic predisposition to disease. Host and virus genetic variation and their influence on KSHV infection and epidemiology remain largely uncharacterized. Here we establish a pipeline to investigate how the genetics of host-virus interactions influence KSHV pathogenesis. Studies have identified a B-cell transcription factor, X-box binding protein-1 (XBP-1) as a key regulator of the lytic switch that links terminal differentiation, B-cell receptor activation and virus production. Thus far, we have identified 40 genetic variants of XBP-1 in whole genome and exome datasets and are currently using PEL cell lines to investigate the biological effects of polymorphisms found in known functional domains. The pipeline used here from identification and annotation of variants to assaying for phenotypic differences can be employed for other host genes important in viral pathogenesis to elucidate function.

**906T**

Studying the effects of pubertal timing-associated gene LIN28B on early vertebrate development using zebrafish as a model. J.T. Leino nen1, Y. C. Chen2, H. Kovula3, P. Panula4, E. Widén4. 1) Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Finland; 2) Neuroscience Center and Institute of Biomedicine, Anatomy, University of Helsinki, Finland.

Sequence variation near LIN28B has repeatedly been associated with pubertal timing in several genome-wide association studies, and animal models suggest that the evolutionarily conserved LIN28 genes encoding for RNA-binding proteins are crucial regulators of development and growth, expressed already during fetal life. Normal puberty requires appropriate fetal development. The mechanism linking LIN28B to puberty is unknown, but we hypothesized it may influence early developmental processes, in particular the formation of the hypothalamic pituitary gonadal axis. Our specific aim is to elucidate the function of the puberty-associated gene LIN28B by assessing 1) its embryonic temporal and spatial expression patterns, 2) the impact of its downregulation and, 3) the impact of its upregulation in developing zebrafish, a model system allowing efficient evaluation of early development and growth. Knockdown- and overexpression-experiments showed that the correct function of LIN28B is essential for proper embryonic development. Both LIN28B knockdown, as induced by morpholino oligonucleotide injection, and overexpression, induced by mRNA injections, caused a dose-dependent variation of phenotypes, with high doses being lethal and lower doses causing severe to mild malformations. The brain morphology associated with LIN28B downregulation was severely compromised. Co-injection of morpholino and mRNA resulted in phenotypic rescue, suggesting that the knockdown phenotypes are specific to LIN28B downregulation. The spatiotemporal expression patterns of LIN28B in wild-type zebrafish assessed by RNA in situ hybridization on whole mount embryos and tissue sections, indicated that while the expression is not spatially restricted during the first 24 hours of development, it subsequently appears most prominent in developing organ primordia, including regions in the telencephalon. Interestingly, the timing of the formation and migration of gonadotropin releasing hormone (GnRH) neurons to the hypothalamus, which is essential for normal sexual development in vertebrates, coincided with LIN28B expression. We therefore intend to investigate whether LIN28B and GnRH3 expression may be co-localized. Taken together, the experiments indicate that changes in LIN28B expression have far-reaching consequences for the development of vertebrate body plan, and we further plan to study the impact of LIN28B on the development of the GnRH neuronal network.

**907S**


A rare missense variant R47H in the TREM2 gene in the TRM2 gene is recently identified strong risk factor for Alzheimer disease (AD). TREM2 is expressed by microglia and plays a role in immune responses in the CNS. We report a large multigeneration family affected with late-onset AD in which multiple affected members (11/15) carry the variant. Neuropathologic examination of 11 available brains from family members found features typical of AD that included neuritic amyloid plaques and neurofibrillary tangles (Braak stages 5-6) in all affected patients. Vascular pathology features were common in the R47H-positive patients (4/7). Results of ongoing assessments of microglia status in TREM2-R47H carriers will be reported.

**908M**

Functional study of a novel unexpected interferon-responsive gene, GRAMD1B, identified in Multiplex MS families. F. Martinelli Boneschi1, A.M. Osiceanu1, F. Esposito1, A. Zauli1, M. Sorosina1, B. Bettagazzi4, D. Cittaro1, E. Maccia1, S. Santoro1, A. Calabria1, D. Lazarovic1, D. Zacchetti1, G. Comi1, E. Stupka2, 1) DEPT NEUROLOGY AND INSPE, SCIENTIFIC INSTITUTE SAN RAFFAELE, MILAN, MI, Italy; 2) CENTER OF TRANSLATIONAL GENOMICS AND BIOINFORMATICS, SCIENTIFIC INSTITUTE SAN RAFFAELE, MILAN, MI, Italy; 3) SAN RAFFAELE TELETHON INSTITUTE FOR GENE THERAPY (HSR-TIGET), SCIENTIFIC INSTITUTE SAN RAFFAELE, MILAN, MI, Italy; 4) CELLULAR NEUROPHYSIOLOGY UNIT, SCIENTIFIC INSTITUTE SAN RAFFAELE, MILAN, MI, Italy.

BACKGROUND: While the role of common genetic variants is clearly established in multiple sclerosis (MS[MIM126200]), the heritability of the disease is still poorly explained, suggesting the existence of rare variants implicated in the susceptibility to the disease. OBJECTIVE: To identify low-frequency and rare genetic variants contributing to MS susceptibility in an Italian multiplex family. DESIGN/METHODS: SNP microarray genotyping and whole-genome sequencing (mean coverage: 20x) in 4 MS patients and 4 unaffected individuals belonging to an Italian multiplex family originating from a first cousin consanguineous marriage have been performed. The Merlin software was used for the linkage analyses and SNPeff and GATK software were applied to prioritize rare variants. RESULTS: Filtering criteria, including a linkage analysis that identified a unique signal showing a LOD score >2.0 at chr11q23.3, narrowed down the list of variants up to a rare functional variant which determines an amino acid change, S601P, in an unexplored gene, GRAMD1B. The mutation is in a context that is highly conserved across species, and it segregates along the family consistent with an autosomal recessive transmission (p=0.02). By performing WGS experiments, we found that the gene was downregulated in affected relatives (p=0.01) with the exception of the only case who was IFNβ treated. We performed a IFNb stimulation of PBMCs isolated from 20 healthy controls, showing an increase in GRAMD1B expression (p<0.001). We also observed a significantly higher expression of GRAMD1B in the brain tissue, and in immune cells, compared to other cells and tissues (p<0.05). Moreover, we observed a significant increase in expression in activated rat microglial cells compared to unstimulated ones (p<0.05), suggesting a putative role of the protein in the modulation of the immune system. Ongoing experiments are aimed to: 1) explore the role of GRAMD1B in the IFNb pathway; 2) identify the localization of the protein; 3) study the effect of the mutation on the expression, localization and/or function of the GRAMD1B protein in MS etiology. CONCLUSIONS: The use of next-generation sequencing approach in an Italian MS multiplex family has been successful in identifying a novel rare variant in an unknown gene recently found to be associated with IgE levels. Further investigations are ongoing to explore the role of the variant.
200T The causal basis of Hirschsprung disease risk: functional consequences of polymorphisms in two RET shadow enhancers. S. Chaf-tanjee, A. Kapoor, A. Chakravarti. Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD.

A significant proportion of the genetic risk of Hirschsprung disease (HSCR) in European ancestry subjects arises from two common variants at the gene encoding the receptor tyrosine kinase RET. We have previously shown that the first variant, rs2435357 (risk/non-risk alleles T/C, T control allele frequency 27%, odds ratio –4), lies within RET intron 1 and disrupts binding of the transcription factor Sox10. We now demonstrate that the second variant, rs2506690 (risk/non-risk alleles G/A, G control allele frequency 40%, odds ratio –2), maps 125kb upstream of RET and affects the binding of the transcription factor, Retinoic acid receptor beta (RARβ). In mouse Neuro2a cells, siRNA mediated knock down of either transcription factor reduces Ret expression, by 8 fold for Sox10 (P< 1.8x10^-7) and 5 fold for RARβ (P=4x10^-4). Surprisingly, when Ret is significantly (P=1.1x10^-3) depleted by siRNA, expression of Sox10 is decreased by 5 fold (P=2x10^-5) but RARβ remained unaffected (P=0.66), suggesting that Sox10 and Ret regulate each other. These results are supported by in vivo gene expression data in Ret homzygous null mice gut at 2 stages in development (E11.5 and E12.5), as compared to wild type mice at the same stages, where early gut neurogenesis occurs: Sox10 expression is reduced 4 fold (E11.5, P=1.2x10^-4) and 8 fold (E12.5, P=1.2x10^-4) but RARβ expression is, as expected, unaffected (E11.5, P=0.64; E12.5, P=0.19). In concert, Gdnt, the Ret ligand, expression is increased 1.5 fold (E11.5, P=0.03) and 3.4 fold (E12.5, P=6x10^-3) as is the expression of the co-receptor Gfra1 (3.7 fold at E11.5, P<0.02; 3 fold at E12.5, P<0.03). These in vitro and in vivo data show the presence of compensatory mechanisms within the Ret gene regulatory network. Specifically, Ret, the major gene for HSCR, is a positive regulator of its own transcription through Sox10. These results explain how common polymorphisms can lead to large genetic effects in HSCR since disruption of the transcriptional activity at the HDL-C GWAS locus.

910T Functional regulatory assessment of the APOL1 kidney disease risk variants. P. An1, J. Kopp2, C.A. Winkler1. 1) Basic Research Laboratory, Leidos Biomedical Research, Inc., Frederick National Lab for Cancer Research, Frederick, MD; 2) Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Disease, Bethesda, MD.

Background: Apolipoprotein L1 (encoded by APOL1) is a serum apolipoprotein bound to high-density-lipoprotein (HDL) particles. Common trypanolytic APOL1 coding variants G1 (S342G:IS84M) and G2 (2 aa deletion) are associated with markedly increased risk to kidney diseases including focal segmental glomerulosclerosis (FSGS) (OR 2-29) in Africa Americans and are correlated with markedly increased risk to kidney diseases. These data suggest a unique and complex molecular interaction that alters transcriptional activity at the APOL1 HDL-C GWAS locus.

Methods: We performed electrophoretic mobility shift assay (EMSA) to assess the binding of APOL1 G1/G2 with nuclear transcription factors in the HEK 293 cell line, real-time PCR to quantify gene expression in the lymphoblastoid cell lines (LCLs) derived from FSGS patients, and gene reporter assay to test the regulatory potential of the variants. Results: We observed that G1 and G2 variants conferred differential transcription factor binding patterns, G2 allele carriers had increased mRNA gene expression, and the 3’ terminal appeared to confer regulation activity. Our preliminary experimental data suggest that the APOL1 G1/G2 variants may influence APOL1 gene expression via differential transcription factor binding and gene regulation. Summary: APOL1 gene expression may be differentially regulated by G1 and G2 variants. Further investigation is required in podocytes and other renal cell types. (Funded by the National Cancer Institute Contract HHSN261200800001E).
An atopy-associated variant in the 11q13.5 locus regulates promoter activity. J. Manz1,2, A. Kretschmer1,2, G. Möller1, A. Peters1,2, J. Adamski1, M. Waldenberger1,2, S. Weidinger1. 1) Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany; 2) Institute of Epidemiology II, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Neuherberg, Germany; 3) Department of Dermatology, Venereology and Allergy, University Hospital Schleswig-Holstein, Kiel, Germany; 4) Institute of Experimental Genetics, Genome Analysis Center, Helmholtz Zentrum Muenchen, Neuherberg, Germany.

Atopic dermatitis (AD) is a common chronic inflammatory skin disorder with a strong genetic component. Genome-wide association studies (GWAS) robustly identified common risk variants at 11q13.5 located in the intergenic region between leucine rich repeat containing 32 (LRRC32) and chromosome 11 open reading frame 30 (C11orf30). The same locus has also been associated with asthma, hayfever and Crohn’s disease (CD). One GWA proxy SNP (rs2155219) is associated with the expression of LRRC32 and C11orf30, but the functional relevance of this SNP has not been investigated yet. The aim of this study was to characterize the functional impact of rs2155219 on cis-regulatory transcriptional activity and differential transcription factor binding. Reporter vector constructs carrying the major or minor allele of rs2155219, including two additional variants naturally occurring in this sequence (rs34455012indel; rs11236797C>A), were tested for regulating transcription activity in HaCaT (human keratinocyte), A549 (human alveolar basal epithelium), HeLa (human cervical cancer) and Jurkat (human T lymphocyte) cells using luciferase assays. In addition, different variant combinations within this haplotype (differing from the natural occurrence of the risk allele composition) were investigated. Differential transcription factor binding was analyzed by electrophoretic mobility shift assays (EMSAs). Reporter vector constructs containing rs2155219 together with the naturally occurring variants acted as enhancer on promoters in HaCaT, A549, HeLa and Jurkat cells. The AD- and CD-risk associated T allele of rs2155219 in combination with rs34455012del and rs12236797A showed significantly higher promoter activity compared to the non-risk allele (in concert with the A1 allele), resulting in A1A1 (12.4%), A1A2 (33.3%) and A2A2 (54.3%) genotypes. The frequency of the A1 allele in obese and eutrophic controls was 34.5%, and 23% respectively (p=0.05; RR=1.3). This allele was also associated with lower triglycerides (TG) levels and increased parents’ BMI. The children were divided into groups according to the reference values of the metabolic variables studied. Significant difference in allelic distribution was observed in children with total cholesterol (TC) <170mg/dl or TC ≥ 170mg/dl. The HOMA β was abnormal in 52.7% of obese and 10% of eutrophic patients. Allele A1 was present in 38.2% of children with altered HOMA β and 24.6% of children with normal HOMA β (p=0.037; RR=1.5). Up to the present, only 5 studies have studied the allele A1 frequency in children. We accessed the nutritional state of 105 children (55 obese/50 eutrophic) based on the definition of the World Health Organization (z score of BMI). Peripheral blood samples were taken to determine the lipid profile, glucose, insulin and for analysis of DRD2 polymorphism through RFLP-PCR. The Homeostatic Model Assessment (HOMA) was calculated. We found A1 and A2 alleles, resulting in A1A1 (12.4%), A1A2 (33.3%) and A2A2 (54.3%) genotypes. The frequency of the A1 allele in obese and eutrophic groups was 34.5% and 23% respectively (p=0.05; RR=1.3). This allele was also associated with lower triglycerides (TG) levels and increased parents’ BMI. The children were divided into groups according to the reference values of the metabolic variables studied. Significant difference in allelic distribution was observed in children with total cholesterol (TC) <170mg/dl or TC ≥ 170mg/dl. The HOMA β was abnormal in 52.7% of obese and 10% of eutrophic patients. Allele A1 was present in 38.2% of children with altered HOMA β and 24.6% of children with normal HOMA β (p=0.037; RR=1.5).
understanding of the both disease-causing repeat mutations, the functions be particularly exacerbated in Huntington's patient brains. This furthers our repeat mutations and aRPA may modulate this, a phenomena that may and processing of slipped-DNAs. However, high levels of aRPA (at aRPA-control brains, while expression of RPA is unaltered. We assessed the ability andive RPA (aRPA) has been shown to be expressed in human tissues at a sessions in the cerebellum versus the larger expansions in the striatum - which lum compared to the striatum correlated with the lower levels of CAG expansion, thereby highlighting a need to understand the expansion process or used somatic CAG expansions can contribute to disease onset and progression.

Posters: Complex Traits and Polygenic Disorders

915T
The International Genomics & Translational Research in Transplanta
Network (iGeneTrain). J. van Sellten1, Y.R. Li23, M.V. Holmes2, S. Balla1, K. Birdwell1, P.I.W. de Bakker1, P. Jacobson1, Y.H. Wang1, A. K. Israni1, B.J. Keating1, E. Schadt1, Y.R. Li2, Y. H. Wang2, A. LaSpada2, C.E. Pearson2. 1) The Hospital for Sick Children, Toronto, Ontario, Canada; 2) University of Pennsylvania, Philadelphia, USA; 2) Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 3) Forest School of Medicine, Winston-Salem, NC; 4) Department of Medicine, University of Maryland, Baltimore, MD; 5) National Institute on Aging, National Institutes of Health, Bethesda, MD; 6) Framingham Heart Study, National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health, Framingham, MA.

916S
Repair activity of primate-specific alternative single-stranded DNA binding protein aRPA may explain brain-region repeat instability in CAG/CTG trinucleotide repeat diseases. A. Shaked1, S. Shaked1, J. S. Lan1, S. Shaked1, Y.H. Wang2, A. LaSpada2, C.E. Pearson2. 1) The Hospital for Sick Children, Toronto, Ontario, Canada; 2) University of Pennsylvania, Philadelphia, USA; 2) Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 3) Forest School of Medicine, Winston-Salem, NC; 4) Department of Medicine, University of Maryland, Baltimore, MD; 5) National Institute on Aging, National Institutes of Health, Bethesda, MD; 6) Framingham Heart Study, National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health, Framingham, MA.

917M
Novel indoleamine 2, 3-dioxygenase (IDO) gene mutation in the Pathogenesis of age-related cataract. P. Gundã1, M. Mamata1, R.K. Kondreddy2, N. Thurapiti3, P. Trinurati1, 1) Dept of Genetics, Osmania Univer-
sity, Hyderabad, India; 2) Sarojini Devi Eye Hospital, Mehdipatnam, Hyderabad, India; 3) Department of Zoology, Osmania University, Hyderabad, India.

Purpose: Exposure to UV light is considered as the major risk factor for the development of age-related cataract. UV light produces during the tryptophan catalysis maintain the transparency of the lens and also protect retina from photodamage. Indoleamine 2, 3 dioxygenase (IDO) is the first rate limiting enzyme in the tryptophan catalysis which is encoded by IDO gene located on chromosome 17q21-22. Mutations in IDO gene can affect synthesis of UV filters qualitatively or quantitatively. Hence the present study was planned to screen for mutations in IDO gene and to evaluate their role in the causation of age-related cataract. Methods: Geno-

918T
Genetic Variants Concomitantly Influence Nonalcoholic Fatty Liver Disease and Correlated Metabolic Traits. M.F. Feidosa1, A.T. Kraja2, E.K. Speight23, M.K. Wojczynski, N.D. Palmer3, L.M. Yerges-Armstrong3, T.B. Harris3, C.S. Fox4, I.B. Borecki3, GOLDC Consortium, CHARGE Adiposity Consortium, GIANT Consortium, MAGIC Consortium. 1) Dept of Genetics, Osmania University School of Medicine, St. Louis, MO; 2) Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 3) Forest School of Medicine, Winston-Salem, NC; 4) Department of Medicine, University of Maryland, Baltimore, MD; 5) National Institute on Aging, National Institutes of Health, Bethesda, MD; 6) Framingham Heart Study, National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health, Framingham, MA.

Nonalcoholic fatty liver disease (NAFLD) is associated with metabolic disorders, including overall and central obesity, dyslipidemia (high levels of triglycerides and low levels of high-density lipoprotein cholesterol), and impaired glucose tolerance (high fasting levels of glucose and insulin). Several genes have been identified by genomewide association (GWA) studies for NAFLD and for these correlated metabolic disorders; however, few efforts have focused on determining whether the discovered genes/loci concomitantly influence these traits. To identify common variants contributing to pleiotropic effects on NAFLD and correlated metabolic traits, we employed a correlated meta-analysis (CMA) assessing the large published GWA meta-analysis Consortium data of CT measured fatty liver (FL; GOLD) with (i) triglycerides and (ii) high density cholesterol (GLGC), (iii) fasting levels of glucose and insulin, (iv) waist circumference (GIANT), and (v) waist circumference (CHARGE). Our methodological approach yielded evidence of genetic variants (e.g. ZNF14126068822, SUGP2-rs3810444, NOTCH4-rs1303320, ATRAID-rs7437, NAT2-rs1495743, SMP3D-rs2863973, and GCK-rs780093) with pleiotro-
ic effects on CT measured FL and these metabolic traits. In the current study, the CMA approach made the opportunity to discover novel variants (APBA1-rs1330326, TWY3-rs11161851, LOC400940-rs2693827 and HLA-DK22-rs2863910) that did not reach the GWA significance level in the primary discovery efforts. The associations between genetic functions, predict signaling pathway networks, and are associated with cell-specific-types and tissues involved in central nervous system, energy balance, glucose homeostasis, lipid metabolism, and fat liver accumulation. Our findings provide insight into the genetic basis of the correlated architecture of NAFLD metabolic pathway.
919S 
Epigenetic effects of environmental enrichment and EGCG treatment on a mouse model of Down Syndrome. C.N. Hör1-2, S. Jhanwar1-2, S. Espeso1,2, C. L. M. Pons-Esparis1-2, M. de Lagran1,2, M. Friedländer1-2, Dierssen1,2, X. Estivill1,2, S. Ossowski1,2. 1) Centre for Genomic Regulation, Barcelona, Barcelona, Spain; 2) Universitat Pompeu Fabra (UPF), Barcelona, Spain.

Epigenetic marks, defined as modifications to the DNA molecule and higher order structure not affecting the nucleotide sequence, have been hypothesized to be the long sought after interface between genes and the environment. The epigenetic state of a locus, i.e. the local combination of epigenetic marks, is thought to be determined partly by the DNA sequence and influenced by environmental cues. To directly explore the effects of genotype and environment under controlled conditions, we subjected transgenic Dyrk1A overexpressing mice (a model of Down syndrome) and their wild-type littermates to an enriched environment and/or treatment with the green tea polyphenol epigallocatechin gallate (EGGC, a Dyrk1A kinase inhibitor), both of which were previously shown to improve the cognitive performance of the transgenic mice. We have generated genome-wide profiles of cytosine methylation, a selection of histone marks and open chromatin in the cerebral cortices of these mice, as well as transcriptome and proteome. The epigenomic data are combined by computational analysis to yield a map of chromatin states, which can then be correlated with transcript and protein levels in each genotype and treatment group. Our analysis can (1) guide the discovery of regulatory elements active in the mouse brain, (2) uncover environment-specific signatures in the epigenome when comparing the profiles of different exposures across one genotype, and (3) elucidate the differential response attributable to the transgene by comparing these changes between two genotypes. Preliminary analysis of a subset of the mouse data show that DNA methylation patterns remain stable upon treatments, while H3K36me3, a histone modification present along active gene bodies, is differentially detected in a set of genes enriched in annotation terms related to long term potentiation, which is important for memory and learning.

920M 
The Molecular Convergence of non-HLA Ankylosing Spondylitis Risk Genes with Autoimmune Diseases. D. O’Rielly1, M. Udani1, M. Husan2, D. Coler2, P. Rahmat1. 1) Faculty of Medicine, Memorial University, St. John’s, Newfoundland, Canada; 2) Genetics and Genomic Biology, The Hospital for Sick Children, Toronto, Ontario.

Genome-wide association studies (GWAS) and subsequent follow-up studies validated a significant number of non-HLA ankylosing spondylitis (AS) risk genes. Although these genes are independently recognized as small contributory risk factors to the pathogenesis of AS, the combined effect is yet to be quantified. In this study, we hypothesize that AS risk genes converge at the molecular level with related autoimmune disease genes. To identify such convergence a systematic approach is required to further investigate and understand AS pathobiology. A protein-protein interaction network-based method was utilized to infer connectivity between AS genes identified in GWAS studies with other autoimmune risk gene sets. We investigate 418 risk genes from ten (10) autoimmune diseases and their complex interactions with AS. The constructed PPI network demonstrated significantly (~p<1.8 X 10^-4 after 50,000 permutations) dense modular connectivity for AS genes with six other autoimmune disease genes including Crohn’s disease (CD), psoriasis (Ps), ulcerative colitis (UC), celiac disease (CeD), multiple sclerosis (MS), and primary biliary cirrhosis (PBC). The most significant connectivity observed (after correcting for gene number) was with CeD and UC, which is consistent with the known pathobiology of AS, where CeD and UC often manifest as comorbid conditions. To better understand the shared pathway and to identify etiological genes, each non-AS risk gene was ranked according to their contributing connectivity to AS associated genes. We have identified STAT3, JAK2, IL12RB2, LSP1, and NFKBIA genes which are highly clustered with AS associated genes. We have developed a systematic approach to infer causal genes that converge at the molecular level to a common pathway involved in disease pathogenesis. The highly connected non-AS genes identified in this study will be prioritized for further investigation to identify the contribution to AS pathogenesis and to better understand the complex relationship of autoimmune diseases.

921T 
Characterizing the Nphp10 (Sdccag8tn(sb-Tyr)2161B.CA1Cove) mouse model. K. Weinbrecht1-2, V. Sheffield1-2, S. Seo2. 1) Dept. of Pediatrics, Univ. of Iowa, Iowa City, IA; 2) Dept. of Ophthalmology, Univ. of Iowa, Iowa City, IA.

A subset of nephropathies genes show normal localization to the cilary-renal complex, creating a link to disorders described as nephrophthisis-related ciliopathies (NPHP-RC). Aside from early end-stage renal failure and kidney cysts, seen in patients with just NPHP, the two other main phenotypes of NPHP-RCs are retinal and cerebellar degenerations. One such disease, Down Syndrome Renal and Cerebellar Dysplasia (Down Syndrome (BBS), with cardinal phenotypes of retinal degeneration, renal abnormalities, obesity, polydactyl, and learning disabilities. One NPHP gene associated with BBS is serologically defined colon cancer antigen 8 (SDCCAG8), later defined as NPHP10. However, the association with this disease was only recently made and little is known about the molecular function and how it leads to the observed phenotypes. BBS patients with NPHP10 mutations exhibit retinal and renal abnormalities, obesity, and learning disabilities. The purpose of this study is to characterize mouse model Sdccag8tn(sb-Tyr)2161B.CA1Cove, generated by a Sleeping Beauty Transposon (SBT) insertion between exons 12 and 13 of Nphp10. The goal in characterizing this mouse model is to determine its efficacy as a BBS or NPHP disease model, as well as to determine its use for functional analysis of NPHP10. NPHP10 has not been shown to interact with known NPHP complex proteins, nor has it been shown as an interactor with the other current BBS genes. Thus, the pathway through which NPHP10 leads to BBS phenotypes is novel and this mouse can be used to study the severity of typical BBS phenotypes in a BBS model. In order to study NPHP10, we perform our analysis in only a subset of patients of NPHP10 that DNA methylation patterns remain stable upon treatments, while H3K36me3, a histone modification present along active gene bodies, is differentially detected in a set of genes enriched in annotation terms related to long term potentiation, which is important for memory and learning.

922S 

Purpose. Left ventricular hypertrophy (LVH) is a potent risk factor for cardiovascular disease. The development of LVH is accompanied by significant changes in gene expression patterns in cardiomyocytes (CMs). However, traditional methods of transcriptome analysis do not analyze the abundance of individual splice forms (SFs), leaving the potential role of differentially expressed isoforms in complex diseases, including LVH, unexplained.

Taking advantage of RNA sequencing, we identify a set of genes that are differentially expressed in hypertrophied CMs and explore the contribution of individual splice-forms to the LVH-transcriptome profile. Methods. We developed a model of human CM hypertrophy using induced pluripotent stem cell derived CMs (iPSC-CMs). Stimulating the iPSC-CMs with isoproterenol induced characteristic changes associated with the development of LVH, evidenced by an increase in cell surface area and two well-established LVH markers, c-fos and BNP. RNA-Seq was performed for three replicates. Transcripts were assembled using TopHat2 and Bowtie2, and analyzed for differentially expressed levels of expression using cufflinks2. Results. Upon hypertrophic induction, 920 genes exhibited significantly differential expression. 433 of these genes have one SF and 487 have two or more (55 genes had ≥ 5 SFs). Of those that are alternatively spliced, 367 (75.4%) showed significant changes at the global level only, 15 (3.1%) at the splice-form level only, and 45 (27.1%) at both. To explore potential functional implications of differentially expressed SFs, we compared the exons that are unique to only some SFs of each gene to those that are common (common exon) to all SFs. On average, unique exons are approximately twice as long as common exons. Further, unique exons are associated with numerous well-characterized functional domains. Functional characterization of differentially expressed isoforms will likely provide novel insights into the etiology of LVH. Conclusions. Our study demonstrates that differential splice form expression is present in the development of LVH and is an important component of the physiology of hypertrophied CMs. The same is likely true for many other complex diseases. Reliance on analysis of global expression patterns alone obscures changes associated with individual splice forms, resulting in a substantial loss of information and potential insights into the developmental and physiological complexity of diseases.
Obesity arises from a complex interaction of genetic predisposition and environment that can accumulate throughout life. Previous work has shown that there are maternal effects on obesity outcomes in humans and nonhuman primates, but few have investigated the impact of maternal effects on obesogenic growth through the lifespan. This research investigates the impact of maternal effects on the genetic underpinnings of obesogenic growth from birth to adulthood in a genetically well-characterized model system under a controlled diet and environment: the African green monkey (Chlorocebus aethiops sabaueus) in the Vervet Research Colony at Wake Forest University School of Medicine. We used growth curve analysis on measures taken thrice yearly on body size and composition - such as body weight (BW) and BMI - in a population of 641 monkeys measured from 2000 through 2013. Of these, 33 individuals, 6 males and 27 females, presented with signs of chronic abdominal obesity defined as having an adult waist circumference above 40.5 cm for at least three successive measurements. Individuals measured < 6 times were excluded from analysis (max measures = 30, mean = 12.6 ± 6.12). Growth was modeled using three-parameter logistic growth curves in nonlinear mixed models, with parameters modeled as fixed effects and subject and sex/obesity status modeled as random effects. We assessed heritability of individual growth parameters using maximum likelihood variance components analysis in SOLAR, with the variance attributed to maternal ID (c2) partitioned from environmental variance to determine maternal effects. We found significant heritability and maternal effects on all static measures of adult body condition (e.g., BW, h2 = 0.86, p = 6.52×10^{-10} and c2 = 0.11, p = 0.04; BMI, h2 = 0.77, p = 1.39×10^{-05} and c2 = 0.12, p = 0.05), and significant maternal effects on parameters of growth (e.g., BW: b2, asymptote of growth, h2 = 0.77, p = 4.93×10^{-21} and c2 = 0.05, p = 0.11). This study suggests that although adult obesity is a developmental process driven in part by heritable obesogenic trajectories resulting in faster and longer growth, those trajectories do not appear to be influenced by maternal effects. A better understanding of how growth can be decoupled from maternal effects on adult obesity outcomes will be necessary to assess early obesity risks and inform potential biomedical interventions.

Posters: Complex Traits and Polygenic Disorders

924T The Genetic Landscape of Hematopoietic Stem Cells. H. Allayee1, A.L. Crow1, J. Hartilaia1, T.J. Spindler1, A. Ghazalpour1, L.W. Barsky1, B.J. Bennett1, B.W. Parks2, E. Eskin1, R. Jain1, J.A. Epstein2, A.J. Lusis3, G.B. Adams4. 1) Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90033; 2) Broad Center for Regenerative Medicine and Stem Cell Research at USC, University of Southern California, Los Angeles, CA 90033; 3) Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095; 4) Department of Genetics and Nutrition Research Institute, University of North Carolina, Chapel Hill, Kannapolis, NC 28081; 5) Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104.

The ability to carry out association studies on a genome-wide basis has revolutionized the field of human genetics. Such GWAS have identified hundreds of novel genes for atherosclerosis, diabetes, cancer, and related quantitative phenotypes. However, human GWAS can be limited since other biomedically relevant phenotypes, such as hematopoietic stem cell (HSC) number/function, are logistically difficult to obtain in large numbers of subjects. To overcome these limitations, we used a GWAS approach with a recently developed panel of inbred and recombinant mouse strains, termed the Hybrid Mouse Diversity Panel (HMDP), to characterize the genetic landscape of HSCs. We used flow cytometry to quantify the frequency of long-term (LT) and short-term (ST) HSCs, as well as hematopoietic progenitor cells (HPCs), in bone marrow (BM) from 111 HMDP strains. These cell populations varied 25 to 80-fold, thus providing strong evidence for a genetic basis to HSC variation. Consistent with this notion, a GWAS analysis identified 8 loci significantly associated with HSCs (p=6.9E-6 - 7.0E-14), including a locus for ST-HSCs on chromosome 5 (p=6.2E-7). Of the genes in this region, highly significant cis eQTLs was identified for homeodomain only protein (Hoxp) in liver (p=3.4E-18) and heart (p=6.2E-19), suggesting that genetic variation at this locus is functional with respect to Hoxp expression. Hoxp expression in BM from 25 HMDP strains was significantly positively correlated (r=0.44; p<0.05) with ST-HSCs. To functionally validate Hoxp, we characterized BMI from Hoxp knockout (Hoxp−/−) mice by flow cytometry and observed a significant 4-fold reduction in ST-HSC frequency compared to wild-type littermates (p=0.01). Additional analyses revealed that Hoxp−/− mice had significantly reduced numbers of quiescent ST-HSCs (60% vs 85%; p<0.05) and increased numbers of cells in the G1 phase (35% vs 10%; p<0.05). Hoxp deficiency specifically affected ST-HSCs since there were no differences in the frequency or cell cycle status of LT-HSCs. These findings are entirely consistent with Hoxp mRNA levels being positively correlated with ST-HSCs and the Hoxp locus only being associated with this HSC subset in the GWAS. Taken together, these results provide compelling evidence that Hoxp is a novel positive regulator of HSC physiology in mice and demonstrate the power of the HMDP for elucidating the genetic architecture of complex traits that would otherwise not be feasible in humans.
925S Genotyping-by-sequencing in outbred CFW mice yields a powerful approach for genome-wide mapping of complex trait loci. P. Carbonatto1, S. Gopakumarshan1, C.C. Parker2, N.M. Gonzalez1, A. Lionakis1, C.L. Ackert-Bicknell3, E.H. Leung4, E. Arrey1, J. Park5, J. Davis6, A.A. Palmer7. 1) Dept. of Human Genetics, University of Chicago, Chicago, IL; 2) Dept. of Biological Sciences, Rice University, Houston, TX; 3) School of Medical Sciences, College of Life Sciences and Medicine, University of Aberdeen, Foresterhill, Aberdeen, UK; 4) The Jackson Laboratory, Bar Harbor, Maine; 5) Dept. of Genetics, Stanford School of Medicine, Stanford, CA.

Despite the successes of genome-wide association studies for augmenting our understanding of many common human diseases and disease-related traits, mice still remain an important complementary resource for studying genetics of complex disease. Advantages include being able to control the environment, and measure gene expression in tissues that are ordinarily inaccessible in humans. (We present gene expression from multiple brain regions in a separate poster.) One important disadvantage of working with traditional mouse lab strains compared to humans is that the mice exhibit long-range correlations on their chromosomes, due to a lack of accumulated recombination, often making it very difficult to assign any functional interpretation to a complex trait locus. We describe our efforts to develop a new resource for mapping complex trait loci from CFW mice, a commercial stock of outbred mice, that permits posterior solution to map complex traits. We show that CFW mice have desirable properties for QTL mapping, including rapid decay of linkage disequilibrium and a high proportion of common alleles. To ascertain whole-genome variation at low cost, we adapt “genotyping-by-sequencing” (GBS) protocols originally developed for plants, and contribute additional technical improvements to this protocol for outbred mice. To address the problem of false positives caused by “cryptic” or “hidden” relatedness, we test for association using the linear mixed model approach, implemented in the software GEMMA. In a cohort of ~1000 CFW mice, we obtain strong support for genetic associations in a variety of complex traits, including bone-mineral density, testes weight, musculoskeletal traits, prepubertal inhibition and methamphetamine sensitivity. In some QTL regions, the association signal isolates only a few genes. We also identify loci that map to known complex trait genes, such as testes weight gene Inha and osteoporosis gene Col1a1. Many of these associations have been replicated in an independent panel of CFW mice developed by our collaborators at University of Oxford. Our findings show that the CFW mouse stock is a practical and powerful resource for identifying loci underlying a variety of complex traits.
929M Nasopharyngeal microbiome composition is associated with lung function in adult Hutterites. C. Igartua1, E.R. Davenport2, J.B. Becker2, S.R. White2, Y. Gilad1, C. Ober1, J.M. Pinto1. 1) Human Genetics, University of Chicago, Chicago, IL; 2) Section of Pulmonary and Critical Care Medicine, Department of Medicine, University of Chicago, Chicago, IL.

The human microbiome contains trillions of microbes and has been associated with disorders of the airway, such as sinusitis and asthma. We predict that host immunity is related to the composition of bacteria in the airways and that ecologic characteristics of these bacterial communities may be predictive of respiratory disease state. We explore this possibility by considering the relationship between the nasopharyngeal microbiome (an accessible, representative measure of airway bacteria) in 142 adult Hutterites, including 24 asthmatics, and FEV1/FVC ratio as a measure of lung obstruction, which is significantly associated with asthma in this population. The Hutterites are a founder population of European decent who live communally, thereby ensuring that environmental exposures are relatively similar among individuals. We sequenced the V4 region of the 16S rRNA gene from bacterial cells collected from the nasopharynx and calculated the relative abundance of each bacterium after bioinformatic classification of QC filtered reads using Mothur. After subsampling a maximum of 550,000 reads from each individual (range 273,064-550,000; median 550,000), standardized bacterial reads present in at least 75% of individuals were fit to a standard normal distribution. At the taxonomic level of order, after regressing out age, sex and date of collection, the first principle component, which accounted for 44.4% of the variance between individuals, was associated with the FEV1/FVC ratio (β=0.008). After removing correlated bacteria (Spearman correlation >0.8), we performed association tests for the relative abundance of the 60 bacteria in the taxonomic level of order using a linear mixed model. For 60 tests (uncorrected p=5.3×10^-1), our study demonstrates that the nasopharynx is an accessible airway site in which to study microbiome-lung disease associations and suggests that microbes in this anatomic location may reflect environmental exposures and/or host mucosal immunity status related to important disease-associated phenotypes.

930T Hematopoietic stem cells target neovascular tissue in a novel preclinical model of proliferative diabetic retinopathy. K. Wert1, V. Mahajan2,3, Y. Yan4, Y. Li4,5, J. Tosi4, T. Nagasaki4,5, T. Nagasaki4,5, M. Grant5,6, S. Teaney5,6,1. 1) Jaensch Laboratory, Whitehead Institute for Biomedical Research, Boston, MA; 2) Department of Ophthalmology & Visual Sciences, University of Iowa, Iowa City, IA; 3) Omics Laboratory, University of Iowa, Iowa City, IA; 4) Department of Pharmacology and Therapeutics, University of Florida, Gainesville, FL; 5) Bernard & Shirlee Brown Glaucoma Laboratory, Columbia University, New York, NY; 6) Edward S. Harkness Eye Institute, Columbia University, New York, NY; 7) Eugene and Marilyn Glick Eye Institute, Department of Ophthalmology, Indiana University School of Medicine, Indianapolis, IN; 8) Department of Pathology & Cellular Biology, Columbia University, New York, NY.

Currently, more than 382 million people have been diagnosed with diabetes, and this number is expected to rise to 592 million people by 2035. Approximately one third of patients diagnosed with diabetes display signs of diabetic retinopathy (DR). Approximately 60% of non-proliferative DR patients progress to the neovascular stage, proliferative DR (PDR). PDR patients will develop vitreous hemorrhage, tractional retinal detachment, neovascular glaucoma and blindness. Most currently available mouse retinopathy models are inadequate for the study of PDR. They only display mild, non-proliferative disease within the lifespan of the animals. An animal model for PDR is critical to study signaling pathways involved in this disease stage. In this study, we examined protein levels from vitreous samples of human patients with PDR. We discovered an increase of hypoxia inducible factor 1 alpha (HIF1a) protein in untreated human PDR patients compared to control groups. During hypoxic conditions, von Hippel Lindau tumor suppressor protein (VHL) releases its normal degradative action on HIF1a. We tested whether the human PDR condition could be modeled by using a knockout of VHL, which would constitutively activate HIF1a and release its inhibition of HIF1a transcriptional activity. Therefore, we created a novel, tissue-specific preclinical model in which VHL was knocked out within the neural retina using a Chx10 cre promoter. This promoter allowed for complete embryonic retinal loss of Vhi gene expression. We found that the mice developed severe proliferative retinopathy that correlated with human PDR patients. Since hematopoietic stem cells (HSCs) are found to localize to sites of angiogenesis, we hypothesized that HSCs delivered into the intravitreal space of the eye might localize to the sites of neovascularization, where angiogenesis is occurring and abnormal and leaking blood vessels are formed. Since the PDR mouse model was treated using transplantation of HSCs into the intravitreal space of the eye. We found that HSCs localized to the sites of damaged retinal vasculature after intravitreal injection into our novel preclinical PDR model. In summary, we developed a new mouse model for the study of human PDR and other ischemic retinal diseases by constitutive down regulation of HIF1 regulator, VHL. Rapid progression of retinopathy in these mutants will expedite the evaluation of therapeutic agents for this group of blinding disorders, including treatment using HSCs.
931S

The Role of Copy Number Variants in Latino Children with Asthma. M.L. Spear1, 2, 3, M. Pino-Yanes2, C.R. Gignoux2, C. Eng2, B.P. Coe4, C.D. Campbell4, E.E. Eichler4, 2, D.G. Torgerson5, E.G. Burchard2, 3 on behalf of the GALA I Investigators. 1) Biomedical Sciences Graduate Program, University of California, San Francisco, San Francisco, CA; 2) Department of Medicine, University of California, San Francisco, San Francisco, CA; 3) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA; 4) Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA; 5) Howard Hughes Medical Institute.

Asthma is a complex respiratory disease influenced by social, environmental, and genetic factors. Twin studies have found the heritability to be between 75% and 92%, suggesting an important genetic contribution to asthma susceptibility. Although variation in over 150 genes has been associated with asthma, they only explain a small fraction of the disease heritability. Copy number variants (CNVs) are an important source of variation in the human genome due to their large size. CNVs have been previously associated with immune-related diseases, however knowledge of their role in asthma is limited. In this study we performed a genome-wide assessment of common CNVs in Latino children from the Genetics of Asthma in Latino Americans (GALA I) study, to identify novel CNVs associated with asthma. CNVs were called from genome-wide SNP genotypes obtained from the Affymetrix 6.0 GeneChip Array (Affymetrix, Santa Clara, Calif) using the Affymetrix Genotyping Console 4.1 software. A total of 733 Puerto Ricans and Mexicans, including 445 cases and 288 controls were included in the study. We limited our analysis to common CNVs present in more than 5% of our case/control population, and greater than 30 kb in length. We performed logistic regression adjusting by genetic ancestry and self-reported ethnicity, to test for an association between asthma and copy number of common CNVs. We identified a significant association between asthma and CNVs located within a genomic region previously associated with atopic dermatitis (OR = 0.85 [0.79-0.92]; p-value = 1.84 x 10^-6). The associated CNV is 300 kb upstream of COL4A3. Overall, our findings suggest an important role of a common CNV at 3q22.1 in asthma susceptibility in Latino children.

932M

Family-based Associations and Parent of Origin analyses reveal novel associations with inflammatory bowel disease (IBD). X. Yan1, M. Dubinsky2, 3, T. Haritunians2, 3, J. Rotter2, K. Taylor2, S. Rich2, S. Chang-Gut-Gumusc2, W. Chen3, X. Guo3, S. Targen3, D. Li3, M. DCO-GEN-ERN1, 1) F. Widjaja Foundation Inflammatory Bowel & Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA; 2) Pediatric IBD Center, Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, CA; 3) Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center; 4) Center for Public Health Genomics, Dept Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA.

Background: Only a minority of the heritability for IBD has been explained despite the identification of >160 susceptibility loci. We aimed to identify additional IBD associations by Transmission Disequilibrium Test (TDT) and to assess for any Parent-of-Origin effect to explain, in part, the missing heritability. Methods: 965 IBD trios with an affected offspring and both parents genotyped by ImmunoChip were included. We tested for association using TDT and Parent-of-Origin test (PLINK). Novel signals were confirmed by checking for LD and conditional analyses on the known 193 IBD-associated SNPs. An independent confirmation cohort of 2904 cases and 4791 controls was used to validate the TDT associations. 10^6 permutation and meta-analysis were performed.

Results: SNPs with TDT association after conditioning on known IBD SNPs (p<0.01 for SNPs within known region or p<0.0001 for SNPs outside known region, and p<5.0e-4 for conditional analyses) and confirmed in the case/control replication cohort are listed as below: IL1RL1(SNP=rs1420101, p_discovery=7.6e-3,p_confirmation=4.5e-5,p_meta=9.2e-5); SLC2A13/LRRK2(SNP=rs75904798,p_discovery=1.0e-3,p_confirmation=2.5e-7,p_meta=1.2e-4); SBNO2(SNP=rs8099951,p_discovery=6.4e-3,p_confirmation=3.4e-4,p_meta=1.4e-4); LOC441108(SNP=rs2158101,p_discovery=4.3e-4,p_confirmation=4.8e-5,p_meta=1.2e-5). Parent-of-Origin analyses identified differential transmission at the following genes, which mapped to a known region, and p<5.0e-4 for conditional analyses) and confirmed in the case/control replication cohort are listed as below: CTSH(SNP=rs78764205, p_discovery=1.43e-5,p_confirmation=4.8e-5,p_meta=1.2e-5); SBNO2(SNP=rs8099951,p_discovery=6.4e-3,p_confirmation=3.4e-4,p_meta=1.4e-4); LOC441108(SNP=rs2158101,p_discovery=4.3e-4,p_confirmation=4.8e-5,p_meta=1.2e-5).

Discussion: TDT analyses have identified additional IBD associations. The Parent-of-Origin analyses identified new IBD signals and if validated may explain some of the missing heritability in IBD.
933T
Parent-of-origin effects of the APOB gene on adiposity in young adults.
1) Braun School of Public Health, Hebrew University-Hadassah, Jerusalem, Israel; 2) Département de Mathématiques, Université de Sherbrooke and Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC, Canada; 3) Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA, USA; 4) Department of Medicine, University of Washington, Seattle, WA, USA; 5) Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA; 6) Department of Statistics, University of Auckland, Auckland, New Zealand; 7) Department of Biostatistics, University of Washington, Seattle, WA, USA; 8) Department of Epidemiology, University of Washington, Seattle, WA, USA; 9) Harvard Medical School and General Medicine Division, Massachusetts General Hospital, Boston, MA, USA; 10) Institute of Human Genetics, University of California, San Francisco, CA, USA; 11) Cardiovascular Research Institute, University of California, San Francisco, CA, USA; 12) Department of Dermatology, University of California, San Francisco, CA, USA; 13) Department of Population Medicine, Harvard Pilgrim Health Care, Boston, MA, USA; 14) Department of Genetics Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO, USA; 15) Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, University Medical Center, Rotterdam, the Netherlands; 16) Institute for Translational Genomics, Howard Hughes Medical Institute, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; 17) Department of Genome Sciences, University of Washington, Seattle, WA, USA; 18) Department of Human Genetics, Hebrew University-Hadassah, Jerusalem, Israel; 19) Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.

Background: Genome-wide association studies (GWAS) of unrelated individuals have identified multiple loci associated with cardiometabolic phenotypes. However, the majority of data type-phenotype pairs is measured in young adults. We propose a method to estimate and locate the loci that account for only a small proportion of the traits' heritability. To date, most association studies have not considered parent-of-origin effects (POEs). Objectives: We sought POEs on adiposity and glycemic traits in young adults. Methods: The Jerusalem Perinatal Study (JPS) sample comprises 1250 young adults (mean age 32) born in Jerusalem, and their mothers, with recently-collected DNA samples and cardiometabolic measurements. We focused on a set of 18 genes (182 tag SNPs) identified by previous GWAS as associated with selected cardiometabolic traits. Using linear regression, we examined the associations of maternally- and paternally-derived offspring minor alleles with body mass index (BMI), waist circumference (WC), fasting glucose and insulin. To replicate and meta-analyze findings, we performed similar analyses in individuals of European ancestry aged≥50 belonging to extended pedigrees from three additional studies: Framingham Heart Study, Family Heart Study and Erasmus Rucphen Family study (Total maximum N=4800). We considered p<2.7×10^-4 as statistically significant to account for multiple testing. Replicated findings were followed-up by also assessing POEs on lipids and blood pressure (BP). Results: A common coding variant in APOB (rs1367117) demonstrated a significantly heritable effect on BMI (β=0.8, 95% CI: 0.4, 1.2, p=0.0003), and WC (β=0.2, 95% CI: 0.1, 0.3, p=0.01), and BMI (β=0.7, 95% CI: 0.5, 0.9, p=7.4×10^-4). There was little evidence to suggest POEs on glucose, insulin and systolic BP were attenuated after further adjustment for BMI.

Conclusions: Our results demonstrate that a common genetic variant in APOB is associated with adiposity through maternal POE, an association that was not previously detected in GWAS of adiposity. These results provide support for a role of POEs in adiposity and related cardiometabolic traits, and motivate further research in this area.

934S
1) Albert Einstein College of Medicine, Bronx, NY; 2) Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 3) Harvard School of Public Health, Boston, MA; 4) Boston University School of Medicine, Boston, MA.

Habitual dietary intake is a complex behavior that has been suggested to have a genetic component, with estimated heritability ranging from 8% to 70% in family and twin studies. However, genetic determinants of dietary intake are poorly understood and it is unknown how much heritability for dietary intake can be captured by common SNPs in genome. Using a linear mixed model to fit all SNPs simultaneously by GCTA software, we estimated narrow sense heritability for dietary intake explained by common SNPs (~300,000 to 700,000 SNPs in the genome) among 37,421 unrelated individuals of European ancestry from three US cohorts: the Nurses’ Health Study, the Health Professionals Follow-up Study, and the Women’s Genome Health Study. Dietary intakes of total energy, macronutrients (percentage of energy intake), and individual foods and beverages were assessed by almost identical validated semi-quantitative food frequency questionnaires across three studies, and two major dietary patterns (prudent pattern and Western pattern) were derived using factor analysis. In our combined results from three studies by meta-analysis, small but significant proportions of variance for intakes of total energy (5.8 [95% CI 2.9, 8.7%], P=0.0001), fat (5.9 [3.0, 8.8%], P=0.0001), protein (4.8 [1.9, 7.7%], P=0.001) and carbohydrate (3.4 [0.5, 6.3%], P=0.02) were explained by all common SNPs. We also found that a number of fats and oils, dairy products, vegetables, whole grains, refined grains, condiments, coffee, alcohol, fruit juice, sugar-sweetened beverages, diet beverages, and water had genetic influences, with significant heritability estimates ranging from 3.1 [P=0.02] to 9.3 [3.6, 14.7%], P<5.3×10^-10. In addition, a Western dietary pattern (characterized by higher intakes of red and processed meats, sweets, desserts, French fries, and refined grains) showed the highest heritability estimate (10.3 [7.2, 13.4%], P<4.1×10^-10) among dietary intake traits. This demonstrates for the first time how genetic factors contribute to complex traits. Genome-wide association studies are warranted to identify genetic variants associated with dietary intake.

935M
The role of STRs in shaping complex traits. T. Willems1,2, A. Gordon1, M. Gymrek1, Y. Eritch1.
1) The Whitehead Institute, Cambridge, MA; 2) Computational and Systems Biology, MIT, Cambridge, MA.

Despite the advent of whole-genome sequencing data and increasingly large sample sizes, genome-wide association studies have failed to fully uncover the genetic basis for the heritability of most complex traits. While various studies have proposed that this “missing heritability” may stem from rare SNPs or epistatic interactions, other variant classes may also be responsible. Recently, a handful of single-gene and genome-wide studies have identified short tandem repeats (STRs), sequences of DNA consisting of repeating patterns of 2-5 base pairs, as modulators of gene expression. To assess whether this class of variant may also influence complex traits, we genotyped ~700K STRs in over 1500 individuals using data from the UK10K Project. In conjunction with a wide range of phenotypes, these genotypes were utilized to perform a genome-wide STR GWAS. The resulting analysis identified a handful of STRs strongly associated with a variety of phenotypes, even after controlling for tagging SNPs. In addition, we utilized linear mixed models to jointly assess the phenotypic variance explained by SNPs and STRs. In aggregate, our results shed new light on the issue of “missing heritability” and suggest that STR variations substantially contribute to complex traits.
936T

Regulatory variants explain much more heritability than coding variants across 11 common diseases. A. Gusev1, S.H. Lee2, B.M. Neale3,4,5, G. Tynkä6,7,8,9, B.J. Vihma10, S. Finucane11, H. Xu12, C. Zang13, S. Ripke14,2, E. Stahl15, SCZ-PGC16, SWE-SCZ17, A.K. Kahler12, C.M. Hultman12, S.M. Purcell18,11,12, S.A. McCarroll11, M. Daly11,14, B. Pasanivic11, P.F. Sullivan11, N.R. Wray13, S. Raychaudhuri11,14,15, A.L. Price1,12, 1) Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA; 2) Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; 3) Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA; 4) Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; 5) Division of Genetics, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, USA; 6) Division of Rheumatology, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, USA; 7) Partners Center for Personalized Genetic Medicine, Boston, Massachusetts, USA; 8) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; 9) Department of Mathematics, Massachusetts Institute of Technology, Massachusetts, USA; 10) Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts, USA; 11) The Department of Psychiatry at Mount Sinai School of Medicine, New York, New York, USA; 12) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 13) Department of Pathology and Laboratory Medicine, Geffen School of Medicine at UCLA, Los Angeles, California, USA; 14) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; 15) Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK. 1) Free haplotype variation has not been estimated on the haploid scale. Efforts at increasing coverage through imputation have yielded only small increases in the heritability explained, possibly due to the relatively poor imputation of low-frequency causal variants. We propose to estimate the heritability explained by a set of haplotypes constructed directly from the study sample ($g_{haplo}$). Our method constructs a set of haplotypes by extending shared segments subject to a 4-gamete test using data phased with the HAPI-UR software package (Williams et al. 2012). The haplotype values (0, 1 or 2 copies per individual) are then used to estimate genetic relationships between individuals. These relationships are used as the covariance structure in a linear mixed model, which is used to estimate the heritability explained by this set of haplotypes. Our simulations show that this approach can yield substantial gains in heritability explained relative to genotyped SNPs ($h^2_{haplo}$). We applied our method to WTCCC2 data consisting of 9K multiple sclerosis cases and 5K controls, genotyped at 452K SNPs. We built 37M haploSNPs with MAF > 1%. The heritability explained by this set of haploSNPs was $h^2_{haplo} = 0.39$ (s.e. 0.02), substantially larger than $h^2_g = 0.26$ in the same data. We also applied our method to PGC2 data consisting of 10K schizophrenia cases and 10K controls of European ancestry, all of which were genotyped at the same set of 348K SNPs. We built 28M haploSNPs with MAF > 1%. The heritability explained by this set of haploSNPs was $h^2_{haplo} = 0.38$ (s.e. 0.02), significantly larger than $h^2_g = 0.33$ in the same data. Overall, our results suggest that haplotypes can explain substantially more heritability than genotyped or imputed SNPs. Additionally, haplotype-based approaches can shed light on the genetic architecture of complex traits and informing strategies for disease mapping.
938M
Heritability Estimates and Genetic Association for 60+ Complex Traits in a Young Healthy Sibling Cohort. Q. Ma1, A.B. Ozel1, D. Siemieniak2, K.C. Deitch3, D. Ginsburg3,5, J.Z. Li1. 1) Departments of Human Genetics, University of Michigan, Ann Arbor, MI 48109; 2) Howard Hughes Medical Institute, Ann Arbor, MI 48109; 3) Pediatrics and Communicable Disease, University of Michigan, Ann Arbor, MI 48109; 4) Internal Medicine, University of Michigan, Ann Arbor, MI 48109.

As genotyping becomes more efficient, sample recruitment and phenotyping remain a major limiting factor. In a GWAS of bleeding and blood clotting traits we sought to increase the utility of the cohort by collecting > 60 self-reported complex traits through web-based questionnaires. The cohort of 1,191 healthy young subjects consists of 509 sibships, 80% Europeans, and age of 14-35 yrs. The traits include 16 quantitative traits (e.g., weight, height, age of menarche, hematological measures RBC, HCT, MCV, MCH, MCHC, RDW, WBC, HGB, PLT, MPV), 21 ordinal traits (e.g., Smoking, BleedingTendency, SkinTags, Acne, TanningTendency, SkinColor, Freckles, DentalCaries, VisionCorrection, EatingSweets, EatingSaltyFood, Athleticability, Aphthousulcers), and 27 nominal traits (e.g., Immunization, ToothExtraction, EyeColor, HairColor, Hairline, EarLobeCreased, EarLobe-Attachment, Dimples, Dyslexia, Migraines, Stuttering, Allergies, Flatfeet, Handedness, PhoticSneleness, BrainFreeze, InterlockingFingers, etc.). We used the known relatedness to estimate heritability using Merlin-regress and found that >1/2 of the traits have heritability >40%. Since the samples have been genotyped over ~800K SNPs in the original GWAS we used SNP data to calculate the actual genetic relatedness, and estimated the variance explained by all the genotyped SNPs using GCTA. With all subjects, pedigree-based estimates were similar to SNP-based estimates; but the latter were more precise. We found that one trait from each of the eight domains outside common SNPs associated with common disease through Genome Wide Association Studies (GWAS) often lie in gene enhancer elements located distal to protein-coding genes. The enhancer-SNPs likely confer disease risk by impacting gene expression, but quantifying their effects on target gene expression has proven challenging. Here, through integrative analyses of haplotype block maps, epigenomic landscapes, and GWAS data from six different autoimmune disorders, we show that 94% of instances in which a GWAS SNP lies in an enhancer predicted to regulate a given gene, there is at least one additional SNP that lies in a different enhancer predicted to target the same gene and is located distal to the LD block from which the GWAS-association arose. Using RNA-seq data matched with genotypy data from 400 individuals, we show that the distal enhancer variants act additively and epistatically with enhancer-SNPs within the GWAS-associated LD block to affect expression of the target gene. Thus for these loci, the transcriptional effect is dictated by SNPs located within and outside of the GWAS-association signal. Additionally, many of the outside variants disrupt transcription factor binding sites and affect chromatin accessibility. These results have important implications for GWAS and eQTL studies as we demonstrate that for some loci, the transcriptional effect is only apparent when the additive and epistatic effects of enhancer variants are considered. Given that outside variants can boost the transcriptional effect but are often not considered, we propose that the effects of GWAS SNPs are often underestimated, which could explain some of the missing heritability for these disorders. We propose utilizing the chromatin structure to identify epistatic interactions between GWAS SNPs, to interpret the effect of common disease alleles on gene expression, and to help refine the relative risk of enhancer variants on disease susceptibility.

940S
The genetics of exceptional human longevity: new clues from big data on disease. K. Fortney1, P. Sebastiani2, T.T. Pets1, S.K. Kim1. 1) Developmental Biology, Stanford University, Stanford, CA; 2) Biostatistics, Boston University School of Public Health, Boston, MA; 3) Medicine, Boston University School of Medicine, Boston, MA.

In order to begin to decipher the genetic basis of human longevity, we are taking a big data approach to scour large genome-wide association (GWA) datasets on age-related disease to narrow the search for new SNPs associated with longevity. Rather than search for longevity loci directly in centenarian GWAS, we can leverage data from many age-related disease GWAS (often with > 10,000 cases each) by assuming that individuals with extreme longevity are likely to be depleted of polymorphisms associated with disease, such as Alzheimer disease, type 2 diabetes or cardiovascular disease. So far, we have acquired data from eight large disease GWAS meta-analyses: Cardiovascular disease, diabetes, late-onset Alzheimer disease, age-related macular degeneration, chronic kidney disease, lung cancer, pancreatic cancer, and rheumatoid arthritis. In total, this preliminary dataset contains > 250,000 SNPs for > 150,000 cases of 8 diseases. We combined these data with the results of a GWAS on exceptional longevity that included 801 subjects enrolled in the New England Centenarian Study (Sebastiani et al. 2012). We applied a new statistical method to identify variants that are enriched in people with disease but depleted in centenarians. In this approach, a strong association signal in disease can boost a weak association for longevity to statistical significance. Our approach is potentially a powerful new method to identify longevity loci. We will validate all the candidate SNPs that we identify in independent cohorts of centenarians, to test whether they are robustly associated with human longevity.

941M
A meta-analysis of genome-wide association scans for nevus count reveals PPARG1B as affecting both moleliness and melanoma risk. N.G. Martin1, D.L. Duffy1, G. Zhu2, G.W. Montgomery3, N.K. Hayward4, M. Falchi4, P. Hysi3, V. Bataille2, T.D. Spector2, G.C. Smith5, D.M. Evans3, H. Nan6, J. Han6, S. Chanock4, D. Hunter4, L. Jacobs5, T.E. Nielsen2, F. Liu4, M.H. Kayser4. 1) Gen Epidemiology, Queensland Inst Med Res, Brisbane, QLD, Queensland, Australia; 2) Dept of Twin Research & Genetic Epidemiology, St Thomas Hospital Campus, Kings College, London, SE1 7EH, UK; 3) Department of Social Medicine, University of Bristol, Bristol BS82PS, UK; 4) Channing Laboratory, Harvard Medical School, Boston, MA 02115, USA; 5) Erasmus MC, University Medical Centre, Rotterdam, The Netherlands.

The main clinical interest in total number of acquired melanocytic nevus on the skin is its strong correlation with melanoma risk. We have previously described genome-wide association scans (GWAS) in Australian and British samples that identified loci influencing both nevus count and melanoma. Here we report a meta-analysis of five nevus GWAS from Australia, the Netherlands, United Kingdom (two studies), and the United States, comprising a total of 23,371 phenotyped individuals. We confirm known loci including MTAP, PLAG2, IRF4 and ASIP and refine the location of the causal variants. Several new regions exhibit suggestive evidence (P ~ 10^-10). One notable SNP, rs251464, is in the PPARG1B gene, previously shown to modify tanning ability via a regulatory effect on MITF, the master melanocyte regulatory gene [Shog et al Mol Cell. 2013, 49:145-57]. Combining the results for this SNP (P=8x10^-14) with results from a melanoma case-control analysis (P=4.6x10^-10) reaches a genome-wide significant level and direction of effect was consistent in both phenotypes, with the C allele decreasing both melanoma risk and mole count. Another SNP in PPARG1B, rs32579, previously flagged as affecting tanning, was shown to not affect nevus count (P=0.19). We now plan to meta-analyse these results with those from a parallel GWAS meta-analysis of melanoma to achieve even greater resolving power for variants influencing early stage oncogenesis.
942T Identification of 4 novel susceptibility loci for intracranial aneurysms in Portuguese using a pooling-based GWAS. P.C.S. Abrantes1,2, M.M. Santos1, J.M. Xavier1,2, I. Sousa1,2, V. Francisco1,2, T. Krug1,2, J. Sobral1,2, M. Matos1,2, A. Jacinto1,2, D. Colteiro3, S.A. Oliveira1,2. 1) Instituto Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisboa, Portugal; 2) Instituto Gulbenkian de Ciência, 2780-156 Oeiras, Portugal; 3) Department of Neurosurgery, Hospital de Santa Maria, Lisboa, Portugal; 4) Centro de Estudos de Doenças Crônicas (CEDOC), Faculdade de Ciências Médicas, Universidade Nova de Lisboa, 1169-056 Lisboa, Portugal.

Subarachnoid hemorrhage is a life-threatening event that most frequently leads to severe disability and death. Its most frequent cause is the rupture of an intracranial aneurysm (IA). Several studies have consistently demonstrated a genetic component to the risk for IA but until now, its exact etiology remains uncertain.

To identify new susceptibility loci for IA, we performed a genome-wide association study (GWAS) using DNA pooling approach on a Portuguese dataset. Pools of 100 IA cases and 92 age- and gender-matched controls were alleletyped in triplicate using the Affymetrix Human SNP Array 6.0. Top SNPs with IRAS diffI >13% were selected for individual genotyping in the same GWAS dataset. Ninety-nine of the 101 SNPs successfully genotyped were technically validated (P<0.05). Replication of validated SNPs was conducted in an independent Portuguese dataset of 100 IA cases and 407 controls.

We identified 4 variants (rs4667622, rs6599001, rs3932338 and rs10943471) associated with IA in both the discovery and replication datasets (individual and in combination). SNPs rs4667622 (Pcombined=4.0E-05, OR95% CI=1.75[1.32-2.33]) and rs6599001 (Pcombined=4.0E-05, OR95% CI=2.00[1.39-2.88]) were the most significant. rs4667622 is located on chromosome 2q31.1 within the regulatory region of the myosin IIIb gene (MYO3B) and rs6599001 maps to chromosome 3p22.2 upstream of the WDR48. The third most significant marker associated with IA was rs10943471 (Pcombined=3.20E-04, OR95% CI=1.81[1.31-2.51]) which is located on chromosome 6q14.1 in an intergenic region upstream of HTR1B (5-hydroxytryptamine (serotonin) receptor 1B). The top peak in the linkage analysis explained 33.9% of the variation. Common variants at or near the fibrinogen gene cluster on chromosome 4 showed significant association with GPF plasma levels (P < 5.0E-8), explaining 22.6% of GPF variance. The top SNP was rs7654093(T) (MAF = 0.23, P = 4.9E-72) at ~19 kb upstream of the FGG (Fibrinogen Gamma Chain). Suggestive evidence of association was found at 8p23.2 (lead SNP: rs7005128(T) with P = 7.9E-7 near the MCPH1 gene). Meta-analysis of the two cohorts confirmed the signals on chromosome 4, with best p-value of 4.9E-199 and the direction of the effects of the top SNPs were in strong agreement. Linkage analysis using the sibling subset of the two cohorts identified significant signals at a ~40 cM interval on 4q28.3-q34.2 (LOD=2.0, P < 0.05) including the fibrinogen gene cluster, and on a novel locus: a ~37 cM interval on 8p23.3-p12 (LOD=2.00, permutation P = 0.03). The top peak in the linkage analysis explained 33.9% of the variation in the GPF levels. To our knowledge, this is the first genome-wide study investigating gamma prime fibrinogen level variation in a healthy population. These results provide new insight into the regulation of hemostasis, may identify novel genetic modifiers of bleeding and thrombosis risk, and suggest that linkage analysis of blood clotting traits may reveal novel loci supported by association studies.

943S A GWAS of Risk Genes for Birth of a Child With Down Syndrome. E. Feingold1, Z. Zeng1, E.G. Allen1, D. Ramachandran1, M.E. Zwick2, S.L. Sherman2. 1) Dept Human Gen, Univ Pittsburgh, Pittsburgh, PA; 2) Dept Human Gen, Emory University, Atlanta, GA.

We have conducted the first GWAS for risk of having a child with Down syndrome (DS) in approximately 700 mothers of children with maternally-derived free trisomy 21. Mothers, fathers, and children were genotyped at the Center for Inherited Disease Research (CIDR) on the Illumina OmniExpress. Genotypes for chromosome 21 in children were called using methods we previously developed. Genotypes for parents and child were then used to establish parent of origin, stage (meiosis I or meiosis II) and meiotic recombination patterns for each child. We used two different designs for the GWAS. The first uses the fathers as controls. In the second, we take advantage of the fact that the etiologies of meiosis I trisomy 21 and meiosis II trisomy 21 appear to be quite different, and perform a GWAS of mothers of meiosis I cases vs. mothers of meiosis II cases. This approach cannot find variants that are common risk factors for both meiosis I and meiosis II errors, but is a well-controlled design for finding variants that are unique to one or the other. Results include genes that are known to affect meiotic recombination, which is not unexpected since the link between altered recombination and nondisjunction is well established both in humans and experimental organisms.

944M Genome-wide Association Study, Meta-Analysis and Linkage Study of Gamma-Prime Fibrinogen Plasma Levels in a Healthy Young Cohort. A. Ozel1, K. Desch2, D. Siemieniak3, D. Ginsburg4, J.Z. Liu1. 1) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 2) Department of Pediatrics, University of Michigan, Ann Arbor, MI; 3) HHMI, University of Michigan, Ann Arbor, MI.

Gamma-prime fibrinogen (GPF) is one of the isoforms of fibrinogen and a component of the blood-clotting complex, making up ~10% of the total fibrinogen levels in blood. Its role in cardiovascular risk prediction has been recently proposed. Plasma levels of fibrinogen are highly variable among healthy people, with ~50% of their variance attributable to inherited factors. We performed genome-wide association studies in a healthy sibling cohort of 1,152 subjects, focusing on the European subset (n = 940), and a second healthy cohort of 2,304 individuals (Desch et al., 2012). Plasma GPF levels were determined using a monoclonal antibody specific to this isoform of fibrinogen. Heritability (h²) was estimated as 61.6% (using Merlin-Regress) and was consistent with results from GCTA (63.3%) and intra-class correlation (65.8%). Common variants across or near the fibrinogen gene cluster on chromosome 4 showed significant association with GPF plasma levels (P < 5.0E-8), explaining 22.6% of GPF variance. The top SNP was rs7654093(T) (MAF = 0.23, P = 4.9E-72) at ~19 kb upstream of the FGG (Fibrinogen Gamma Chain). Suggestive evidence of association was found at 8p23.2 (lead SNP: rs7005128(T) with P = 7.9E-7 near the MCPH1 gene). Meta-analysis of the two cohorts confirmed the signals on chromosome 4, with best p-value of 4.9E-199 and the direction of the effects of the top SNPs were in strong agreement. Linkage analysis using the sibling subset of the two cohorts identified significant signals at a ~40 cM interval on 4q28.3-q34.2 (LOD=2.0, permutation P < 0.05) including the fibrinogen gene cluster, and on a novel locus: a ~37 cM interval on 8p23.3-p12 (LOD=2.00, permutation P = 0.03). The top peak in the linkage analysis explained 33.9% of the variation in the GPF levels. To our knowledge, this is the first genome-wide study investigating gamma prime fibrinogen level variation in a healthy population. These results provide new insight into the regulation of hemostasis, may identify novel genetic modifiers of bleeding and thrombosis risk, and suggest that linkage analysis of blood clotting traits may reveal novel loci supported by association studies.
non-diabetic controls. In micro-dissected glomerular (n=17) and tubule (n=2,494 controls (p = 0.295). We ultimately interrogated transcriptomic analysis stage DN cohorts (FINNDIANE and UK-ROI) composed of 2,142 cases and 779 patients with diabetes for more than 15 years duration.

eery cohort comprised 683 proteinuric patients with or without renal failure.

In order to identify novel susceptibility genes for diabetic nephropathy (DN), we performed a genome-wide association study using 1,000 Genomes-based imputation in individuals with type 1 diabetes. The discovery cohort comprised 683 proteinuric patients with or without renal failure (cases) and 779 patients with diabetes for more than 15 years duration and no evidence of renal disease (controls). None of the single nucleotide polymorphisms (SNPs) tested reached genome-wide statistical significance. The top forty-six SNPs with p-value < 10^{-5} were brought for initial replication in 820 cases and 885 controls part of the US Gokind study. Two SNPs in strong linkage disequilibrium with each other, located in the SORB1 gene, were consistently and significantly associated with DN (p = 7.87 \times 10^{-5} in the discovery, p = 1.32 \times 10^{-4} in the initial replication). In the combined samples, the minor allele of the candidate SNP was less frequent in cases than in controls (37% vs 45%) and associated with a decreased risk for DN of OR = 0.74 [0.67 - 0.82]. However, this association was not observed in a second stage DN cohorts (FINNDIANE and UK-ROI) composed of 2,142 cases and 2,494 controls (p = 0.295). We ultimately interrogated transcriptomic analysis stage DN cohorts (FINNDIANE and UK-ROI) composed of 2,142 cases and 779 patients with diabetes for more than 15 years duration.

To extend a thorough catalog for FPG loci, we conducted meta-analyses for additional genome-wide association studies, each copy of the risk allele increased the odds of SDR by 25% (OR=1.26, 95%CI: 1.15-1.38, p=6e-7). NLRP3 is a member of the inflammasome complex which serves as a platform for the canonical and noncanonical activation of caspase-1 in immune response. SNPs within NLRP3 (not in or near PDK1-RAPGEF4, KANK1 and IGF1R. Our results could provide additional insight into the genetic variation implicated in fasting glucose regulation.
A Genome-wide Association Study of Apnea-Hypopnea Index in Children with Obstructive Breathing, T. Pellegrino et al.

Objective: We examined single nucleotide polymorphisms (SNPs) identified in a Genome-wide Association Study (GWAS) of the adult apnea-hypopnea index (AHI) and SNP-analyte level for association with related traits such as waist circumference and C-reactive protein. The potential of neonatal genetic and metabolic profiles to predict risk for associated with increased C10 in adults, and C6 and C8 have been found to be associated with decreased insulin sensitivity, and higher C4 levels respectively.

Conclusions: The C allele of rs2014355 has previously been shown to be associated with biomarkers that are important in Type 2 Diabetes (T2DM) and a variable fraction of persons develop T2DM and T1DM. The disease is characterized by habitual snoring, disturbed sleep and problems with daytime neurobehavioral functioning. OSAS appears to result from diverse gene–gene interactions and associations with environment changes. There have been no published GWAS in children for OSAS. The identification of genetic variants associated with increased risk for OSAS could potentially translate into earlier recognition and treatment with reduced morbidity, and may also serve to identify potential targets for novel therapies. Here, we present a genome-wide association study of the Apnea-Hypopnea Index measured in a cohort of children referred to the sleep clinic at the Children’s Hospital of Philadelphia for suspected obstructive breathing. Our aim was to identify common genetic variants that increase OSAS severity in children with sleep difficulties. A total of 2,473 children participated in the sleep study. The primary reason for referral was for suspected OSA and they had a Polysomnography (PSG) exam performed. 1,782 children in the sample were given a potential diagnosis of OSA prior to the sleep study. Also, a total of 1201 children were listed as having at least one mental disorder. All participants were genotyped using either the Illumina HumanHap550 or 610 Quad arrays. The association between the natural log of the Apnea-Hypopnea Index (AHI) and SNP was assessed using a linear model in PLINK. Principal components for each individual were extracted using EIGENSTRAT. Subsequently, we used GCTA to account for population stratification and relatedness of the sample. Although none of the variants reached the genome-wide significance (p < 5 x 10-8), we found several suggestive association signals located in interesting candidate genes, which mapped to pathways involved in the cell cycle, microtubule organisation, cell-cell interactions and the immune response. Our strongest individual association is located within the CHR11 region, which may play a role in mechanisms behind the disease.

Background: The human function of lycopene remains undetermined, this nutrient has been both positively and inversely associated with the risk of several chronic diseases. The inconsistencies in lycopene-disease association studies may stem from a lack of knowledge about the genetic variation in the synthesis, metabolism, and deposition of transport and binding proteins, which potentially influence serum lycopene concentrations. Objective: Here we examined the association between variation across the genome and serum concentrations of lycopene in a multi-ethnic population. Design: Participants included African (n=914), Hispanic (n=464), and European American (n=1,203) postmenopausal women from the Women’s Health Initiative (WHI). We analyzed ~7 million single nucleotide polymorphisms (SNPs). Linear regression models were used to assess associations between each SNP and serum concentrations (log-transformed, continuous) of lycopene, adjusting for age, BMI, and population substructure. Models were run separately by ethnicity and then combined in a trans-ethnic fixed effects meta-analysis. Results: In the meta-analysis, the SCARB1 gene, which encodes for a plasma membrane receptor for high-density lipoprotein cholesterol, significantly associated with lycopene concentrations (rs1672879, P-meta < 2.68 x 10-9). Here each additional G allele resulted in a 12% decrease in lycopene levels for African Americans, 20% decrease for Hispanics, and a 9% decrease for European Americans. In addition, two regions significantly associated with serum lycopene levels in African Americans were found. One of these regions, which included the OSAR gene, is involved in cell cycle regulation and may play a role in regulating the potential of neonatal genetic and metabolic profiles to predict risk for later life chronic conditions will be of paramount importance for personalized medicine, and may provide insight into the pathogenesis of complex metabolic disorders.
952S Systematic genome-wide microbiome association analysis in the Northern-German population identifies genetic variation that impacts microbial diversity in the gut. A. Franke1, J. Skjerveiviciiene1, J. Wang2, P. Rausch3, U. Noetling3, W. Lieb4, T.H. Karlsen5, J. Baines6, The PopGen Microbiome Study Group. 1) Inst Clinical Molec Biol, University of Kiel, Kiel, Germany; 2) Max-Planck Institute of Evolutionary Biology, Plien, Germany; 3) Institute for Ermahungs- und Liihtsiseen Medizin (IEL), Friedrich-Wilhelms-Universitat Bonn, Bonn, Germany; 4) Institute of Epidemiology, The PopGen Biobank, University of Kiel, Kiel, Germany; 5) NoPSC Study Group, Rikshospitalet, University of Oslo, Oslo, Norway.

The bacteria of the gut are a community that has co-evolved with the host and confers beneficial effects on human physiology and nutrition, and is crucial for human health. Increasing evidence suggests that the host’s genetic variation might influence bacterial composition in addition to other direct (secretions of bile and defensins, control of gut motility etc.) or indirect (food and lifestyle) factors. An effect of host genotype on bacterial composition has been demonstrated by our group and others for single genes (MYD88, NOD2, TLPIs, immunoglobulin A (IgA), and the human leukocyte antigen (HLA) region. In order to systematically evaluate the relationship between variation in the human genome to variation in the human microbiome, large-scale GWAS (genome-wide microbial association studies) studies of population-representative cohorts are needed. Here, we performed a GWAS of a cross-sectional, population-representative study cohort of Northern Germany (n=775). Ultra-dense SNP genotype information, available from different genotyping platforms, detailed phenotypic and dietary data, 16S rDNA microbiome data (V1-V2 region, 454 Roche platform) were generated for this Northern German cohort. Systematic association studies with phenotype data allow for study of inter-individual microbial variability. From the screening study, an independent study cohort was recruited from the same geographic region within Northern Germany (n=866). Again, dense SNP genotype data, phenotypic, dietary, and 16S rDNA microbiome data (again V1-V2 region) were collected and association analysis was performed. The discovery-to-the discovery cohort, were collected for the replication cohort. Our benchmarking study revealed that the discovery cohort can be used to introduce the sequencing method. We detected that genetic variation of the host contributed more than 4% to the inter-individual variability. Moreover, we have managed to replicate several genetic loci that are implicated in determining microbial diversity.

953S Identification of Susceptibility Loci for Crohn’s Disease in Koreans through Immunochip. M. Hong1, H. Choi1, W.J. Yun2, J. Baek3, T. Haritunians3, S.H. Park2, B.D. Ye4, J. Liu5, D.P.B. McGovern5, S. Yang6, K. Song6, 1) Dept. Biochemistry and Molecular Biology, Ulsan university College of Medicine, Ulsan, South Korea; 2) The F. Widjaja Foundation Institute of Medical Sciences, Kyushu University, Fukuoka, Japan; 3) J. Liu, H. Choi, M. Esaki, S. Nakamura, T. Matsuji, T. Tanaka, M. Motoya, Y. Suzuki, Y. Kiyohara, Y. Katozono, S. Nakamura, M. Kubo. 1) Laboratory for Genotyping Development, Center for Investigative Medical Science, RIKEN, Yokohama, Japan; 2) Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; 3) Laboratory for Statistical Analysis, Center for Investigative Medical Science, RIKEN, Yokohama, Japan; 4) Department of Medicine, Division of Gastroenterology, Social Insurance Chuo General Hospital, Tokyo, Japan; 5) Department of Gastroenterology, Fukuoka University Chikushi Hospital, Fukuoka, Japan; 6) Department of Gastroenterology, Sapporo Kosei Hospital, Sapporo, Japan; 7) Department of Internal Medicine, Faculty of Medicine, Toho University, Chiba, Japan; 8) Department of Environmental Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; 9) Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Crohn’s disease (CD) and ulcerative colitis (UC) are the two major forms of inflammatory bowel disease (IBD), induced by multiple factors. Meta-analysis of genome-wide association studies (GWAS) identified 163 susceptibility loci for IBD from European descendants. However, there was very little report of IBD GWAS from other ethnic groups. Recently, the human genetic resource has been growing over and it can make to impute missing genotypes. Therefore, we performed three GWAS (CD, UC and IBD) using imputed sequencing data of 3,124 subjects which was a total of 376 CD cases, 376 UC cases and 3397 controls as previously reported. After applying stringent quality control for samples and SNPs, we performed the association analyses of 4,062,308 SNPs in CD GWAS, 4,059,167 SNPs in UC GWAS and 4,077,852 SNPs in IBD GWAS with accuracy (Rsq) > 0.9 and MAF > 0.05 of both cases and controls. However, we could not find significant association in each GWAS except for 2 reported loci (TNFSF15 and MHC region). To identify new susceptibility loci, we performed the replication studies with independent cases and controls. We selected 1,605 SNPs (596 CD specific, 606 UC specific and 403 IBD shared) with P values < 5 x 10^{-5} after excluding SNPs located in previously reported loci. Among them, we picked up 137 tagging SNPs (53 SNPs as CD specific, 47 SNPs as UC specific and 37 SNPs as IBD shared) and genotyped 1,648 subjects using 948 CD patients, 361 UC patients and 416 controls. By combined analysis, we identified 3 suggestive loci with P value less than 1 x 10^{-4}. The two candidate loci as CD specific were in RUNX3 on chromosome 1p36 (rs8767105: 4.01 x 10^{-7}, odds ratio (OR) = 0.81) and in the chromosome 11q21 (rs4365033: 4.45 x 10^{-4}, OR = 1.37). The third locus was in PAR6Dgene 18q23 (rs4798947: 5.39 x 10^{-7}, OR = 1.51) as UC specific. Though there was no suggestive locus shared IBD, the strongest association was shown in TNK on chromosome 3q25 (rs852209: P = 5.10 x 10^{-6}, OR = 1.18) as the strongest candidate locus for IBD and 1 candidate locus for UC in Japanese population. Further studies will be required what these loci play role in the pathogenesis of IBD.
956M

Genome-wide and eQTL study of subcutaneous and visceral adipose tissue reveals novel gender-specific adiposity loci in Hispanic Americans: The Insulin Resistance Atherosclerosis Family Study (IRASFS). C. Gao1,2,3, J. Ziegler1,2,4, K.D. Taylor2, J.M. Norris4, Y.D.I. Chen2, J.I. Rotter2, L.E. Wagenknecht7, C.D. Langefald4, N.D. Palmer2,8,9, 1 Program of Molecular Genetics and Personalized Medicine Research; Wake Forest School of Medicine, Winston-Salem, NC; 2 Center for Public Health Genomics; Wake Forest School of Medicine, Winston-Salem, NC; 3 Center for Biostatistical Sciences; Wake Forest School of Medicine, Winston-Salem, NC; 4 Institute for Translational Genomics and Population Sciences and Department of Pediatrics; Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA; 5 Department of Epidemiology, Colorado School of Public Health; University of Colorado, Aurora, CO; 6 Department of Public Health Sciences; Wake Forest School of Medicine, Winston-Salem, NC; 7 Center for Diabetes Research; Wake Forest School of Medicine, Winston-Salem, NC; 8 Department of Biochemistry; Wake Forest School of Medicine, Winston-Salem, NC.

Obesity is a major public health concern with strong correlations to numerous metabolic diseases. Moreover, regional fat deposition has been shown to increase disease risk above and beyond total adiposity. Here we report a genome-wide association study (GWAS) in 994 Hispanic Americans (Nmas=408, Nfemale=586) from IRASFS for ~8 million SNPs (genotyped and imputed). In addition, evaluation of low-frequency/rare variants was facilitated using the exome chip (81,560 variants). Association analyses were performed using computed tomography (CT) measures: subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT), VAT relative to BMI (VAT_BMI), and visceral/subcutaneous adipose tissue ratio (VSR). As fat deposition appears to be gender specific, gender-stratified analyses were performed. Our results identified a total of 12 SNPs from five loci reaching genome-wide significance (P<5.0x10^-8). SNP rs12657394 (MAF=19%), an intrinsic variant of SRBP1 on chromosome 5, showed strong association with VSR in males (PMale=3.32x10^-8; Pfemale=0.0025). SRBP1, also named as p49/STRAP, encodes a Serum Response Factor Binding Protein. It has been shown to be involved in biosynthesis and regulation of glucose homeostasis. Upstream of SNPs rs12657394 (MAF=42%), downstream of SNX3 (Sorting Nexin 13), was also associated in males with VAT_BMI (PMale=2.30x10^-5; Pfemale=0.28); rs117206355 (MAF=1%), located downstream of p49/STRAP, was strongly associated with VAT_BMI in females (PMale=1.34x10^-8; Pfemale=1.18x10^-6). In addition, two highly correlated intronic SNPs (rs2142795 and rs2425494; MAF=9%), in PTPRT showed strong association with VSR in males (PMale=2.30x10^-5; Pfemale=0.30); rs12123452 (MAF=6%), downstream of RIMS3, showed a strong signal for VAT in males (PMale=4.35x10^-8; Pfemale=0.12). To evaluate gender heterogeneity in fat deposition, interaction analyses were performed and significant gene-gender interactions were detected for all five significant loci with the most significant interaction at rs117206355 (P=8.20x10^-5). These results revealed five novel adiposity loci displaying strong gender specificity. These observations provide genetic evidence for a differential mechanistic basis of fat deposition between genders and warrant further replication in larger cohorts.

957T

Meta-analysis of macronutrient intake in over 64,000 individuals using 1000 Genomes imputed genotypes confirms the association of FGF21 with composition of dietary intake and suggests potential tissue-specific effects in liver and skeletal muscle. A.Y. Chu1,2, M. Graff2, K.E. North8, R.N. Lemaire3, J. Zhao3, L. Yuan4, R.A. Scott5, N. Ts樱ovska5, M. Perola6, J.S. Ngwa6, L.A. Cupples2, M.A. Nalls4, D.K. Houston12, J. Huang11, Q. Qi13,12, T.S. Ahuwalia13, T.I.A Sorensen13, C. Schutz14, M. Ortho-Melander14, A.C. Frazier-Wood15, T. Chen16, P.S. de Vries16, F.J.A. van Rooij16, J.C. Kiefe16, S. Kanoni17, G. Dedousis17, T. Lehtimäki18, O. Raitakari19, D.I. Chasman1, T. Tanaka20, CHARGE Nutrition Working Group; DietGen Consortium. 1 Brigham & Women’s Hospital; Boston, MA, USA; 2 University of North Carolina, School of Public Health, Chapel Hill, Chapel Hill, NC, USA; 3 University of Washington School of Medicine, Seattle, WA, USA; 4 MRC Epidemiology Unit, University of Cambridge, Cambridge, UK; 5 Institute of Molecular and Cell Biology, University of Tartu, Estonia; 6 National Institute for Health and Welfare (THL), Helsinki, Finland; 7 Howard University, Washington DC, USA; 8 Boston University School of Public Health, Boston MA, USA; 9 National Institute on Aging, National Institute on Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, National Institute of Arthritis and Musculoskeletal and Skin Diseases, and National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; 10 Karolinska Institute, Stockholm, Sweden; 11 Albert Einstein College of Medicine, Bronx NY, USA; 12 University of Copenhagen, Copenhagen, Denmark; 13 University of Lund, Lund, Sweden; 14 Department of Clinical Imaging, Fimlab Laboratories and School of Medicine, University of Tampere, Tampere, Finland; 15 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; 20 National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health, Bethesda, MD, USA.

While dietary proportions of macronutrients, namely carbohydrate, protein, and fat, are relatively constant across populations, inter-individual variation of these quantities has been associated with chronic conditions such as obesity and diabetes. Prior genome-wide investigations based on common SNPs in FGF21, RASIP1 (RAS-like GTPase, p49/STRAP, also named as SRBP1), and/or processing of SRFBP1 (Sorting Nexin 13) suggests a role in regulating glucose homeostasis. SNP rs13247968 (Glucose Transporter Type 4) in adipocytes, and/or processing of SRFBP1 with composition of dietary intake, estimated from food frequency questionnaires as mean intake of at least one macronutrient (p<1e-5) we found a greater proportion of caloric intake from carbohydrate, protein, or fat. To further investigate the genetics of dietary intake, including associations from regions of low imputation quality in MapQ2 and lower frequency variants (MAF<5%), we performed a combined genome-wide inverse-variance weighted meta-analysis including >64K participants from 18 cohorts of European ancestry (ARIC, CHS, EPIC-Norfolk, FamHS, Fenland, FINRISK, FHS, H300, HealthABC, InChianti, Rotterdam Study, I.I & II, WHGS, and YFS) using imputed genotypes interrogated to the 1000 Genomes phase 1 release 13 reference panel (March 2012). A total of 11.2 million variants remained after filtering on imputation quality (<0.4) and MAF (<0.01%), of which 43% had MAF<5%. In combined analysis, we confirmed the association of the minor allele for synonymous variant rs838133 (MAF=42%) at the FGF21 locus with higher carbohydrate (β=0.3%) and lower fat (β=0.2%) intake at genome-wide significance (p<5e-8) and at suggestive significance with lower protein intake (β=-0.1%, p=1.2e-6). We identified two missense SNPs in FGF21 (MAF=36%) and RASIP1 (MAF=49%) in moderate LD (r2=0.6) with rs838133: only the RASIP1 SNP is predicted to be damaging based on 3 of 4 protein prediction algorithms. Conditional analyses of rs838133 attenuated associations for both SNPs (p>0.05). No other variants attained genome-wide significance. However, among variants associated with intake of at least one macronutrient (p<1e-5) we found a greater proportion than expected that overlap with chromatin signatures for active gene regulation (H3K4me3) in liver, skeletal muscle and immune cells (p<0.05). An ongoing effort to include 6 additional cohorts (~46K participants) will increase the sample size to >100K making it the largest genome-wide investigation of dietary intake. The larger sample has potential to identify novel loci and therefore pathways for dietary habits and diet-disease relationships.
Genome-Wide Association Study Identifies Novel Genetic Determinants of Emphysema Distribution Patterns. A. El Boueziz1,2,3, S.M. Lutz2, R.P. Bowlby4, M.H. Cho4,5,6, M.L. McDonald1,3, N.M. Laird1,3, T.H. Beatty4, J.D. Crapo5, E.K. Silverman1,2,3, P.J. Castaldi1,3, D.L. DeMeo1,2,3, 1) Channing Division of Network Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA; 2) Pulmonary and Critical Care Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA; 3) Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA; 4) Department of Biostatistics, Colorado School of Public Health, University of Colorado, Aurora, CO, USA; 5) Division of Pulmonary Medicine, Department of Medicine, National Jewish Health, Denver, CO, USA; 6) Harvard School of Public Health, Boston, MA, USA; 7) Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA.

Background: Emphysema is a chronic obstructive lung disease (COPD) with an irreversible destruction of lung parenchyma. The distribution of this destruction throughout the lung varies considerably. Upper lobe emphysema has emerged as an important predictor of response to lung volume reduction surgery. Yet, the determinants of apical versus basal emphysema distributions remain largely enigmatic. Alpha 1-antitrypsin deficiency is often associated with basilar emphysema whereas polymorphisms in xenobiotic enzymes (GSTP1 and EPHX1) in non-alpha 1-antitrypsin deficient smokers have been related to apical predominance. These findings suggest the presence of genetic influences on emphysema distribution patterns. To investigate this hypothesis and identify additional genetic markers of emphysema distributional phenotypes, we performed a GWAS in the COPDGene study. Methods: From the full cohort of smokers in COPDGene, 6,094 non-Hispanic white (NHW) and 2,869 African American (AA) subjects with complete genotype and CT densitometry data were included in this analysis. Genotyping was performed on the Illumina Omni Express platform with additional markers imputed using 1,000 Genomes reference data. Standard Genetic model adjusting for age, gender, pack-years of smoking, and genetic ancestry, each of the four CT scan emphysema distribution phenotypes (apical percent emphysema, basal percent emphysema, difference, and Log ratio between the two) was tested for genetic associations. Separate analyses in NHW and AA subjects were followed by a meta-analysis. Results: In the meta-analysis, markers from 7 distinct loci were associated with one or more of the emphysema distribution phenotypes at genome-wide significance, 3 loci (4q31 near HHIP; 15q25 near CHRNA3/CHRNA5; 1q41 near TGFBR2) have been previously associated with COPD susceptibility, and 4 loci are newly identified (TLE1; RANBP17; TOR1; NR2F1-AS1). Of these 7 loci, 4 were nominally significant in both NHW and AA subjects (CHRNA3/CHRNA5; RANBP17; TOR1B; NR2F1-AS1) and 3 in NHW subjects only (HHIP; TGFBR2; TLE1). Conclusion: This GWAS in NHW and AA individuals identified new loci that may represent risk loci for upper lobe predominant versus diffuse emphysema. Our results support the notion that genetic factors may impact emphysematous destruction. Grant Support: The COPDGene study is supported by NIH R01 HL089897 and R01 HL089856.

Heritability and locus susceptibility in age-related macular degeneration varies by clinical phenotypes. L. Shen1, T. Hoffmann2,3, R. Melles4, L. Sakoda5, M. Kvala6, Y. Banda6, N. Risch7, C. Schaefer1, E. Jorgenson1, 1) Division of Research, Kaiser Permanente Northern California, Oakland, CA 94612; 2) Institute for Human Genetics, University of California San Francisco, San Francisco, CA 94143; 3) Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA 94143; 4) Department of Ophthalmology, Redwood City Medical Center, Kaiser Permanente Northern California, CA.

Age related macular degeneration (AMD) is a complex, late onset vision disorder that progresses in stages. Majority prior studies focused on advanced AMD, and the most recent genome-wide association study (GWAS) meta-analysis identified 19 independent loci. Other AMD related clinical phenotypes have been less extensively examined. Here, we conducted a GWAS of overall and subgroups of AMD based on imputation to the 1000 Genomes Project reference panel. The analysis included a total of 5,762 overall AMD cases (2,504 with advanced AMD) and 45,929 controls of non-Hispanic white ancestry from the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort as part of Kaiser Permanente Research Program on Genes, Environment and Health. We observed large differences in genetic susceptibility across AMD stages and subtypes. The estimated variance explained by all autosomal SNPs was lower for overall AMD (32%) than advanced AMD (40%), lower for geographic atrophy (GA) (36%) than choroidal neovascularization (NV) (47%), and highest for cases with both NV and GA (52%). For the 19 previously reported risk loci, the proportion of variance explained, area under the ROC curve, and odds ratios also vary by AMD phenotypes. GWAS of the five AMD categories identified 20 new risk loci, with 7 in NHW and 14 in AA subjects. While conditioning on the top SNPs at each risk locus, we identified a non-synonymous SNP in HLA-DQ1 to be associated with the risk of overall AMD (p = 1.64x10-12). These findings suggest that shared genetic factors underlie the risk of AMD of all clinical forms. The strength of the effects vary by disease severity and subtype.

Genome-Wide Association Study of Serum Sodium Concentration in Han Chinese Population residing in Taiwan. I. Song1, J. Yang1, C. Chen1, C. Sung2, S. Lin2, Y. Chen1, J. Wu1, 1) Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; 2) Division of Nephrology, Department of Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan.

Water is the most important human body component that counts for 60-70% of body weight. Water homeostasis is the essential vital balance which maintains human survival. Serum sodium concentration is the index reflecting the systemic water balancing condition. The normal serum sodium concentration is between 135 to 145 mM. Hypernatremia (>145 mM) was reported to associate with increase mortality and hyponatremia (<135 mM) was suggested to be a risk factor for osteoporosis, myocardial infarction, and attention deficits. Recently, high heritability of serum sodium concentration was demonstrated in non-Hispanic Caucasian, African American, and American Indian populations. However, the genetic influences on serum sodium concentration in Han Chinese Taiwan have not yet been studied. Therefore, in current study, we aimed to identify the loci affecting serum sodium concentration. We performed GWAS on 375 individuals with relative high sodium concentration (138-153 mM) and 394 individuals with relative low sodium concentration (85-133 mM). All subjects carry no severe medical illness. We found that chromosomes 14q31.1 (p<10-7) and 17p13.3 (p<10-6) showed strong association. The replication with another independent cohort is in processed. We hope the findings from this study can help in early identifying subject that predisposing to spontaneous or iatrogenic hyper/hyponatremia.
Y chromosomal degradation and loss is a known phenomenon in male aging and has been the focus of recent studies. We explored the hypothesis that Y chromosome variants contribute to Y chromosome stability and also to successful aging in the Long Life Family Study (LLFS). We used data on 2182 men, of whom 2135 have 450 Y chromosome genetic markers and intensity data from the Illumina HumanOmni2.5 GWA chip. LLFS was designed to examine the genetics of healthy aging and families were recruited consisting of siblings who showed exceptional longevity, their children, and spouses where available. The mean age of male subjects on first examination, when DNA was collected, was 71. The mean age of men with Y chr intensity < 0.9 (indicating substantial Y chromosome degradation/loss) was 91, consistent with prior studies. Using a minor allele persistence algorithm, we have identified several SNPs with significant differences in age-specific allele frequencies, possibly associated with Y chromosome loss (p-values < 1.1×10^{-4}, significant for 450 Y chromosome SNPs—Bonferroni alpha correction). Of these, rs17316547 is a variant within a predicted snRNA and is within 30kb of PRKY. Additionally, of the 231 individuals who carry this allele in our data, the youngest recorded death was 83 years old. The PRKY gene is similar to the protein kinase, X-linked gene in the pseudoautosomal region of the X chromosome. The gene is classified as a transcribed pseudogene because it has lost a coding exon that results in transcripts being candidates for nonsense-mediated decay (NMD) and unlikely to express a protein. Abnormal recombination between this gene and a related gene on chromosome X is a frequent cause of XX males and XY females, making it a plausible biomarker for genomic instability in age related diseases.

Genetic admixture and proliferative diabetic retinopathy in Latinos. X. Gao, W.J. Gauderman, P. Marjoram, M. Torres, Y.I. Chen, K.D. Taylor, J.J. Roundtree, R. Varama. 1) Dept Ophthalmology & Visual Sci, Univ Illinois, Chicago, Chicago, IL; 2) Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) USC Eye Institute, Department of Ophthalmology, University of Southern California, Los Angeles, CA; 4) Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA, Torrance, CA.

Diabetic retinopathy is a leading cause of blindness in working age adults, and proliferative diabetic retinopathy (PDR) is the most advanced stage of the disease. Epidemiology studies have observed that Latinos have a higher prevalence of severe diabetic retinopathy than whites. The purpose of this study is to test the association between genetic admixture and PDR in Latinos with type 2 diabetes mellitus (T2DM). We conducted a case-control study using 647 T2DM subjects (552 controls and 95 PDR cases) selected from the tails of the diabetic retinopathy distribution from the Los Angeles Latino Eye Study. Genotyping was performed on the Illumina OmniExpress BeadChip (730,525 markers). We estimated genetic ancestry in Latinos using STRUCTURE with the HapMap reference panels. Univariate and multivariate logistic regression analyses were used to test the relationship between the proportions of genetic ancestry in Latinos and PDR. Native American ancestry in Latino T2DM subjects is significantly associated with PDR (P = 0.004) in our univariate analysis. The association remained significant after adjusting for age, sex, hemoglobin A1c, body mass index, systolic blood pressure, education and income. Risk factors for PDR in Latinos in our multivariate analysis include duration of diabetes (P < 0.0001), hemoglobin A1c (P = 0.0003), systolic blood pressure (P = 0.0004), and Native American ancestry (NAA, P = 0.006). The empirical p-value for NAA from permutation tests was 0.005. We also validated the NAA estimates in Latinos using ABIDE and 1000 Genome Project reference panels and obtained consistent results. Our results demonstrate for the first time that NAA is a significant risk factor for PDR in Latinos. We are further performing local ancestry analysis, which may identify specific genomic regions associated with PDR in Latinos.
965M
Genetic determinants of healthspan - analysis of Wellderly dataset. W. Sikora-Wohlfeld1, M. Sirola1, E. Scott2, A. Torkamani2, E. Topol2, A.J. Butte1. 1) Department of Pediatrics, Stanford University, Stanford, CA; 2) Scripps Translational Science Institute, La Jolla, CA.

Understanding the mechanisms of aging would potentially enable detecting and preventing debilitating diseases and expanding the life expectancy. Numerous efforts have focused on searching for longevity genes. Development of genotyping and genome sequencing techniques enabled direct investigation of the genetic background of longevity. In recent years a number of studies have been conducted, where the genomes of people who have lived to a very advanced age in good health, were analyzed to search for variants for longevity. However, the mechanisms governing long healthspan remain to be explained.

In this study we have analyzed genome sequences of 534 participants of the Wellderly project, who have lived to 80 years or beyond. We compared this cohort to a control group, derived from the 1000 Genomes Project, assuming that the 1000 Genomes participants represent a sample from an ordinary population with average life expectancy. Using principal component analysis to estimate genetic ancestry, we identified a group of 416 Wellderly individuals who matched 174 individuals from CEU and GBP populations in the 1000 Genomes Project.

We performed an association test for a selected set of SNPs known to be associated with age-related diseases. We derived this list of SNPs from VARIMED, a database of human disease-SNP associations. At the time of this work, VARIMED contained 466,890 unique SNPs associated with 6,691 disease and related phenotypes, manually curated from 17,088 publications.

Among the top variants differentiating the Wellderly and 1000 Genomes cohorts, we identified four markers in the T cell receptor alpha constant gene (TRA). In addition, we found a marker in the immunoglobulin heavy focus gene (IGH). A link between aging and immunity has been established before and the performance of the immune system has been suggested to affect healthspan. While still preliminary, our findings suggest an association between immune pathway genetics and longevity, as well as enabling the exploration of other longevity-related markers.

966T

Adjustment for covariates in genome-wide association studies (GWAS) has a dual purpose: (a) to account for potential confounding factors that can bias SNP effect estimates, and (b) to improve statistical power by reducing residual variance. Recently, researchers have conducted GWAS of human traits and diseases while adjusting for other heritable covariates with another motivation: identifying genetic variants associated only with the primary outcome. We show that this objective is fulfilled when the tested variants have no effect on the covariate or when the correlation between the covariate and the outcome is fully explained by a direct effect of the covariate on the outcome (i.e. mediation). For all other scenarios, an unintended bias is introduced with respect to the primary outcome as a result of the adjustment, and this bias may lead to false positives. We illustrate this point by providing examples from published genome-wide association studies, including a large meta-analysis of waist-to-hip ratio (WHR) adjusted for body mass (BMI) index from the GIANT consortium, where genetic effects may be misestimated as a result of adjustment for BMI.

We first show that the expected bias of the effect estimate in the covariate adjusted analysis is proportional to the correlation between the covariate and the outcome, and the association between the covariate and the genetic variant tested. We then derive a statistical test to evaluate the presence of such bias using GWAS summary information and the correlation between the covariate and the phenotype. When applied to the GIANT summary statistics, we observed that half of the reported associations with WHR adjusted for BMI are likely influenced by a genetic association with BMI. Finally, we show that heritability estimation of covariate adjusted traits is also subject to the same bias. In particular, we demonstrate that the heritability of an outcome adjusted for a heritable covariate may not necessarily represent an outcome specific, covariate-specific and shared genetic component.

In summary we highlight that adjusting for heritable human traits such as BMI to remove mediated genetic effect through these variables, as commonly done in many genetic studies, should be done with care as it can induce false signal.

967S
Meta-analysis of genome-wide association studies in alopecia areata reveals new susceptibility loci and resolves HLA associations. R. Betz, L. Petukhova1,3, S. Ripke2,5, S. Huang2,6, S. Redler1,5, T. Stecker2,5, S. Hellmann1,2,4, T. Yamany1, M. Duvic10, M. Hordinsky11, D. Norris12, V. Price13, J. Mackay-Wiggan1, A. Menelau14, G.M. Destano15, M. Bohn16, U. Blume-Peytavi1, H. Wolff17, G. Lutz18, R. Kruse19, C.I. Amos20, A. Leer21, P.K. Gregersen22, B. Blauemeyer23, D. Althoff22,6, P.I.W. de et Nkoo24,15, M.M. Nothen12,4, R. Clynes2,4, M. Daly6,7, A.M. Christianson2,5. 1) Institute of Human Genetics, University of Bonn, Bonn, Germany; 2) Department of Dermatology, Columbia University, NY, NY; 3) Department of Epidemiology, Columbia University, NY, NY; 4) Department of Medicine, Columbia University, NY, NY; 5) Department of Genetics & Development, Columbia University, NY, NY; 6) Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA; 7) Stanford Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA; 8) German Center for Neurodegenerative Diseases, Bonn, Germany; 9) Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany; 10) Department of Dermatology, MD Anderson Cancer Center, Houston, TX; 11) Department of Dermatology, University of Minnesota, Minneapolis, MN; 12) Department of Dermatology, University of Colorado, Denver, CO; 13) Department of Dermatology, UCSF, San Francisco, CA; 14) Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands; 15) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; 16) Department of Dermatology, University of Münster, Münster, Germany; 17) Clinical Research Center for Hair and Skin Science, Department of Dermatology and Allergy, Charité-Universitätsmedizin Berlin, Berlin, Germany; 18) Department of Dermatology, University of Munich, Munich, Germany; 19) Dermatological Practice, Hair and Nail, Wesseling, Germany; 20) Dermatological Practice, Paderborn, Germany; 21) Community and Family Medicine and Genetics, Dartmouth College, Hanover, NH; 22) The Feinstein Institute for Medical Research, Manhasset NY; 23) Department of Dermatology, University Medical Center Utrecht, Utrecht, The Netherlands; 24) Department of Genomics, Life & Brain Center, University Bonn, Bonn, Germany.

Alopecia areata (AA) is one of the most prevalent autoimmune diseases, with ten known susceptibility loci so far. Here, we performed a meta-analysis in AA, combining data from two GWAS, and including replication data from Immunochip supplemented with Sequenom genotyping, for a total of 2,807 cases and 6,975 controls. The strongest region of association was the MHC, where we fine-mapped 4 independent effects, all implicating HLA-DR as a key etiologic driver, three of which alter protein sequence in the binding cleft of HLA-DR1. Outside the MHC, we identified two novel loci that exceeded statistical significance, containing ACOXL/BCL2L11 (BIM) (2q13; rs3789129, p=1.5x10−8, ORA=1.3); LRRC32 (GARP) (11q13.5; rs2155219, p=1.25x10−8, OR=1.2). A third region achieved nominal significance: SH2B3 (LINK)/ATXN2 (12q24.12; p=1.3x10−7), which has been observed in other autoimmune diseases. Expression analysis of genes in these three regions provide biologic plausibility to these findings and further support the causal role of aberrant immune processes in AA, with contributions from both immune cells as well as the end organ (hair follicle). Finally, we integrated a cross phenotype meta-analysis integrating our data with data from seven other autoimmune diseases, providing insight into the molecular taxonomy of autoimmune diseases and the alignment of AA within this class of disorders. Importantly, as GWAS help to resolve disease mechanisms and identify pathogenic pathways perturbed in AA and autoimmunity in general, these approaches advance the field towards precision medicine in autoimmunity.
Posters: Complex Traits and Polygenic Disorders

968M
Examining the genetic basis of variation in sitting height ratio (SHR) using population cohorts. Y. Chan1,2,3, E.T. Lim1,2,3, D. Strachan4, G. McMahon4, G. Davey-Smith5, R.M. Salem1,2, J.N. Hirschhorn1,2,3, 1 Endocrinology, Boston Children’s Hospital, Boston, MA; 2 Medical and Population Genetics, Broad Institute, Cambridge, MA; 3 Genetics Dept, Harvard Medical School, Boston, MA; 4 Genetics and Genomics, Boston Children’s Hospital, Boston, MA; 5 Population Health Research Institute St George’s University of London, London, UK; 6 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

Sitting height ratio (SHR) is the ratio of sitting height to total height and is a measure of body proportion. Studying SHR is important because it can be a more sensitive phenotype to detect mild growth disorders than height. While sitting height is heritable, with approximately 70% of variation within a population attributable to genetic factors, the heritability of SHR has not been well studied. Nonetheless, there can be large differences in the average SHR between populations; for example, the average SHR of European ancestry individuals is about 1 standard deviation greater than the average SHR of African-ancestry individuals. This difference is one of the largest reported for anthropometric traits. Here, we report the results from the first genome-wide genetic studies of SHR in large population cohorts of different ancestries. Using data from dbGAP, we first examined the association of SHR with European ancestry in 3,069 African-Americans. We observed a significant peak near PTPRM, which is on the X-chromosome and is associated with SHR only in women. This locus was recently reported to also be associated with height and is known to escape X-inactivation. Interestingly, these three loci do not account for differences in SHR between individuals of European and African ancestries, and there were no significant peaks from admixture mapping, suggesting that the genetic basis for the difference in SHR between populations is largely polygenic. Finally, we tested 421 independent SNPs known to be associated with height and observed that most of these SNPs were either more strongly associated with SHR than expected by chance (N=49, P=2×10^-5) or of the same height-decreasing alleles, 31 are associated with increased SHR and 18 are associated with decreased SHR, suggesting that different height loci have specific effects on the different components of body proportion. This study highlights the importance of SHR as a more sensitive phenotype to detect mild growth disorders than height.

969T
Six novel loci associated with VEGF circulating levels identified by a meta-analysis of genome-wide association studies. S. Choi1,2,3, D. Ruggiero4,5,6, H. Picard4,5,6, T. Nulle7, C. Song8,9, N. Nataye9, M. Stathopoulos9, C. Barbieri7, C. Bellenguez10,11,12, M. Concas13, P. Fitzgerald14, V. Guillaume15,16, A. Leutenegger17,18, E. Ingelsson9, P. Kovacs19,20, V. Lagou21,17,18, J. Lamont22,3, L. Lind23, M. Gaestrelle24, M. Pirastu25, C. Sala26, A. Smith17,18, D. Toniolo27, A. Törjesen28, M. Traglia29, R. Vasan30,10, A. DeStefano1,2,3, S. Visvikis-Siest30,1,2,3, S. Seskind30,1,2,3, M. Ciullo30,1,2,3,1 Department of Neurology, Boston University School of Medicine, Boston, MA; 2 Department of Biostatistics, Boston University School of Public Health, Boston, MA; 3 National Heart, Lung and Blood Institute’s Framingham Heart Study; 4 Institute of Genetics and Biophysics, National Research Council of Italy, Naples, Italy; 5 Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 6 Institute of Population Genetics; National Research Council of Italy; Sassari, Italy; 7 Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy; 8 UMR INSERM U 1122; IGE-PCV ‘Interactions Gène-Environnement en Physiopathologie Cardio-Vasculaire’.Faculté de Pharmacie, Université de Lorraine, Nancy, France; 9 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden; 10 Institut Pasteur de Lille, Lille, France; 11 InsERM, U744, Lille, France; 12 Université Lille-Nord de France, Lille, France; 13 InsErM, U946, Paris, France; 14 Univ Paris Diderot, Sorbonne Paris Cité, IUH, Univ Sorbonne, Paris, France; 15 Cardiovascular Research Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 16 Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom; 17 Icelandic Heart Association, IS-201 Kopavogur, Iceland; 18 University of Iceland, 101 Reykjavik, Iceland; 19 insulin phenotype programme, ISF Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 20 University of Leipzig, Department of Medicine, Leipzig, Germany; 21 University of Leipzig, IFB Adiposity, Leipzig, Germany; 22 Randox Laboratories, Crumlin, United Kingdom.

Vascular Endothelial Growth Factor (VEGF) is the most important proangiogenic factor, implicated in both physiological and pathological angiogenesis. A previously published GWAS had identified four loci independently associated with VEGF circulating levels at genome-wide significance (p<5×10^-8). Of the 49 height-decreasing alleles, 13 loci were found. To identify independently associated variants within the genome-wide significant genomic regions, conditional analyses were carried out. These analyses revealed 10 independent signals (1 on 5q14.3, 4 on 6p21.1, 1 on 8q23.1, 3 on 9p24.2, and 1 on 10q21.3). Further, 13 loci contained variants suggestively associated at 5×10^-5. The lead variant for each of the 10 independent signals and the 13 suggestive loci was carried forward to in silico and de novo replication in ~2800 individuals from 6 cohorts using genome wide association data imputed to the 1000genomes v3 panel. A GWA of VEGF levels was performed in each cohort and the results were meta-analyzed. Five chromosomal regions (5q14.3, 6p21.1, 8q23.1, 9p24.2, 10q21.3) containing SNPs associated with VEGF levels at genome-wide significance (p<5×10^-8) were found. To identify independently associated variants within the genome-wide significant genomic regions, conditional analyses were carried out. These analyses revealed 10 independent signals (1 on 5q14.3, 4 on 6p21.1, 1 on 8q23.1, 3 on 9p24.2, and 1 on 10q21.3). Further, 13 loci contained variants suggestively associated at 5×10^-5. The lead variant for each of the 10 independent signals and the 13 suggestive loci was carried forward to in silico and de novo replication in ~2800 individuals from 6 additional independent cohorts. Ten signals, 8 out of the top 10 independent variants and 2 out of the 13 suggestive signals in the discovery sample, were successfully replicated (p<5×10^-5) in the meta-analysis of the combined discovery and replication samples. Overall, we confirmed the association of 4 already known loci and found 6 new signals, 4 located in novel chromosomal regions (5q14.3, 10q21.3, 16q24.2, and 18q22.3) and 2 in previously identified chromosomal regions (6p21.1 and 9p24.2). These 10 variants explain about 51% of the variability of the circulating levels of VEGF. The Ingenuity Pathway Analysis software (IPA) was used to explore the relationships between all genes associated with circulating VEGF levels. Twenty-four genes located in the regions identified by the 10 replicated variants were selected as focus genes. Out of them, 17 genes were connected within a unique network. This analysis revealed strong functional relations of the gene network with embryonic and organism development, and cardiovascular system development and function. Further analyses will be necessary to identify the functional variants in the identified loci.
970S

Genome-Wide Association Study in Different Stages of Clinical Progression of Alzheimer Disease. J. Chung1, L. Farrer1,2,5,4, G. Jun1,2,3,4,5.
1) Biomedical Genetics, Boston University Medical Campus, Boston, MA; 2) Biostatistics, Boston University, Boston, MA; 3) Ophthalmology, Boston University, Boston, MA; 4) Epidemiology, Boston University, Boston, MA; 5) Neurology, Boston University, Boston, MA.

Background: Amyloid-beta 42 (A42) and phosphorylated tau (p-tau) levels in cerebrospinal fluid (CSF) are key biomarkers for Alzheimer disease (AD) and known to become abnormal before the onset of clinically diagnosed AD. However, the discovery of new genetic risk factors in AD using CSF biomarkers is a major challenge in AD. In this study, we conducted genome-wide association (GWA) analyses separately in these subgroups and in the total sample. Methods: A total of 783 subjects comprising with 112 AD, 483 MCI, and 188 CN participants with GWA data were available from two sets of the Alzheimer’s Disease Neuroimaging Initiative (ADNI) Study. The residuals of a quantitative trait of A42/p-tau ratio after adjusting for age and sex were normalized by randomization. GWA markers using the normalized trait in each set were analyzed in a linear regression model after accounting for population substructure. The results across sets were combined using meta-analysis. Results: We observed genome-wide significant (GWS, p≤5×10−8) association with SNPs in the APOE region and three novel loci including MAST4 (rs80222306, minor allele frequency [MAF]=14%, meta-analysis p [meta-p]=2.9×10−8), and FBXO16 (rs10112794, MAF=4%, meta-p=1.1×10−8) in CN subjects, and CCR5/CORT (rs12566199, MAF=5%, meta-p=2.3×10−8) in MCI subjects. Associations with several top SNPs in CADM1 approached the GWS threshold (rs10891864, MAF=39%, meta-p=5.9×10−8) in AD cases. Rs80222306 was also strongly associated in separate analyses of CSF A42 (meta-p=8.1×10−8) and p-tau (meta-p=6.0×10−8) levels in CN subjects. None of these genome-wide or near genome-wide significant SNPs outside the APOE region were even nominally significant in other subgroups (meta-p>0.05) and less significant in the total sample (meta-p=4.7×10−8). Conclusion: Our results suggest that markers of AD pathology are influenced by multiple loci which exert their effects at different stages of clinical progression toward AD. Further studies in large independent datasets are warranted to identify the functional variants and understand the mechanisms underlying the stage-specific genetic associations with CSF biomarkers of AD.

972T

Genetic determinants of normal human facial variation. J.B. Cole1, M. Manyam2, J. Larson2, D.K. Libert2, T.M. Ferrara3, S.L. Riccardi2, J.R. Shaffer6, J.A. Lyon7, L. Zuo7, W. Mio7, S.A. Santorico2, B. Halgrimsson2, R.A. Spitz1, J.B. Cole1,2,3,4,5,6,7,1, Human Medical Genetics and Genomics Program, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO; 2) Department of Cell Biology & Anatomy, University of Calgary, Calgary, AB; 3) Department of Mathematics, Florida State University, Tallahassee, FL.

Facial shape exhibits perhaps the greatest range of variation of any normal human trait, yet at the same time represents the most recognizable human characteristic. Facial shape comprises a multiplicity of complex multifactorial traits with clear genetic components. Nevertheless, little is known about normal human facial development. To identify genetic determinants of normal human facial shape variation, we carried out a large genome-wide association study (GWAS), replication study, and meta-analysis in 6300 normal African Bantu children from Tanzania. We used advanced 3D imaging technology, a revolutionary automated landmarking method, and geometric morphometrics to capture and analyze quantitative facial shape phenotypes, followed by genome-wide analysis of over 2.5 million SNPs. Our GWAS identified a number of loci with association surpassing criterion for genome-wide significance in an African population (P<2.5E-08). These include a gene desert at chr11q23.1 associated with nasal ala length, AKAP13 associated with ptiltrum width, DRAM2 associated with upper facial height, PTCH1 associated with allometry, DHX34 associated with upper facial depth, SCHIP1 associated with centroid size, and SRC associated with nasal width. AKAP13 encodes an A-kinase anchor protein. DRAM2 encodes an inducer of autophagy. PTCH1 encodes a receptor for sonic hedgehog, and PTCH1 mutations result in holoprosencephaly. DHX34 encodes an A-kinase anchor protein with limited functional information. SRC has an essential role in bone remodeling, and Src-knockout mice have broad, abnormal craniofacial growth. PTCH1 encodes a schwannomin interacting protein with limited functional information. SNPs with respect to the loci associated with facial shape in an African population did not confirm previously-reported association of mid-nasal shape with PACS3 in European-derived whites, did not identify loci that have been associated with non-syndromic cleft lip/palate, and largely did not identify loci that are associated with facial shape in an African population. Most of the loci that were associated with facial shape in our GWAS encode proteins that either participate in basic cellular processes or for which functions remain largely unknown. Our results thus provide both novel candidate genes and novel insights into early development of the human face.

971M

Genome-wide association studies for dental caries in African American and Latino populations: Novel genes, heterogeneity, and replication. J. Colavincenzo1, J.R. Shaffer1, E. Peingold1, C. Sanchez1, T. McHenry1, F.W.B. Deleyiannis2, D.W. McNeil3, R. Crout4, J.R. Weyant1,2,3,4,5, J.B. Cole1,2,3,4,5,6,7,1, 1) Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3) Center for Craniofacial and Dental Genetics, Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 4) Department of Surgery, Plastic and Reconstructive Surgery, School of Medicine, University of Colorado, Aurora, CO; 5) Dental Practice and Rural Health, West Virginia University, Morgantown, WV; 6) Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV; 7) Department of Dental Public Health and Information Management, University of Pittsburgh, Pittsburgh, PA; 8) Clinical and Translational Sciences Institute, and Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA.

Motivation: Dental caries is the most common chronic disease worldwide, affecting a majority of children and adults. Disease burden and related co-morbidities (pain, tooth loss, missed work/school, difficulty eating, hearing, and sleeping, and emergency room visits) are concentrated in vulnerable populations, including minority groups and those living in poverty or rural areas. Though dental caries experience is highly heritable, few specific risk loci have been identified and rigorously validated. To date, multiple genome-wide association studies (GWAS) in whites have nominated plausible caries risk genes, however, no GWAS in other races or ethnic groups have been performed. Methods: We performed four GWAS scans for dental caries in child and adult African American and Guatemalan cohorts. Results: Several novel caries genes were nominated, including IGFB1 (insulin-like growth factor 1; p-value=2.7E-8), which is thought to be important for tooth development and saliva production. Additionally, loci previously implicated in GWAS of whites were evidence of replication, including LPO (lactoperoxidase), an oral bacterial enzyme. Conclusions: This study provides further support for the role of previously nominated genes, and suggests that studying cohorts with different ancestry and environment may help discover a wider range of caries genes. Grants: R01-DE014899, U01-DE018903, R01-DE016148.
973S Contribution of common polygenic variation captured by the ImmunoChip to celiac disease heritability in an independent Irish population. C. Coleman1, EM. Quinn1, AW. Ryan1, RJL. Anney2, V. Trimble3, DW. Morris2, G. Donohoe2, J. Conroy3, G. Trynka4, C. Wijmenga5, S. Ennis5, R. McManus1. 1) 1Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, St. James’s Hospital, Dublin 8, Ireland; 2) Department of Psychiatry, Trinity College Dublin, Dublin, Ireland; 3) Conway Institute, University College Dublin, Ireland; 4) Genetics Department, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA; 5) Genetics Department, University Medical Center and University of Groningen, The Netherlands.

Celiac disease (CD) is a chronic immune-mediated disease with a prevalence of ~1% in European populations. Following the large ImmunoChip study of Trynka et al. (2011) the HLA and 39 other CD susceptibility loci are known. In an independent Irish CD case-control study we examined whether reported risk alleles were similar in direction of effect and whether a weighted burden of risk alleles (or polygenic risk score) could be used to distinguish case status. The polygenic risk score (Purcell et al. 2009) has been used to examine the influence of risk alleles on mass effect disease susceptibility in several complex disorders. Following stringent quality control we analyzed 143,074 markers genotyped on Immunochip in 425 cases and 453 controls. To examine concordance in the observed direction of effect in our sample, we performed a binomial sign-test for LD independent markers identified as genome-wide significant by Trynka et al. Secondly, for LD independent markers we calculated the polygenic risk score for each individual in our study. Regression was performed for disease status adjusting for marker-count-per-score (missingness), gender and population covariates. Binomial test indicated there was significant concordance in direction of effect between studies. 83% (122/147) of genome-wide significant SNPs show effect in the same direction (P (K)>122=7.9x10 -17). When restricted to non-HLA markers 10/11 (91%) show effect in the same direction (P (K)=12)=0.0059). The polygene analysis showed that polygenic risk scores were significantly associated with coeliac case-control status across a range of p threshold values. Including the HLA markers up to 36% of the variance was explained by the polygenic score (SNPs P<0.001; P=7.86x10 -67), 12% of the variance could be explained by the non-HLA SNPs alone (SNPs P=0.001; P=9.56x10 -22). We have replicated the findings of a large CD association study in an independent Irish population. Use of the polygene risk analysis allows for highly significant findings even in a reduced sample size. Polygenic scores explained a significant proportion of the variance in coeliac disease confirming the contribution of common SNPs to CD susceptibility.

974M Genetic insights into primary biliary cirrhosis - an international collaborative meta-analysis and replication study. H.J. Cordell1, Y. Han2, Y. Li2, G.F. Meels2, G.M. Hirschfelder2, G. Xie3, B. Juran4, M.E. Gershwin2, P. Invernizzi5, K. Lazaridis6, C.A. Anderson3, M.F. Seldin3, C. Amos2, R.N. Sandford2, K. Siminovich5, Canadian-US, Italian and UK-PBC Consortia. 1) Newcastle University, UK; 2) Dartmouth College, USA; 3) University of Cambridge, UK; 4) University of Birmingham, UK; 5) Mount Sinai Hospital, Toronto, Canada; 6) Mayo Clinic, USA; 7) UC Davis, USA; 8) Humanitas, Italy; 9) Wellcome Trust Sanger Institute, UK.

Primary biliary cirrhosis (PBC) results from an interaction of genetic and environmental factors. To date, four genome-wide association studies (GWAS) and two Illumina Immunoarray studies of PBC have helped delineate the genetic architecture of this disease. These studies have confirmed associations at the human leukocyte antigen (HLA)-region and identified 27 non-HLA susceptibility loci. Candidate genes are enriched in the IL-12 signaling cascade. To identify additional risk loci for PBC, we have undertaken genome-wide meta-analysis (GWMA) of discovery datasets from the North American, the Italian and the UK GWASs of PBC, with a combined, post-PC sample size of 2745 cases and 9902 controls. Genome-wide imputation of each discovery dataset was undertaken in Mach using HapMap3 as reference panel; GWMA was undertaken using ProbABEL and META. Following meta-analysis, the index single nucleotide polymorphisms (SNPs) at selected loci with P_GWMA<2x10-5 were genotyped in a validation cohort consisting of 3716 cases and 4261 controls. To prioritize candidate variants and genes at confirmed risk loci, we used the ENCODE and the 1000Genomes datasets to identify SNPs within regulatory elements and non-synonymous (ns) SNPs in strong linkage disequilibrium (LD) with the index variant. (r2>0.8). We identified several ns risk loci for PBC. Functional annotation of these loci revealed SNPs within regulatory elements that are predicted to affect expression of DGKQ (4p16), PAM (5q14 and 11q21) (16p12), that are strongly-correlated to the index variant. Other candidate genes identified included the IL-12, the IL-12Rβ1 and the IL-12Rβ2 signalling cascade, and CCL20 (2q36), which is involved in chemo-atraction of lymphocytes and dendritic cells towards epithelia and is expressed by TH17 cells originating from Foxp3+ T cells. Pathway analysis identified several highly plausible gene sets associated with PBC, including the IL-12 and JAK-STAT signalling pathways, and implicated several other immune processes in the pathogenesis of PBC, including innate immune processes (e.g. IFN-γ signaling). Conclusion: This uniquely powered international collaborative GWMA and replication study confirms additional immunologically relevant loci that are associated with the risk of developing PBC.

975T A Genome Wide Association Study of peanut sensitisation in the Manchester Asthma and Allergy Study. J.A. Curtin1, A. Custovic1, A. Simpson1, E.N.C. Mills2. 1) Centre for Respiratory Medicine and Allergy, Institute of Inflammation and Repair, University of Manchester, Manchester, United Kingdom; 2) Institute of Inflammation and Repair, Manchester Academic Health Sciences Centre, Manchester Institute of Biotechnology, University of Manchester, United Kingdom.

Background: Very little is known about the molecular mechanisms underlying food allergy, including the model food allergen peanut. Twin studies indicate a significant genetic influence on peanut allergy. In the study estimated the heritability of peanut allergy at 81.6%. Within the setting of a prospective birth cohort study, the Manchester Asthma and Allergy Study (MAAS), we investigated if a Genome-wide association study (GWAS) could help explain some of the heritability of peanut sensitisation. Methods: DNA was genotyped using Illumina 610 quad chips. Following standard QC we imputed additional genotypes (IMPUTE version 2.2.2) with the “1000 Genomes Phase I integrated variant set” reference genotypes; we excluded SNPs with INFO <0.4 and MAF <0.01. Genome-wide association study (GWAS) was performed for children that were skin prick positive to peanut at either at age 8 or 11 (46 cases, 554 controls). Results: We identified 66 regions that were associated with peanut sensitisation at genome-wide significance (P<5x10 -8). Some of the most significant association regions included the CLDN14 (P<4x10 -22), DACH1 (P<7x10 -23) and CBFAP2T2 (P=8x10 -23) genes. CLDN14 is part of the claudin family which is an important component of tight junctions. DACH1 is a chromatin-associated protein that helps regulate gene expression during development. CBFAP2T2 forms part of a fused gene that can be important in myeloid leukemias. CBFAP2T2 is a parologue of DEAF1 which is reported to control the expression of genes encoding peripheral tissue antigens in type 1 diabetes. In drosophilia DEAF1 is reported to be a regulator of the innate immune response. Interestingly CCL5 (P=1x10 -13) was also associated with peanut sensitisation. CCL5 is a chemokine that is involved in immunoregulatory and inflammatory processes.

Conclusion: This is the first GWAS of peanut sensitisation in children and we provide additional evidence that peanut sensitisation is heritable. We have identified several loci associated with peanut sensitisation that contain plausible candidate genes. Replication of these results will provide further evidence of these associations. Further studies will also be required using cohorts of individuals with either challenge-confirmed food allergy or a clear history of severe reactions linked to peanut consumption.
A genome-wide association study identifies a LEPR gene as a novel predisposing factor for childhood FPG. M. GO, J. HWNAG, L. Heo, T. Park, B. Kim. KNIH, Oosong, South Korea.

To date, genome-wide population-based association studies have been predominantly conducted on genetic predispositions in adult individuals. In uncovering childhood FPG susceptibility, a combined meta-analysis identified a novel LEPR locus rs17407594 reaching genome-wide significance (n = 1,260, P = 4.98 × 10^-8). We also observed an association with T2D risk in the AGEN consortium (n = 18,817, P = 2.06 × 10^-2). In conclusion, our findings might expand understanding of genetic architecture contributing to glucose regulation and T2D risk.

Comprehensive curation and visualization of ethnicity information from published genome-wide association studies (GWAS): an improved GWAS Catalog. L.A. Hindorff¹, J.A.L. MacArthur², J. Morales³, E.H. Bowler⁴, P. Hall⁵, K. Kimmel⁶, H. Junkins¹, T. Burdett⁷, D. Weitner⁸, T. Manolio⁹, H. Parkinson¹. ¹ Division of Genomic Medicine, NHGRI, NIH, Bethesda, MD; ² European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; ³ Division of Policy, Communication and Education, NHGRI, NIH, Bethesda, MD. Genome-wide association studies (GWAS) have been prolific, with over 1,900 publications identifying 13,400 genetic variants associated with a broad range of human diseases and traits. The GWAS Catalog, available at http://www.genome.gov/gwastudies and http://www.ebi.ac.uk/gpft/gwas/, has served as a repository and visual summary of published GWAS findings since 2008. For each published GWAS study the Catalog includes publication, trait and SNP-trait association information, including SNP identifier, gene and risk allele, P-value, and a brief cohort description. To date the majority of GWAS studies have primarily been carried out in populations of European ancestry. This raises the question of how generalizable these trait-associations are across diverse populations. GWAS in diverse populations are now becoming more common, raising the importance of systematically curating detailed ancestry information to allow this question to be answered. Efforts to curate in greater detail the race/ancestry information from Catalog papers and to integrate this expanded information into a user-friendly interface will be described. Building upon an existing framework that relies largely on author-reported descriptors of ethnicity and ancestry, the more detailed ancestry data are based on systematic criteria and extracted as semi-structured fields. Information is curated at various levels of granularity - broad ancestral categories (e.g. “European”, “Hispanic/Latin American”), specific countries of origin and recruitment, and finer geographic descriptors where available (e.g. city and state). The public availability of these data will coincide with the release of an improved web interface that will enable visualization and searchability of ancestry data. A preview of the curation criteria and visualization of the ancestry data, with opportunity for detailed comment, will be presented and available at http://www.ebi.ac.uk/gpft/gwas/.

The incorporation of detailed ancestry information, with associated search and visualization features, will enable users to investigate the generalizability of trait-associations across diverse populations and identify those limited to specific ancestries.

Variants within ADAMTS9-AS2 influence fingerprint patterns. Y.YW. Ho¹, ², D.M. Evans³, ⁴, G.W. Montgomery³, A.K. Henders⁴, ⁵, J.P. Kemp², ⁴, ⁵, N.J. Taylor⁶, A. Pounou⁷, T. Pournac⁸, D.Z. Loesch⁹, G.D. Smith¹, ², N.G. Martin¹, S.E. Medland¹. ¹ QIMR Berghofer, Herston, QLD, Australia; ² School of Psychology, The University of Queensland, St. Lucia, QLD, Australia; ³ MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; ⁴ School of Social & Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK; ⁵ University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia; ⁶ Department of Psychology, La Trobe University, Melbourne, Australia.

Dermatoglyphics (the scientific study of dermal ridges on the distal phalanges, palms, and soles of primates) is a complex quantitative trait, which provides a model for genetic investigation. The current study sought to identify the genetic influences on fingerprint patterns across all 10 fingers, utilizing genome-wide association and meta-analysis. Data were collected from three samples, consisting of 2296 participants from the QIMR Brisbane Adolescent Twin Study, 1859 from the QIMR health and lifestyle study, and 5339 from the Avon Longitudinal Study of Parents and Children (ALSPAC). Results of meta-analyses across samples identified an effect for rs1523452, within ADAMTS9-AS2 for whorls on the little fingers (left, p = 3.43 × 10^-27; r = .124; right, p = 6.25 × 10^-15; r = .089) and ring fingers (left, p = 7.18 × 10^-13; r = .083; right, p = 3.87 × 10^-11, r = .076) of both hands. Post-hoc TATES multivariate analyses were also conducted on the QIMR and ALSPAC samples, which showed strong association at this locus (p = 2.51 × 10^-08, p = 7.42 × 10^-19 respectively). As genetic variants within ADAMTS9 have previously been associated with type 2 diabetes and waist-hip ratio, these results suggest variants influencing prenatal growth, evidenced by their effect on formation of ridge patterns, have ongoing effects on later development.

Genome-wide association and local ancestry analyses of high-altitude adaptations in Tibetans. C. Jeong¹, ², B. Basra³, ⁴, G. Childs³, S. Craig³, D. Wiltonsky¹, C. Beall², ⁴, Di Rienzo², ⁴. ¹ Department of Human Genetics, University of Chicago, Chicago, IL; ² Oxford University Clinical Research Unit, Patan Hospital, Kathmandu, Nepal; ³ Department of Anthropology, Washington University in St. Louis, St. Louis, MO, USA; ⁴ Department of Anthropology, Dartmouth College, Hanover, NH, USA; ⁵ Department of Anthropology, Case Western Reserve University, Cleveland, OH, USA.

Indigenous human populations in the Tibetan plateau show a set of physiological traits distinct from those of acclimatized lowlanders, e.g. unevolved hemoglobin concentration (Hb) up to 4,000 m altitude and extremely low arterial blood oxygen saturation level (SaO₂). Several population genomic studies of Tibetans independently identified two candidate genes, EGLN1 (endothelial PAS-domain containing protein 1), which harbor extreme allele frequency divergence and signatures of positive selection. However, the genetic basis of the distinctive Tibetan high-altitude physiology remains poorly understood. We performed a genome-wide association study of two key physiological traits, Hb and SaO₂, in a group of 880 ethnic Tibetan women born and raised at altitudes ranging from 2,982 m to 4,052 m in Nepal. Both phenotypes were controlled for known covariates. No genome-wide significant association was found applying the Bonferroni correction for the 363,954 SNPs tested. However, SNPs around EPAS1 gene were associated with Hb (linear mixed model (LMM) p ≥ 2.27 × 10^-10) and those around HIF1A are associated with SaO₂ (LMM p ≥ 2.72 × 10^-10), thus confirming a role for oxygen homeostasis systems in Tibetan adaptations. Next, we conducted a local ancestry analysis across 337 unrelated Tibetans, to find out loci with excess high-altitude ancestry represented by the Sherpa in this study. Both EPAS1 and EGLN1 show marked enrichment in high-altitude ancestry: 73.1% (+6.1 standard deviation; SD; the top signal) for EPAS1 and 60.1% (+3.3 SD; top 0.2%) for EGLN1. The genome-wide high-altitude ancestry is 45.1% with SD = 4.6%. Last, we explored if SNPs with excess high-altitude ancestry were preferentially associated with high-altitude phenotypes. High-altitude ancestry proportion shows a marginally negative correlation with Hb association p-value (Spearman’s rank correlation p = 0.054). However, without EPAS1 SNPs, this correlation becomes weaker (p = 0.100). Our study presents the first genome-wide association study for two physiological traits in Tibetans and our local ancestry analysis raises the possibility that phenotypes other than hemoglobin level and oxygen saturation are important in Tibetan physiology.
GWAS meta-analysis of primary sclerosing cholangitis identifies new disease loci and further clarifies the genetic relationship with inflammatory bowel disease. S. Ji1, B.D. Jurann1, E. Melom2, S. Mucha2, J.Z. Liu1. Analyses comparing our data to those from the largest GWAS of IBD showed 9-39 fold sibling relative risk and marked comorbidity with inflammatory bowel disease (IBD). Previous PSC genome-wide association studies (GWASs) have identified six non-HLA risk loci. Newly associated loci include those previously reported to be associated with other autoimmune disorders and genes linked with alcohol dependence. In addition we found that in both EA and Asian ancestry data. LD cluster analysis and stepwise model selection for the 145 BBJ associated loci found evidence for secondary signals in 31 loci. To identify candidate causal variants, we tested the 265 known novel loci with genomic annotations and classified them based on the level of evidence for protein-coding changes (PolyPhen2) or regulatory function (transcription factor binding motifs, DNaSe hypersensitivity, promoter, 5'UTR, 3'UTR, micro-RNA target sites). Using association strength and functional class, we extracted 1,333 causal candidate SNPs and 470 genes; only 2.5% of candidate SNPs impacted coding sequence. To identify a smaller set of candidate causal genes, we used STRING-DB 9.1 data to perform a gene network randomization analysis and step-wise gene prioritization. The final gene network highlights an interconnectedness between a number of pathways related to growth and development including Estradiol synthesis and estrogen receptor signaling, Insulin-like growth factor/Growth hormone axis, and Indian Hedgehog signaling. These findings significantly add to our current understanding of genes and pathways that impact variation in human height, and likely represent variants that are important for other common human diseases and phenotypic variation.

Genetic epidemiology of alcohol and nicotine dependences. J. Jung, H. Zhang, B. Grant. Laboratory of epidemiology and biometry, National Institute on Alcohol Abuse and Alcoholism/National Institutes of Health, Bethesda, MD. Alcohol and nicotine dependences are common psychiatric disorders that are often developed together. Both disorders are genetically influenced with heritability of 55% and 80% respectively, and co-occurrence of both lifetime dependence is influenced by a substantial genetic correlation between both disorders. We analyzed Study of Addiction: Genetics and Environment (SAGE) to perform a genome-wide association study of a comorbidity of both alcohol dependence (AD) and nicotine dependence (ND), which were diagnosed by the Diagnostic and Statistical Manual of Mental Disorders-IV. Samples in SAGE were genotyped by illumine Human 1M BeadChip Array. European American (EA) samples of 1180 cases and 1378 controls of AD and 1163 cases and 1392 controls of ND were analyzed. Correlated binary model with a bivariate odd ratio were utilized to test association of co-occurrence with SNPs after controlling for age, sex and two population stratification scores. We found that rs12439549 in Gamma-aminobutyric acid (GABA) receptor, gamma 3 (GABR3) is associated with co-occurrence of both disorders (p-value = 1.5e-05). We replicate the results by African American (AA) samples with 709 cases and 516 controls of AD and 659 cases and 692 controls of ND (p-value = 9.5e-04). The gene GABR3 is known to be linked with alcohol dependence. In addition we found that in both EA and AA groups the additional common 109 genes are associated with the comorbidity by a threshold of 10e-02. This finding is very significant to understand the etiology of the co-occurrence of both disorders.
Mapping variation in response to vitamin D in the immune system.

S. N. Kanuki, C. Jeong, J. Maranville, S. Baxter, D. Witonsky, A. D. Rienzo. Department of Human Genetics, University of Chicago, Chicago, IL, USA.

The active hormonal form of vitamin D, 1,25-dihydroxyvitamin D (1,25(OH)2D3) is a potent modulator of immune response. Inhibition of lymphocyte proliferation is one of the effects of 1,25(OH)2D3 on the immune system and varies across individuals. Common genetic variation in the vitamin D receptor (VDR) gene and inclusion of only high imputation accuracy (r2 > 0.2) SNPs remained after quality control (QC) checks, which included allele frequency comparisons between controls genotyped on different chips and inclusion of only high imputation accuracy (r2 > 0.9) SNPs, in addition to standard QC criteria. Associations with SNPs at a locus on 13p21.31 were associated with AD at genome-wide levels of significance (p < 1.8 × 10-8). This locus is in a gene desert and >1Mb from the closest genes, PCDH20 and PCDH40. SNPs at four additional loci had p-values < 1 × 10-6, including SNPs at 15q24.3 that are eQTLs for STIP1 in monocytes after treatment LPS (Fairfax BP et al. Science 2014;343:1246949) and at 6p22.33 in the THEMIS gene. STIP1 is involved in aberrant inflammation in the skin, a primary feature of AD, and THEMIS regulates regulatory T cell function, a key cell type in inflammatory responses. We also identified SNPs at 2p24.3 and 15q24.3, which are enriched for genes involved in the pathophysiology of AD.

Comparing of GWAS data for the personality in four Korean cohort.

B. Kim1, H. Kim1, H. Cho1, S. Roh2, M. Lee2, S. Yang3, N. H. Cho4, C. Shin3, J. Sung2, H. Kim2. 1) Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Korea; 2) Complex Disease and Genetic Epidemiology Branch, Department of Epidemiology and Institute of Environment and Health, School of Public Health, Seoul National University, Seoul, Korea; 3) Department of Preventive Medicine, Ajou University School, Suwon, Korea; 4) Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Korea University Hospital, Ansan, Korea.

Personality is a determinant of behavior and lifestyle associated with health and human diseases. Despite high heritability ranging from 33 to 60%, the understanding of the genetic origins of personality trait variation is extremely limited. To identify common genetic variants associated with each of the five dimensions of personality, we performed a genome-wide association study (GWA) meta-analysis of four cohort, which are the rural Ansan community (1,126 adults) and urban Ansan cohort (1,683 adults), healthy Twin Study (979 adults) and young women cohort (1,581 adults) in Korea. Personality traits were measured using the Revised NEO Personality Inventory for the five-factor model of personality, using the Korean short version of the original NEO-PI-R, a 90-item measure of the five factors of personality. Genomic DNA was extracted from whole-blood samples using a commercial isolation Kit according to the manufacturer’s protocols. We used PLINK for the quality control procedure and PLINK, BEAGLE, SHAPEIT, and IMPUTE2 for SNP imputation, due to different chip platform. After imputation, we used PLINK and GenABEL for association analysis and METAL software for meta-analysis of Ansan and Ansan cohort (1,387,466 SNPs), young women cohort (1,581,609 SNPs), and healthy Twin Study (1,387,466 SNPs). There are no SNPs reached genome-wide thresholds of 5 × 10-8 for statistical significance. In the result of this study, however, we found the association between Neuroticism and DRD1 (r2 = 0.02), Extraversion and CDKAL1 (r2 = 0.02), Openness to experience and ERBB4 (r2 = 0.02), Agreeableness and FAM110B (r2 = 0.02), and Conscientiousness and IGF2BP3 (r2 = 0.02). Among five-factor personality, we speculated that Neuroticism could associate with D1 subtype of the dopamine receptor. It will explain the relationship of personality and some behavioral responses. In addition, Openness to experience could associate with cancer and Conscientiousness could associate with insulin-like growth factor II and RNA synthesis or metabolism. But these suggestions should be replicated and confirmed by further studies. This research was supported by a grant of Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI14C0072).

1. The Center for Applied Genomics, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA; 2. Medical Scientist Training Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3. Genes, Environment and Complex Disease, Murdoch Children’s Research Institute, Parkville, Victoria, Australia; 4. Department of Pediatrics, Emory University School of Medicine and Children’s Health Care of Atlanta, Atlanta, GA, USA; 5. Division of Clinical Pharmacology, Children’s Mercy, Kansas City, MO, USA; 6. Unit of Gastroenterology, Department of Medical and Surgical Specialties, Careggi University Hospital, Florence, Italy; 7. Division of Pediatric Allergy and Immunology, University of Miller School of Medicine, Miami, FL; 8. Yorkhill Hospital for Sick Children, Glasgow, Scotland; 9. Paediatric Gastroenterology and Nutrition, Royal Hospital for Sick Children, Edinburgh and Child Life and Health, University of Edinburgh, EH1 2UW, UK; 10. Mount Sinai Hospital IBD Centre, University of Toronto, 441-600 University Avenue, Toronto, Ontario M5G 1X5, Canada; 11. Department of Immunology, Children’s Hospital Boston, Boston, MA, USA; 12. Texas Scottish Rite Hospital for Children, Dallas, TX, USA; 13. Department of Pediatrics, Pediatric IBD Center, Cedars Sinai Medical Center; 14. Department of Translational Medical Science, Section of Pediatrics, University of Naples “Federico II”, Naples, Italy; 15. Sarah M. and Charles E. Seay Center for Musculoskeletal Research, TX Scottish Rite Hospital for Children, Dallas, TX, USA; 16. Nuffield Department of Medicine, University of Oxford, UK; 17. Institute of Immunology and Department of Medicine, Mount Sinai School of Medicine, New York, NY; 18. Section of Immunology, Allergy, and Rheumatology, Department of Pediatric Medicine, Texas Children’s Hospital, Houston, TX; 19. The Hospital for Sick Children, University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; 20. Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK; 21. Department of Pediatrics, Nemours Children’s Hospital, Orlando, FL, USA; 22. Departments of Pediatrics and Human Genetics, McGill University, Montreal H3H 1P3, Quebec, Canada; 23. Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 24. Division of Gastroenterology, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA; 25. Department of Pathology and Lab Medicine, Perelman School of Medicine University of Pennsylvania, Philadelphia, PA, USA; 26. Department of Biostatistics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 27. Division of Pulmonary Medicine, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA; 28. ICRCS ‘Casa Sollievo della Sofferenza’, Division of Gastroenterology, San Giovanni Rotondo, Italy.

Autoimmune diseases affect seven to ten percent of individuals living in the Western Hemisphere, and represent a significant cause of chronic morbidity and disability. High rates of comorbidity and familial clustering suggest an underlying genetic predisposition to autoimmune disease susceptibility, and genome wide association studies (GWAS) have identified hundreds of susceptibility genes associated with autoimmune diseases with some shared across clinically-distinct disease groups. To investigate the genetic architecture of pediatric autoimmune diseases (pAIDs), we performed a heterogeneity-sensitive GWAS (hsGWAS) across 10 pAIDs in a nested case-control study including over 5,200 cases and 11,000 controls. We identified 86 independent pAID association loci reaching GWS (P < 5 × 10-8) and 61 marginally significant association loci (P < 10-6), including lead SNPs mapping to candidate genes with established immunoregulatory functions (e.g., CD40LG; P < 3.08 x 10-11 and NFATC3; P < 1.18 x 10-10). Of the 147 lead GWS (86) and marginally significant (61) loci, 97% were supported by functional (n=30), regulatory (n=55), conservational (n=30), which is enriched as compared to that observed at random across the genome (p<0.021), particularly for DNAse hypersensitivity sites (p < 0.01) as well as for genetic association annotations, eQTLs, and coding variants (p<0.001). In addition, we extensively characterized the expression profiles of the candidate genes mapped by the lead loci in human tissues and murine immune-specific cell typing, providing evidence to support a disease-specific gene expression signature across subsets of immune cell lineages. Integration of multiple in silico analytical approaches identified highly shared autoimmune signals (e.g., IL2-IL21 P < 6.24 x 10-12) and converging roles for JAK-STAT, innate, and TH1-TH2/TH17 mediated T-cell signaling across pAIDs and as molecular pathways representing attractive pharmacological targets for pAIDs.

Grimbacher, L. Padyukov, Salzer, 1988S


The new loci, in addition to the previously identified risk loci, will help improve risk profiling for glaucoma with better opportunities for management of high-risk individuals.
INTRODUCTION Asthma affects 300 million people worldwide and asthma exacerbations are a major cause of morbidity and medical cost. The objective of this study was to identify genomic predictors of exacerbations in subjects with asthma. METHODS In the Childhood Asthma Management Program (CAMP) and Childhood Asthma Research and Education (CARE) trials, acute asthma exacerbation was defined by treatment with a five-day prednisone course. We first performed a GWAS in CAMP, and to verify that the top hits from the GWAS were not artifacts of a non-normally distributed phenotype, we further performed permutation testing on the top 50 GWAS SNPs. We combined p-values from the CAMP permutation tests with the GWAS results from 205 participants of CARE. From BioVOU, the DNA biobank for Vanderbilt University Medical Center, we identified a cohort of 786 subjects who had asthma, with 191 experiencing an asthma-related exacerbation, defined as a hospitalization or emergency department visit. We used CD4+ lymphocyte genome-wide mRNA expression profiling from the ABRIDGE project to identify associations of top SNPs with mRNA abundance of nearby genes. RESULTS In CAMP and CARE we identified a locus in CTNN3A3 reaching genome-wide significance (rs7895190, p = 4.44E-09); we also prioritized other top hits for replication (rs746578 in TTN, p = 2.15E-06; rs993512 in SEMA3D, p = 7.63E-04). In replication in BioVOU, rs993512 was significant and associated with exacerbations (p = 0.06). Results from the GWAS in strong LD with a variant, rs10997296, identified an eQTL for CTNN3A3 in CD4+ cells from the ABRIDGE cohort (p = 0.00079). DISCUSSION The CTNN3A3 SNP, rs7915690, is intronic. It is in strong LD (r² > 0.97) with a predicted regulatory variant, rs10997296, that is a strong enhancer in human mammary epithelial cells (ENCODE); and in a DNase I hypersensitivity site in small airway epithelial cells (ENCE). This variant was identified as an eQTL for CTNN3A3. The SEMA3D SNP, rs993512, is intronic and in strong LD (r² = 0.99) with a regulatory SNP, rs55834466. The ENCODE project identified rs55834466 as a strong enhancer in epidermal keratinocytes and human mammary epithelial cells, in addition to SEMA3D being relatively overexpressed in the lung (illumina Human BodyMap). CONCLUSIONS We identified two exacerbation-associated regulatory SNPs from GWAS of asthma clinical trials. One met genome-wide significance thresholds while the other replicated in a clinical Biobank database.
Further replication studies are required to confirm our novel findings. Loci that may be missed due to lower LD in African ancestry populations. Analyses to fine-map loci and identify causal genes and variants are ongoing.

Encodes the platelet isoform of phosphofructokinase that is a key regulator in glycolysis. Follow-up analyses to fine-map loci and identify causal genes and variants are ongoing. Our results suggest that a high-density reference panel including low-frequency variants improves the coverage of the genome and unmasks SNPs in other African ancestry populations. PPKF encodes the platelet isoform of phosphofructokinase that is a key regulator in glycolysis. Follow-up analyses to fine-map loci and identify causal genes and variants are ongoing. Longitudinal data includes a yearly single-item survey question over 20 years, and a more detailed follow up study at year 17. A pilot GWAS in this population at year 10 showed association of ED with two SNVs near ALCAM (activated leukocyte cell adhesion molecule). Over the 20 years, 445 individuals had at least one report of ED, and 290 had none. In this population, ED is significantly associated with age, BMI, and HBA1C (all p < 0.01, t-test). We plan to use genetic data and new phenotype information from these men to replicate the association of variants that have previously been found to be significant or sub-significant. We will perform a cross-sectional case-control analysis, as well as a time-to-event analysis using a Cox proportional hazards model. The models will stratify for BMI, age, HBA1C, and other factors that associate with ED status, and use PCA to correct for ancestry. Identification of variants associated with ED risk will be useful for prediction and early treatment in diabetic men.

Common genetic variation explains a substantial fraction of nicotine and cotinine glucuronidation in multiple populations. Y.M. Patel1, D.O. Stram1, L. Le Marchand3, S.S. Hecht2, C.A. Haiman1, S.E. Murphy2. 1) Department of Preventive Medicine and Norris Comprehensive Cancer Center, Keck School of Medicine University of Southern California Los Angeles, CA 90032; 2) University of Minnesota VFW Cancer Research Center 406 Harvard Street SE Minneapolis, MN 55455; 3) University of Hawaii Cancer Center 701 Iloa St #600, Honolulu, HI 96813. Evaluating the role of genetics in the metabolism of nicotine may be important in understanding racial/ethnic differences in lung cancer risk among smokers. Noted racial disparities in risk may be related to differences in internal carcinogen dose resulting from faster metabolism including the glucuronidation of nicotine and cotinine. We conducted a genome-wide association study (GWAS) in search of common genetic variants that may be predictive of nicotine and cotinine glucuronidation in a multiethnic sample of 2,239 current smokers representing 5 ethnic populations. The phenotypes analyzed were cotinine and nicotine glucuronidation. Cotinine glucuronidation is a ratio derived from the difference of total and free cotinine, divided by total cotinine. Nicotine glucuronidation is similarly created. In an analysis of 11,892,802 genotyped and imputed variants appearing in the thousand genomes project, 1,241 were found to be strongly associated with cotinine glucuronidation at p<5×10-8. The vast majority were within the region 4q13, near the previously reported gene UGT2B10. Fifteen different SNPs in 9 different regions were globally significant and contributed independent information. The trait was significantly associated with cigarette smoking (p=5×10-10). Together, these 15 SNPs explain 33.2% of variation observed in cotinine glucuronidation which ranged from 55% for African Americans to 19% among Asian Americans. The strongest single SNP association was rs115765562 (p=1.60×10-15), near UGT2B10, which is highly correlated with a known splice site variant, rs1116294140, in UGT2B10. This splice variant in combination with an Arg667Trp nonsynonymous SNP rs61750900 explain 24.3% of variation in cotinine glucuronidation, indicating effects seen through putative functional missense and splice site variants might be the underlying explanation for the associations. Our analysis has determined that a high fraction of individual variation in cotinine glucuronidation is explained by genetic differences. We also noted that in our data African Americans and Japanese Americans have the lowest cotinine glucuronidation levels and they also tend to smoke fewer cigarettes per day and have lower nicotine equivalents than do Europeans. However, their lung cancer risks (higher in African Americans and lower in Japanese Americans) are quite different. This observation therefore does not strongly support an important role of glucuronidation in determining the factors underlying these risk differences.
996T

Search for new risk gene for Stevens-Johnson Syndrome independent of HLA risk allele. H. Sawai, M. Ueta, Y. Hitomi, C. Sotozono, S. Kinoshita, K. Tokunaga. 1) Department of Human Genetics, The University of Tokyo, Tokyo, Japan; 2) Department of Ophthalmology, Kyoto Prefectural University of Medicine, Kyoto, Japan; 3) Research Center for Inflammation and Regenerative Medicine, Faculty of Life and Medical Sciences, Doshisha University, Kyotanabe, Japan.

Stevens-Johnson syndrome (SJS) and its severe form, toxic epidermal necrolysis (TEN), are acute inflammatory vesiculobullous reactions of the skin and mucous membranes including the ocular surface, oral cavity, and genitalia. The skin and mucous membranes are very rashes, desquamation, and coagulative necrosis of drug-induced agents. We previously revealed the association between a specific HLA-A allele (HLA-A:02:06) and SJS/TEN with severe ocular surface complications (SOC) in Japanese population. We also performed genome-wide association study (GWAS) using Japanese SJS/TEN with SOC cases and healthy controls, and found that the HLA-A region showed the strongest association with susceptibility to SJS/TEN with SOC (odds ratio [OR] = 4.40, p = 3.5 x 10^-14). To identify host genetic factor(s) contributing to dependence of HLA-A:02:06 to SJS/TEN with SOC, we conducted a GWAS with HLA-A:02:06 negative individuals.

We genotyped a total of 820 samples including 118 Japanese SJS/TEN patients and 702 Japanese healthy controls using the Affymetrix Axiom Genome wide AS1 1 Array. We excluded one case sample with an overall call rate less than 97%, and recalled the remaining 819 samples by using Genotype Console v4.1.4 software. All samples used for GWAS passed a heterozygosity check, and five related samples identified by descentant testing were excluded. A principal component analysis found six outliers to be excluded. The GWAS test of all 117 cases and 691 controls formed a single cluster with the HapMap JPT samples but not with CHB samples. We further excluded samples with HLA-A:02:06 and finally used 59 cases and 597 controls for the GWAS. The chi-square test was applied to an allele frequency model. Four loci showed genome-wide marginal associations with SJS/TEN (OR = 5.75, p = 1.19 x 10^-7 for IKZF1; OR = 2.80, p = 2.95 x 10^-7 for Gene A; OR = 2.71, p = 2.07 x 10^-7 for Gene B; OR = 5.34, p = 5.20 x 10^-7 for Gene C). One of four loci (IKZF1) showed a significant association with SJS/TEN with SOC in the previous GWAS using all the patients and controls followed by a replication study using Korean samples (Ueta et al.).

997S

Investigation of genetic variation underlying central obesity among Indian Asians. W. Scott, W. Zhang, M. Loh, S.-T. Tan, B. Lehne, U. Atzil, P. Elliott, J. Chambers, J. Koone, 1) Epidemiology and Biostatistics, Imperial College London, Norfolk Place, London, UK; 2) Imperial College Healthcare NHS Trust, DuCane Road, London, UK; 3) MRC-PHE Centre for Environment and Health, Imperial College London, Norfolk Place, London, UK; 4) Ealing Hospital NHS Trust, Southall, Middlesex, UK; 5) National Heart and Lung Institute, Imperial College London, Hammersmith Hospital, DuCane Road, London, UK.

Indian Asians are 1/4 of the world’s population and have increased susceptiblity to central obesity and related metabolic disturbances. Previous studies investigating the genetic basis of central obesity have been carried out in predominantly European populations. Few studies have investigated the contribution of genetic variation to increased susceptibility to central obesity in Indian Asians. We studied Indian Asian and European men and women, aged 35-75 years, participating in the London Life Sciences Prospective Population (LOLIPPOP) study. Waist hip ratio (WHR) was higher in Indian Asians compared to Europeans (0.94 [95% C.I. 0.86-1.02] v. 0.91 [95% C.I. 0.83-0.99] P=0.0001). To investigate genetic variation underlying central obesity among Indian Asians we examined: i) known genetic variants, comprising the 46 SNPs associated with body mass index and WHR in European genome-wide association studies (n=10,318); ii) common genetic variants through genome-wide association (n=10,318); iii) exonic variants through illumina HumanExome BeadChip (n=9,637). Our sample size provides 80% power to detect SNPs explaining >0.3% of trait variance at P<5x10^-8.

We observed no systematic differences in risk allele frequencies or effect sizes of the 46 known adiposity SNPs in Indian Asians compared to European adults (Sign test P=0.185, P=0.145 respectively). Combined these variants explained 1% of population variation in WHR among Indian Asians. Next we carried out a genome-wide association study of WHR in Indian Asians. Variants from 1000 Genomes haplotypes were imputed into available genome-wide SNP data. >6M SNPs (MAF >5%) passed quality control. No new variant was associated with WHR at P<5x10^-8. We then evaluated exonic variants (247,870 SNPs; 97.4% call rate) for association with WHR in Indian Asians. Again we found no novel associations with WHR (P<2x10^-7; Bonferroni corrected for the number of independent tests). These findings argue against a role for cosmopolitan variants underlying central obesity in Indian Asians. Future efforts aimed at identifying the genetic basis of central obesity in Indian Asians should be focused on other mechanisms, including population-specific variants and transregenerational epigenetic modifications.

998M

Common variants at c11orf30 and CAPN14 are associated with eosinophilic esophagitis. P.M.A. Slieman, A. K. Mostaghel, J. Spergel, S. Acevedo, N. Gonsalves, K. Nadeau, G.T. Furuta, J. Speigel, H. Nakanoarson, 1) The Center for Applied Genomics, The Children’s Hospital of Philadelphia, PA, USA; 2) Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania Philadelphia, PA, USA; 3) Division of GI, Hepatology, and Nutrition, The Children’s Hospital of Philadelphia, PA, USA; 4) Division of Allergy and Immunology, The Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 5) Division of Allergy, Immunology, 9500 Gilman Drive MC-0760, Department of Pediatrics and Medicine, University of California, San Diego and Rady Children’s Hospital, San Diego, CA, USA; 6) Division of Gastroenterology & Hepatology, Northwestern University - The Feinberg School of Medicine, Chicago, IL, USA; 7) Stanford University School of Medicine, Lucile Packard Children’s Hospital, Stanford Hospital and Clinics, Division of Allergy, Immunology, and Rheumatology, CA, USA; 8) Digestive Health Institute, Section of Pediatric Gastroenterology, Hepatology and Nutrition, Children’s Hospital Colorado, Gastrointestinal Eosinophilic Diseases Program, Department of Pediatrics, Mucosal Inflammation Program, University of Colorado.

Eosinophilic esophagitis (EoE) is an inflammatory disorder of the esophagus histologically characterized by accumulation of eosinophils in the esophageal epithelium. Clinical symptoms of EoE include dysphagia, failure to thrive, vomiting and epigastric or chest pain. A diagnosis of EoE is made following endoscopy and biopsy upon finding isolated eosinophils in the esophagus having ruled out gastroesophageal reflux. The stringent diagnostic criteria for EoE, that include biopsy proven eosinophilic infiltration of the esophagus, result in a phenotypically homogenous case series that is well powered for GWAS. To investigate the genetic basis of EoE, we expanded GWAS totaling 248 cases and 1541 controls in an imputed dataset powered for GWAS and a potentially powerful model to study the genetics of EoE. To investigate genetic variation underlying central obesity and related metabolic disturbances, we conducted a GWAS with HLA-A:02:06 negative individuals.

We genotyped a total of 820 samples including 118 Japanese SJS/TEN patients and 702 Japanese healthy controls using the Affymetrix Axiom Genome wide AS1 1 Array. We excluded one case sample with an overall call rate less than 97%, and recalled the remaining 819 samples by using Genotype Console v4.1.4 software. All samples used for GWAS passed a heterozygosity check, and five related samples identified by descendent testing were excluded. A principal component analysis found six outliers to be excluded. The GWAS test of all 117 cases and 691 controls formed a single cluster with the HapMap JPT samples but not with CHB samples. We further excluded samples with HLA-A:02:06 and finally used 59 cases and 597 controls for the GWAS. The chi-square test was applied to an allele frequency model. Four loci showed genome-wide marginal associations with SJS/TEN (OR = 5.75, p = 1.19 x 10^-7 for IKZF1; OR = 2.80, p = 2.95 x 10^-7 for Gene A; OR = 2.71, p = 2.07 x 10^-7 for Gene B; OR = 5.34, p = 5.20 x 10^-7 for Gene C). One of four loci (IKZF1) showed a significant association with SJS/TEN with SOC in the previous GWAS using all the patients and controls followed by a replication study using Korean samples (Ueta et al.).

We observed no systematic differences in risk allele frequencies or effect sizes of the 46 known adiposity SNPs in Indian Asians compared to European adults (Sign test P=0.185, P=0.145 respectively). Combined these variants explained 1% of population variation in WHR among Indian Asians. Next we carried out a genome-wide association study of WHR in Indian Asians. Variants from 1000 Genomes haplotypes were imputed into available genome-wide SNP data. >6M SNPs (MAF >5%) passed quality control. No new variant was associated with WHR at P<5x10^-8. We then evaluated exonic variants (247,870 SNPs; 97.4% call rate) for association with WHR in Indian Asians. Again we found no novel associations with WHR (P<2x10^-7; Bonferroni corrected for the number of independent tests). These findings argue against a role for cosmopolitan variants underlying central obesity in Indian Asians. Future efforts aimed at identifying the genetic basis of central obesity in Indian Asians should be focused on other mechanisms, including population-specific variants and transregenerational epigenetic modifications.
1000T

Genome-wide association study imputed to 1000 genomes identifies novel loci associated with lung function. J. S. Seo, J. Hoffman, I. Sayers, D. Strachan, I. P. Hall, M. D. Tobin. UK BILEVE consortium, SpiroMeta consortium. 1) Departments of Health Sciences and Genetics, Adrian Building, University of Leicester, Leicester, UK; 2) National Institute for Health Research (NIHR) Leicester Respiratory Biomedical Research Unit, Glenfield Hospital, Leicester; 3) MRC Human Genetics, MRC IGMM, University of Edinburgh, Edinburgh, Scotland, UK; 4) Division of Respiratory Medicine, University Hospital of Nottingham, Nottingham, UK; 5) Division of Population Health Sciences and Education, St George’s University of London, London, UK; 6) Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea; 7) Department of Biochemistry, Ewha Womans University, School of Medicine, Seoul, South Korea; 8) Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA; 9) Department of Environment and Health, School of Public Health, National University of Science and Technology, Hanoi, Vietnam; 10) Department of Internal Medicine, University Hospital of Coimbra, Coimbra, Portugal; 11) Hospital Infante D. Pedro, Aveiro, Portugal; 12) Hospital de Santa Maria, Lisboa, Portugal; 13) Instituto Português de Saúde e da Transplantação, Centro Regional de Sangue de Lisboa, Lisboa, Portugal; 14) Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; 15) Hospital de Vila Nova de Gaia, Vila Nova de Gaia, Portugal; 16) Universidade de Évora, Évora, Portugal; 17) Hospital de Santo António (Centro Hospitalar do Porto), Porto, Portugal; 18) Centro de Pneumologia, Faculdade de Medicina da Universidade do Porto, Porto, Portugal; 19) Hospital da Luz, Lisboa, Portugal; 20) Instituto Português do Sangue e da Transplantação, Centro Regional de Sangue de Lisboa, Lisboa, Portugal; 21) Hospital de Santa Maria, Lisboa, Portugal.

Primary spontaneous pneumothorax (PSP) is characterised by the presence of air in the pleural cavity that occurs without preceding trauma or known cause in individuals with no lung disease. Despite elevated incidence and recurrence rates, little is known about its aetiology. So far, the genetics of PSP remain largely unresolved and virtually no research has been dedicated to the identification of genetic factors for risk or recurrence of sporadic PSP.

To identify genetic variants contributing to sporadic PSP risk, we conducted the first PSP genome-wide association study (GWAS). Two replicate pools of cases and controls were generated for the PSP case controls with the same age and sex-matched controls for in silico genotyping. Two single nucleotide polymorphisms (SNPs) were defined by the replication study and were validated by individual genotyping (P < 5.00 × 10^-8) and a genome-wide association study. These findings are in accordance with previous reports and studies are warranted to uncover the function of these polymorphisms.
A genome-wide meta-analysis of hyper- and hypothyroidism and thyroid function. A. Teumer, A. Medicci, H. Völzke, R. Peeters, the CHARGE thyroid function working group. 1) Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany; 2) Department of Internal Medicine, Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands.

Thyroid hormones play a key role in cellular growth, development and metabolism. Overt thyroid dysfunction leads to common endocrine disorders affecting approximately 10% of individuals over their life span. Milder changes in thyroid function are associated with weight changes, coronary heart disease, and osteoporosis, among many other disorders. Circulating concentrations of thyroid hormones have a strong heritable component. Using genome-wide association studies (GWAS) in predominantly population-based cohorts, we and others have identified 26 genetic loci associated with serum TSH and free T4 levels. However, the variance explained by these loci is still low, i.e. 5.64% and 2.30% for TSH and free T4, respectively. To find novel loci associated with thyroid function, we have increased our sample size substantially to conduct a new GWAS in more than 33,000 individuals. In contrast to former studies, these analyses will be performed by imputing against a recent 1000 genomes panel, which gives us the opportunity to fine-map known loci and reveal additional loci due to better imputation and tagging of causal SNPs. Sex-stratified analyses of TSH and free T4 will be performed in individuals with TSH values within the reference range, allowing the analysis of interactions by sex. Additionally, we are performing a GWAS for hypo- and hyperthyroidism which will address genetic influences for more overt forms of thyroid disorders. Preliminary results in up to 1800 cases showed that 11 of the 26 SNPs known to be associated with serum TSH and free T4 levels were also associated with hypothyroidism and hyperthyroidism. Two loci near PDE8B and PDE10A, are also associated with hypothyroidism (p < 4E-4) or hyperthyroidism (p < 4E-5). In this 1000 genomes based GWAS, we expect to reveal additional loci influencing thyroid function as well as hypo- and hyperthyroidism.

A genome-wide Meta-analysis on Myopia and Refractive Error in CREAM and 23andMe. V.J.M. Verhoeven, R. Wojciechowski, P.G. Hysi, F. Klein, N. Eriksson, C.J.C. Hammond, N.A. Frikha, C.C.W. Klaver, 1,2 Consortium for Refractive Error and Myopia (CREAM), 1) Department of Ophthalmology, Erasmus Medical Center, Rotterdam, the Netherlands; 2) Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands; 3) Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore; 4) Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, Netherlands; 5) Department of Twin Research and Genetic Epidemiology, King’s College London School of Medicine, London, United Kingdom; 6) Institute for Community Medicine, Erasmus Medical Center Rotterdam, Rotterdam, the Netherlands; 7) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States; 8) Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore; 9) 23andMe, Mountain View, California, United States; 8) Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore; 9) 23andMe, Mountain View, California, United States.

Myopia is widely recognized as a multifactorial, complex genetic disorder. Recently, multiple loci for refractive phenotypes were identified separately by the Consortium for Refractive Error and Myopia (CREAM) and investigators from 23andMe, Inc. We aimed to identify additional genetic loci that explain the genetic architecture of refractive error using higher power and denser imputation by meta-analyzing data from CREAM and 23andMe. Methods: We conducted a genome-wide association study (GWAS) meta-analysis of refractive error including 63,697 individuals (49,808 Caucasians; 13,899 Asians) from the CREAM and 23andMe dataset. We applied a linear regression model to age-at-onset of myopia for 104,294 individuals from the 23andMe dataset. We conducted a meta-analysis of >35,000 individuals (the GABRIEL and EVE asthma consortia), have identified more than ten highly-reproducible asthma susceptibility loci. We can now start to look beyond single-marker main effects on asthma to epistasis and the role of functional regulatory haplotypes, namely the joint effects of pairs of regulatory variants in regulatory non-coding sequences (SNPs) on disease susceptibility. Such associations may give further insight into potential underlying mechanisms at asthma GWAS loci.

We have performed a pilot genome-wide haplotype association analysis of asthma susceptibility in 359 cases and 359 controls from the Childhood Asthma Management Program (CAMP) along with 846 healthy Illumina iControlDB controls. We constrained our analyses to a set of 526 genes with significant evidence within 50 kb windows of expression quantitative trait loci (eQTL) for each gene. We performed an H - degrees of freedom, degree-of-freedom test for each of the SNP pairs (where H = 4 or 8 degrees of freedom). Among 2,583 omnibus tests, we identified 149 associations significant at nominal P < 0.05, but no associations significant at FDR < 5%. A visual inspection of the distribution of omnibus associations suggested a moderate departure from the expected distribution, though there was no substantial evidence for non-normal distribution. Using tests with P < 0.005 (FDR < 0.005) and 71 new genome-wide significant associations with refractive error in myopia. These SNPs explain 12% of the variance of all common SNPs for SE. The results confirm over-representation of known pathways, such as extracellular matrix, ion channel activity, and glutamate signaling but also suggested potentially new mechanisms, including signaling of calcium, VEGF and TGF-β, mitochondrial function, and apoptosis. Conclusions: This study is the largest meta-analysis on refractive error known to date. This large catalogue of genetic variants opens up new insights in myopigenesis.
1007M
Genetic analysis of central serous chorioretinopathy and polypoidal choroidal vasculopathy. J. Ahn1,2, N.K. Ryoo1,3, S.J. Woo1,2, H.S. Cheong1, K.H. Park1,2
1) Department of Ophthalmology, Seoul National University College of Medicine, Seoul, Korea; 2) Department of Ophthalmology, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Korea; 3) Department of Ophthalmology, Seoul National University Bundang Hospital, Seongnam, Korea; 4) Department of Genetic Epidemiology, SNP Genetics Incorporation, Seoul, Korea.

Central serous chorioretinopathy (CSCR) and polypoidal choroidal vasculopathy (PCV) are both common retinal disorders affecting the macula. Although considered a single age-related macular degeneration (AMD), CSC and PCV share many clinical characteristics such as choroidal vascular hyperpermeability, choroidal thickening demonstrated on enhanced-depth imaging optical coherence tomography and inner choroidal vasculature abnormalities seen on indocyanine angiography. In this study, we performed genetic analysis of CSC and PCV using genome-wide association study (GWAS) and exome chip data to compare the genetic makeup of the two disease entities. This was a case-control analysis consisting of 168 PCV, 167 CSC patients and 543 control subjects. A subgroup analysis of 50 CSC patients diagnosed as the “chronic” phenotype, with persistent or progressive visual symptoms for more than 6 months, was performed in addition. PCV and control subjects were genotyped using Illumina OmniExpress and CSC patients were genotyped with the HumanExome beadchips. Integrated analysis of GWAS and exome array data was done and a total of 12,777 SNPs were included. The well-known ARMS2 rs10490924 SNP showed significant association with PCV compared to the control (P=1.78E-13). A novel SNP rs2455512 had found to be associated with CSC and PCV patients compared to controls (P=2.13E-11 and 6.92E-11, respectively). There were 3 SNPs (rs8098316, rs1363098, rs1214752) that showed borderline significant association with all CSC patients compared to control (P=1.85E-05, 9.40E-05, and 8.74E-05, respectively). In the analysis of chronic CSC, 1 SNP showed borderline significance to control group, 4 SNPs (rs12843815, rs8643141, rs11158685, rs4800452) reached borderline significance (P=1.46E-05, 7.18E-05, 7.19E-05, and 8.89E-05). This is the first study to perform comparative genetic analysis of CSC and PCV and we found differences in genetic makeup for each disease entity compared to the control group. The novel SNPs that were found to be associated with CSC require further replication studies.

1008T
Known Age-Related Macular Degeneration Risk Variants Are Not Associated with Rapid Disease Progression or Good Treatment Response. M.D. Courtenay1, W. Cade1, G. Wang1, S.G. Schwartz2, J.L. Kovach3, A. Agarwal4, M.A. Bramley5, W.K. Scott1, J.L. Haines6, M.A. Pericak-Vance7
1) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 2) Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL; 3) Ophthalmology and Visual Sciences, Vanderbilt University School of Medicine, Nashville, TN; 4) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH.

Age-related macular degeneration (AMD) is the leading cause of irreversible vision loss among older adults in developed countries. Genome-wide association studies (GWAS) have implicated common variations in 19 genes as AMD risk factors. Given the strong genetic influence on the risk of developing AMD, and the substantial variation in progression to advanced AMD and in treatment response, we investigated whether these clinicopathological endpoints are also influenced by the known genetic variant. In an existing dataset, 16 of the 19 GWAS loci were used to calculate a genetic risk score (summed number of risk alleles weighted by effect sizes) for each individual as outlined in Fritsche et al., 2013. To maximize our power for discovery, we tested two clinically extreme case/control datasets with the following characteristics: 1) Rapid clinical progression of 36 individuals with early AMD or geographic atrophy to neovascular AMD within one year vs. 31 early AMD patients who did not progress over five or more years, and 2) 20 patients with bilateral neovascular AMD who were treated with intravitreal injections of anti-vascular endothelial growth factor agents for one year and demonstrated bilateral good treatment response (three lines of vision gain on the Snellen vision chart, prolonged or persistent [at least 6 months] absence of subretinal fluid [SRF] and absence of intraretinal fluid [IRF] on OCT, presence of unchanged small intraretinal cysts with no recurrence of SRF) vs. 22 neovascular AMD patients who demonstrated a poor response in one or both eyes (three lines of vision loss, persistent SRF and IRF, formation of disciform scar). The genetic risk score was tested for association with rapid AMD progression or good treatment response with logistic regression adjusting for age and sex. Risk score was not associated in either dataset in our study (progression P=0.43 and treatment P=0.63). Therefore, the genetic effects of the 19 known AMD risk loci are not associated with rapid progression or good treatment response in this small dataset, but perhaps a larger sample size would have sufficient power to demonstrate an effect. Nevertheless, unknown genetic factors may still be regulating variability in progression and treatment response. Whole exome sequencing is currently underway in this dataset to test if novel rare or common variants are associated with these clinical endpoints.
Identification of novel and rare coding variants associated with free fatty acids and serum fatty acid profile. X. Sim, R.P. Welch, A.U. Jackson, H.M. Stringham, T.M. Teslovich, P.S. Chines, N. Narisu, C. Fuchsberger, J.R. Huyghe, A.E. Locke, M. Cannon, M. Ala-Korpela, M. Boehnke, K.L. Mohlke, M. Laakso, 1) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 2) Genome Technology Branch, National Human Genome Research Institute, Bethesda, Maryland, USA; 3) Center for Biomedicine, European Academy of Bozen/Bolzano (EURAC), Bolzano, Italy; 4) Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA; 5) Computational Medicine Research Group, Institute of Health Sciences, University of Oulu, Finland; 6) Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland.

Fatty acid (FA) levels play a critical role in insulin sensitivity, and are often elevated in obese or individuals with type 2 diabetes. Previous GWAS studies have implicated common genetic variants with FAs, however, genetic contributions at coding and low frequency variants remain largely unknown. We assayed exonic variation in 6,373 non-diabetic Finnish males from population-based METSIM study with Illumina HumanExome array (N=242,000) and Illumina OmniExpress genotypes imputed to G012D reference panel (N=-19.3M). Free FA (FFA) was measured on enzymatic colorimetric method while FA profile (including n-3, n-6, n-7, n-9, mono/poly-unsaturated, phosphatidylcholines, and sphingomyelins) were measured using proton nuclear magnetic resonance spectroscopy. Inverse normalised residuals were generated after adjusting for age, BMI, and smoking as a primary model. We identified genome-wide associations at 19 unique loci with multiple FAs. An intergenic variant (MAF=17%, P=2x10^{-10}) was associated with FFA, near PTPFRB, a glucagon-like regulatory subunit regulating liver glycogen metabolism. The remaining 18 loci included previously identified FAs (e.g. FADS1-2-3, CPT1A), lipids and glycemic loci. As FAs are correlated with lipids and glycemic traits, we repeated association analyses on an adjusted model, conditioning these traits from FA levels. We identified 2 additional loci with the adjusted model only: (i) common intronic variant rs4985154 at PDXDC1 associated with polyunsaturated FAs other than linoleic acid (MAF=34.4%, P=5x10^{-4}), and (ii) rare missense variant rs199717050 in CPT2A. We also observed associations with phosphatidylcholines and other cholines (PC) and total phosphoglycerides (TotPG) (MAF=46%, P=7x10^{-10}, P_{TotPG}=3x10^{-5}). We also note that associations at 9 loci were substantially diminished after conditioning on lipids and glycemic traits, suggesting that they were mediated by the major classes of lipproteins and insulin measures. Finally, we aggregated rare (<1%) protein-altering variants and carried out gene-level association analysis using SKAT-O. We identified gene-level associations at LIPG with PC (P=2x10^{-4}) and TotPG (P=5x10^{-5}), combining evidence across 3 missense variants and a splice acceptor on the primary model. We demonstrate that rare coding genetic variants influence FAs, and could be informative about biological pathways in lipid metabolism and insulin sensitivity.

The PhenX (consensus measures for Phenotypes and eXposures) Toolkit (https://www.phenxtoolkit.org/) is a well-established, Web-based catalog of measures for collaborative biomedical research. PhenX Working Groups (WGs) consider measures that are suitable for a variety of study designs, including longitudinal and clinical studies as well as genomic studies. Expert Review Panels are being convened to ensure that the current Toolkit measures are responsive to the evolving needs of the biomedical community. PhenX phase 1 focused on measures of phenotypes and exposures relevant to genomic studies of common complex diseases and released measures for 21 research domains (including Demographics, Nutrition, Diabetes, and Environmental Exposures) in the PhenX Toolkit. PhenX phase 1 also added depth to the Toolkit in the area of Substance Abuse and Addiction by adding one Core collection and six Specialty collections. PhenX phase 2 expands the scope of the Toolkit to address four new domains and also adds depth in the areas of Mental Health Research and Tobacco Regulatory Research. The first new domain to be addressed is Rare Conditions; preliminary measures and Toolkit annotations will be presented. "PhenX Measures for Mental Health Research," funded by the National Institute of Mental Health, aims to add several Specialty collections (Suicide, PTSD, and Eating Disorders) and one Core collection to the Toolkit. "PhenX Measures for Tobacco Regulatory Research," funded by the Tobacco Regulatory Science Program, aims to add several Specialty collections (Suicide, PTSD, and Eating Disorders) and one Core collection to the Toolkit. Preliminary measures for the Suicide and Social/Cognitive WGs will be presented. To support investigators who want to collect data via the Web, Web-based versions of PhenX protocols will be made available as Research Electronic Data Capture (REDCap) modules. Study variables in the database of Genotypes and Phenotypes (dbGaP) are being classified as identical, comparable, or related to PhenX variables. The dbGaP advanced search tool uses these mappings to help investigators discover variables and studies that are comparable, or related to PhenX variables. Funding provided by U01 HG007050, 3U01 HG004597-02S1, and studies.

1013M Analysis of Interleukin 10 haplotypes with soluble IL-10 levels and autoantibody production in Mexican patients with primary Sjögren’s syndrome. M. Vázquez-Villamar1, CA. Pailaas-Sánchez1,2, JF. Muñoz-Velarde1, Y. Valle1, D. Salazar-Camarena1, BA. Treviño-Talavera1, G. Orozco-Barocio1, E. Oregon-Romero1. 1) Instituto de Investigación en Ciencias Biomédicas (IICB), Universidad de Guadalajara, Guadalajara, Mexico; 2) Servicio de Reumatología, Hospital General de Occidente, Zapopan, Mexico.

Background: Primary Sjögren’s syndrome (pSS) is a systemic autoimmune disease characterized by lymphocytic infiltration of salivary and lachrymal glands, leading to keratoconjunctivitis sicca (dry eye) and xerostomia (dry mouth) as well as sicca symptoms (dryness). Infiltrating monocytes produce interleukin 10 (IL-10) that is capable of promoting the activation and differentiation of B cell as well as autoantibodies production, such as anti-Ro and anti-La which have an important role in the pathogenesis of pSS. Three polymorphic sites in positions -1082(A/G), -819(C/T), and -592(G/A) in the IL10 promoter can form haplotypes that may influence the IL-10 production. Objective: To determine the frequency of IL10 haplotypes and serum levels of IL-10 (sIL10), anti-Ro, and anti-La antibodies in patients with pSS. Methods: The study included 111 patients with pSS, from the Department of Rheumatology (HGO, SSJ, Zapopan, Jalisco, Mexico). As a control group, 260 subjects were included. IL10 polymorphisms were genotyped by PCR/RFLP technique. The haplotype frequencies were inferred by EmapFre software. Levels of sIL-10, anti-Ro, and anti-La autoantibodies were performed by ELISA. Rheumatoid factor (RF) levels were determined by nephelometry. Statistical analysis was performed with Stata 9.0 and GraphPad Prims 6 software. Results: The polymorphisms were in Hardy-Weinberg equilibrium and showed a high linkage disequilibrium (100%, pc=3673 16). The most frequent haplotypes were: ACC (44%), ATA (23%) and GTA (16%). The ATC haplotype was associated with increased risk for pSS (OR 3.42, CI:0.16-11.04, p=0.0395). Higher sIL10 levels were observed in patients with pSS than in the control group (3.0 vs 2.4 pg/mL, p<0.0001), sIL-10 correlated with anti-Ro (r=0.3615, p=0.0053) and anti-La (r=0.2, p=0.0401) autoantibodies. Patients positive to anti-Ro showed higher levels of sIL-10 than negative patients (3.2 vs 2.9 pg/mL, p=0.0243). Similar results were found to pSS positive to anti-La and IL-10. Conclusions: Our results suggest that carriers of the haplotype ATC have 3.42-folds more likely to develop SSp, however, due to their low frequency in our population lacks clinical utility. Serum IL-10 levels are increased in pSS patients and are positively correlated with autoantibodies. IL-10 may have potential role as a pathogenic cytokine in pSS. However, the low associations of IL-10 with the clinical parameters suggest that it should not be considered as a biomarker of the disease.

1014T Large scale meta-analysis of 1000G imputed genotypes in 95,061 subjects reveals 7 novel loci for loss of kidney function. M. Gorski1,2, the CKDGen Consortium. 1) Department of Genetic Epidemiology, University of Regensburg, Regensburg, Bayern, Germany; 2) Department of Nephrology, University Hospital Regensburg, Regensburg, Germany.

Meta-analyses of genome-wide association studies (GWAS) using Hapmap imputed genotypes have successfully revealed several susceptibility loci for estimated glomerular filtration rate (eGFR). 1000G imputed genotypes promise to reveal novel loci and to give additional insights into the genetic architecture of kidney function measured by eGFR. A large scale meta-analysis of 1000G imputed genotypes has not yet been performed. The objective was to answer if a meta-analysis of 1000G imputed GWAS can identify loci for the quantitative measure eGFR, and to evaluate how potentially new loci compare to previous meta-analysis based on Hapmap imputed genotypes. GWAS was estimated by the four-variable MDRD Study Equation. GWAS were performed on genotype dosages, imputed with reference panels of up to 30 million variants in 32 studies including 95,061 subjects. We identified 7 new susceptibility loci for reduced eGFR (p values range from 2.89E-08 to 6.21E-11). The loci were in or near the genes LPHN2, ACVR2A, KIAA1715, ARL15, SLC17A1, ASTN2 and SLC7A6. The lead variants of the novel loci were either new in the 1000G reference panel or were previously included in Hapmap panels but did not show genome wide significance. One of the loci, SLC7A6, was suggested in previous work or were previously included in Hapmap panels but did not show genome wide significance. One of the loci, SLC7A6, was suggested in previous work.
1015S
Association analysis of PPARγ (p.Pro12Ala) polymorphism with type 2 diabetic retinopathy in patients from north India. N. Kaur, V. Vanita. Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: To determine whether p.Pro12Ala polymorphism in PPARγ is associated with type 2 diabetic retinopathy in patients from north India. Material & Methods: Present case-control association study included total 1613 individuals, of which 717 were diagnosed as diabetic retinopathy (DR) patients due to presence of neovascularisation, microaneurysms and hemorrhage in their retina. 608 T2DM patients without any sign of retinopathy (CDR) and 288 healthy controls (NDC) without diabetes and any sign of retinopathy, were also collected and analyzed. The mean duration of diabet-es in DR and CDR group was 12.44±6.356 and 8.050±5.772, respectively. The genotyping for p.12Ala polymorphism of PPARγ in these total 1613 cases and controls was performed by Taqman Drug Metabolism Genotyping Assays using Real time PCR. Further, few genotypes were confirmed by direct DNA sequence analysis of the amplified products. Statistical analysis was performed using SPSS for windows version 16 (SPSS Inc). Binary logistic regression analyses were used to determine odds ratio (OR). Results: The genotype distribution and allele frequency of p.12Ala (PPARG) polymorphism did not differ significantly (p>0.05) between DR and CDR. However, statistically significant differences in genotype (p=0.002) and allele distribution (p=0.0326) were observed on comparing DR with NDC group. A significantly higher frequency of heterozygous CC genotype was observed in DR group as compared to NDC group (18.86% vs 10.76%). On comparison of CDR group with NDC group, statistically significant differences in the genotype distribution were observed (p=0.0005) whereas distribution of allele frequency did not differ significantly (p=0.091) among these two groups. The homoygosity of dominant allele CC was found to be more common in NDCs as compared to cases (86.11% vs 78.94%) which suggested the evidence of the associa-tion of this genotype in dominant mode of inheritance (OR: 1.65, 95% CI: 1.72-2.43, p=0.009).

Conclusion: These findings suggest a statistically significant association of PPARγ (p.12Ala) polymorphism with T2DM. Genotype distribution and allele homozygosity of CC were found to be significantly higher in cases as compared to controls. These findings suggest lack of association of -106 C/T polymorphism in AKR1B1gene with diabetic retinopathy.

1016M
To study the association of -106 C/T polymorphism in the aldose reductase (AKR1B1) gene with diabetic retinopathy, Vanita, Kumar, N. Kaur. Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: Present study aimed to assess the association of -106C/T polymorphism of the aldose reductase gene (AKR1B1) with retinopathy in type 2 diabetic cases from North India. Material and Methods: The present study included 213 patients diagnosed with diabetic retinopathy and 233 type 2 diabetes mellitus (T2DM) cases without any sign of retinopathy, taken as controls. All the cases and controls were recruited from Dr. Daljit Singh Eye Hospital, Amritsar. All the participating individuals underwent complete ophthalmologic examination that included visual acuity testing, Humphrey’s perimetry, optical coherence tomography (OCT) and fundus examination after pupil dilation. The genotyping was performed by direct DNA sequence analysis of the amplified products. Statistical analysis was performed using SPSS for windows (Version. 16.0). Genotype distributions and allele fre-quencies were evaluated using statistical calculator (StatPac V. 3.0). Med-Calc (V.9.3.9.0) was used to determine the odds ratios (OR) and 95% confidence intervals (CI). Results: The alleles (p=0.8842) and genotypes frequencies (p=0.6304) showed no statistical differences when DR cases were compared with controls (T2DM). OR analysis also revealed no signifi-cant co-relation of -106 C/T polymorphism with DR. Conclusions: Present findings suggest lack of association of -106 C/T polymorphism in AKR1B1 gene with diabetic retinopathy in tested patients from north India.

1017T
Menopausal age shares a common genetic background with diabetes and lipid traits: a study on 13,484 Finns. A. Joensuu1,2, J. Kettunen2,3, S. Ripatti1,4, J. Sinisalo5, M.S. Nieminen5, M. Loks6, A. Jula2, J.G. Eriksson7, V. Salomaa5, M. Perola1,2,3, K. Auro6,1, Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland; 2 National Institute for Health and Welfare (THL), Helsinki, Finland; 3 University of Oulu, Finland; 4 Helsinki University, Helsinki, Finland; 5 Heart and Lung Center HUCH, Helsinki University Central Hospital, Finland; 6 Haartman Institute, University of Helsinki, Finland; 7 National Institute for Health and Welfare (THL), Turku, Finland; 8 Department of General Practice and Primary Health Care, University of Helsinki, Finland. Background: The average age for menopause in western countries is 51 years, but individual variation is high with natural menopause occurring between 40-60 years of age. Later age at menopause has been associated with decreased risk of cardiovascular diseases and osteoporosis as well as increased life expectancy. However, associations between menopausal age and diabetes are controversial and no genetic links have been reported. Smoking has been shown to have a lowering effect on menopausal age whereas obesity raises the age at natural menopause, possibly due to the secretion of hormones from adipose tissue. These and other confounding environmental effects might disturb some of the reported associations between menopausal age and diseases. Methods: Stolik et al. performed a large meta-analysis of genome-wide association studies reporting 17 genetic variants which associate with age at natural menopause (Nat.Genet). We constructed a genetic score of menopausal age by summing the reported effect sizes (years/allele) of these polymorphisms in five Finnish cohorts (FINRISK1997, PredicCVD, Health2000, Corogene and HBCDS) and studied its association with nondiabetic retinopathy (NDC) and diabetic retinopathy (DR). Results: One-year increase in the genetic score (range -2.30 - 3.13 years) associated significantly with future diabetes in women (N=2831, hazard ratio=1.46, P=0.0083) but not in men (P=0.3). Interestingly, the genetic score associated nominally with prevalent diabetes in men (N=6249, odds ratio=0.86, P=0.016) but no association was seen in women (P=0.7).

1018S
eSNP regulators of genes underlying Mendelian diseases are enriched among GWAS-assessed loci. T. A. Torres. A. Cox. Medicine/Genetic Medicine, The University of Chicago, Chicago, IL, IL.

Thousands of associations between Mendelian and complex diseases have been recently detected through extensive data mining of medical records in large meta-analysis of genome-wide association studies reporting 17 genetic variants which associate with age at natural menopause (Nat.Genet). We constructed a genetic score of menopausal age by summing the reported effect sizes (years/allele) of these polymorphisms in five Finnish cohorts (FINRISK1997, PredicCVD, Health2000, Corogene and HBCDS) and studied its association with nondiabetic retinopathy (NDC) and diabetic retinopathy (DR). Results: One-year increase in the genetic score (range -2.30 - 3.13 years) associated significantly with future diabetes in women (N=2831, hazard ratio=1.46, P=0.0083) but not in men (P=0.3). Interestingly, the genetic score associated nominally with prevalent diabetes in men (N=6249, odds ratio=0.86, P=0.016) but no association was seen in women (P=0.7). Conclusions: The timing of menopause may share a common genetic background with cardiovascular disease endpoints. The timing of menopause may share a common genetic background with cardiovascular disease endpoints.
Large scale genome-wide association studies (GWAS) identified polymorphisms reproducibly associated with asthma. Most GWAS were performed with a broadly defined asthma phenotype. The role of specific variants/genes largely undefined. We recently identified four homogeneous subgroups of asthma patients defined by factor and cluster analysis: 1) older patients with low atopy and low lung functions (n=104), 2) patients with high atopy and low lung functions (n=125), 3) non-smoking patients (n=362), and 4) patients with smoking history (n=105). The goal of this study is to evaluate if GWAS-nominated variants are more strongly associated with asthma patients sharing the same clinical characteristics, which may help reveal the role of recently identified genes. Genotyping of GWAS-nominated variants were performed in the Quebec City Case-Control Asthma Cohort (QCCAC). The QCCAC consists of 982 unrelated French Canadian subjects (566 cases and 416 controls) with data collected on lung function and a broadly defined asthma phenotype, leaving the role of susceptibility variants more strongly associated with type 2 diabetes for some American Indian populations.

1020T Asthma susceptibility genetic variants are more strongly associated with phenotypically similar subgroups of patients. E. Lavoie-Charland1, J.-C. Bérubé2, A. Marinelarena1, S. Kobes1, S. Coles2, R.L. Hanlon3, W.C. Knower1, C. Bogardus1, L.J. Baier1. 1) NIH, Phoenix, AZ; 2) Texas Biomedical Research Institute, San Antonio, TX; 3) Department of Biomedical Chemistry, Seville, Spain. We performed a meta-analysis of published asthma GWAS studies. 133,561 asthma cases and 978,552 controls were included. A total of 135,717 SNPs were genotyped in 6817 cases and 7448 controls. In addition, 2,795,804 novel SNPs were genotyped in 1754 cases and 10,028 controls. We performed a trans-ethnic analysis of 58,936 cases and 116,664 controls. We also performed a replication study of 2535 cases and 2535 controls. The results show that 5 highly concordant exonic SNPs, rs75403593, rs75418186, rs13342692, and rs11776867, in the SLC16A11 gene that were associated with type 2 diabetes in European Americans were also associated with type 2 diabetes in American Indians. Whole genome sequencing data for 234 Pima Indians showed that all 5 of the SNPs were in near complete concordance (r2=0.99), therefore, rs75493953 was selected as the representative SNP for genotyping in 7,710 American Indians (3,625 were full-heritage Pima Indians and the remaining 4,085 were on average 1/2 Pima and 3/4 American Indian). No association with diabetes was observed (P=0.38, OR=1.04 [0.95-1.14] adjusted for age, sex, heritage, and family relationship). To gain more power, rs75493953 was genotyped in two additional samples of American Indians consisting of various tribes who were participants of the Family Investigation of Nephropathy and Diabetes study (n=3,095) and Strong Heart Study (n=2,421). Analysis of all American Indian samples combined (n=12,811) showed a nominal association between rs75493953 and type 2 diabetes (P=0.001, OR=1.1 [1.04-1.19] adjusted for age, sex, and study center). To define a possible mechanism for the association with diabetes, genotypic data for rs75493953 was merged with genome-wide adipose tissue and skeletal muscle gene expression data obtained in a prior study in Pima Indians using the Affymetrix Human Exon 1.0 ST Array. The top genome-wide expression quantitative trait loci (eQTL) for both adipose tissue and skeletal muscle was RNASEK located 33 kb upstream of SLC16A11 (adipose, n=187, P=0.0004; skeletal muscle, n=196, P=0.001). The association between rs75493953 and RNA-SEK expression levels in adipose tissue was validated by qRT-PCR in a subset of the microarray subjects (n=159, P=0.01). These results suggest that rs75493953 or a highly concordant variant may affect RNASEK transcription. In conclusion, our study shows that either SLC16A11 or RNASEK could potentially play a minor role in type 2 diabetes for some American Indian populations.
1023T
New mutations in the MYOC gene in patients with juvenile open-angle glaucoma. P.C. Vasconcellos1, P.V. Svindick2, C.A. Braghini2, M.B. Melo1. 1) Dept Ophthalmology, Univ Campinas - UNICAMP, Campinas SP, Brazil; 2) CBMEG, Univ Campinas - UNICAMP, Campinas SP, Brazil.

Glaucoma is the leading cause of irreversible blindness and is characterized by progressive optic disc cupping with corresponding visual field loss. Both intraocular pressure (IOP) and positive family history are risk factors for the development of the disease. Juvenile open angle glaucoma (JOAG) is a more severe form of open angle glaucoma characterized by an early onset (10 to 35 years of age). Mutations in the myocilin gene (MYOC) account for most cases of JOAG. This study was approved by the Ethics Committee of the University of Campinas. DNA samples from 98 unrelated Brazilian patients with JOAG and 92 normal individuals were screened for mutations in the three exons and intron/exon junctions of the MYOC gene through direct sequencing. In order to evaluate clinical aspects, ophthalmologic examination included evaluation of the visual field and optic disc, intraocular pressure measurement, and gonioscopy. Mutation screening revealed two new glaucoma causing mutations: the frame-shift mutation 1536delT and X506W. Besides, the previously described glaucoma-causing mutations: 343R (27 patients), Q368X (2 patients), P370L (1 patient) and K423E (1 patient) were also identified. Among neutral variants, we observed the previously reported G122G (1 patient), R76K (14 patients), P13P (1 patient), R13Q (1 patient), 540g (1 patient) and the unreported variant 15 b/a (2 patients), 540g (1 patient) and the unreported variant 15 b/a (2 patients). This study provides an overview of MYOC mutation spectra and frequencies in JOAG Brazilian patients. These findings may be important for the increase of MYOC mutations panel and their effects on phenotype. Financial Support: Fapesp 02/11575-0.

1024S
Association of ADIPOQ rs266729, rs17300539, rs2241766 and rs17846866 with Type 2 Diabetes and Diabetic Retinopathy in North-West Indian population. A. Bhanwer1, R. Sikka1, K. Matharoo1. Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

ADIPOQ is an adipokine gene present on chromosome3 (3q27.3) and is known to be associated with Type 2 diabetes (T2D) and metabolic syndromes. It is specifically and abundantly expressed in the adipose tissue and sensitizes the body to insulin. Adiponectin act as an anti-inflammatory adipokine with anti-diabetic properties. However, the role of ADIPOQ polymorphism has not been clarified but it seems to play a role in vascular damage by interfering with anti-inflammatory and anti-diabetic effects of adiponectin, which may contribute to worsening or acceleration of microvascular complications. Polymorphisms in ADIPOQ gene influences the protein function leading to acceleration of microvascular complications through vascular damage. Therefore, in the present study, we aimed to explore the role of ADIPOQ rs266729, rs17300539, rs2241766 and rs17846866 with the aetiology of T2D and diabetic retinopathy (DR). A total of 454 individuals (207 DR and 247 T2D) and 241 gender matched above 50 years of age healthy control were recruited from the region of Amritsar. Genotyping was done by using PCR-RFLP for rs266729, rs17300539, rs2241766 and PCR-ARMS for rs17846866. rs266729 CG+GG and GG genotypes tend to give 1.8 and 1.9 fold risk, respectively towards DR when compared with T2D cases and 1.7-1.9 fold risk towards DR when compared to healthy control. For rs17300539, AA genotype tends to give 4 fold risk towards DR when compared with T2D cases however; AA+GA and GA genotypes tend to give 2-4 fold risk towards DR when compared with healthy controls. In rs2241766, wild genotype TT gives 2 and 2.5 fold risk towards DR when compared with T2D and controls, respectively. However, TG genotype tends to give protection towards DR when compared with T2D [OR=0.39 (0.21-0.73)] and healthy controls [OR=0.35 (1.9-0.65)]. In case of rs17846866, GG+TG and GG genotypes tend to give fold risk towards DR when compared with T2D and healthy controls. However, no association was observed for rs266729, rs2241766 and rs17846866 with T2D. In conclusion, rs266729, rs17300539, rs2241766 and rs17846866 seem to be associated with risk of DR when compared with T2D cases as well as with healthy controls in the population of Punjab.

1025M
More evidence for association of a rare TREM2 variant (R47H) with Alzheimer’s disease risk. S.L. Rosenthal1, M.N. Bamshad2, X. Wang3, S. Bernstein2, B.E. Snitz2, W.E. Klunk2, R.A. Sweeney2, F.Y. Demirci2, O.L. Lopez2, M.I. Kambholj1. 1) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; 2) Department of Neurology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; 3) Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.

Over twenty risk loci have been identified for late-onset Alzheimer’s disease (LOAD). Among them, only one, APOE, produces a strong effect size for this neurodegenerative disease. Individuals homozygous for the APOE-ε4 allele have a 60% increased lifetime risk of developing LOAD. Recent sequencing efforts have detected a rare variant in TREM2, R47H, that has an effect size similar to that of APOE. In our study, we genotyped this variant in a case-control population of Caucasian descent derived from two study cohorts: 1283 cases (mean age-at-onset, AAO=73.26 ± 6.76) and 996 controls (mean age=75.67 ± 6.37) from the University of Pittsburgh Alzheimer’s Disease Research Center, and 338 cases (mean AAO=84 ± 3.98) and 1950 controls (mean age=83.15 ± 9.93) from the Ginkgo Evaluation of Memory study. In the total sample, carriers of the R47H variant had a significantly increased risk of having LOAD (OR=7.40, p=3.66E-06). We also examined the effect of this variant on psychosis in LOAD and Aβ deposition (as measured by Pittsburgh compound B uptake) and found no significant association between this variant and either phenotype. We have successfully replicated the association of the TREM2 R47H variant with LOAD risk in Caucasians. Combined with other studies, our results provide further evidence for the involvement of TREM2, specifically the R47H variant, in modulating LOAD risk.

1026T
The IL10 ACA haplotype is associated with rheumatoid arthritis in patients from Western Mexico. J. Hernandez-Bello1, M. Vazquez-Vilar2, E. Oregon-Romero1, Y. Valle1, J.R. Padilla-Gutierrez1, G. Martinez-Bonilla1, J.F. Muñoz-Valle1. 1) Instituto de Investigación en Ciencias Biomedicas, CUCS, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico; 2) Servicio de Reumatología, Hospital Civil de Guadalajara, Fray Antonio Alcalde, Guadalajara, Jalisco Mexico.

Background: Rheumatoid arthritis (RA) (MIM 180300) is a chronic autoimmune disease of unknown etiology affecting approximately 1% of the population worldwide. RA is characterized by a cartilage and bone erosion resulting from tissue destruction and have been reported as a risk factor for autoimmune diseases. Individuals homozygous for the AA genotype have a 60% increased lifetime risk of developing LOAD. Recent sequencing efforts have detected a rare variant in TREM2, R47H, that has an effect size similar to that of APOE. In our study, we genotyped this variant in a case-control population of Caucasian descent derived from two study cohorts: 1283 cases (mean age-at-onset, AAO=73.26 ± 6.76) and 996 controls (mean age=75.67 ± 6.37) from the University of Pittsburgh Alzheimer’s Disease Research Center, and 338 cases (mean AAO=84 ± 3.98) and 1950 controls (mean age=83.15 ± 9.93) from the Ginkgo Evaluation of Memory study. In the total sample, carriers of the R47H variant had a significantly increased risk of having LOAD (OR=7.40, p=3.66E-06). We also examined the effect of this variant on psychosis in LOAD and Aβ deposition (as measured by Pittsburgh compound B uptake) and found no significant association between this variant and either phenotype. We have successfully replicated the association of the TREM2 R47H variant with LOAD risk in Caucasians. Combined with other studies, our results provide further evidence for the involvement of TREM2, specifically the R47H variant, in modulating LOAD risk.

Background: Rheumatoid arthritis (RA) (MIM 180300) is a chronic autoimmune disease of unknown etiology affecting approximately 1% of the population worldwide. RA is characterized by a cartilage and bone erosion resulting from tissue destruction and have been reported as a risk factor for autoimmune diseases. Individuals homozygous for the AA genotype have a 60% increased lifetime risk of developing LOAD. Recent sequencing efforts have detected a rare variant in TREM2, R47H, that has an effect size similar to that of APOE. In our study, we genotyped this variant in a case-control population of Caucasian descent derived from two study cohorts: 1283 cases (mean age-at-onset, AAO=73.26 ± 6.76) and 996 controls (mean age=75.67 ± 6.37) from the University of Pittsburgh Alzheimer’s Disease Research Center, and 338 cases (mean AAO=84 ± 3.98) and 1950 controls (mean age=83.15 ± 9.93) from the Ginkgo Evaluation of Memory study. In the total sample, carriers of the R47H variant had a significantly increased risk of having LOAD (OR=7.40, p=3.66E-06). We also examined the effect of this variant on psychosis in LOAD and Aβ deposition (as measured by Pittsburgh compound B uptake) and found no significant association between this variant and either phenotype. We have successfully replicated the association of the TREM2 R47H variant with LOAD risk in Caucasians. Combined with other studies, our results provide further evidence for the involvement of TREM2, specifically the R47H variant, in modulating LOAD risk.
1027S
Recent reports on genome2wide association studies focusing on serum uric acid concentrations identified several novel loci in European descent samples. In the current study, we aimed to evaluate the association between these loci and concentrations of serum uric acid in the Chinese. We genotyped fourteen single nucleotide polymorphisms (SNPs) from eleven loci mapped in or near PDZK1, GCKR, LRP2, SLC2A9, ABCG2, LRRC16A, SLC17A1, SLC17A3, SLC22A11, SLC22A12 and SF1 in the Shanghai Chinese, including 2329 participants from a community based study. As uric acid metabolism differs between genders, all of the variants were analyzed for gender differences. SNPs including rs780094 in GCKR, rs11722228 in SLC2A9 and rs606458 in SF1 showed association to serum uric acid levels in males after adjustment for age and BMI as confounders (p=0.016, 0.0011 and 0.030, respectively); SLC2A9 rs3775948 demonstrated a trend towards association to uric acid (p=0.071). In females, only rs606338 in SLC22A12 was detected to potentially associate with uric acid (p=0.0569). After combined analysis between males and females, we only detected that SLC2A9 rs3775948 and SF1 rs606458 were related to uric acid (p=0.036 and 0.043, respectively). The interaction between rs11722228 in SLC2A9 and gender on serum uric acid levels existed in our samples. In conclusions: We showed that GCKR, SLC2A9 and SF1 variants modulated uric acid levels in the Shanghai Chinese. Moreover, gender difference between GCKR polymorphisms and uric acid levels was also found in our study. The findings provide a perspective that genetic variation is one of the key influences upon regulation of serum uric acid levels in humans.

1028M
Transferrin receptor and hereditary hemochromatosis gene variants interact to modify childhood leukemia risk. A.E. Kennedy1, E. DePtycke2, G.Y. Lau3, K.Y. Kamdar2, P.J. Lupo4, M.F. Okcu5, M.E. Scheurer6, M.K. Baum7, D. Seminara8, M.T. Dorak2, 1) Epidemiology and Genomics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, NIH, Rockville, MD; 2) Section of Hematology-Oncology, Department of Pediatrics, Baylor College of Medicine, Houston, TX; 3) Department of Dietetics and Nutrition, Robert Stempel College of Public Health & Social Work, Florida International University, Miami, FL; 4) School of Health Sciences, Liverpool Hope University, Liverpool, United Kingdom.
An interaction between the transferrin receptor (TFRC) rs3817672 (S142G) and hereditary hemochromatosis gene (HFE) rs1800672 (C282Y) is associated with increased risk for multiple myeloma, breast, colorectal and hepatocellular cancers as well as childhood acute lymphoblastic leukemia (ALL) risk in European studies. HFE C282Y and H65D correlate with body iron levels. Since HFE and TFRC proteins biologically interact in iron transfer across membranes, the observed statistical interaction suggests that the involvement of HFE variants in cancer risk modification is mediated via their synergistic effect on body iron levels. We sought to replicate the HFExTFRC interaction in childhood ALL in a US-based study, and used bioinformatic tools to assess the contribution of TFRC S142G to this association. Genotyping was conducted in a multi-ethnic sample from Houston, TX (161 incident cases with childhood ALL and 231 controls, all <18 yr). Being positive for either HFE variant yielded an elevated odds ratio for childhood ALL risk in males (1.4, 95% CI=0.8 to 2.4), which increased to 3.0 (95% CI=1.3 to 6.8) in the presence of the S142G allele A homozygous (HFE C282Y and H65D) allele A were identified by silver nitrate. Results: The general population is in Hardy Weinberg Equilibrium. We found only one individual in the group of patients without the variant submitted NL homozygous genotype (Ala12Ala). The genotype and allele frequencies in the two groups are shown in Table 1. Statistical analysis was performed using the Finetti diagram and only statistically significant differences were found when comparing the alleles in SLE patients with or without NL, where the risk allele is 12Ala. (OR = 9.6, C.I 2.115-43.583 (p=0.00197), *) Conclusions: This is the first report of an association analysis of PPARG Pro12Ala gene polymorphism with SLE in Mexican population and the results suggest that the presence of the variant allele (Ala) confers risk for LN; however it is necessary to increase the sample size to confirm this association.
1030S
Quantification of Hirschsprung disease susceptibility from common polymorphisms in relation to gender, segment length of aganglionosis and family history. A. Kapoor1, Q. Jiang1,2, C. Bernoth1, A. Chakravarti1.
1) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 2) Department of Medical Genetics, Capital Institute of Pediatrics, Beijing, China.

Hirschsprung disease (HSCR), a congenital developmental defect of the enteric nervous system, is a multifactorial disorder in which common, low penetrance, noncoding variants at RET, NRG1 and CAV3 underlie disease risk, primarily in male, short segment simplex (~70%) cases. Rare, high penetrance, coding variants in at least a dozen different genes underlie disease risk in ~25%, enriched for female, long segment multiplex cases. We have used the largest collection of HSCR cases and relatives of European ancestry ascertained in the USA (~1100 individuals: 365 probands, 40 affected relatives and 710 unaffected relatives) to directly assess the joint contribution of RET (rs2435357, rs2506030), NRG1 (rs7835688, rs16879552) and CAV3 (rs15833147, rs21707682, rs17166001) common variants and risk covariates to HSCR susceptibility. Case-control analysis in 342 cases and 379 controls, showed significant associations at RET at rs2435357 (odds ratio (OR) = 3.7 and P<1.2×10^{-10}) and at rs2506030 (OR = 1.9 and P=3.5x10^{-9}), and at CAV3 rs17166001 (OR=1.5 and P=0.002). These results were corroborated by transmission disequilibrium tests in 243 trios confirming that these associations were not due to cryptic population structure (OR=4.7, P=5.1x10^{-5}; OR=1.9, P=3.6x10^{-7}; OR=1.7, P=0.002 for rs243537, rs2506030 and rs17166001, respectively). We failed to find any evidence of a reported NRG1 effect, indicating that this association is likely restricted to Asians where it was first reported. HSCR disease risk was directly associated with having at least one 3+ risk alleles at rs243537, rs2506030 and rs17166001 (OR=1.85, P=0.001; OR=3.29, P=5.95x10^{-7}; OR=10.35, P=6.56x10^{-5} for 3, 4 and 5+ alleles, respectively); fewer alleles were protective. These risks translated into a significantly higher incidence (penetrance) by allelic dosage: 24.6%, 42.2% and 141.3% cases per 100,000 live-births as compared to the population average of 15.100,000. Common allele dosage-based disease risk was significantly higher in males (P=0.005), and higher but not significant in short segment and simplex cases. Finally, despite our over-ascertainment of familial collections, siblings’ recurrence risk is 29% in sibs of probands with <3 risk alleles as compared to 19% when probands had ≥3 risk alleles. Consequently, these common polymorphisms delineate a class of frequent HSCR cases that have lower recurrence rates within sibs, a feature that should inform genetic counseling.

1031M
Association between common variant near CAV1 and CAV2 genes and phenotypic features of primary open angle glaucoma. F. Mathur1,2, Sakurada1, S. Yoneyama1, K. Kashiwagi1, Z. Yamagata1, H. Iijima1. 1) Dept Ophthalmology, Univ Yamanashi, Chuo, Yamanashi, Japan; 2) Dept Health Sciences, Univ Yamanashi, Chuo, Yamanashi, Japan.

Purpose: It was reported that common variants near cavinol 1 (CAV1) and cavinol 2 (CAV2) genes were associated with primary open-angle glaucoma (POAG). This study was performed to assess the association between the common variant near these genes and the phenotypic features in patients with POAG, including non-POAG, primary angle closure glaucoma, and tension glaucoma (HTG). Methods: Four hundred and one Japanese patients with POAG, including 176 patients with NTG and 225 patients with HTG, and 191 control subjects without glaucoma were analyzed for the common variant (rs1052990) near CAV1 and CAV2 genes. The genotype and allele frequencies were compared between the patients with NTG or HTG and the control subjects. Demographic and clinical features, including age at diagnosis of glaucoma, gender, family history of glaucoma, refractive error, maximum intraocular pressure (IOP), visual field, glaucoma visual field test and history of glaucoma surgery, were compared between the genotypes in patients with POAG. A multiple linear regression analysis was carried out with the maximum IOP as a dependent variable and age, gender, and the common variant as independent variables. Results: There was a significant difference (P = 0.024, analysis of variance) of the maximum IOP among GG (21.8 ± 4.6mmHg, mean ± standard deviation), GT (22.7 ± 6.6mmHg), and TT (24.9 ± 9.2mmHg) genotypes in patients with POAG, although no significant differences of the genotype and allele frequencies could be found between the patients with NTG (GG: 10%, GT: 40%, TT: 50%, P = 0.90; G allele: 22.2%, T allele: 77.8%, P = 0.99) and HTG (GG: 4.9%, GT: 31.1%, TT: 64.0%, P = 0.76; G allele: 20.4%, T allele: 79.6%, P = 0.61) and the control subjects (GG: 4.7%, GT: 34.6%, TT: 60.7%, G allele: 22.0%, T allele: 78.0%). Baseline visual field and visual field change in IOP (P = 0.05), age (P = 0.028), and gender (P = 0.001) were significant when compared between the common variant and the maximum IOP (P = 0.12, standard error = 0.70, P = 0.017). Conclusion: The association of the common variant near CAV1 and CAV2 genes with a maximum IOP in patients with POAG suggests that this locus may contribute to the pathogenesis of POAG as a genetic risk factor associated with IOP elevation.

1032T
GSTM1 and GSTM1 gene frequency in Punjabi population exposed to pesticides. M. Ahluwalia, A. Kaur. Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

The glutathione-S-transferase (GST) enzyme system play an important role in defence mechanism against environmental chemicals and endogenous substances. Many of these enzymes show variability among individuals due to their polymorphic character and thus show difference in predisposition to various diseases such as cancers, neurological disorders, endocrine system disruption, and etc. The aim of the study was to determine GSTT1 and GSTM1 gene deletones in 69 individuals from different regions of Punjab involved in agricultural activities such as preparation, loading, spraying and storage of pesticides, harvesting of crops etc. Genotyping was done by Multiplex Polymerase chain reaction approach. The deleted GSTT1 and GSTM1 genes show frequency of 14.5% and 50.7% respectively. The results suggest an increased tendency of exposure related diseases due to complete gene deletion of either of the two genes. However, 5.7% population has shown deletion of both the genes and are expected to have significantly reduced detoxification tendency and thus an increased risk to diseases caused by exposure to environmental carcinogens. The study thus provides the basis for gene-environment association studies to be carried out. The detection of oxidative stress in serum by biochemical methods is in progress.

1033S
Relationships between genetic ancestry and nicotine and tobacco carcinogen metabolisms in the Multiethnic Cohort. H. Wang1, S. Park2, C. Haiman3, S. Murphy2, S. Hecht4, B. Henderson2, L. Kolonel1, D. Straw2, L. Le Marchand1. 1) UH Cancer Center, Honolulu, HI; 2) Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) Masonic Cancer Center, University of Minnesota, Minneapolis, MN.

It has been reported in the Multiethnic Cohort (MEC) study that major (2- to-5-fold) differences exist among US racial/ethnic minorities in the risk of lung cancer associated with cigarette smoking, even after taking into account self-reported dose and duration. Native Hawaiians and African Americans have a higher risk for lung cancer than European Americans, who in turn have a higher risk than Japanese Americans and Latinos. Here we tested the hypothesis that genetic ancestries are related to nicotine metabolism (up to 80% through C-oxidation by CYP2PA6) and internal dose of tobacco smoke carcinogens, thus contributing to the observed differences in lung cancer risk across populations. Genetic ancestries were estimated using STRUCTURE based on 121 ancestral informative markers for 2,170 MEC African Americans, European Americans, Japanese Americans, Latinos, and Native Hawaiians, along with 5 HapMap reference populations, namely, ASW, CEU, YRI, JPT and MEX. Percent ancestries, i.e. African, American Indian, Native Hawaiian and East Asian ancestries were obtained and categorized according to deciles. Nicotine metabolism (total trans-3-hydroxycytidine: free cotinine) and exposure to and detoxification of nitrosamine 4-(methyltrinitrosamino)-1-(3-pyridyl)-1-butanone (NNK), a major tobacco-specific lung carcinogen, were regressed on percent ancestors in a linear regression model, adjusting for age, sex, assay batch, creatinine and internal dose of nicotine. African ancestry was related to a lower level of CYP2PA6 activity (P = 0.12) and a higher level of exposure to (P = 0.03) and non-significant lower level of detoxification of (P = 0.93) NNK. East Asian and Native Hawaiian ancestries were related to lower levels of CYP2PA6 activity (P ≤ 0.002), lower levels of exposure to (P’s ≤ 0.04) and of detoxification of (P’s <0.0001) NNK. American Indian ancestry (in Latinos) was related to a higher level of CYP2PA6 activity (P < 0.0001), a non-significant lower level of exposure to (P = 0.81) and significant lower level of detoxification of (P = 0.002) NNK. These results are consistent with and could partly explain the observed differences in lung cancer risk due to smoking in African American, East Asian, Latinos, and comparison to European Americans, but could not explain the high disease risk of Native Hawaiians, suggesting more factors contribute to risk in this recently admixed population.
1034M Haplotypic association and synergic effect of Renin-angiotensin aldosterone system gene polymorphisms causing susceptibility to essential hypertension in Indian patients, C. Bhupatiraju, MV. Upuluril, P. Gunda, D. Pandaparthurar, P. Tirunilv. 1 Dept of Genetics, Osmania University, Hyderabad, India; 2 Gandhi Medical College and Hospital, Hyderabad, India.

The present study investigates association of essential hypertension (EHT) with candidate gene polymorphisms involved in vasoconstriction (REN, AGT, ACE, AGTR1 and ced95-ADR) and vasodilation (BDKRB2); haplotype association and epistatic interactions between them. Evaluation of 316 hypertensive and 200 healthy individuals for various epistatic interactions suggests high risk for developing EHT for a) CT genotypes and allele C carriers of g.-5434C>T and g.18-83G>A of REN; b)-17G>A of AGT; c)-5434C>T and g.18-83G>A of BDKNB2 (Danilo) and d) +9bp allele carriers of BDKNB2 +9bp especially in males. The positive association of g.-217A variant and +9bp variant with EHT may be correlated with various transcriptional activities of the respective genes resulting in vasoconstriction. Analysis for epistatic interactions suggest a high risk for individuals with i) g.-5434CC and g.18-83GA, ii) g.-5434TT and g.18-83GA and iii) g.-5434TT and g.18-83GG genotypic combinations of REN gene causing susceptibility to EHT. High risk with respect to epidemiological factors was observed for females with positive family history of EHT. The gene-environmental interaction analysis suggests an interaction between both the above-mentioned genes affecting EHT with respect to gender. Estimates of the degree of linkage disequilibrium (LD) revealed a complete linkage disequilibrium between g.-152G>A and p.M235T and between p.T174M and g.-5312C>T polymorphisms in both hypertensives and controls. The strong evidence of linkage disequilibrium between g.-152G>A and g.-58C>T, g.-20C>A and g.1166A>C, p.T174M and g.1166A>C, g.-5434C>T and g.18-83G>A polymorphisms, while LD between these polymorphisms was absent or weaker among the controls. In general the study suggests the need for evaluating lifestyle modifications in male and female patients and diet at an early age and monitoring the BMI levels regularly. The genes showing polymorphic significant high risk of developing EHT may provide clues for designing drugs whose mode of action would be genotype based. Any such effort to ameliorate the suffering due to the disease improves the quality of life of affected.

1035T Genetic variation and Insulin Resistance in middle age Chinese men. B. Villegas, X. Shu, M. D. Ritchie, H. E. Abboud, B. L. Goring, P. Pendergrass, F. Peissig, C. A. McCarty, M. D. Ritchie. 1) Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA; 2) Marshfield Clinic, Marshfield WI; 3) Essentia Rural Health, Duluth, MN.

Gene-environment interactions (GEs) and insulin resistance and susceptibility to diabetes are defied in an interactive manner. The underlying mechanism for this phenomenon is not clear and elucidation of this would be a start for prevention and intervention. The current study is an attempt to make some progress towards this end. We used two datasets, one from a case control study of type 2 diabetes (T2D) and the other from a GWAS study of insulin resistance (HOMA-IR) where we measure the genetic variation in a large set of SNPs, their interaction with environmental factors, and their association with diabetes and insulin resistance. We have studied interactions amongst a list of candidate genes that are involved in insulin resistance. In order to identify these candidate genes we used the following resources: 1) the medical genetic literature, 2) the genome-wide interaction database (HAPGEN2) and 3) the GWAS dataset of HOMA-IR scores from the Marshfield Clinic. We found 3 interactions with a P<0.05 with fasting insulin. Conclusion: Our data support the hypothesis that the IGFI gene variation is associated with insulin resistance among middle age Chinese men. We also found interactions between some of the study genes with physical activity, BMI and WHR with insulin resistance.


In an earlier study we have presented clues for designing drugs whose mode of action would be genotype based. Here, we postulate that the use of a genotyping and GWAS data resource such as GTEX can be tested for GxE to further explain the trait. Using PLATO software, we performed a dietary EWAS by analyzing 57 dietary exposures from the National Cancer Institute developed Dietary History Questionnaire using logistic regression, adjusting for age, gender, and type 2 diabetes (T2D) status in 2,629 samples (932 controls, 348 males (M); 1,697 cataract cases, 711 M) of European descent from the Marshfield Clinic Personalized Medicine Research Project, part of the Electronic Medical Records & Genomics (eMERGE) Network. Several dietary measures were predictive of cataract (p < 0.05). A monounsaturated omega-9 fatty acid known as erucic acid (FA22:1) (p=5.5x10^-4) met a Bonferroni corrected p-value threshold (p<0.05). We then tested FA 22:1 for GxE using 498,829 SNPs in a subset of samples for whom genetic data was available (831 controls, 314 M; 1,511 cases, 633 M) using logistic regression adjusted for age, gender, and T2D status. Twenty SNPs-FA22:1 models were statistically significant (p<1.0x10^-7). The most significant Gx model was between FA22:1 and rs726712, an intronic SNP in LPP (p=2.9x10^-4). The erucic acid-cataract association is novel; although significantly depleted levels of two polyunsaturated fatty acids have been found in cataractous human lenses when compared to mature typical lenses. LPP encodes a protein involved in cell-cell adhesion. A variety of studies have previously demonstrated the importance of cell-cell adhesion in regulating cataract formation. Given the role of cataract formation with this specific factor. These findings indicate the role of GxE in susceptibility to cataract and demonstrate the utility of EWAS as a data-driven filtering method for investigating the genetic and environmental interplay of common, complex diseases.
1038T
MS Risk Conferred by Obesity may be Independent of Predisposing Genetic Factors for Obesity: Results from the Kaiser Permanente MS Research Program. M. Gianfrancesco1, X. Shao1, B. Rhead1, L. Shen1, H. Quach1, A. Bernstein2, C. Schaefeli2, L.F. Barcellos1,2. 1) School of Public Health, Dept of Epidemiology, UC Berkeley, Berkeley, CA; 2) Kaiser Permanente Division of Research, Oakland, CA; 3) Palm Drive Hospital, Sebastopol, CA. Rheumatoid arthritis (RA) is a prototype of complex disease that involves multiple genetic and environmental factors. The individual effects of the known genetic risk factors are most often very moderate with highest impact for seropositive RA from the shared epitope (SE) alleles of the HLA-DRB1 gene. We previously have described significant interactions between the SE and some non-HLA genetic variants (i.e. PTPN22) in RA that indicated relatively strong effects in the susceptibility to seropositive RA. To provide more general view on gene-gene interactions in RA risk we tested the hypothesis regarding enrichment of the interactions between the SE alleles and previously shown genetic variants in association with RA. The study cohort includes 1151 seropositive RA patients and 1079 healthy controls from the Swedish EIRA study. Genome wide genotyping was performed using the IlluminaHumanHap300 BeadChip and subsequent imputation was done with the 1000 Genomes Phase I (v) Europeans as reference panel. All the individuals were genotype by low resolution SSP for the HLA-DRB1. Finally, 49 SNPs that previously were shown in association with RA and 48 SNPs selected from the study of genetic basis for reading and writing ability were analysed for interaction with SE alleles. SNPs from the HLA and PTPN22 loci and SNPs with MAF<0.01 were excluded from the analysis. The GEIRA algorithm provides measures of interaction including the attributable proportion (AP) due to interaction. We compared the distribution of the AP values observed in two groups of SNPs, using the Kolmogorov-Smirnov test. Overall 92 tests for interaction were valid in our study due to statistical constraints and we observed an enrichment of p-values <0.05 for AP>20% in RA association compared to control group of SNPs (10 vs. 3). Distribution analysis of p-values reveals significant difference between the two groups (p<0.05). Additionally we noticed that only 17 out of 49 tested SNPs for RA exhibited a significant association in our study, but 7 of these and rare variant data for each candidate gene were obtained through genome wide association profiling and imputation. Preliminary results show oGRS to be weakly correlated with childhood body size, body size in 20’s, body mass index (BMI) in 20’s and BMI in 30’s (r<0.10). IV estimates demonstrated no association between oGRS and MS (causal odds ratio (OR) = 1.09, 95% CI 0.92, 1.29); however, the OR increased when restricting to individuals diagnosed with MS at age 30 or older (OR = 1.17, 95% CI 0.97, 1.42). MS risk conferred by obesity may be independent of predisposing genetic factors for obesity, suggesting that alternative mechanisms may mediate disease onset. Obesity acquired due to environmental factors rather than genetics may have greater influence on MS development, or interact with other risk factors to increase risk of disease.

1039S

The BTBR T+ Ifpr39/J (BTBR) inbred mouse strain is predisposed to obesity and pre-diabetes symptoms compared to the commonly studied C57BL/6J (B6) inbred strain. We have previously mapped loci affecting these traits to a congenic strain in which a ~6 Mb region of B6 genome had been transferred into the BTBR strain background. These congenic mice have reduced body weight, body fat, food intake and pre-diabetes compared to their pure BTBR littermates. Analysis of this congenic strain has revealed a significant gene-environment interaction that modifies the effect of this QTL. Here we show that under standardized phenotyping procedures the genetic effect on body weight was markedly influenced by their housing environment (from a 0 to 17% difference in body weight between homozygous B6 and BTBR congenic mice across 3 facilities; P geno < 0.0001, P env < 0.0001, P geno×env<0.0006). One major difference between facilities was ongoing rodent construction nearby. This is known to cause stress in rodents and to affect food intake. In the facility with the most adjacent construction and where the genetic effect was completely overcome, the mice were never observed to be sleeping. In light of this, we examined the orexin (Hcrt) pathway. Orexin is a neuropeptide that promotes wakefulness and food intake and is known to be affected by stress. Orexin expression was 5 to 10-fold higher (P < 0.001) in mice housed in the facility with the most adjacent construction. We discovered a significant genotype-environment interaction in orexin receptor 1 (Hcrt1) expression in the brain (P=0.02), consistent with the changes in body weight between genotypes and facilities and suggesting that genotype at this locus modifies the influence of stress on orexin signaling. When the effects of stress were directly tested in the congenic mice we demonstrated a significant genotype-stress interaction (P=0.02) on body fat. Thus these studies have identified stress as a significant environmental modifier of obesity risk that is subject to genetic modulation. Discovery of the gene underlying the QTL will reveal a mechanism by which the orexin-mediated effects of stress on obesity can be modulated, which may prove useful for the design of novel strategies to aid weight loss.
4041T Smoking-dependent genetic effects on obesity traits: the GIANT (Genetic Investigation of Anthropomorphic Traits) Consortium. V.A. Fisher, A.E. Justice, L. Winkler, T.N. Heard-Costa, K.L. Young, J. Czajkowski, M. Graff, X. Deng, T.S. Ahluwalia, Q. Qi, L. Qi, A. Mahajan, D. Hadley, A.Y. Chi, L. Barata, J.S. Nguwa, R.A. Scott, L. Xue, T. O. Kipliläinen, C.T. Liu, D.J. Chasman, K. Mohlke, R.J.F. Loos, C. Beresford, C.E. North, L.A. Cupples, 1. G.I.A. Consortium. 1) Department of Biostatistics, Boston University School of Public Health, Boston, MA; 2) Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC; 3) Department of Genetic Epidemiology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany; 4) NHLBI Framingham Heart Study, Framingham, MA; 5) Boston University School of Medicine, Boston MA, USA; 6) Department of Genetics, Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO; 7) The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 8) Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY; 9) Department of Nutrition, Harvard School of Public Health, Boston, MA; 10) Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 11) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 12) Division of Population Health Sciences and Education, St. George’s, University of London, London, UK; 13) Division of Preventive Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, MA; 14) MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK; 15) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 16) The Genotype-Omics and Related Traits Program, The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY.

Obesity and cigarette smoking (SMK) are important risk factors for cardiovascular disease. They are related: average body weight is lower among smokers. To investigate obesity as body size and distribution of body fat, we studied body mass index (BMI), waist circumference adjusted and waist-to-hip ratio adjusted for confounding variables. There was association between smoking status and urate levels (P=0.06 mmol/L). 1, P=0.01), urate-lowering effect (P=0.02 mmol/L) (in men and respectively) was modified by the genotype increasing urate levels in the un-stratified analysis (P=0.02 mmol/L, P<0.001). Subject to replication these data provide insight into the mechanisms involved in urate control and may lead toward personalised dietary advice in strategies to reduce urate concentrations.

4042S Tomato consumption, an anecdotal trigger of gout flares, interacts with three urate transporters (ABC2G2, SLC22A12 and SLC22A7) in a non-additive fashion to influence serum urate. T. Flynn, M. Cadzow, N. Dalbeth, L. Stamp, R. Topless, T. Merriman. 1) Department of Biochemistry, University of Otago, Dunedin, Otago, New Zealand; 2) Department of Medicine, University of Auckland, Auckland, New Zealand; 3) Department of Medicine, University of Otago, Christchurch, Canterbury, New Zealand.

Raised urate levels, the main cause of gout, arise through a complex combination of genetic and environmental factors. Thirty genetic loci and several dietary exposures have well-established effects on urate levels. Interactions between these genetic and environmental factors have been reported. Urate is ideal for investigating gene-environment interactions as causal environmental exposures can be measured concurrently with phenotype. This study aimed to test two hypotheses: that tomato consumption, an anecdotal trigger of gout flares, is associated with urate levels and whether any of 7 genes encoding uric acid transporters interact with tomato consumption. 12,720 European individuals from the Atherosclerosis Risk In Communities (n=7517), Cardiovascular Health Study (n=2511) and Framingham Heart Study (n=3052) cohorts were used. All analyses were adjusted for confounding variables. There was association between tomato intake and urate levels (β=0.66 mmol/L, 1, P=0.01), which was modified by ABC2G2: rs2321142, SLC22A12: rs3825018 and SLC22A7: rs414178 genotypes (Pinteraction=0.04, 0.02 (in men) and 0.049, respectively). In the presence of the urate-increasing genotypes of rs2321142 (TT) and rs3825018 (GG) consumption of tomatoes showed a urate-lowering effect (β=−0.81 mmol/L, P=0.01, β=−3.94 mmol/L, P=0.02 (in men), respectively) compared to the urate-increasing effect shown when no genotypic stratification was performed. In the presence of the rs414178 urate-increasing genotype (GG) consumption of tomatoes showed a significantly higher increase in urate levels than in the un-stratified analysis (β=2.92 mmol/L, 1, P=0.03). Subject to replication these data provide insight into the mechanisms involved in urate control and may lead toward personalised dietary advice in strategies to reduce urate concentrations.


Background: Studies have suggested significant associations between variants in the bradykinin receptor B2 gene (BDKR2B) and the development of hypertension. The bradykinin is one of the most important kinins involved in the modulation of the blood pressure (BP) and many of its physiological functions are mediated by the B2 receptors. In addition, high sodium and potassium intake has also been related to changes in the BP. Objective: The objective of this study was to investigate the interaction between the polymorphism rs1799722 of the BDKRB2 gene and the environmental factors - sodium and potassium intake - on the variation in the BP in a sample of healthy adults from a private university in the South of Brazil. Methods: The participants were investigated for food consumption and BP by professional in the Nutrition Ambulatory from the university. The measurements of sodium and potassium intake were assessed by the 24-hour recall method (24-H-RQ), using the DietWin Professional software. BP values were determined by the mean of three measurements with the digital apparatus Omron. The interaction was assessed using a 1-way ANOVA with post-hoc Bonferroni correction. The allele frequencies were estimated by direct counting and the Hardy-Weinberg Equilibrium (HWE) calculated by the chi-square test. The gene-nutrient interactions were tested using multiple linear regressions with the nutrient and stepwise elimination. Results: The sample was composed by 267 women and 71 men. The average age of the sample was 25.5 years. Compared to women, men showed significantly higher values in the systolic BP (SBP) (men, 124.4 ± 11.0 mmHg; women, 113.7 ± 10.2 mmHg) and in the sodium (men, 2403.8 ± 1725.6 mg 24-H-RQ; women, 1758.5 ± 684.9 mg 24-H-RQ) (all P<0.001). The allele frequencies for the polymorphism were 0.59 (C) and 0.41 (T) and the genotype frequencies did not reveal a significant deviation from the expected values for the Hardy-Weinberg Equilibrium. The main results did not differ significantly from the model with no interaction in the investigated outcomes. However, we detected an interaction between the TT-genotype and the dietary potassium intake on the SBP (p<0.001). The increased consumption of potassium, in carriers of the TT-genotype, was associated with an increasing on the SBP in the TaqMan SNP genotyping assay. A preliminary indication of a gene-nutrition interaction on the BP in a sample of healthy adults.
404T

Genome-wide gene-physical activity interaction study of BMI and waist-hip ratio in 180,418 individuals. T.O. Kipliainen1, M. Graff2, P.A. Scott3, A. Justice4, T.W. Wang5, L. Xue6, A. Mahajani6, D. Hadley7, T.G. Ahluwalla7, F. Renström8, T. Workukenah9, M. den Hoed10, A.Y. Chu11, M.L. Grove12, N. Heard-Costa13, L. Quaye14, J. CZajkowski15, J. Ngwaa16, K. Young17, O. Qi18, O. Pedersen19, K. Mordt14, K.E. North15, I.B. Borecki11, R.I.F. Loos16, L.A. Cupples20, The GIANT Consortium. 1) Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Denmark; 2) Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, NC; 3) RRHE Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK; 4) Department of Genetic Epidemiology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, Germany; 5) Department of Biostatistics, Boston University School of Public Health, MA; 6) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 7) Division of Population Health Sciences and Education, St. George’s, University of London, UK; 8) Department of Clinical Science, Genetic and Molecular Epidemiology Unit, Lund University, Sweden; 9) Department of Epidemiology, Harvard School of Public Health, MA; 10) Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Sweden; 11) Division of Preventive Medicine, Brigham and Women’s Hospital, Boston, MA; 12) Human Genetics Center, University of Texas Health Science Center at Houston, TX; 13) Department of Neurology, Boston University School of Medicine, MA; 14) Department of Twin Research and Genetic Epidemiology, King’s College London, UK; 15) Division of Statistical Genomics, Center for Genome Sciences & Systems Biology, Washington University School of Medicine, MO; 16) Department of Preventive Medicine, Albert Einstein College of Medicine, Bronx, NY; 17) Department of Genetics, Washington University School of Medicine, MO; 18) The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, NY.

The global obesity epidemic underscores the importance of gaining a deeper understanding of the biology of weight regulation to develop better preventive strategies. Obesity has a strong genetic component, but lifestyle factors, such as physical activity (PA), may modify the impact of genetic susceptibility. Recently, a candidate-based gene-meta-analysis confirmed that the body mass index (BMI)-increasing effect of FTO is attenuated by PA. To identify novel loci that interact with PA on BMI or waist-hip ratio adjusted for BMI (WHRadjBMI), we performed a meta-analysis of 64 studies with genome-wide or metaobochip data, including up to 180,418 individuals of European descent. We standardized PA by categorizing it into a dichotomous variable (inactive vs. active) in each study. Overall, 23% of individuals were categorized as inactive. We pooled the results from individual studies using fixed effects meta-analysis of the SNP main effect and SNP×PA interaction. In the genome-wide meta-analysis of SNP×PA interaction alone, we identified a novel locus near CDH12 (P=3×10⁻⁸) which showed a BMI-increasing effect in the inactive group, but a BMI-decreasing effect in the active group. In the joint meta-analysis, 69 loci reached P<5×10⁻⁸ (43 loci for BMI, 26 for WHRandBMI), four of which have not been previously identified (near FLJ30838 or MRAS for BMI; near HHAT or PMS2P5 for WHRandBMI). For each of the 69 loci, the joint test association was driven by the SNP’s main effect, rather than by its interaction with PA. When we examined the SNP×PA interaction effects of the 69 loci, separately from main effects, the strongest interaction occurred at the FTO locus (P=5×10⁻⁸); the BMI-increasing effect was 29% smaller in the active than in the inactive group. In this sample of 180,418 individuals, we identify a novel locus near CDH12 interacting with PA on BMI and replicate the PA-interaction of FTO. Furthermore, while accounting for PA in the model, we identify four novel loci associated with BMI or WHRandBMI. We are currently performing a detailed follow-up of the loci to elucidate their potential functional roles. Our study may yield important novel insights into the role of gene-lifestyle interactions in the etiology of obesity.

404S

Association of physical activity with lower type-2 diabetes incidence is weaker in those with high genetic risk. Y.C. Klimentidis1, Z. Chen1, A. Arora2, C. Hsu2, O. Pedersen2, M.L. Grove3, N. Heard-Costa2, L.A. Cupples2, K. Young2, O. Qi2, K. Mordt1, K.E. North1, I.B. Borecki1, R.I.F. Loos1, L.A. Cupples2, The GIANT Consortium. 1) Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Denmark; 2) Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, NC; 3) RRHE Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, UK; 4) Department of Genetic Epidemiology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, Germany; 5) Department of Biostatistics, Boston University School of Public Health, MA; 6) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 7) Division of Population Health Sciences and Education, St. George’s, University of London, UK; 8) Department of Clinical Science, Genetic and Molecular Epidemiology Unit, Lund University, Sweden; 9) Department of Epidemiology, Harvard School of Public Health, MA; 10) Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Sweden; 11) Division of Preventive Medicine, Brigham and Women’s Hospital, Boston, MA; 12) Human Genetics Center, University of Texas Health Science Center at Houston, TX; 13) Department of Neurology, Boston University School of Medicine, MA; 14) Department of Twin Research and Genetic Epidemiology, King’s College London, UK; 15) Division of Statistical Genomics, Center for Genome Sciences & Systems Biology, Washington University School of Medicine, MO; 16) Department of Preventive Medicine, Albert Einstein College of Medicine, Bronx, NY; 17) Department of Genetics, Washington University School of Medicine, MO; 18) The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, NY.

We examined whether the association of physical activity (PA) with type-2 diabetes incidence differs according to several types of genetic susceptibility. In a large prospective cohort with 838 incident cases, we examined interactions of physical activity with 1) each of 65 type-2 diabetes-associated single nucleotide polymorphisms (SNPs), 2) a genetic risk score (GRS) comprised of all 65 SNPs, 3) two GRSs comprised of SNPs implicated in insulin resistance and insulin secretion, and 4) GRSs for fasting insulin and glucose. We find a significant interaction of physical activity and the type-2 diabetes GRS (p=0.019), suggesting a weaker protective effect of physical activity in those at high genetic risk. Based on the interactions observed with the insulin resistance GRS (p=0.044) and the fasting insulin GRS (p=0.060), it appears that this overall type-2 diabetes GRS interaction most likely occurs through genetic susceptibility to insulin resistance, as opposed to insulin secretion. Furthermore, this interaction was more pronounced in women (p=0.0046) than in men (p=0.47). No single SNP stood out as displaying a strong interaction with physical activity. We conclude that although physical activity appears to have an overall protective effect on type-2 diabetes, this putative effect is weakest in women with high genetic risk for type-2 diabetes and insulin resistance.

404M

Genome-wide scan for context-dependent marker SNP effects in coronary heart disease. S.M. Raj1, C.F. Sing2, G. Dyson3, A.G. Clark4, 1) Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY; 2) Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA; 3) Department of Oncology, Wayne State University, Detroit, MI, USA.

The existence of epistatic and genotype-environment interaction effects may uncover new genetic pathways that contribute to disease pathogenesis. The role of context in determining the utility of genomic information has been understudied. We hypothesize that loci not detected by a GWAS may have allelic effects in the context of subgroups of the population who have the highest risk to disease. We have carried out the first genome-wide scan to detect such context-dependent locus effects. We used a modified Patient Rule-Induction Method (PRIM) to assess which combinations of SNP marker and traditional risk factors identified subgroups with the highest cumulative incidence of coronary heart disease (CHD). The PRIM was applied to 84,000 autosomal SNPs and eight risk factors collected on 7,589 European-American adults from the Atherosclerosis Risk in Communities Study. We also carried out a GWAS on this dataset to evaluate similarities and differences in the genetic signals captured by the PRIM. The PRIM-based scan for context-dependent effects of marker SNPs identified approximately three times more hits than the GWAS, with females having a proportionally higher number of PRIM associations. These two approaches yielded very little overlap in associated genetic variants. We found that SNPs showing the strongest associations with CHD in a particular sub-group context defined by a traditional risk factor did not differ significantly between females and males, but did vary among contexts, with hypertension being the most common measure of context. These results suggest that consideration of context-based SNP associations with a complex disease may uncover new genetic pathways that contribute to disease pathogenesis. The role of context in determining the utility of genomic information has powerful implications for the clinical and population based management of the risk of a common disease having a complex multifactorial etiology.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1047T
Utility of the rhesus macaque (Macaca mulatta) as a novel genetic model for spontaneous human inflammatory bowel disease (IBD): sexual dimorphism and gene-by-sex effects on chronic diarrhea. A. Vinson1,2, M.J. Raboin1, K.D. Prongay1, E.R. Sundhaussen1, 1) Div. of Neuroscience, Oregon National Primate Research Center, Beaverton, OR; 2) Div. of Comparative Medicine, Oregon National Primate Research Center, Beaverton, OR.

Inflammatory bowel disease (IBD) is one of the most prevalent gastrointestinal diseases in the U.S., and is known to have a strong genetic component. Moreover, recent studies support the role of female bias in genetic risk for IBD. However, no animal model of spontaneous IBD currently exists that will support genetic investigation of this disease, including sex-specific genetic effects. Based on our observation that the rhesus macaque displays chronic diarrhea at a high prevalence in captive colonies, our goal was to explore the utility of this species as a genetic model of spontaneous human IBD. Of the 233 animals affected with chronic diarrhea in a captive, pedigreed population of Indian-origin rhesus macaques, we then scored animals based on the number of days of diarrhea, per the total number of days present in the colony. After accounting for differences in housing, medication, and covariates of age and sex, we found low, but statistically significant heritability (h² = 0.06, P = 4.7 x 10^{-10}; N = 1,676), and much lower, although statistically significant, heritability among males (h² = 0.16, P = 5.4 x 10^{-4}; N = 997). Consistent with the phenotypic sexual dimorphism suggested by covariate effects of sex (P-value = 1.5 x 10^{-7}), initial tests of gene-by-sex interaction confirmed the significant effects of sex on genetic variance influencing chronic diarrhea in this pedigree (P₀ = 0.02; test for P₀ = 1.0). Our findings provide initial support for the utility of the rhesus macaque as a novel genetic model for sex-biased effects on spontaneous human IBD.

1048S
Genome wide gene-vitamin D interaction analysis suggests potential role for melanoma related genes in Parkinson disease. L. Wang1,2, L. Maldonado1, D. Ritchie1, M. Evatt1, J.L. Haines1, M. Pericak-Vance1,2, G.W. Beecham1,2, E.R. Martin1,2, J.M. Vance1,2, W.K. Scott1,2, 1) Hussman Institute for Human Genetics, University of Miami, Miller School of Medicine, Miami, FL; 2) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miller School of Medicine; 3) Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA; 4) Department of Neurology, Emory University School of Medicine, Atlanta, GA; 5) Department of Epidemiology and Biostatistics, Case Western Reserve University, School of Medicine, Cleveland, OH.

Despite extensive genome-wide association studies (GWAS) in Parkinson disease (PD), much of the estimated genetic risk has yet to be detected. One source of the "missing heritability" may be accounted-for gene-environment interactions. Vitamin D (Vit D) deficiency has been associated with increased PD risk in multiple studies. As Vit D regulates expression of a wide array of genes, we sought to identify genes involved in modifying the association between Vit D and PD by analyzing genetic interactions with Vit D deficiency. We imputed up to 7.2 M SNPs in 477 PD cases and 430 controls, whose Vit D metabolites were measured in stored plasma samples using mass spectrometry. Vit D deficiency was defined as having plasma 25(OH)D concentration <20 ng/ml and was strongly associated with PD (Odds Ratio (OR) = 2.7, P = 0.0001). Joint tests of gene-environment interactions were conducted by comparing a full model containing SNP dosage, Vit D deficiency, an interaction term, and covariates age, sex, and sampling season to a restricted model with only Vit D deficiency and covariates. While joint tests produced genome-wide significant results at P < 5 x 10^{-8}, the most significant interaction terms were detected with SNPs in two genes associated with melanin production and risk of melanoma. This is intriguing given that PD is characterized by a loss of melanin-positive, dopaminergic neurons in the substantia nigra and an increased incidence of melanoma in PD patients has been reported in several studies. Further evidence for interaction was found at rs7312710 in FBRSL1 (P = 2 x 10^{-6}). The effect of this SNP on PD depended on Vit D status: the minor allele is associated with increased risk in Vit D deficient individuals (OR = 2.3, P = 0.0004) and with decreased risk in Vit D non-deficient individuals (OR = 0.7, P = 0.036). The second strongest interaction was found in C10orf11 with a similar pattern. We hypothesize that PD and melanoma share biological pathways (e.g. melanin production) that when perturbed modify risk to both diseases. This is supported by our previous pathway analysis using GWAS and microarray data suggesting that variation in the melanogenesis pathway contributes to PD risk. Our study demonstrates Vit D-gene interactions and nominates new genes that were not implicated in previous GWAS in PD. Our findings illustrate that examining gene-environment interactions can identify novel genes underlying the "missing heritability" in PD.

1049M
Smoking Then and Now: What Can the Aggregate of Genome-wide SNPs Tell Us About the Correspondence of Genetic Factors Influencing Cigarette Smoking Initiation Between Birth Cohorts. A.G. Wills1,2, G. Carey1,2, M.C. Keller1,2, 1) Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, USA; 2) Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA.

Twin studies have drawn attention to the importance of birth cohort when examining the genetic etiology of cigarette smoking behavior. For the smoking phenotype of initiation, we used whole-genome SNP data from a combined sample of 9189 European Americans that had participated in the Atherosclerosis Risk in Communities Study (ARIC) and the Multi-Ethnic Study of Atherosclerosis (MESA). The majority of these individuals were born between 1920 and 1950. Using Genome-wide Complex Trait Analysis1, we estimated the SNP heritability for individuals born before SNP h² = .16) and after SNP h² = .23) the combined sample’s median birth year of 1934. In the bivariate analysis, the correlation between the set of SNPs that explained variance in those born before 1934 and those born after was moderate (rₛ = .23) and significantly different from 1 (p = .048). Further, in a model that accounted for a gene x cohort interaction parameter, we found a marginally significant interaction (.11, p = .06) between birth year and the conglomerate of SNP effects. To better understand the differing genetic contribution based on year of birth, we divided our sample into birth decades, and the largest genetic effect was found for those born in the late 1930s. Due to disparity in sample sizes and the potential for unexamined gender effects, we remain cautious in interpreting this result. However, these findings highlight the potential for etiological heterogeneity among smokers of different eras.
1050T
Risk prediction and Type II Diabetes. N. Furlotte, S. Dandekar, R. Smith, N. Eriksson, D. Hinds. 23andMe, Mountain View, CA.

As the prevalence of Type II Diabetes (T2D) continues to increase worldwide, the ability to accurately assess the risk for developing this disease is becoming increasingly important in clinical practice. Accounts vary as to the clinical utility of genetic risk prediction models for T2D. Estimates of heritability vary widely, but it is generally accepted that environmental factors such as food consumption, exercise frequency and body mass index play a much larger role in the development of this disease than the genetic variations implicated through genome-wide association studies. Furthermore, the effects of genetic variations may only be apparent under particular environmental conditions - so-called gene-by-environment interactions. We explore these issues in a cohort of over 300,000 23andMe customers. We evaluate the predictive power of the most recent and robust genetic variations implicated in the etiology of T2D through association analysis and compare their predictive ability with environmental predictors related to individual behavior. In addition, we search for genetic variations exhibiting environmentally specific genetic effects and quantify the proportion of the total trait variation attributed to these gene-by-environment interactions.

1052M

1) Department of Cell Biology, Embryology and Genetics, Federal University of Santa Catarina, Florianopolis, Brazil; 2) Department of Genetics, Federal University of Pernambuco, Brazil; 3) GENYO; Centre for Genomics and Oncological Research; Pfizer; University of Granada; Andalusian Regional Government; Granada, Spain and Arthritis and Clinical Immunology Program, Oklahoma. 4) Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden.

Background: Systemic lupus erythematosus (SLE) as well as other autoimmune diseases show a strong gender bias, where the great majority of patients are women. Although controversial, the effect of estrogen has been implicated with SLE in several ways. However, the mechanism on how estrogen contributes to the development of SLE is still unknown.

Objective: To contribute to the knowledge of the mechanisms behind the strong sex-biased prevalence of SLE, we investigate the role of estrogen on the expression of one of the strongest associated gene with SLE, the interferon regulatory factor 5 (IRF5), in human immune cells. Material and methods: IRF5, as well as IRF3, IRF4, IRF7 and IRF9 expression was measured in PBMCs, LCLs, monocytoids and monocyte-derived macrophages from both male and female origin. Cells were treated with different concentrations of estrogen and gene expression was measured by quantitative PCR.

Results: We found that the initial levels of IRF5 in PBMCs were almost 2-fold higher in women that in men, although not reaching statistical significance. After 12 h in culture the IRF5 levels became roughly equal in both sexes, and further stimulation with estrogen lead to up-regulation of IRF5 levels in both PBMCs and LCLs in both women and men. No difference was seen for IRF3, IRF4, IRF7 and IRF9 expression, and none of the genes analyzed was up-regulated in LCLs, upon estrogen treatment, regardless of the gender. The later could be explained by the fact that we could not detect any expression changes after estrogen treatment in any of the cell lines.

Conclusions: We could show that in human PBMCs and monocytes from healthy individuals IRF5 expression can be regulated by exogenous estrogen. The effect in PBMCs was seen in both gender but was more pronounced in women. This feature might be specific to IRF5 since four other IRF genes tested did not show any up-regulation in these cells. This data could to some extent explain the sex bias in the development of SLE as well as other autoimmune diseases.

1051S
Investigate cytokine levels of cord blood samples in related to maternal allergic status. H. Tsai1,2,3, Y. Huang4, Y. Tsai2, Y. Ho4, W. Hisao5, J. Wang5,6.

1) Dept Pediatrics, Northwestern Univ Sch Med, Chicago, IL; 2) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Taiwan; 3) Department of Medical Genetics, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 4) Department of Pediatrics, National Cheng Kung University Hospital, Tainan, Taiwan; 5) Institute of Basic Medical Sciences, College of Medicine, National Cheng Kung University Hospital, Tainan, Taiwan.

Introduction: The increasing prevalence of early childhood allergic traits such as eczema, wheezing, asthma and allergic rhinitis has been a substantial public health burden in Taiwan and worldwide. Previous studies have reported certain kinds of cytokines and chemokines are involved in the pathophysiology of asthma and/or atopy. Based on our previous work, we hypothesize that maternal allergic status during pregnancy may play an important role influencing cytokine levels of their fetus, that is, may have an effect of immune dysregulation in-utero. Methods: A total of 55 maternal participants was recruited in this study. In addition to maternal biospecimen, we also collected their cord blood samples. We stimulated cord blood mono-nuclear cells (CBMCs) with lipopolysaccharide (LPS) and specific allergen from Der. Pteronyssinus (Derp), separately, and measured cytokine levels, specifically, IL-6, IL-8, IL-10, IL-23 and Tumor necrosis factor alpha (TNF-alpha), from the stimulated cord blood samples. We then applied Student’s t test and Wilcoxon rank sum test, separately, to examine whether examined cytokine levels differed related to maternal allergic status. Results: We examined 5 different cytokine levels in medium (not stimulated), LPS and Derp stimulated CBMCs in related to maternal allergic status. We only observed that there was borderline significance in TNF-alpha between maternal sub- jects with and without allergy (p=0.06). Beyond that, the results indicated that there was no difference in the examined cytokine levels between maternal subjects with and without allergy, no matter medium, LPS or Derp stimulated cord blood samples. Conclusions: The present study indicated cytokine levels of IL-6, IL-8, IL-10 and IL-23 in cord blood were not different in related to maternal allergic status. Further investigation will be warranted to better understand maternal influence on TNF-alpha levels in-utero.
A new GATK framework for RNA-seq variant discovery identifies differential A-to-I RNA editing in autistic brains. A. Eiran1,2, A. Levy-Moonshine1, R. Anner1, N. Milo1, E. Bachmati1, R. Zhang1, J.B. Lit1, D.M. Margulis1, L.M. Kunkel2, I.S. Kohane1. 1) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 2) Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA. 3) Department of Computer Science, Ben Gurion University of the Negev, Beer Sheva, Israel. 4) Department of Genetics, Stanford University, Stanford, CA; 5) Department of Genetics, Boston Children’s Hospital and Harvard Medical School, Boston, MA.

Adenosine-to-inosine (A-to-I) RNA editing is an epigenetic mechanism that fine-tunes synaptic function in response to environmental stimuli. Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder commonly associated with synaptic dysfunction. Targeted RNA sequencing studies have identified consistent A-to-I editing alterations of candidate genes in ASD, suggesting that A-to-I RNA editing may play a role in the disorder. However, transcriptome-wide editing profiles of individuals with ASD are still unknown, and the overall contribution of A-to-I editing alterations to the ASD phenotype remains to be discovered. Here we present a comprehensive framework for accurate SNP and indel detection in RNA-seq data, and apply it to conduct a first transcriptome-scale survey of A-to-I RNA editing in postmortem prefrontal cortex and cerebellum from individuals with ASD. The Genome Analysis Toolkit’s RNA-seq variant analysis framework is based on local de-novo assembly of RNA haplotypes. It begins with RNA-seq data pre-processing, including spliced alignments, duplicate marking, splicing and indel realignment, and the recalibration of base and mapping scores. It continues with de-Bruijn assemblies of variable regions, their likelihoods, splicing coordination, and isoform counts. It ends with variant filtering and the production of a SNP and indel call set in a VCF file. We evaluated the performance of this approach using whole exome sequences from the same tissues, and applied it for a first genome-scale comparative survey of A-to-I RNA editing in autistic brains. We compared cerebellar to prefrontal cortex editing levels, in synaptic vs. non-synaptic transcripts, across isoforms, genes, and gene families. We identified differential editing of bona fide autism genes and pathways, and characterized their relations to expression and splicing patterns. We then validated these findings using an independent deep targeted sequencing approach, mmPCR-seq. These results improve our understanding of the role of A-to-I editing in ASD and demonstrate the effectiveness of the GATK’s RNA-seq variant analysis framework.
1058M Genetic influences common to bronchial asthma and pulmonary tuber-
culosis present targets for intervention. R. C. McEachin1, M. B. Freidin2, E. Bragin3, V. Chertkov4, L. A. Konova1,2, D. C. M. S. Van Hout5. University of Vermont, Burlington, VT, USA; 2) Research Institute of Medical Genetics, SB RAMS, Tomsk, Russia.

Purpose: Bronchial asthma and pulmonary tuberculosis (TB) are both serious health conditions worldwide. Although asthma is an allergic disease and TB is an infectious disease, both show evidence of genetic influences on susceptibility. Interestingly, epidemiological data show their co-occurrence is rare, relative to the expected. Historically, TB caused more deaths in industrialized societies than any other disease over the past 200 years in Westernized societies. While TB is now rare in the progeny of survivors of this epidemic, asthma is prevalent. Equally, in populations that currently have high rates of TB, asthma is rare. One explanation for these observations is that differences in hygiene may influence differential TB and asthma rates. We hypothe-
size that there are also genetic influences shared across asthma and TB, including risk variants for one disease that are protective for the other. In this work, we test this hypothesis and develop a model that a) places the results into biological context and b) presents gene/drug targets for interven-
tion.

Methods: Gene expression is a primary cellular response to the environ-
ment, so we looked for genes differentially expressed (DE) in both asthma and TB, but in opposite directions. We assessed DE in T-cells from asthma patients vs controls, and from TB patients vs controls, as well as in lungs of non-infected TB-resistant vs TB-susceptible mice (fold change $> 2.0$ AND FDR $< 0.05$). We used GeneGo to model interactions among DE genes that responded in opposite directions in human TB and asthma, emphasizing transcriptional regulation as the mechanism of action. We used DAVID to assess enrichment for specific pathways or ontologies for the newly discovered 14 genes.

Results: Consistent with our hypothesis, 14 human genes were DE in oppo-
site directions in asthma and TB. The network shows how these 14 genes interact, and how a given variant could be a risk factor for one disease and protective for the other. Two genes in the network are known targets for 35 drugs currently in use for a wide range of disorders, including lung disorders. The mouse TB model also yields results consistent with the hypothesis, with “asthma” as the most significantly enriched annotation for the set of DE genes in FDR $< 1.7E-08$. Conclusion: The proposed hypothesis and novel hypothesis on how TB and asthma responsive genes could have opposite effects in predisposing the two diseases and presents immediate targets for follow on testing in asthma and/or TB.

1059T Identification of a Common Pathogenesis for Chronic Kidney Disease: Perspectives from Gene Ontology Analysis. W. Wu1, J. Guo1, R. Yen-
chek2, J. Abraham3, A. Cheung3, L. Jorde1. 1) Department of Human Genetics, University of Utah, Salt Lake City, UT, United States; 2) Department of Internal Medicine, University of Utah, Salt Lake City, UT, United States; 3) Department of Oncological Sciences, University of Utah, Salt Lake City, UT, United States.

Background: Chronic kidney disease (CKD) affects more than 13% of people in the United States, and is one of the leading causes of death. Genetic studies have provided molecular insights into CKD, but there is substantial locus heterogeneity. Yet the clinical manifestations of CKD are often shared across various etiologies. Is there a common pathogenesis that may explain these heterogeneous diseases? Here we employed a bioinformatics approach to identify one or a few common groups of genes for CKD pathogenesis. Methods: We first identified known causative genes associated with CKD curated by the Online Mendelian Inheritance in Man (OMIM). These genes were mapped to the Gene Ontology (GO) networks, and we traversed these trees to trace back to the root nodes. The GO terms dictated by these genes and their parent-child relationships were parsed. A statistical hypothesis test with theoretical hypergeometric distribution was carried out to test if the genes are significantly over-represented in each GO node. Simulation was performed to correct for multiple comparisons. These GO terms and their relationships were structured and further visualized for grouping for possible common pathogenic candidates. Results: From OMIM, we identified 84 causative genes for CKD. When mapping back to the GO networks, 2,728 GO terms were hit, among which 2,158 were in the biological process (BP) domain, 263 in the cellular component (CC) domain, and 306 in the molecular function (MF) domain. However, after testing for over-
representation using various statistics for expression values and fold changes, only 141 GO terms were statistically significant, among which 117 were in the BP domain, 23 in the CC domain, and 1 in the MF domain. Structure and visualization of these GO nodes and edges showed a diverse pattern across different modules of the network. Based on an alignment of different lists of ontologies to be ruled in or ruled out for future studies of CKD. We also provided a quantitative prior probability for each GO term that can be integrated into a computational model for pathogenesis prediction.

1060S Pathway Burden Analysis Identifies Genes Underlying Complex Human Limb Disorders. D. Alvarado1, G. Haller1, P. Yang1, C. Cruchaga2, M. Harms3, M. Dobs2, C. Gurpide2. 1) Orthopaedic Surgery, Washington Univ, St Louis, MO, USA; 2) Psychiatry, Washington Univ, St Louis, MO, USA; 3) Neurology, Washington Univ, St Louis, MO, USA; 4) Shriners Hospital for Children, St Louis, MO, USA; 5) Pediatrics, Washington Univ, St Louis, MO, USA.

Cluefoot, vertical talus and hip dysplasia are congenital birth defects with complex inheritance patterns. The genetic and morphological basis of all three lower limb disorders is poorly understood. Although each disorder describes a different clinical phenotype, their common occurrence within families and individuals suggests a common genetic developmental etiology. To identify the genetic risk factors associated with complex human limb disorders, we performed exome sequencing and genome-wide rare variant pathway burden analysis on 119 patients (88 families) with human lower limb malformations and 298 (249 families) unaffected controls. Using fam-
SKAT, we calculated a collapsed gene burden analysis for all novel coding variants not previously identified in dbSNP137. An enrichment of muscle development, extracellular matrix, hindlimb morphogenesis and chondrocyte differentiation GO term genes was found using the top 100 rank ordered genes. Genotype-wide collapsed pathway burden analysis of novel coding variants for all non-redundant GO term gene pathways revealed an association of lower limb malformation cases with collagen catalysis/metabolism and hindlimb morphogenesis GO term genes ($p=1.22x10^{-5}$ and $p=1.39x10^{-4}$). Based on these results, candidate genes associated with collagen or hind-
limb morphogenesis GO terms with collapsed famSKAT $p < 0.05$ were selected for re-sequencing in a larger cohort. Using Multiplexed Direct Genotyping (MDiGS), a newly described BAC capture approach for CNV and SNP/INDEL detection, we sequenced candidate genes in an additional 400 cases. Overall, these results suggest that cluefoot, vertical talus and hip dysplasia are highly heterogeneous disorders related to each other through an increased frequency of novel variants in genes responsible for hindlimb morphogenesis and extracellular matrix organization.
1062T
The Brainstorm project; a cross-phenotype analysis of 14 brain disorders by heritability-, constraint- and pathway-based methods, using genome-wide association data from 500,000 samples. V. Antilla1,2,3, W. Ripke1,2,3, K. Malig3,5, T.H. Pers4,5, K.-H. Fath2,9,7, D. Looman1,2,3, A. Briller4,1,8,9,10, P. Lee1, K. Kendler11,12, J. Scharf1, J. Smoller1, A. Palotie2,3,13, M. Daly4,6, J. Rosand6,7,10, B. Neale1,2,3,1, Analytical and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, MA; 2) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA; 3) Harvard Medical School, Boston, MA; 4) Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-University, Munich, Germany; 5) Munich Cluster for Systems Neurology (Synergy), Munich, Germany; 6) Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children’s Hospital, Boston, MA; 7) Psychiatric and Neuropsychological Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, MA; 8) Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA; 9) Division of Neurocritical Care and Emergency Neurology, Department of Neurology, Massachusetts General Hospital, Boston, MA; 10) J. Philip Kistler Stroke Research Center, Massachusetts General Hospital, Boston, MA; 11) Departments of Psychiatry and Human Genetics, Virginia Commonwealth University, Richmond, VA; 12) Virginia Institute for Psychiatric and Behavioral Genetics, Richmond, VA; 13) Institute of Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland.

Many neurological and psychiatric diseases have considerable comorbidity, and the strength of the etiological boundaries is a topic of active debate. However, studying these co-morbidities in the real world has proven difficult, and new approaches are needed. To this end, hypothesis-free methods using large genetic datasets may provide a way to study these co-morbidities by examining the genetic correlation through orthogonal approaches. Our aim is to provide insights into the molecular basis of these phenotypic co-morbidities and leverage large-scale genome-wide association data to reveal more of the underlying pathophysiology of an array of brain disorders in a statistically robust fashion. To this end, we have set up collaborations with a number of disease-specific genetics consortia, to obtain summary statistics from roughly 150,000 cases from 14 different diseases and corresponding in silico data. In this study, we set out to integrate the best available neurological and psychiatric genome-wide association studies (GWAS), which evaluates the significance of the multivariate (partial) effect on a given trait for all tissue-mark pairs using a three-pronged approach to identify obesity-related gene networks and tested whether these networks were dysregulated in SMS. First, we utilized Human Phenotype Ontology and identified 156 genes that are known to cause a significant increase in body weight. Second, we took a global genomics approach and identified 723 RA11-regulated genes by performing gene expression microarrays using HEK293 cells haploinsufficient for RA11 generated by siRNA knockdown. Third, we narrowed our focus to genes that are neuronally regulated by RA11 and performed an additional gene expression microarray using neuroblastoma cell lines haploinsufficient for RA11 generated by siRNA knockdown and found 549 differentially expressed genes. Using Ingenuity Pathway Analysis, we analyzed and integrated all three gene lists to generate obesity-related co-expression gene networks. Expression analysis of genes within each network in both SMS patient fibroblasts and across multiple SMS mouse model tissues (brain, liver, subcutaneous fat, and omental fat) showed that RA11 regulatory genes such as RAI1, CMKLR1, BBS4, and BMP2K bioinformatics analyses revealed multiple RA11 consensuses among the dysregulated genes, suggesting possible regions of direct regulation. To test one of the genes for responsiveness to RA11 transcriptional regulation, we employed a luciferase assay using the putative binding region in CMKLR1. Reporter assays revealed that RA11 acts as an enhancer of CMKLR1 expression, confirming expression data in both SMS patient fibroblasts and mouse tissues and the genetic interaction of RA11 at the consensus site. These data support RA11 as a regulator of multiple obesity-related gene networks and implicate RA11 as a contributor to early onset obesity.

1063S
RA1 is a multi-hit regulator of obesity gene expression networks. J.T. Alamo1, S.V. Mulligan1, R. Tahir2, J. Beach3, J. Cruz4, E. McMullen5, B. Burns6, S.H. Elsea4,5, 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA.

Smith-Magenis syndrome (SMS, OMIM 182290) is a complex neurobehavioral disorder characterized by a deletion of chromosome 17p11.2 and is due to haploinsufficiency of retinoic acid induced 1 (RAI1). Affected individuals have an array of phenotypic features that include sleep disturbance, varying degrees of intellectual disability, skeletal and craniofacial abnormalities, neurological, and behavioral issues, and obesity. Individuals with SMS are obese by early adolescence, and SMS mouse model work has further implicated RAI1 in both obesity and diet induced obesity. We used a three-pronged approach to identify obesity-related gene networks and tested whether these networks were dysregulated in SMS. First, we utilized Human Phenotype Ontology and identified 156 genes that are known to cause a significant increase in body weight. Second, we took a global genomics approach and identified 723 RA11-regulated genes by performing gene expression microarrays using HEK293 cells haploinsufficient for RA11 generated by siRNA knockdown. Third, we narrowed our focus to genes that are neuronally regulated by RA11 and performed an additional gene expression microarray using neuroblastoma cell lines haploinsufficient for RA11 generated by siRNA knockdown and found 549 differentially expressed genes. Using Ingenuity Pathway Analysis, we analyzed and integrated all three gene lists to generate obesity-related co-expression gene networks. Expression analysis of genes within each network in both SMS patient fibroblasts and across multiple SMS mouse model tissues (brain, liver, subcutaneous fat, and omental fat) showed that RA11 regulatory genes such as RAI1, CMKLR1, BBS4, and BMP2K bioinformatics analyses revealed multiple RA11 consensuses among the dysregulated genes, suggesting possible regions of direct regulation. To test one of the genes for responsiveness to RA11 transcriptional regulation, we employed a luciferase assay using the putative binding region in CMKLR1. Reporter assays revealed that RA11 acts as an enhancer of CMKLR1 expression, confirming expression data in both SMS patient fibroblasts and mouse tissues and the genetic interaction of RA11 at the consensus site. These data support RA11 as a regulator of multiple obesity-related gene networks and implicate RA11 as a contributor to early onset obesity.

1064M
Uncovering the pairs of tissue-epigenetic mark which are relevant for a trait. T. Bigdeli, D. Lee, B.P. Riley, A.H. Fanous, K.S. Kendler, S-A. Bacanu. Virginia Commonwealth University, Richmond, VA.

Background: For many phenotypes there is a strong working assumption regarding which tissues are biologically relevant to the underlying pathophysiology (e.g. brain tissues for psychiatric disorders). While network and gene-set methods often support these a priori expectations, they largely analyze each tissue univariately. However, only a multivariate analysis of all tissues can empirically assess the likelihood of these assumptions being true. Moreover, such a joint analysis might provide much stronger evidence for other-otherwise unexpected effects on tissue expression for those with unexpected tissue expression patterns.

Method: We develop a method based only on summary statistics from genome-wide association studies (GWAS), which evaluates the significance of the multivariate (partial) effect on a given trait for all tissue-mark pairs using a three-pronged approach to identify obesity-related gene networks and tested whether these networks were dysregulated in SMS. First, we utilized Human Phenotype Ontology and identified 156 genes that are known to cause a significant increase in body weight. Second, we took a global genomics approach and identified 723 RA11-regulated genes by performing gene expression microarrays using HEK293 cells haploinsufficient for RA11 generated by siRNA knockdown. Third, we narrowed our focus to genes that are neuronally regulated by RA11 and performed an additional gene expression microarray using neuroblastoma cell lines haploinsufficient for RA11 generated by siRNA knockdown and found 549 differentially expressed genes. Using Ingenuity Pathway Analysis, we analyzed and integrated all three gene lists to generate obesity-related co-expression gene networks. Expression analysis of genes within each network in both SMS patient fibroblasts and across multiple SMS mouse model tissues (brain, liver, subcutaneous fat, and omental fat) showed that RA11 regulatory genes such as RAI1, CMKLR1, BBS4, and BMP2K bioinformatics analyses revealed multiple RA11 consensuses among the dysregulated genes, suggesting possible regions of direct regulation. To test one of the genes for responsiveness to RA11 transcriptional regulation, we employed a luciferase assay using the putative binding region in CMKLR1. Reporter assays revealed that RA11 acts as an enhancer of CMKLR1 expression, confirming expression data in both SMS patient fibroblasts and mouse tissues and the genetic interaction of RA11 at the consensus site. These data support RA11 as a regulator of multiple obesity-related gene networks and implicate RA11 as a contributor to early onset obesity.
1066T
Network analysis to identify human genes influencing susceptibility to Mycobacterium tuberculosis and Nontuberculous mycobacteria infection. E.M. Lipner1,2, M. Strong1. 1) University of Colorado, School of Medicine, Department of Pharmacology, Aurora, CO, 80045, US; 2) National Jewish Health, Integrated Center for Genes, Environment, and Health, Denver, CO 80206, US.

Mycobacterial diseases, including tuberculosis (TB) and nontuberculous mycobacterial (NTM) infections, result in over 1.4 million deaths per year. In order to investigate and identify human genes that may influence control or susceptibility to mycobacterial disease, we compiled gene-disease relational information from various database resources. We then performed network and enrichment analysis to identify relevant and enriched human genes and pathways. We hypothesized that such an approach would provide a more comprehensive understanding of the biological pathways involved in mycobacterial infection and disease. We identified two gene sets that have not been previously identified via conventional genetic analytic methods. In order to identify genes associated with TB and NTM phenotypes, we searched three publically available databases; OMIM, HuGe Navigator, and CTD. Our search identified 50 TB-associated and 15 NTM-associated genes, with literature support. Based on these genes, we employed network and functional enrichment analysis using bioinformatic tools including Ingenuity Pathway Analysis (IPA) and DAVID in order to build and examine networks of TB-associated genes, NTM-associated genes, and identified TB and NTM gene sets. For TB, the highest ranking network included 14 query genes in a 35 gene network, including IL12B (MIM 161561), TLR1 (MIM 601194), TLR4 (MIM 603030), TLR6 (MIM 605403), TLR9 (MIM 605474), TRAP (MIM 606252), as well as the putatively associated genes IRAK (MIM 602563) and TICAM (MIM 606121). For NTM, e.g. BglB from M. tuberculosis, the network contained 14 query genes in a 35 gene network, including IL12B (MIM 161561), TLR2 (MIM 603028) and STAT1 (MIM 600555), and putatively associated genes NFKB complex (MIM 164401), ERK1/2 (MIM 176948) and MAPK4 (MIM 600208). This approach allowed us to investigate the interaction of genes and pathways that may imply similar susceptibility and control mechanisms for TB and NTM. Additionally, these networks suggest candidate genes that may have biological relevance to TB and NTM infection susceptibility for further study.

1066S
Reevaluating the clinical delineation of inflammatory bowel disease using genetic and subphenotype data. L. Jostins1,2,3,4, International IBD Genetics Consortium. 1) Statistical Genetics, Wellcome Trust Centre Human Genetics, Oxford, United Kingdom; 2) Medical Genomics, Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

Inflammatory bowel disease (IBD) is generally considered to consist of two major diseases, Crohn’s disease (CD) and ulcerative colitis (UC). However, both show significant clinical heterogeneity, and the delineation between the diseases remains the subject of significant controversy. Recent developments in IBD genetics (with >160 loci described to date) may help shed light on the clinical characteristics of this heterogeneous disorder.

Clinicians across 54 centers collated IBD subphenotypes on 34,819 patients (19,713 CD, 14,683 UC) according to the Montreal Classification. All subjects harbored genetic data generated by Illuminmpth and data was QC’d and collated at a central data coordinating center. The clinical data recapitulated known epidemiological results, such as the static nature of CD location (colonic vs ileal) over time, and the dependence of disease progression on location at diagnosis.

We confirmed known subphenotype loci (at NOD2 for CD location, p<10^{-34}, and MHC for multiple traits, p<10^{-22}), and identified a novel signal (at MST1 for age of onset, p<10^{-11}). Aggregate risk scores of other IBD risk loci showed association with all subphenotypes (e.g. CD vs UC score with CD location, p<10^{-12}), suggesting a role for most all IBD risk loci in disease subphenotype. Reverse regression suggested that the majority of the outcome subphenotypes in CD were explained by factors that are present at diagnosis (the location of the disease and the age of onset).

Different risk variants and scores predict different aspects of disease subphenotype (including age of onset and disease outcome), but the dominant axis appears to separate progressively restricted ileal disease from more stable colonic disease. UC and ileal CD share a high degree of genetic and clinical overlap. In this study, we identified a genetic profile that is distinct from, and intermediate between, the two. Blind rephenotyping by clinicians of outlying patients (e.g. CD patients with UC-like genetic profiles) demonstrated diagnostic blur at the extremes of this axis. 17% of outlying patients had a diagnosis of UC and ileal CD, compared to 8% of randomly selected patients (p<0.001). These results, combined with epidemiological data, suggest that IBD should be considered a spectrum of disorders, with Crohn’s colitis as a distinct disease intermediate between UC and ileal CD.

1067M
Integration of diverse genomic datasets identifies novel pathways and key regulatory genes for type 2 diabetes and related traits. L. Shu, V. Mäkinen, X. Yang. Department of Integrative Biology and Physiology, University of California Los Angeles, Los Angeles, CA.

Type 2 diabetes (T2D) is a highly prevalent metabolic disease and is among the top leading cause of death in developed countries. Recent genome-wide association studies revealed 119 susceptibility loci associated with T2D and its relevant traits. However, the molecular mechanisms underlying many of these loci remain unclear and together these loci only explain ~10% of the genetic heritability. In this study, we utilized an integrative genomics approach to compile genome and genomics datasets from humans and mice including 1) GWAS of T2D and related traits such as HbA1c and beta-cell function indices from DIAGRAM and MAGIC consortium, 2) knowledge-driven biological pathways, 3) data-driven, tissue-specific gene co-expression networks, and 4) data-driven, tissue-specific Bayesian networks. The integration of these diverse datasets allowed us to identify tissue-specific gene subnetworks (i.e., specific parts of gene regulatory networks) that are enriched for risk variants of T2D and relevant traits, as well as the key regulators of these subnetworks. Our results confirmed that subnetworks for beta-cell regulation and glucose transport are shared among the different T2D traits. More interestingly, we captured multiple trait-specific subnetworks such as incretin, adipocytokine and vesicle budding subnetworks for T2D, and adipocytokine and protein folding subnetworks for HbA1c trait, and bile acids and iron transport subnetworks for beta-cell function traits. Finally, by utilizing the gene network topology of the subnetworks, we pinpointed various tissue-specific key regulators, both known (e.g. ACSL1 in adipose and FASN in liver for T2D; ATF3 in islet for beta-cell function) and novel (e.g. BIRC5 and GLO1 in adipose and RGS12 in muscle for T2D; ZBTB11 in islet for beta-cell function). Our integrative analysis provides deeper insights into the mechanisms underlying T2D development and reveals novel key regulators, which could serve as potential targets for T2D intervention when experimentally validated.
1069S Clusters of urate transporter genes as genetic biomarkers in the early detection, diagnosis and prediction of gout. C. Chung1, 2, H. Tu1, A. Ko3, S. Lee4, H. Lai5, C. Lee6, C. Lee6, C. Huang7, Y. Ko2, 7. 1) Graduate Institute of Clinical Medical Science, China Medical University, Taichung, Taiwan; 2) Environment-Omics-Disease Research Center, China Medical University Hospital, Taichung, Taiwan; 3) Department of Public Health and Environmental Medicine, School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 4) Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig D-04103, Germany; 5) Division of Plastic Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 6) Division of Rheumatology, Allergy and Immunology, Department of Internal Medicine, Chang Gung Memorial Hospital-Kaohsiung Medical Center, Kaohsiung, Taiwan; 7) Department of Public Health, Kaohsiung Medical University; Kaohsiung, Taiwan; 8) Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; 9) Division of Immunology and Rheumatology, China Medical University Hospital, Taichung, Taiwan.

Objective: Gout is the most prevalent inflammatory arthritis in men. We aimed to evaluate the contribution of clusters of urate transporter genes and environmental factors to provide predictive biomarkers for gouty arthritis progression. Methods: We took a pathway based candidate gene approach to investigate the association between genetic variants in clusters of urate transporter genes and gout. Four single nucleotide polymorphisms (SNPs) in urate transporter pathway (rs2321142 in ABCG2 gene, rs3733591 and rs1014290 in SLC2A9 gene, and rs475688 in SLC22A12 gene) were genotyped in 558 male Taiwanese Hans. Results: Per copy increment in the Q141K [T], rs1014290 [A] and rs475688 [C] independently predicted risk of tophaceous gout (OR=3.33, P=9.76×10-7; OR=2.18, P=1.08×10-2; OR=0.995, P=0.47). Somewhat surprisingly, the complement and coagulation signaling cascade (CCSC) pathway was not significant (p=0.47). We compared these results to a non-Amish European-American dataset of 354 cases and 379 controls. The p-values for the HSBS, MAPK, and FCELL pathways were 0.995, 0.10, and 0.91 respectively while the CCSC pathway was highly significant (p=5.0×10-3). The results for the HSBS pathway are interesting since heparin sulfate plays a role as an inactivator in the complement alternative pathway through cleavage of C3b, also a target of complement factor H (CFH). From this preliminary analysis we observe signals in pathways previously shown to be associated with AMD. Interestingly, the complement pathway, one of the most highly significant pathways, is shown to be important in the early detection, diagnosis and prediction of gout.


Neural tube defects (NTD), including anencephaly and spina bifida, are among the most common and severe birth defects affecting nearly 1 in 1000 births. They are caused by partial or complete failure of neural tube closure in the rostro-caudal axis during embryogenesis. The etiology of NTD is complex involving genetic and environmental factors that remain largely unknown. While preconceptional intake of folic acid protects against 50-70%; of NTDs, they still affect thousands of families urging the need for improved prevention and counseling strategies. Here we suggest and approach aimed to evaluate the contribution of clusters of urate transporter genes and environmental factors to provide predictive biomarkers for gouty arthritis progression. Methods: We took a pathway based candidate gene approach to investigate the association between genetic variants in clusters of urate transporter genes and gout. Four single nucleotide polymorphisms (SNPs) in urate transporter pathway (rs2321142 in ABCG2 gene, rs3733591 and rs1014290 in SLC2A9 gene, and rs475688 in SLC22A12 gene) were genotyped in 558 male Taiwanese Hans. Results: Per copy increment in the Q141K [T], rs1014290 [A] and rs475688 [C] independently predicted risk of tophaceous gout (OR=3.33, P=9.76×10-7; OR=2.18, P=1.08×10-2; OR=0.995, P=0.47). Somewhat surprisingly, the complement and coagulation signaling cascade (CCSC) pathway was not significant (p=0.47). We compared these results to a non-Amish European-American dataset of 354 cases and 379 controls. The p-values for the HSBS, MAPK, and FCELL pathways were 0.995, 0.10, and 0.91 respectively while the CCSC pathway was highly significant (p=5.0×10-3). The results for the HSBS pathway are interesting since heparin sulfate plays a role as an inactivator in the complement alternative pathway through cleavage of C3b, also a target of complement factor H (CFH). From this preliminary analysis we observe signals in pathways previously shown to be associated with AMD. Interestingly, the complement pathway is not significant in this analysis; a major genetic risk factor for AMD. These data further support the hypothesis that the genetic architecture in the Amish is different than in the general European-descent population.

1071T The Genetic Architecture of Age-Related Macular Degeneration in the Amish. J.D. Hoffman1, J.N. Cooke Bailey1, R.J. Sardelli1, A.C. Cummings2, L.N. D’Aoust1, W.K. Scott1, M.A. Perciak-Vance1, J.L. Haines2, 2. 1) Center for Human Genetics Research, Vanderbilt University School of Medicine, Nashville, TN; 2) Department of Epidemiology & Biostatistics, Case Western Reserve University School of Medicine, Cleveland, OH; 3) Institute for Computational Biology, Case Western Reserve University, Cleveland, OH; 4) John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL.

The Amish are a genetically and culturally isolated founder population descended from Swiss and German Anabaptists who emigrated from Western Europe to North America in the 1700 and 1800s. The Amish also live a more homogeneous lifestyle, thus reducing the variability of environmental factors typically seen in complex diseases. Previous work demonstrated that the cumulative genetic risk across known age-related macular degeneration (AMD) loci is significantly lower in Amish AMD cases as compared to non-Amish European American cases, although the prevalence of AMD is similar. In this study we perform a pathway analysis using data from an Affymetrix 6.0 SNP chip to explore the potential role of functional pathways in AMD in the Amish population. Single-variant association analysis corrected for relatedness was performed on 92 Amish cases and 640 Amish controls using the MQLS program. Pathway analysis was performed on the results of the single variant tests using the Pathway Analysis by Randomization Incorporating Structure (PARIS) algorithm, and restricted to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database, which contains 199 defined pathways. Results of the AMD pathway analysis showed nominal associations with the heparan sulfate biosynthesis (HSBS; p=8.0×10-4), MAPK signaling (p=0.08), and FCELL pathways (p=0.001). We compared these results to a non-Amish European-American dataset of 534 cases and 379 controls. The p-values for the HSBS, MAPK, and FCELL pathways were 0.995, 0.10, and 0.91 respectively while the CCSC pathway was highly significant (p=5.0×10-3). The results for the HSBS pathway are interesting since heparin sulfate plays a role as an inactivator in the complement alternative pathway through cleavage of C3b, also a target of complement factor H (CFH). From this preliminary analysis we observe signals in pathways previously shown to be associated with AMD. Interestingly, the complement pathway is not significant in this analysis; a major genetic risk factor for AMD. These data further support the hypothesis that the genetic architecture in the Amish is different than in the general European-descent population.
1072S

Late-Onset Alzheimer’s disease (LOAD) is a complex neurodegenerative disease with a substantial genetic component. Despite the success of GWAS studies, much of the heritability of LOAD remains unexplained. We hypothesize that evaluation of the genetic and epigenetic mechanisms that influence gene expression in the brain will lead to identification of novel LOAD genes and improved understanding of vulnerable molecular pathways. We previously performed a gene expression GWAS that assessed genome-wide genotypes and mRNA levels of ~24,000 transcripts in two brain regions (temporal cortex and cerebellum) for ~200 AD subjects and ~200 subjects with other pathologies (Non-AD). Gene expression levels were measured using Illumina’s WG-DASL array and genotypes were generated using the Illumina Hap300 chip. We have now performed transcript profiling analysis comparing gene expression levels in AD vs. non-AD subjects in both brain regions, using linear regression carried out in R, and conducted pathway enrichment analysis on these data using the functionalities in Metacore. In addition, we have collected Cpg methylation data (methylome), using reduced representation bisulfite sequencing, on a subset of the temporal cortex samples (55AD, 55 NonAD) to evaluate the influence of DNA methylation on gene expression in the brain. DNA methylation levels at promoter regions were tested for association with gene expression levels using linear regression carried out in R. Transcript profiling analysis identified 743 targets in the cerebellum and 2839 targets in the temporal cortex (uncorrected p<0.05). Transcript profiling analysis identified more significant associations than expected by chance alone. Many of the top genes identified in this analysis were also found to be differentially expressed in the transcript profiling study indicating that epigenetic changes that influence gene expression may be associated with disease status. Further integrative analysis of genome-wide genotype, methylation and expression data in this cohort provides an opportunity to enhance our understanding of the role of genetic and epigenetic changes on brain gene expression and risk for LOAD.

1073M
Heterogeneous Network Link Prediction Prioritizes Disease-Associated Genes. D. Himmelstein, S. Baranzini. 1) Biological and Medical Informatics, University of California, San Francisco, San Francisco, CA; 2) Department of Neurology, University of California, San Francisco, San Francisco, CA.

The first decade of Genome Wide Association Studies (GWAS) has uncovered a wealth of disease-associated variants. To go beyond links we will be transcribing this information into a multiscale understanding of pathogenic variants, and increasing the power of existing and future studies through prioritization. We show that link prediction on heterogeneous networks—gene expression with multiple node- and edge-type annotations—given a heterogeneous network comprised of diverse information sources, our technique elegantly learns the importance of each connection type. We demonstrate the method’s distinctive performance and apply it to infer essential genes in a complex disease phenotype, multiple sclerosis (MS). First we constructed a network with 18 node types—genes, diseases, tissues, pathophysioligies, and 14 MsiIDB collections—and 19 edge types from high-throughput publicly-available resources. From this network we extracted features describing the topology between specific genes and diseases. Using a machine learning approach that relied on GWAS-discovered associations as positives during model training, we predicted the probability of association between each protein-coding gene and each of 23 diseases. These predictions achieved a testing AUROC of 0.814 and a 285-fold enrichment in precision at 10% recall. We quantified the performance of each network component and identified pleiotropy, transcriptional signatures of perturbations, pathways, and protein interactions as the most predictive domains. The full model greatly outperformed any individual domain, highlighting the benefit of integrating interactive and non-interactive networks. Our prediction browser, available at <a href=http://het.io>http://het.io</a>, that decomposes a prediction to show the contribution of each component. Since the entirety of GWAS knowledge is incorporated into every prediction, our method performed well for diseases with few seed genes—where most existing methods fail. We masked the largest MS GWAS from our network leaving only 15 MS-associated genes. Nonetheless, our method predicted the 36 genes discovered by the masked study with an AUROC of 0.78. Finally, we predicted 5 novel susceptibility genes for which (JAK2, REL, TNFAP3, RUNX3) achieved Bonferroni validation on the 9,772-case masked GWAS. Regions containing two of these genes were uncovered in a recent MS Immunochip-based study, highlighting our ability to identify the causal gene within a locus.

1074T

Background: Whole-transcriptome RNA expression analysis provides new insight into disease mechanisms. Interpretation of data from studies of differential gene expression in complex human diseases based on RNA-seq data is sensitive to study design and requires rigorous evaluation. Integration of genetic association data with transcriptome analysis facilitates cross validation of expression and GWAS data and may provide additional clue for identification of new candidate genes and gene-networks. We performed this analysis in a common autoimmune disease, rheumatoid arthritis (RA).

Materials and methods: RNA-seq was performed on whole-blood RNA samples from 12 healthy controls, 6 untreated RA patients, and 6 patients receiving standard treatment with Methotrexate. RNA-seq was done on the Illumina HiSeq 2000 platform with TruSeq library construction and selection for polyadenylated mRNA, with 28 M paired-end reads per sample. TopHat and Cufflinks software packages were used for RNA-seq data analysis. Pathway analysis was carried out with Ingenuity Pathway Analysis software.

Results: We address the expression of genes containing markers previously associated with RA in GWAS. Out of 109 genes, 39 genes were differentially expressed (DE) between any two of the three groups (Cuffdiff p<0.05) - naive RA patients, treated RA patients and healthy controls. These DE genes were grouped by pathway analysis into 4 functional protein networks, involving 75 additional “connector” molecules. Out of these 75 corresponding genes suggested by pathway analysis, the information for 37 were present in our RNA-seq data, and 15 were differentially expressed (Cuffdiff p<0.05). The majority of these new candidate genes (9/15) coded for proteins from a single network, where tumor suppressor protein p53 (TP53) was centrally connected with other 8 genes. Clustering samples based only on the expression of the 9 members of the TP53 network produced reasonable grouping of healthy controls and patients in our data. Conclusion: An integration of RNA-seq data with prior information from association research allowed us to infer new candidate genes for rheumatoid arthritis using pathway analysis. Our data implies, that TP53 could be a hub in the network that could contribute to the molecular pathological mechanisms behind rheumatoid arthritis.
1075S
Stratified enrichment test for dissecting colocализed genomic annotations to fine-map complex trait variants. G. Trynka1,2, B. Han1,2, K. Stowkolowski1,2, J.H. Westra1,2, H. Xu1,2, M. Daly1,2, X.S. Li1,2, R. Klein1, S. Raychaudhuri1,2,1. 1) BWH, Harvard Medical School, Boston, MA; 2) Broad Institute of Harvard and MIT, Boston, MA, USA; 3) Dana-Farber Cancer Institute, Boston, MA, USA; 4) Harvard School of Public Health, Boston, MA, USA; 5) Massachusetts General Hospital, Boston, MA, USA; 6) Icahn School of Medicine at Mount Sinai, New York NY USA; 7) University of Manchester, Manchester, UK.

Motivation: Enrichment tests offer the opportunity to translate genetic associations to biological mechanisms by integrating trait-associated variants with genomic annotations. However, genomic annotations highly colocализ with each other; regulatory elements cluster in the proximity of genes, DNase hypersensitive sites (DHS) colocализ with exons and transcription factors co-occur with enhancers. This is a critical issue because, if the colocализation of annotations is not accounted for, the observed enrichment might reflect inappropriate annotations rather than the most biologically relevant annotations. This might obscure fine-mapping efforts. Method: We developed a stratified enrichment test (GoShifter) that uses a local annotation shifting method to robustly quantify the enrichment of a given annotation while taking into account co-localized background annotations. Results: (1) We tested GoShifter using simulated GWAS SNPs, where associated SNPs tag loci at exons of DHSs, as these are known to colocализ. Standard test that does not correct for colocализation between these annotations resulted in 48% higher enrichment, while GoShifter reported enrichment in only 10% of time, reflecting appropriate false positive rate. (2) We applied GoShifter to SNPs influencing gene expression (eQTL) in whole blood. The unstratified test resulted in significant enrichment for DHS, genes, proximity to transcription start (TSS) (p<0.01), many of which colocализ. Stratified analysis with GoShifter revealed that the DHS enrichment best explains eQTL associations. (3) Finally, we applied GoShifter to NHGRI GWAS Catalog variants and observed that DHS were the most strongly enriched annotation (p<0.1) above and beyond TSS proximity or genes. H3K4me3 and H3K4me1 showed less significant, but independent signals, from DHS enrichment (p<0.05). We estimate that 20-40% of GWAS Catalog SNPs for traits in DHS, a much lower proportion than the widely reported 80%. Conclusions: GoShifter can dissect colocализed annotations to distinguish the true driving annotations from spuriously enriched annotations due to annotation correlations. We expect that it will be widely used to understand the molecular mechanisms underlying common traits associations.

1077T
Analysis of Endosomal Trafficking and Protein Recycling Genes in Parkinsonism. E.K. Gustavsson1,2, J. Trinh1, I. Guella1, H.E. Sherman1, C. Szu-Tuf1, S. Appel-Cresswell1, A. Stoeoss1, A. Rajput1, A.H. Rajput1, J.O. Aasly1,2, M.J. Farrer1. 1) Medical Genetics, University of British Columbia, BC, Canada; 2) Department of Neuroscience, Norwegian University of Science and Technology, Trondheim, Norway; 3) Department of Neurology, St. Olav’s Hospital, Trondheim, Norway; 4) Pacific Parkinson’s Research Centre, Department of Medicine (Neurology), University of British Columbia Vancouver, BC, Canada; 5) Division of Neurology, University of Saskatchewan and Saskatchewan Health Region, Saskatoon, SK, Canada.

Pathogenic mutations in genes involved in endosomal protein recycling, trafficking and lysosomal autophagy have been implicated in autosomal dominant late-onset Parkinson’s disease (PD). One key component of protein recycling and trafficking is the retromer complex, which consists of SNX-BAR proteins (SNX1, SNX2, SNX5 and SNX6), and a cargo recognition trimer (VPS26, VPS29 and VPS35). DNAJC13 also regulates endosomal trafficking and protein recycling by coating and uncoating clathrin-coated vesicles. In this study, we aim to assess rare missense variants (MAF<0.01) in these genes. Next Generation Sequencing (NGS) of VPS26, VPS29, VPS35, SNX1, SNX2, SNX5 and DNAJC13 were performed in 411 patients with PD, 224 atypical parkinsonism patients and in 371 controls. Patient and control subjects are unrelated and of multi-ethnic background. Identified rare missense variants (MAF<0.01) were Sanger sequenced and subsequently genotyped in two PD cohorts (Canada, 970 cases and 470 controls; Norway, 950 cases and 660 controls). For carriers with family history segregation analysis was performed when DNA was available. Seven variants were observed across VPS26A, VPS29 and VPS35, five of which were not seen in the control population. Furthermore, 22 variants were observed across the sorting nexin genes (SNX1, SNX2, SNX5 and SNX6), twelve of which only existed in patients. Twelve mutations were identified in DNAJC13, eleven of which were only seen in patients. Segregation analyses in families were performed for these variants. Subsequent genotyping of DNAJC13 mutations revealed p.Glu1740Gln and p.Leu2170Trp to be more frequent in patients (OR=7.65, p=0.041 and OR=4.12, p=0.007, respectively), and p.Pro336Ala, p.Val722Leu, p.Nsn855Ser, p.Arg1266Gln and p.Pro1431Arg to be more frequent in controls (OR=0.41, p=0.007, OR=0.41, p=0.013, OR=0.46, p=0.005, OR=0.35, p=0.007, respectively). These findings suggest that DNAJC13 mutations in patients with PD might impact RNA expression using STAR alignment and variant calling with VQSR by the GATK v3.1. In summary, we provide an in depth transcriptional presentation, t0. We conducted expression analysis with TopHat and Cuffdiff of a patient with SIRS suggestive of possible sepsis at emergency department presentation, t0. We conducted transcriptional sequencing of 132 patients with SIRS suggestive of possible sepsis at emergency department presentation, t0. We conducted expression analysis with TopHat and Cuffdiff links pipeline using HTSeq-Tools. A distinct transcriptional response was identified in sepsis survivors, sepsis nonsurvivors, and controls. Differentially expressed genes were analyzed to determine related biological pathways and functions. Expression patterns were also analyzed based on pathogenic organisms (staphylococci vs gram negative) and ability to use variable calls from RNA-seq to examine genomic variation that might impact RNA expression using STAR alignment and variant calling with VQSR by the GATK v3.1. In summary, we provide an in depth transcriptional analysis of patients with SIRS, and provide evidence of differential gene expression patterns capable of delineat- ing the three groups.
1079M
Liver-specific long non-coding RNAs and its association with liver disease. J. Fu1, B. Atanasovski1,2, M. van der Sijde1, S. Rensen2, P. Dreesen1, D. Zhernakova1, L. Dragoni1, V. Kramer1, L. Frank1, V. Kumar1, S. Withoff1, C. Wijmenga1, M. Hofker1, 1) Dept Genetics, UMCU, University of Groningen, Groningen, Netherlands; 2) Department of Paediatrics, Molecular Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Introduction: Long noncoding RNAs (lncRNAs) have been shown to play a significant role in the regulation of gene expression. However, the role of liver-specific lncRNAs in the pathophysiology of liver disease is largely unknown. Recently, an increasing number of studies have shown that lncRNAs can be involved in various biological processes and their dysregulation can lead to the development of complex human diseases. However, the role of liver-specific lncRNAs remains largely unknown.

We focused on the expression of 2,359 microarray probes that co-localize with IncRNA and assessed their correlation with NASH phenotypes. We further validated the correlation using quantitative RT-PCR (qRT-PCR) in a subset of 34 samples and investigated the association between lncRNA expression and genetic variation. Results: The correlation analysis based on microarray data identified 15 significant correlations for 12 IncRNAs at P<2.1x10^{-5}, corresponding to Bonferroni-corrected 0.05 level. We observed that one probe (rs1448304) was correlated with four different NASH phenotypes: ALAT (r=0.50, P=3.00×10^{-6}); Kleiner score (r=0.56, P=1.39×10^{-7}); Lobe inflammation (r=0.47, P=7.02×10^{-6}); Steatosis (r=0.55, P=1.39×10^{-7}). We performed qPCR analysis of the 196 trans-regulated genes annotated in the Database for Annotation, Visualization and Integrated Discovery (DAVID) and found that 56 (61%) of these genes have cis-eQTLs. Among these SNPs, 131 (associated with 231 genes, and representing 589 eQTL-gene pairs) also have cis associations with genes residing within 1 megabase. The mediation testing suggested that for 116 of these 131 eQTLs, the disease-associated SNPs may alter expression of remote (i.e., trans) genes via their cis associations with genes near the eQTL (P<0.005 for 1,000 permutations). The cis-eQTLs may provide clues regarding underlying mechanisms by which genetic variation influence disease susceptibility and can be subjected to genetic regulation. Further functional studies are warranted.

1080T
The Effects of Genetic Perturbation on Networks of Phenotypes in Complex Diseases. J. Chu1,2, P.J. Castaldi1,2, M.H. Cho1,2, B.A. Raby1,2,3, E.K. Silverman1,2, P.M. McQuillan1,2, C. Liu1,2,3, C. Yao1,2,3, A.D. Johnson1,2, P. Courchesne1,2, P. Misonson1,2, D. Levy3, 1) Channing Division of Network Medicine, Brigham & Women's Hospital, Boston, MA; 2) Department of Medicine, Brigham & Women's Hospital, Boston, MA; 3) Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA.

Complex diseases are often assessed using an array of disease-related phenotypic variables, which may have subtle, hidden relationships not captured by standard epidemiological analyses. Investigation of the potentially nonlinear relationships among phenotypes and between phenotypes and specific genetic variants is a challenging problem. To examine the effects of genetic perturbations on the relationships between disease-related phenotypes, we developed a novel network-based approach using the Gaussian graphical model (GGM) to infer the connections between each pair of phenotypic variables and create a phenotypic network based on partial correlations. First, for each SNP we constructed GGM phenotypic networks for groups of subjects defined by their genotypes. Then, for each pair of phenotypic variables, we test for differences in the network connectivity between different genotypes. We applied this method to COPDGene, a large, well-characterized set of smokers from a study of chronic obstructive pulmonary disease (COPD). We selected 10 key quantitative COPD-related phenotypes to build the COPDGene networks, and the effects of perturbations at 467 top loci (P < 0.001 from GWAS) were tested on 5,471 non-Hispanic White subjects. Although most SNPs have little or no effect on the phenotypic networks, there are many SNPs with significant effects on multiple interactions, disrupting a large part of network (15-20 out of 45 possible edges), and these SNPs can provide clues regarding underlying mechanisms by which genetic variant influence disease susceptibility and outcome. For example, in the networks characterized by rs7671167 (near FAM13A), which was previously associated with COPD in genome-wide association studies, we found that measures of lung function and emphysema were negatively correlated in the group with COPD-protected alleles but not correlated in the group with COPD-associated alleles, which suggests that FAM13A may lead to reduced lung function through mechanisms other than increased emphysema. Our results are based on the genetic effects on networks of multiple quantitative phenotypes within one complex disease and would not have been observed using traditional multivariate approaches. More importantly, the results can help understand the relationships between these disease-related phenotypes and provide novel insight into disease susceptibility, disease severity, and genetic mechanisms.

1081S
Mediation effect of eQTLs reveals trans-regulation of gene expression in complex disease traits. C. Yao1,2, R. Joehanes1,2,3, T. Huan1,2, B.H Chen1,2, X. Zhang1,2, C. Liu1,2, A.D. Johnson1,2, P. Courchesne1,2, P. Misonson1,2, D. Levy3, 1) The National Heart, Lung and Blood Institute; the NHLBI's Framingham Heart Study, National Institutes of Health, Bethesda, MD, USA; 2) Population Sciences Branch, National Institutes of Health, National Heart, Lung, and Blood Institute, Bethesda, MD, USA; 3) Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, MD, USA.

Expression quantitative trait locus (eQTL) analyses have shed light on transcriptional regulation in numerous complex diseases. An eQTL can be classified as cis or trans acting, based on its association with nearby or remote gene expression, respectively. Compared to cis-eQTLs, the underlying regulatory mechanisms responsible for remote effects of trans-eQTLs are still largely unknown. We hypothesize that expressed genes of cis-eQTLs can act as trans regulators. Using 11,672 published trait- or disease-associated SNPs for 427 complex traits, we explored the gene expression associations of trait-associated SNPs from GWAS that are eQTLs in 5,257 individuals from the Framingham Heart Study. At a false discovery rate (FDR) <0.05 (p<10^{-6}), we identified 177 trans-eQTLs that are associated with expression of 546 genes representing 1162 eQTL-gene pairs. Among these eQTLs, 131 (associated with 231 genes, and representing 589 eQTL-gene pairs) also have cis associations with genes residing within 1 megabase. Mediation testing suggested that for 116 of these 131 eQTLs, the disease-associated SNPs may alter expression of remote (i.e., trans) genes via their cis associations with genes near the eQTL (P<0.005 for 1,000 permutations). Examples of cis mediated effects on remote genes include RS1362082 on FADS1 that is mediated by eQTLs on chromosome 19 (FDR <0.05, p<10^{-6}) and SNPs on chromosome 6 (FDR <0.5×10^{-6}). The cis-eQTLs may provide clues regarding underlying mechanisms by which genetic variation influence disease susceptibility and can be subjected to genetic regulation. Further functional studies are warranted.
Beyond GWAS: Probing the landscape between pathway associations, genome-wide associations and protein-protein interaction networks in Chronic Obstructive Pulmonary Disease. M. McDonald, M. Mattheisen, Y. Liu, B. Harshfield, C. Hersh, P. Bakke, A. Gulsvik, C. Lange, T. Beaty, E. Silverman on behalf of the GenKOLS, COPDGene and ECLIPSE study investigators. 1) Channing Division of Network Medicine, Brigham & Women’s Hosp/Harvard Sch Med, Boston, MA, USA; 2) Department of Biomedicine and Centre for integrative Sequencing (iSEQ), Aarhus University, Aarhus, Denmark; 3) Division of Pulmonary and Critical Care Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA; 4) Department of Clinical Science, University of Bergen, Bergen, Norway; 5) Harvard School of Public Health, Boston, MA, USA; 6) Johns Hopkins School of Public Health, Baltimore, MD, USA.

In a traditional genome-wide association study (GWAS), thousands to millions of markers are tested; however, only a small fraction of markers reaches a level of genome-wide significance (GWS). As such, we sought to apply a traditional pathway approach and an integrative systems biology approach to maximize information from GWAS results to gain insight into Chronic Obstructive Pulmonary Disease (COPD). Specifically, we performed gene-based and gene-set associations, defined from gene-sets in the molecular signatures database (MSigDB), with FORGE software. For discovery, we analyzed GWAS results from two well-characterized COPD cohorts, COPDGene and GenKOLS. As replication, we used a third well-characterized COPD case-control cohort (ECLIPSE). Next, we used an integrative systems biology approach that integrates GWAS results with protein-protein interactions (PPI) to identify COPD disease modules with the R library dmGWAS. Using the traditional pathway approach, no gene-sets reached experiment-wide significance in either discovery population. Using the systems biology approach, we identified a consensus network of 10 genes identified in modules by integrating GWAS results with PPI that replicated in COPDGene, GenKOLS, and ECLIPSE. We found enrichment of members of four gene-sets among these 10 genes: (i) lung adenocarcinoma tumor sequencing genes, (ii) IL7 pathway genes, (iii) kidney cell response to arsenic, and (iv) CD4 T cell responses. Further, several genes, including KCNK3 and NEDD4L, have also been associated with pathophysiology relevant to COPD. In particular, KCNK3 has been associated with pulmonary arterial hypertension, a common complication in advanced COPD. Thus, we report a set of new genes that may influence the etiology of COPD that would not have been identified using traditional GWAS and pathway analyses alone.

Expression quantitative trait locus (eQTL) analysis is a well-established approach to discover novel biological mechanisms underlying genome-wide association study (GWAS) associations. Recent studies have shown that eQTLs may affect expression levels of multiple nearby and distal genes indicating that eQTLs may impact the transcriptional regulation of biological pathways. To search for common variants with systemic effects on expression levels, we developed an approach to identify eQTLs that regulate modules of co-expressed genes (transcriptional components). We used whole blood expression data from 814 genotyped, Caucasian individuals to construct 775 transcriptional components using independent component analysis (ICA). We annotated each transcriptional component by assessing enrichment of highly loading genes in a diverse set of pre-defined gene sets, including canonical pathways from the REACTOME database, Gene Ontology terms, and sets of genes that upon knock-out in mice result in the same phenotype. We then used individuals’ scores on each of the transcriptional components as phenotypes in an eQTL association analysis based on 2,933 independent SNPs that have previously been associated with diseases and complex traits at genome-wide significance. Finally, we applied our approach on Framingham Heart Study expression and genotype data from 2,324 unrelated individuals to replicate our findings. We identified 152 significant eQTLs, which associated with 106 independent transcriptional components (at false discovery rate <10%). Majority of them exerted systemic effects on multiple genes. The strongest eQTL was between a variant previously associated with inflammatory bowel disease (rs1363907; Spearman r=0.8, P<10^-500) and a transcriptional component enriched for immune cell response and inflammatory bowel disease at genome-wide significance. The variant was in high linkage disequilibrium with a splice site variant (rs2248374), which previously has been reported to result in non-sense mediated decay of the immune system related ERAP2 gene transcripts. The ERAP2 gene was the highest loading gene on the associated transcriptional component but more than 250 other genes loaded high as well (removal of the ERAP2 gene from the transcriptional component did not abolish the association). We replicated this association along with 33 other eQTLs in the Framingham data. Our work supports the idea that genetic variants can have broad biological effects.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Pathway analyses of extreme age-related macular degeneration phenotypes using whole exome sequencing data. R.J. Sardell, W.K. Scott, G. Wang, W. Cade, J.N. Cooke Bailey, M.D. Courtenay, S.G. Schwartz, J.L. Kovach, A. Agarwal, M.A. Brantley, J.L. Haines, M.A. Pericak-Vance. 1) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 2) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 3) Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL; 4) Department of Ophthalmology and Visual Sciences, Vanderbilt University School of Medicine, Nashville, TN.

Age-related macular degeneration (AMD) is a leading cause of vision loss among older adults in the US. Although multiple common genetic variants as well as several environmental risk factors for AMD have been identified, a substantial portion of the variance in disease risk and heritability remains unexplained and may reflect rare variants, gene-gene or gene-environment interactions. We used whole exome sequencing of 74 phenotypically extreme individuals to identify rare AMD related risk variants. We hypothesized that these extreme affected AMD individuals are more likely to carry rare variants of large effect. A genetic risk score was calculated based on 16 common (>1% minor allele frequency) single nucleotide variants (SNVs) known to be associated with risk of advanced AMD. We then selected 38 individuals with bilateral neovascular AMD, the lowest genetic risk score, and youngest age at examination, and 36 unaffected controls with no or very few small drusen, the highest genetic risk score, and oldest age at examination for sequencing. Variants were annotated for predicted function using SeattleSeq. We used a pathway based analysis approach to combine information on multiple genes within biological pathways using WebGestalt (WEB-based GEne SeT AnaLysis Toolkit). P-values from gene-based tests, run using SKAT were used to identify genes of interest for input (those with p<0.05). The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to identify pathways. Genes that contained SNVs used to define the risk score were excluded from analyses. Considering all common and rare SNVs, we did not identify any pathways significantly associated with the extreme phenotype. Considering rare (minor allele frequency <5%), damaging variants only, we identified one enriched pathway that remained significant (p=0.05) in our original SKAT analysis: FUT2, FUT3, FUT7 and ABO. Sphingolipid metabolism has been linked to retinal pigment epithelium cell inflammation and death, suggesting this pathway may be relevant to AMD disease process. Although FUT3 is nearby the previously identified AMD-associated gene, C3, variants at these two loci were not in strong linkage disequilibrium. These initial results highlight a pathway that may be associated with extreme AMD phenotypes and support the validity of this approach to identify new risk genes.

In contrast to Mendelian disorders, family-based linkage analysis has had limited success in identifying genetic variants relevant to complex traits such as cardiovascular disease and type 2 diabetes (T2D). Current efforts to identify high effect, low frequency coding variants have renewed interest in linkage methods. Data from an Illumina Infinium HumanExome BeadChip array, designed to assay less common coding variants, was available in the Diabetes Heart Study (DHS), a study of subclinical cardiovascular disease in T2D-enriched families (T2D prevalence 83.7%). Two-point linkage analysis was performed in 1190 European-American individuals from 468 families (1318 sibpairs). We hypothesized that linkage analysis might highlight variants and regions of interest for cardiometabolic traits, including measures of vascular calcified plaque, dyslipidemia, inflammation, renal function, and adiposity, and could be aligned with single variant association analysis results. Both two-point linkage and association analyses were performed using the Sequential Oligogenic Linkage Analysis Routines (SOLAR) program adjusting for age, sex, and T2D status. Across 26 traits, 174 LOD scores over 3.0 were observed; the highest was located intergenically on chromosome 4 (rs10939444, LOD=5.09, minor allele frequency (MAF)=0.35) with plasminogen activator inhibitor-1 levels. Notable evidence of linkage was observed for several traits on chromosome 6, including C-reactive protein (max LOD=3.97, rs2395730, MAF=0.43), high-density lipoprotein cholesterol (max LOD=3.53, rs4441254, MAF=0.23), albumin/creatinine ratio (max LOD=3.97, rs2395730, MAF=0.43), apolipoprotein B (max LOD=2.94, rs9370418, Ser387Pro in GFRAL, MAF=0.24), and waist circumference (max LOD=3.72, rs1264344, MAF=0.47). These variants were distributed over a 9 to 49 Mb region, depending on the traits included, suggesting a possible pleiotropic influence on related cardiometabolic traits. Linkage and association analyses also highlighted a variant in cadherin 13 (rs72807847, Asn39Ser, MAF=0.006) with evidence for both association (p=6.23 x 10^-6) and linkage (LOD=2.62) with abdominal aortic calcification. Cadherin 13 has previously been associated with blood pressure, adiponectin levels, and myocardial infarction risk. A combined linkage and association analysis approach may be effective for identifying loci, including loci for late onset diseases such as cardiometabolic traits, in sibpair studies without parental data.
1088M
Candidate high-penetration locus for myopia identified on chromosome 7 using linkage and family-based association analyses of exome chip data in 3 U.S. populations. J.E. Bailey-Wilson1, C.L. Simpson1, L. Portas1,2, F. Murgia1,2, S.S. Yee3, E.B. Ciner4, D.D. Stambolian5. 1) Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD; 2) Institute of Population Genetics, National Research Council of Italy, Sassari, Italy; 3) Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA; 4) Pennsylvania College of Optometry, Salus University, Elkins Park, PA.

Myopia, the most common cause of visual impairment worldwide, is one of several related refractive error traits and is thought to have both environmental and genetic causes. Current understanding of the genetic basis of myopia, based on genetic linkage and genome-wide association studies, suggest that there are many genetic loci that influence variation of mean spherical equivalent (MSE) and risk for myopia (MSE ≤ -1Diopter). Some of these loci appear to have common risk alleles with smaller effects on trait heritability while other loci may have rarer variants with higher penetrance and larger effects on trait heritability. The current study analyzed Exome Chip (Illumina WGHum-ExomePlus) data in extended pedigrees from Ashkenazi Jewish (64 families), Old Order Amish (40 families), and African American (103 families) ethnic groups where multiple individuals were diagnosed with myopia, in a search for higher penetrance susceptibility alleles. This genome-wide chip includes a set of 4761 GWAS tag SNPs with common minor allele frequencies (MAF), 3468 ancestry informative SNPs, > 230,000 SNVs (some with more rare MAFs) that occur in exons and 30,000 custom SNVs that were chosen to provide denser coverage of chromosomal regions that had been previously reported to show linkage or association to myopia and/or MSE. After extensive quality control procedures, allele frequencies of all variants were estimated from pedigree founders in each ethnic group separately. Parametric linkage and family-based association analyses were performed in each ethnic group separately using the PARAMLINK and FBAT programs, respectively. All three sets of families showed evidence of linkage of myopia to the same region of chromosome 7p (maximum two point LOD scores were 3.38 at 54.9cM, 2.45 at 56.9cM and 3.35 at 44.1cM in the Ashkenazi, Amish and African-American samples). Both the Ashkenazi and Amish families showed evidence of association in this linkage region to variants in the gene Homo sapiens oxysterol binding protein-like 3, OSBPL3 (different variants in each population, p=2.51E-6 at 42.32cM and 2.76E-6 at 42.18cM respectively). OSBPL3 encodes a protein that is implicated in the regulation of cell adhesion and organization of the actin cytoskeleton. These results give strong evidence of a myopia susceptibility locus on chromosome 7p15. Follow-up of variants in the linkage regions is ongoing to determine which variants best explain the linkage signal in each linked family.

1089T
Linkage-based Analytical Approaches with GWAS Data to Localize Variants Underlying Complex Traits. N.D. Palmer1,2, J.N. Hellwege2, K.D. Taylor1, J.R. Norris3, C. Lorenzo4, J.L. Rotter5, C.D. Langefield6, D.W. Boden2,1. 1) Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC; 2) Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC; 3) Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver, Aurora, CO; 4) Department of Medicine, University of Texas Health Science Center, San Antonio, TX; 5) Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA; 6) Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC.

In the era of Next Generation Sequencing, novel approaches will be needed to assess the contribution of low frequency, functional variants postulated to explain the missing heritability in common complex disease. We have computed two-point linkage analysis of genome-wide association study (GWAS) and exome chip data in the IRAS Family Study Hispanic American cohort in an effort to map low frequency, high impact variants contributing to novel, diabetes-related phenotypes. This approach reduces the stringent multiple comparison penalty in association testing by highlighting linked genomic regions for focused association and burden testing, with the advantage of being easily aligned. Analysis of 687,095 SNPs across six diabetes-related quantitative traits in 88 families (1034 individuals) revealed 1170 nominal (LOD>3) and 19 significant (LOD>5) signals. Consistent evidence of linkage was observed for acute insulin response (AIR; beta-cell function) on 1q24.3-25.2 (LOD=5.1-6.4). Among the genes implicated in association of common and rare variants was RALGPS2/ANGPTL1 (rs86995283, P=4.2x10⁻⁴) although family-based burden testing did not enhance the evidence of association (P=0.20). A more defined interval on 2p13.1 was also linked to AIR (LOD=5.3) encompassing ALMS1 which underlies Alstrom Syndrome; an autosomal recessive syndrome including insulin resistance, hyperinsulinemia, and type 2 diabetes. Analysis of common and rare variants revealed association with a missense SNP (S2027P, P=0.0042). Family-based burden testing weighted for rare variants more significantly supported these results (P=1.5x10⁻⁶). Among other phenotypes, insulin sensitivity (S1) was linked to 7q11.23 (1CA; LOD=3.1). Regional evaluation of candidates in the interval using association analysis revealed signals in CALN7, a member of the calmodulin family. Association analysis revealed common variants associated with S1 (rs844762, P=0.0091) while only a single rare variant was captured by the exome chip. Family-based burden testing at this locus revealed nominal association (P=0.048) likely due to down weighting of common variants. Family-based linkage is a powerful approach in contemporary array-based genetic datasets to prioritize genomic intervals for association analysis and burden testing. Whole genome sequence data should add to the power of this approach. Thus, a combined linkage and association testing framework using high density genetic data has the potential to reveal novel loci.

1. The Wellcome Trust Centre for Human Genetics, Oxford, UK; 2. Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands; 3. Center for Applied Genomics, The Children’s Hospital of Philadelphia, PA; 4. Population Health Research Institute, St George's University of London, London, UK; 5. The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 6. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; 7. Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; 8. Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; 9. Center for Biomedical Research on Internal Medicine, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, Lausanne, Switzerland; 10. COPSSAC and the Danish Pediatric Asthma Center, Copenhagen University Hospital, Copenhagen, Denmark; 11. Helmholtz Zentrum München, Munich, Germany; 12. School of Women’s and Infants’ Health, The University of Western Australia, Perth, Australia; 13. Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK; 14. Life Science Institutes, National University of Singapore, Singapore, Singapore; 15. University of Leicester, Leicester, UK; 16. Dept of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands; 17. Department of Physiology, Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland; 18. Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland; 19. Department of Genetics, University of Oxford, Oxford, UK; 20. Department of General Medicine, School of Public Health, Imperial College London, London, UK; 21. University of Exeter Medical School, Royal Devon and Exeter Hospital, Exeter, UK; 22. Oxford Centre for Diabetes, Endocrinology & Metabolism, Oxford, UK; 23. Department of Biostatistics, University of Liverpool, Liverpool, UK.

Lower birth weight (BW) is associated with increased risk of future type 2 diabetes (T2D) and cardiovascular disease. Based on HapMap 2 imputed European-genome-wide association (GWAS) studies, we predicted 7 loci associated with BW, of which two (ACDYS, CDKAL1) have been implicated in T2D and one (ADRB1) in hypertension. Here we report analyses based on an increased sample size (55,729 full term singletons) from multiethnic groups (22 European and 6 non-European of African-Americans (AA), Philippines, Moroccon, Surinamese, Turkish and Chinese descent) and imputation up to 20.8M SNPs from the more dense 1000 Genomes Project reference panel. Using inverse-variance fixed-effects meta-analysis, GWAS summary statistics between standardized sex-specific Z-scores of BW and each SNP were combined across studies and complete 8 months less education. Contrary to earlier reports in much smaller samples, no evidence of an association was seen for the burden of homozygosity on blood pressure traits or LDL cholesterol, or on other cardiovascular risk factors tested (classical lipid, glycaemic and adiposity traits). Even for the traits where we observed an effect, we estimate the proportion of phenotypic variance explained in an outbred European population to be less than 0.001. We therefore consider it unlikely that population cardiovascular health or the Flynn effect can be materially explained by background inbreeding or secular changes in the level of outbreeding. Inbreeding depression is predicted to arise in evolutionary fitness-related traits. Thus, our study suggests that increased stature and cognitive function (as measured by educational attainment and the general cognitive factor, g). In all four cases the effect sizes were small (-0.0075, -0.0012, -0.00126, -0.00160 phenotypic standard deviations per megabase of SROH). The effect of inbreeding on fitness traits in plants and animal populations has been studied since the time of Darwin. Close inbreeding is also associated with increased risk of Mendelian diseases in humans, however the effect of the distant inbreeding found in present day populations on complex traits is less well understood. Our study analysed 16 health-related quantitative traits in up to 350,000 individuals from multiple continental ancestries in 102 cohorts. Within each ethnically homogeneous sub-cohort, we modelled the effect on each trait of increased inbreeding as measured by the summary statistic for genome-wide association (GWAS) studies. For each liability trait we adjusted for age, sex, and at least 3 principal components of genetic variation. We then used fixed effect inverse-variance meta-analysis to combine results. We found that SROH is associated with reduced stature (as measured by height and reflected in a measure of fund function), and reduced cognitive function (as measured by educational attainment and the general cognitive factor, g). In all four cases the effect sizes were small (-0.0075, -0.0012, -0.00126, -0.00160 phenotypic standard deviations per megabase of SROH).

John P. Hussman Institute for Human Genomics, University of Miami, Miami School of Medicine, Miami, FL; 2) Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland OH.

As the population of industrialized nations ages, the health problems associated with aging will become a major public health concern. The Amish of Indiana and Ohio provide an ideal population to study aging because of their geographic stability and cultural isolation that minimizes environmental variability. The pedigree of the Amish is too large and complex for conventional linkage (e.g., lod score) analysis. For linkage analysis in the Amish, the pedigree must be cut into smaller pedigrees, which may result in false negative results. An alternative approach is to use long range chromosomal phasing (LRP), which allows the detection of segments shared identical-by-descent (IBD) derived from dense SNP genotypes; this avoids the computational complexity of conventional linkage analysis approaches. Population-based linkage using LRP was assessed for successful aging (SA), a composite phenotype consisting of lack of cognitive impairment, lack of depression (as determined by the Geriatric Depression Scale), high performance on four scales: Activities of Daily Living, Instrumental Activities of Daily Living, Nagi (a test of musculoskeletal function), and Rosow-Breslau (a test of lower-extremity function). 74 individuals met all of these criteria and formed primarily cousin-pair relationships. Amish participants (n=813) were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0. Missing genotypes were imputed in BEAGLE, and IBD segments of at least 250kb were identified using the GERMLINE unphased algorithm. The configuration of these IBD states was quantified for each as the proportion of affected pairs IBD for the marker. Statistical significance for each marker was determined by repeated sampling without replacement of an equivalent number of individuals from the population and assessing their IBD configuration. The p-value was the proportion of samples where the IBD configuration was as or more extreme than that of the successfully aged individuals. 

Suggestive evidence of linkage (p<1.7×10^{-3}; Lander and Kruglyak 1995) was determined by repeated sampling without replacement of an equivalent number of individuals from the population and assessing their IBD configuration. The p-value was the proportion of samples where the IBD configuration was as or more extreme than that of the successfully aged individuals. Suggestive evidence of linkage (p<1.7×10^{-3}; Lander and Kruglyak 1995) was determined for 46025X : Even oxidative stress, making it an excellent candidate gene in the aging process.
Linkage and association mapping for osteoarthritis progression in the Genetics of Generalized Osteoarthritis Study. M.S. Yau, L.M. Yerges-Armstrong, J.R. O’Connell, R.D. Jackson, M.C. Hochberg, B.D. Mitchell, V.B. Kraus. 1) University of Maryland, Baltimore, MD; 2) The Ohio State University, Columbus, OH; 3) Duke University, Durham, NC.

Osteoarthritis (OA) is the most common form of arthritis and is a major cause of morbidity. To identify genetic risk factors associated with structural knee OA progression, we undertook a genome-wide linkage scan using data from the longitudinal component of the Genetics of Generalized Osteoarthritis (GOGO) study (N=1,245, 79% women). We defined knee OA progression as worsening of Kellgren-Lawrence (KL) grade, osteophyte grade, or joint space narrowing grade in individuals with radiographic evidence of OA (KL grade=2 or 3) at baseline (N=703). Individuals who progressed to total joint replacement were also considered cases. We have previously reported OA progression to two chromosomal regions, Xp21.25 (LOD=4.3, P<0.01) and 18q21 (LOD=2.0, P<0.01). To fine map these regions, we conducted association analyses using markers genotyped on the Illumina 550K platform and replicated our most significant associations in the Osteoarthritis Initiative (OAI), a multi-center natural history study of knee OA (N=1,682, 58% women). Within the X chromosome linkage region, the most significantly associated SNP was rs437033 (OR=1.6, \textit{P}=9.6E-5), which is located in SLC25A43. This did not replicate in the OAI, although the most significant SNP across the chromosome (rs6528082; OR=1.8, \textit{P}=9.5E-5) located on Xp22 between SMS and PHEX showed nominal association with OA progression in the OAI (OR=1.5, \textit{P}<0.03). Within the chromosome 18 linkage region, the most significantly associated SNP was rs470497 (OR=1.6, \textit{P}=6.1E-4), located near GAREM. This SNP did not replicate in the OAI, although the most significant SNP across the chromosome on 18q12 (rs6680959; OR=1.7, \textit{P}=5.8E-5), which is closest to GAREM. Interestingly, though not in LD with rs6680959, another SNP in this gene, rs76809900, was one of the most highly associated SNPs with OA progression in the OAI (\textit{P}=3.2E-7). GAREM encodes a GBM2 associated regulator of MAPK1 that is involved in the epidermal growth factor-receptor-mediated signaling pathway. In summary, while our association studies did not identify a replication of related genetic risk factors on chromosome 18, \textit{i.e.}, GAREM. Further bioinformatics analyses will be needed to contribute the association of these SNPs to the linkage regions and the specific role that these genetic markers may play in the pathogenesis of OA progression.


Insulin resistance is linked to high risk of type 2 diabetes, hypertension and coronary artery disease but the underlying mechanisms are poorly understood. Studies of monogenic forms of insulin resistance, however, have established mechanistic paradigms of potential relevance to the common form of the disease. Genome wide association studies (GWAS) have found 19 common genetic variants associated with fasting insulin based measures of insulin resistance but the variants have subtle effects and most have unclear mechanisms. We aimed to combine the distinct characteristics of monogenic forms of insulin resistance and findings of GWAS to get mechanistic insights. We used results from GWAS of 8 metabolic traits that can distinguish between different forms of monogenic insulin resistance and used hierarchical clustering to group the 19 insulin resistance variants. We analysed the effect of each group against type 2 diabetes (12,171 cases), coronary artery disease (40,365 cases) and blood pressure (69,828 individuals). Hierarchical clustering identified a group of 11 variants among 19 insulin resistance variants associated with a metabolic profile resembling rare form of insulin resistance secondary to lack of adipose tissue under the skin – lipodystrophy. Collectively the 11 insulin resistance risk alleles were associated with higher triglycerides (\textit{p}=0.018; \textit{p}=4\times10^{-13}), lower HDL-C (\textit{p}=0.020; \textit{p}=7\times10^{-37}), greater hepatic steatosis (\textit{p}=0.021; \textit{p}=3\times10^{-9}), higher alanine transaminase (\textit{p}=0.002; \textit{p}=3\times10^{-9}), lower SHBG (\textit{p}=0.010; \textit{p}=9\times10^{-13}) and lower adiponectin (\textit{p}=0.015; \textit{p}=2\times10^{-9}). The same risk alleles were associated with lower BMI (per-allele \textit{p}=0.008; \textit{p}=7\times10^{-9}), and increased visceral-to-subcutaneous adipose tissue ratio (\textit{p}=0.015; \textit{p}=6\times10^{-4}). Individuals carrying \textit{>17} fasting insulin raising alleles were slimmer (0.30 kgm\textsuperscript{-2}) but at higher risk of insulin resistance and metabolic traits. Our results provide genetic evidence for an as yet unrecognized association between reduced fat deposition and the risk of insulin resistance and metabolic traits. Our study gives an example of how monogenic studies can complement GWAS to understand mechanisms underlying complex traits/diseases.
Copy number of the salivary amylase gene AMY1 modulates serum amylase levels and is associated with the metabolic profile. J.S. El-Sayed Moustafa1, T. Martin2, M. Beaumont2, L. Yengo1,4,5, P. Maboudou1, C. Menni2, T. Brousseau6, T.D. Spector2, P. Froguel1,3,4,5, M. Falchi1,2. 1) Department of Genomics of Common Disease, Imperial College London, London, London, United Kingdom; 2) Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK; 3) CNRS UMR 8199, Lille Pasteur Institute, Lille, France; 4) Lille 2 University, Lille, France; 5) European Genomics Institute for Diabetes (EGID), Lille, France; 6) CHRU de Lille, Pôle de Biologie-Pathologie-Génétique, UF Biochimie automatisée, Lille, France.

We have recently shown low AMY1 copy number to be strongly associated with increased body mass index (BMI) and risk of obesity (Falchi et al, Nature Genetics, 2014), suggesting that salivary amylase copy number may play an important role in energy balance and metabolism. We have measured both AMY1 genomic copy number and salivary serum amylase levels in 219 female twins from the UK TwinsUK cohort. Salivary amylase levels were highly heritable (h² = 0.32; P < 2.2 x 10⁻¹⁶), BMI was negatively associated with salivary amylase levels (β = -0.062; P = 4.15 x 10⁻⁴) and also inversely associated with fasting insulin levels (β = 0.08; P = 0.02). Association analysis with blood metabolomic profiles generated using mass spectrometry in the TwinsUK sample suggest that genetically-determined low salivary amylase levels are accompanied by dysregulation of lipid and glucose metabolism. Alterations observed in lipid metabolism, bile acids and amino acid catabolism may also reflect changes in the gut microbiome. Salivary amylase is a digestive enzyme involved in the metabolism of dietary starch. Our analysis of serum amylase levels suggests low salivary amylase levels to be a genetic trait associated with increased risk of metabolic abnormalities. In order to further investigate its role in human metabolism, we are currently expanding our sample size by measuring serum amylase levels in a further 1,500 subjects from the TwinsUK cohort as well as over 2,000 subjects from the French DESIR cohort for whom AMY1 copy number measurements and metabolomics data are also available.

**1097M**

Accurate molecular prediction in inflammatory bowel disease. H. Huang1,2, J.B. Essers2, B.M. Neale1,2, S. Ripke1,2, M. Vu+, R. Xavier2, A. Ananthakrishnan5, P. Fleschner1, M. Dubinsky2, T. Haritunians3, S.R. Targan4, D.P.B. McGovern5, M.J. Daly1,2. 1) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 2) Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA; 3) Inflammatory Bowel Disease Center, Children’s Hospital Boston, Boston, MA; 4) Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA; 5) Gastrointestinal Unit, Massachusetts General Hospital, Boston, MA; 6) F. Widjaja Family Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA; 7) Department of Pediatrics, Cedars-Sinai Medical Center, Los Angeles, CA.

Inflammatory bowel diseases (IBD) are chronic inflammatory diseases of the gastrointestinal system. Ulcerative colitis (UC) and Crohn’s disease (CD) are two major types of IBD. While some IBD therapies are universally effective, others show greater clinical efficacy in either CD or UC. Accurate classification of IBD can lead to optimal treatment choices and reduce disease complications and relapses. Due to the heterogeneous presentation of IBD, diagnosis can be challenging in up to 20% of patients using the conventional classification. Here we propose a molecular based risk model as a novel clinical tool to discriminate CD from UC by aggregating genetics, serology and tobacco smoking information.

The genetics component includes polygenic risk scores for CD and UC using independently associated SNPs from 163 IBD loci [Jostins et al]. IBD-associated serum immune responses to ASCA, ANCA, anti-Cib1, anti-OMPC and anti-I2 were generated by ELISA. Smoking status at diagnosis was uniformly collected and categorized as never, ex-smoker and current. The study cohort, including 1363 CD and 501 UC patients, was split evenly into training and validation groups. The overall risk model achieved an area under the curve (AUC) of 0.80 for discriminating CD from UC. Individually: genetics AUC=0.72; serology AUC=0.79; and smoking AUC=0.55. This genetics model outperformed the genetics component of the commercially available IBD sgI diagnostic by PROMETHEUS (AUC=0.55). Serology alone showed the best AUC but is uninformative in over half of patients. Importantly genetics performs similarly well when serology is uninformative, establishing that genetics and serology provide independent information. We identified 42 patients who developed ‘denovo’ CD after surgery with a UC diagnosis. We assigned our CD risk scores to these patients and correctly classified 26 (62%) of them. We achieved the same success in other cohorts using only the genetics component. Out of 31 patients whose diagnoses changed from UC to CD, our model gave correct diagnoses to 24 (77%). In addition to diagnosis, the risk score from our model was also associated with a younger age of onset in CD (p=3E-12).

In conclusion, we have shown that our molecular-based model combining genetics, serology and smoking information could complement current diagnostic strategies and is particularly useful when conventional diagnosis of IBD is challenging.
1098T
Improving the Power of Genetic Association Tests with Imperfect Phenotype Derived from Electronic Medical Records. J.A. Sinnott, W. Dai, K.P. Liao, S.Y. Shaw, A.N. Ananthakrishnan, V.S. Gainer, E.W. Karson, S. Churchill, P. Szolovits, S. Murphy, T. Kohane, R. Plenge, T. Cai 1. 1) Biostatistics, Harvard School of Public Health, Boston, MA; 2) Division of Rheumatology, Immunology and Allergy, Brigham and Women’s Hospital, Boston, MA; 3) Center for Systems Biology, Massachusetts General Hospital, Boston, MA; 4) Division of Gastroenterology, Massachusetts General Hospital, Boston, MA; 5) Research Computing, Partners Healthcare, Charlestown, MA; 6) (b2b National Center for Biomedical Computing, Boston, MA; 7) Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA; 8) Laboratory of Computer Science, Massachusetts General Hospital, Boston, MA; 9) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 10) Merck Research Laboratories, Boston, MA

To reduce costs and improve clinical relevance of genetic studies, there has been increasing interest in performing genetic studies in hospital-based cohorts by linking phenotypes extracted from electronic medical records (EMRs) to genotypes assessed in routinely collected medical samples. A fundamental difficulty in implementing such studies is extracting accurate information about disease outcomes from large numbers of EMRs. Recently, many algorithms have been developed to infer phenotypes from EMRs, and although these algorithms are quite accurate, they typically do not provide perfect classification due to the difficulty in inferring meaning from text. Some algorithms produce for each patient a probability that the patient is a disease case. This probability can be thresholded to define case-control status, and this estimated case-control status has been used to replicate known genetic associations in EMR-based studies. However, using this dichotomized outcome (denoted DO) in place of the true disease status results in outcome misclassification, which can diminish test power and bias odds ratio (OR) estimates. We propose to instead directly model the algorithm-derived probability of being a case. We demonstrate how our approach (denoted pD) improves test power and OR estimation in simulation studies: improvements in test power are small but consistent, but improvements in odds ratio estimation are dramatic (e.g., a bias of 34% using DO vs. 3% using pD). We also demonstrate our method’s performance in an EMR-based study of rheumatoid arthritis (RA). For some SNPs, our approach recovers the OR observed in a recent meta-analysis: e.g., for rs6457620 in HLA, the meta-analysis OR is 2.35 and the pD OR is 2.28, while the DO OR is 2.03. For other SNPs (e.g., rs6679677 in PTPN22), both EMR-based estimates differ from the meta-analysis estimate. We discuss possible causes for these discrepancies, which may impact other EMR studies: our cases are likely to have more severe disease than a random sample of RA cases, while our controls are likely to have more comorbidities than typical "healthy controls." We also provide power and sample size calculations in terms of algorithm accuracy measures for use in planning future studies. Our work provides an easily implemented solution to a major practical challenge that arises when using EMR data, which can facilitate the use of EMR infrastructure for more powerful, cost-effective, and diverse genetic studies.

1099S

Diverticulosis is defined by small outpouchings formed through weakening of the colonic mucosal wall muscles that when inflamed (diverticulitis) can cause significant abdominal pain and morbidity. As part of the electronic Medical Records and Genomics (eMERGE) network, we identified diverticulosis and diverticulitis cases and controls for GWAS using data captured through routine clinical care across six institutions with different electronic medical record systems. All participants had to have an abdominal CT scan or a colonoscopy to be included. Using natural language processing of these procedure reports, cases had a positive mention of diverticula, diverticulosis or diverticulitis; while controls had ≤1 colonscopy and no mention of any type of diverticular. This approach yielded an 88-100% positive predictive value and an 89-100% negative predictive value for identification of diverticular cases and controls compared against clinician review when validated across three eMERGE sites. We identified 5,285 cases and 3,761 controls for diverticulosis and 569 cases and 8,477 controls for diverticulitis, all of European ancestry. Previously genotyped data associated with these participants were collaboratively quality-controlled, merged, and imputed to 1000Genomes reference across all six institutions. Logistic regression analyses were performed for case-control status under an additive model, adjusting for age, sex, BMI, smoking status, study site, and the first two principal components for ancestry from the genotype data. For diverticulosis, we observed a robust association signal spanning CAV1/CAV2 on 7q51.2 lead by rs17887966 in 1.7×10^-6. Another association signal approaching genomewide significance was observed (chr3:168780737:d:p=1.7×10^-8) near EVI1. For diverticulitis we observed a genome-wide significant association with rs148547269 (p=4.4×10^-8) on 12q22. The most significant association occurred in an upstream region of EVI1, in a highly conserved region containing a CTCF zinc finger transcription factor binding site. Multiple association signals approaching genome-wide significance were also observed including rs28646272 (p=1.1×10^-5) in PALLD on 4q32.3, rs458451 (p=1.4×10^-5) in HPS1 on 10q24.2, and rs12347739 (p=1.7×10^-5) in DENND4C on 9p22.1. Our results identify a number of novel associations for follow-up fine-mapping in order to elucidate the genetic architecture of diverticular disease in European ancestry populations.
1100M

Comparative Analysis of Electronic Health Record (EHR)-driven and Conventional Cohort-driven Genomic Research. V. Thaker$^1$, T. Lingren$^1$, B. Namjou$^2$, C. Perry$^1$, N. Crimmins$^2$, C. Brady$^2$, I. Solt$^2$, G. Savova$^1$, I. Kohane$^1$, J. Harley$^2$. 1) Boston Childrens Hospital, Boston, MA; 2) Cincinnati Childrens Hospital and Medical Center, Cincinnati, OH.

BACKGROUND: Cost-effective and time-efficient systems to perform genetic research require phenotypic data extraction from EHRs. Within the eMERGE network (electronic MEDical Records and GENomics, an NHGRI funded initiative), Boston Children’s Hospital (BCH) and Cincinnati Children’s Hospital Medical Center (CCHMC) collaborated to develop an algorithm for Early Childhood Obesity. We describe the process of case selection, gaps identified, logistical and cost analysis of potential solutions. METHODS: The phenotypic identification of cases and controls was performed using a validated algorithm with structured and unstructured data using Natural Language Processing (NLP). Biorepository databases were searched for identified cases and controls. New systems targeted for biological sample collection were defined to bridge gaps in availability. BCH established the conventional method of cohort-driven biorepository. CCHMC used EHRs for subject recruitment. RESULTS: Both institutions are recruiting subjects at weight management and primary care clinics. BCH is recruiting a cohort for family study including probands and first-degree relatives. Using a study algorithm, 194 unique study subjects have been approached with an enrollment of 35 (18%). The IRB approval required 7 months. The cost of recruitment of each subject is ~$400.00 including permission from physician and patients, enrollment, sample collection at Clinical Study Unit. After IRB approval (12 months), CCHMC is recruiting new cases using EHRs. Clinical research coordinators evaluate patient data with the algorithm and flag cases. Flagged patients provide saliva for DNA extraction. The cost of recruitment is ~$53 per subject including screening, approach and enrollment. Current consent rates (73.9%, 17/23) indicate support of previous claims made about the advantages of EHR-based recruitment. CONCLUSION: The selection of a common phenotype does not guarantee availability of DNA samples for genetic studies. For genomic studies using EHR, the burden of development of a phenotype algorithm, identifying cases, collecting samples satisfying quality thresholds, improving coverage of known and candidate respiratory regions. Nested case-control analyses were undertaken, using genotyped and imputed variants within the heavy-smoking and never-smoking strata. We reconfirmed associations at previously identified loci and identified additional novel signals of association, notably a large ~1.5Mb region of association on chromosome 17 centred on a directly genotyped SNP, rs2696671, upstream of KANSL1 (in never smokers, airway obstruction phenotype (low %FEV$_1$) versus high %FEV$_1$, $P=7.20 \times 10^{-10}$). This signal spans a number of genes in a well characterised inversion (17q21.31) and overlaps previously identified signals of association for multiple traits including idiopathic pulmonary fibrosis and Parkinson’s disease. We will compare the genetic architecture of extreme lung function phenotypes in smokers to that in never-smokers, and report other emerging novel findings from UK BiLEVE. The genotype data for the UK BiLEVE study will be made available to the scientific community via UK Biobank from 2014, as will subsequent genome-wide genotyping from a similar array in the remaining UK Biobank participants as this becomes available from 2015. This research has been conducted using the UK Biobank Resource.
1102S
APOL1-associated kidney disease risk and hypertension management in primary care - A project of the IGNITE Network (Implementing Genomic medicine In pracTicE). N.S. Abul-Husn, S. S. Ellis, S. Sanderson, E. Maddox, H. Junkins, S. E. Kimme, E. Bagiella, D. Hauser, N. Calman, C. R. Horowitz, E. P. Bottinger, Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY; Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; The Institute for Family Health, New York, NY; Department of Family Medicine and Community Health, Icahn School of Medicine at Mount Sinai, New York, NY.

African Americans are more likely to develop kidney disease than European Americans and are at higher risk for progression to kidney failure. Two alleles in the last exon of the APOL1 gene confer a five-fold increased risk for kidney disease in hypertensive, non-diabetic individuals of African ancestry: the missense variants S342G and I384M (in almost complete linkage disequilibrium; referred to as the G1 allele) and the 6 base pair deletion N388/Y389 (referred to as the G2 allele). Approximately 14% of African Americans carry two APOL1 risk alleles (vs. < 0.5% of European Americans), contributing to the high burden of kidney disease in this population. Management of hypertension in African Americans with high risk APOL1 genotypes constitutes an important opportunity for genome-guided medicine. Here, we describe a randomized controlled trial studying the effects of incorporating APOL1 genotype information into the primary care management of African ancestry patients with hypertension. We will recruit hypertensive, non-diabetic adults with no kidney disease and self-reported African ancestry from ten primary care practices in New York City and randomize them to ten primary care practices in New York City and randomize them to either to baseline (intervention) or at one year (control). We will educate providers about genome-guided medicine and APOL1 risk alleles in kidney disease through training sessions, electronic medical records-enabled decision support, and educational materials. We will return results to tested patients, and provide them with educational materials suitable for a range of literacy levels. During patient encounters, providers will receive point-of-care clinical decision support to alert them of their patients’ APOL1 genotype and risk status, and to provide recommendations. At three and twelve months, we will assess primary outcomes (blood pressure reduction and appropriate renal surveillance) and secondary, psychosocial-behavioral outcomes, comparing three patient groups: increased genetic risk, no increased genetic risk, and genetic risk not yet assessed. This study will help to establish the effective implementation of APOL1 risk-informed management of hypertension in African Americans who are at high risk of kidney disease, and will provide a framework for future endeavors to implement genome-guided medicine in clinical practice.

1103M
Mendelian Randomization study of body mass index/waist hip ratio-associated SNPs and five cancer types. C. Gao, C. Patel, S. Lindstrom, B. Pierce, P. Kraft on behalf of the Genetic Mechanisms in Oncology (GAME-ON) initiative. 1) Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA; 2) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 3) Department of Health Studies, The University of Chicago, Chicago, IL.

Body mass index (BMI) has been shown to be associated with multiple cancer types, but whether the observed associations are causal is unclear due to possible confounding and reverse causality. Mendelian Randomization is a technique that uses genetic predictors of a putative causal factor to eliminate reverse causality and reduce confounding bias. Here, we carried out Mendelian Randomization analyses to assess the causal relationship between BMI and waist-hip-ratio (WHR) and risk of breast, ovarian, prostate, colorectal and lung cancers. BMI and WHR genetic risk scores were derived from 32 GWAS-significant BMI-associated variants and 14 WHR-associated variants. We also derived sex-specific scores based on twelve GWAS significant WHR-associated SNPs for women and four GWAS significant WHR-associated SNPs for men. We tested the association between these genetic scores and risk for each of the five cancers using summary statistics from the Genetic Associations and Mechanisms in Oncology (GAME-ON) Consortium, which conducted GWAS meta-analyses involving between 7,272 and 62,528 cases and controls (depending on cancer type). The test was performed by calculating a weighted average across SNPs of the allelic log odds ratio for cancer, where the weights were defined using the allelic effect on BMI (or WHR) and the standard error of the estimated log odds ratio. We found that the genetic risk score of BMI was significantly associated with decreased risk of breast cancer (OR=0.89; 95%CI: 0.85, 0.93; pvalue=3.37×10-7); conversely, the genetic risk score for BMI was associated with increased risk of ovarian cancer (OR=1.08, 95%CI: 1.00, 1.15; pvalue=0.047). Positive associations, though not significant, were observed between the genetic score for WHR and risk of all cancer types except breast cancer. We observed a significant association between the male-specific genetic score for WHR and decreased risk of breast cancer (OR=0.41; 95%CI: 0.24, 0.72; pvalue=0.002). Our findings implicating opposite risk profiles between adiposity and breast cancer may reflect the confounded relationship between menopausal status, adiposity, and cancer. By presenting BMI-associated risks across five different cancer types, findings from this study provide additional insights to help understand the relationships between adiposity and cancer risk.
1104T
Distinct differences in HLA genotypes for latent autoimmune diabetes in adults (LADA) and type 1 diabetes within the same extended pedigree, K.J. Basile¹, V. Guy², S. Schwartz³, D.S. Monos³, S.F.A. Grant⁴-⁶. ¹Division of Human Genetics, The Children’s Hospital of Philadelphia, Philadelphia, PA; ²Main Line Health System, Wynnewood, PA; ³Department of Pathology and Laboratory Medicine, The Children’s Hospital of Philadelphia, PA; ⁴Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA.

Latent autoimmune diabetes in adults (LADA) shares characteristics of both type 1 diabetes (T1D) and type 2 diabetes (T2D), thus often referred to as “type 1.5 diabetes”. According to the World Health Organization, LADA is a slowly progressing form of T1D resulting from autoimmunodestruction of pancreatic beta-cells. However, LADA also shares characteristics with T2D, such as adult age onset, non-insulin dependence and strong association with variants in the T2D-implicated locus, TCF7L2. Symptoms of LADA are often indistinguishable from early stage T2D, and it has been estimated that 8-10% of patients diagnosed with T2D are in fact positive for specific circulating auto-antibodies. In order to further clarify differences between these forms of diabetes, we performed HLA genotyping of members of a large family consisting of four LADA cases (age of onset >25 yrs), four T1D cases (age of onset <15 yrs) and eight unaffected members, all derived from the National Disease Research Interchange. Of particular note was the generational split for the two diseases, with LADA being all in one generation and T1D all occurring in the next generation. We elected to perform HLA class I and II genotyping (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1, HLA-DQA1) for all DNA samples obtained from this family using next-generation sequencing tools. Interestingly, the HLA genotypes of the family members with T1D were distinctly different, while the HLA genotypes of the family members with LADA were more similar to unaffected family members. All the LADA cases had HLA genotypes that were identical to at least one unaffected family member. However, none of the family members diagnosed with T1D had any genotypes in common with an unaffected family member and exhibited more classical HLA genotypes (e.g. one T1D case exhibited DRB1*0401 homozygosity, DOB1*0302 homozygosity and DQA1*0301 heterozygosity while two other cases had a distinct DRB1 heterozygous combination). We hypothesized that LADA may have, at least in part, arisen from genetic influences beyond the HLA locus in this family and may reflect a more general mechanism underpinning the genetic etiology of this disease.

1105S
Like Mother, Like Daughter: Analysis of Parent-Child Phenotypic Correlations for Hundreds of Medical and Behavioral Traits. E. Pierson¹, D. Hinds¹, A. Kleinman¹, N. Eriksson², ¹23andMe, 1390 Shorebird Way, Mountain View, CA; ²Coursera, 381 E Evelyn Ave, Mountain View, CA.

We examined correlations between the phenotypes of parents and children in 15,683 trios for 2,235 medical and behavioral traits. While parent and child phenotypes were almost always positively associated, and no phenotypes showed significant negative associations, for most traits that showed significant parent-child associations (58%) the mother’s phenotype was more strongly associated with the daughter’s than was the father’s. There was no disparity for sons. We observed this effect regardless of whether we examined the correlation between parent and child phenotype or used a regression model that controlled for the age, sex, and genetic principal components of the child and included the mother’s and father’s phenotypes as covariates. When we filtered for phenotypes in which the disparities in parental effect sizes were especially significant (p < .05), the mother-daughter effect became more pronounced: the mother showed a stronger association with the daughter than did the father for 35/40 traits. We observed the opposite effect, to a weaker degree, in sons: the father showed a stronger association with the son than did the mother for 24/35 traits. We found specific phenotypes for which children showed significantly more association with one parent after multiple hypothesis correction. Standardized BMI in daughters showed a stronger association with mother’s standardized BMI, consistent with previous results; there was no disparity for sons. Smoking behavior in sons (as measured by years smoked) was more strongly associated with father’s smoking behavior. We estimated the proportion of these disparities that were attributable to genetics as opposed to environment using SNPs previously associated with the phenotypes and compared parent-sibling correlations with heritability estimates from distant relatives.

1106M
ROBUST MicroRNA EXPRESSION UPREGULATION EXISTS IN INFLAMMATORY BOWEL DISEASE. S. Ben-Shachar¹, H. yanai², L. Baram³, H. Eliad³, H. Sherman Horev³, E. Brzowski⁴,⁵, H. Tulchinsky⁶,⁷, M. Pasmanik-Chor⁸, N. Shomron⁹, I. Dotan²-⁶, ¹Genetic Institute, Tel Aviv Medical Center, Tel Aviv, Israel; ²IBD Center, Department of Gastroenterology and Liver Diseases, Tel Aviv Medical Center, Tel Aviv, Israel; ³Department of Pathology, Tel Aviv Medical Center, Tel Aviv, Israel; ⁴Proctology Unit, Department of Surgery, Tel Aviv Medical Center, Tel Aviv, Israel; ⁵Bioinformatics unit, G.S.W. Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel; ⁶Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

Introduction: The etiology of inflammatory bowel diseases (IBD), Crohn’s disease (CD) and ulcerative colitis (UC) is yet unknown. Similarly, it is not clear why some patients with UC undergoing large bowel resection and creation of ileal pouch (pouch surgery), develop small intestinal inflammation (pouchitis) reminiscent of CD. We have previously shown that gene expression profiles are associated with ileal inflammation and disease behavior in patients with CD and pouchitis. As microRNAs (miRNAs) regulate gene and protein expression, we hypothesized that miRNAs may have a major role in IBD. Methods: Patients with CD ileitis, UC (unoperated and after pouch surgery) and normal controls (NC) were recruited. Pouch patients were stratified according to disease behavior into three groups: normal pouch (NP), chronic pouchitis (CP), and Crohn’s-like disease of the pouch (CLDP). miRNAs expression in ileal and pouch biopsies was analyzed using parallel massive sequencing (next generation sequencing-NGS). Effect of inflammation on miRNAs expression was evaluated in Caco-2 and HCT-116 epithelial cell lines after stimulation with IL1β, INFγ, and TNFα. miRNAs expression in Caco-2 and HCT-116 cell lines were normalized using real time PCR. Results: Fifty-Six subjects: 10 CD, 12 NP, 16 pouchitis (12 CP and 4 CLDP) and 11 unoperated UC as well as 7 NC were recruited. The ileum of unoperated UC patients was comparable to NC. Eight, 25 and 124 significant miRNAs in chronic pouchitis were identified when compared to normal pouch and ileal biopsies, respectively. miRNAs expression upregulation of 8/13 (62%) of tested miRNAs. Conclusions: The robust increase in miRNAs expression may result from the inflammatory process in IBD. The consequence may be down regulation of mRNA transcripts and proteins, further affecting disease course.
Meta-Analysis of Glaucoma Genome-Wide Imputed Datasets. J.N. Cooke Bailey1, S.J. Loomis2, M.A. Hauser3,5, L.R. Pasquale2,5, J.H. Kang4, R.R. Allingham4, R.N. Weinreb4, J.L. Wiggs6, J.L. Haines7, NEIGHBOURHOOD Consortium. 1) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 2) Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA; 3) Channing Division of Network Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA; 4) Department of Ophthalmology Duke University Medical Center, Durham, NC; 5) Department of Medicine, Duke University Medical Center, Durham, NC; 6) Department of Ophthalmology, Hamilton Eye Center, University of California, San Diego, SD, CA.

Glaucoma is a phenotypically and genetically complex neurodegenerative disease that is the second leading cause of blindness worldwide. Although genetic risk factors are known, such as primary open angle or primary angle closure phenotypes, their interaction with environmental factors is associated with it (e.g., intraocular pressure, cup-to-disc ratio, optic nerve parameters, and central corneal thickness), identified risk loci fail to fully account for the genetic component of glaucoma. To extend the proportion of the genome testable for association with glaucoma as well as the power to detect associations at numerous loci, we evaluated imputed genome-wide data in the NEIGHBOURHOOD Consortium dataset, which consists primarily of individuals of European descent. Eight datasets, typed on different genome-wide SNP arrays, were combined to the March 2012 version of the 1000 Genomes data using IMPUTE2 and/or MaCH/miMac. Each dataset was individually evaluated for genomic inflation and then filtered for minor allele frequency (MAF≥0.05) and imputation quality (r2 or info metric ≥0.7). We then performed study-specific dosage analysis (using ProbABE) evaluating risk estimates and estimated genotypic probabilities from the imputed 1000 Genomes step. The study-specific logistic regression models included as covariates, age, gender, and significant eigenvectors, along with, in some cases, study-specific covariates. These study-specific results were meta-analyzed applying the inverse variance method in the METAL software, with the exception that for the control group we applied the control correction. The meta-analyzed dataset includes 3,873 POAG cases and 33,642 controls, and a total of 7,301,525 variants that passed quality control parameters in at least one dataset. Significant associations were found for genomic regions previously associated with POAG (TMCO1, CDKN2BAS, SIX6) as well as novel regions on 6p and 14q. Imputation of genome-wide array data extends the genomic coverage beyond what can be interrogated by a single or multiple arrays. Meta-analyzing multiple imputed datasets can successfully be used to study common diseases with complex inheritance, such as glaucoma, and will help fully define the underlying genetic architecture.
1110T
ABC2G dysfunction causes renal underexcretion hyperuricemia as well as renal overload hyperuricemia. H. Matsuo, A. Nakayama, M. Sakaiyama, S. Shimizu, H. Nakashima, H. Nakao, K. Waka, T. Ito, K. Yamamoto, Y. Sakurai, K. Ichida, T. Shimizu, N. Shinomiya. 1) Department of Integrative Physiology and Bio-Nano Medicine National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan; 2) Department of Preventive Medicine and Public Health, National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan; 3) Department of Integrated Genetics, National Institute of Genetics, 1111 Yata, Mishima, Shizouka 411-8540, Japan; 4) Department of Preventive Medicine, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan; 5) Department of Internal Medicine, Self-Defense Forces Central Hospital, 1-2-24 Ikijiri, Setagaya-ku, Tokyo 154-8522, Japan; 6) Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan; 7) Department of Pathophysiology, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan; 8) Midori-gaoka Hospital, 3-1-1 Makai-cho, Takatsuki, Osaka 569-1121, Japan.

Background and objectives
Gout is a common disease which results from hyperuricemia. We have reported that the decrease in extra-renal (intestinal) urate excretion by ABC2G dysfunction induced renal urate overload (ROL hyperuricemia). This mechanism, however, does not give a sufficient explanation for all ABC2G dysfunction cases as a major cause of hyperuricemia and gout, because renal underexcretion (RUE) hyperuricemia is the most prevalent subtype of hyperuricemia. In this study, we focus on the involvement of ABC2G dysfunction in RUE hyperuricemia. Design, setting, participants & measurements: We performed genotyping for 2,267 Japanese male participants, who consisted of 644 hyperuricemia cases with AE model as the best fitting model with low to moderate (a = 0.21 - 0.55) control. Variations in secular change in the rest of the phenotypes have CE model as the best fitting model with low to moderate (c = 0.25 - 0.46) control. Results: Variations in secular change in 3 lipids (total cholesterol; triglyceride; LDL), weight and blood pressure (SBP, DBP) have AE model as the best fitting model with low to moderate (a = 0.25 - 0.46) control. Variations in secular change in the rest of the phenotypes show limited genetic control. Our results emphasize the special importance of unique environment in determining individual change in metabolic phenotypes in adult Chinese.

1112M
Secular change in 13 metabolic phenotypes: A Chinese longitudinal twin study. S. Li, Z. Pang, D. Zhang, H. Duan, Q. Tan, J. Hjelmborg, T. Kruse, K. Kyv¨e. 1) Human Genetics, Institute of Clinical Research, University of Southern Denmark, Odense, Odense, Denmark; 2) Qingdao Center for Disease Control and Prevention, Qingdao, China; 3) Dept. of Public Health, Qingdao University Medical College, Qingdao, China; 4) Epidemiology, Biostatistics and Biodemography, Institute of Public Health, University of Southern Denmark; 5) Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark.

Aims: The genetic and environmental influences on metabolic phenotypes have been intensively studied by twin modeling in different populations. However, twin studies on secular change in metabolic phenotypes have been rare due to high expenses, losses of follow up, and long waiting time in prospective investigations. Based on Chinese twin data collected from Danish-Chinese collaboration research, we perform twin modeling on 13 metabolic phenotypes (total cholesterol; triglyceride; high density lipoprotein (HDL); low density lipoprotein (LDL); uric acid (UA); glucose; weight; body mass index (BMI); waist and hip circumference and ratio (WHR); systolic blood pressure (SBP); diastolic blood pressure (DBP)) measured longitudinally in a period of 7 years in 254 middle aged Chinese twins (128 monzygotic; 128 dizygotic twins). Methods: Univariate ACE, ADE models and their nested models were fitted to the secular changes in each of the 13 phenotypes with best fitting model selected based on model performance. Age and sex were included as covariates in the models to adjust for their effects on secular trend. Results: Variations in secular change in 3 lipids (total cholesterol; triglyceride; LDL), weight and blood pressure (SBP, DBP) have AE model as the best fitting model with low to moderate (a = 0.25 - 0.46) control. Variations in secular change in the rest of the phenotypes have CE model as the best fitting model with low to moderate (c = 0.21 - 0.55) control. Univariate ACE, ADE models and their nested models were fitted to the secular changes in each of the 13 metabolic phenotypes show limited genetic control. Our results emphasize the special importance of unique environment in determining individual change in metabolic phenotypes in adult Chinese.
Latent variable adjustment of NIH Epigenomics Roadmap ChIP-seq data for utilization in tissue-specific polygenic analysis of type II diabetes in the DIAGRAM3 GWAS meta-analysis. A.L. Dobbyn, B.F. Voight, P. Roussos, A. Morris, S. Raychaudhuri, M. McCarthy, E.A. Stahl, the DIAGRAM consortium. 1) Icahn School of Medicine at Mount Sinai, New York City, NY; 2) Departments of Pharmacology and Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 4) Partners HealthCare Center for Personalized Genetic Medicine, Brigham & Women’s Hospital, Harvard Medical School, Boston, MA; 5) Broad Institute, Cambridge, MA; 6) Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY.

INTRODUCTION: Previous analyses have shown that tissue specific chromatin marks can be used to inform association studies in order to dissect the genetic architecture of complex disease. Enrichment for the H3K4me3 chromatin mark in GWAS identified variants has implicated causal tissues in the etiology of complex disease, including both Liver and Pancreatic Islets for type II diabetes (Trynka 2013). However, we have detected similar levels of enrichment in non-causal tissues in separate analyses, and observe a lack of correlation between H3K4me3 peaks from samples of the same tissues, consistent with the presence of latent confounding variables. We outline a method for the normalization of ChIP-seq data that adjusts for confounders, allowing for improved detection of enrichment signal in downstream analyses, and then utilize the normalized data in polygenic risk score analysis of Type 2 Diabetes.

METHODS: We constructed a matrix of the union of all peaks in all samples from the NIH Epigenomics Roadmap Project (82 samples, 44 tissues/cell types), merging peaks with any overlap, resulting in a total of 214,704 peaks. We then conducted surrogate variable analysis. We used H3K4me3ChIP-seq peaks to define tissue-specific H3K4me3 enriched sets of genes, and selected SNPs within 20kb of each gene’s transcription start site for use in polygenic SNP analysis. We then conducted polygenic risk score analysis of T2D using discovery SNPs from the DIAGRAM3 GWAS meta-analysis (Morris 2012, 12,562 subjects, 11,622 controls), and tested polygenic risk scores in the WTTCC (2007; 1,924,293 cases/controls). RESULTS: After normalization, maximum enrichment (observed/expected variance explained) for Liver and Pancreas SNP sets, based on a small number of tissue-specific peaks, increased from 4.37 to 4.90, and 3.07 to 3.09, respectively. Pancreas within-tissue correlations increased from R_Liver=0.785 and R_Pancre=0.570 to R_Liver=0.858 and R_Pancre=0.859. CONCLUSIONS: We have implemented surrogate variable analysis for the normalization and adjustment of ChIP-seq data, reducing the effects of hidden covariates and improving detection of tissue-specific compartments of polygenic signal. The adjustment used is applicable to all downstream analyses that utilize tissue specific ChIP-seq data.


1) Medical Sciences, University of Trieste, Trieste, Trieste, Italy; 2) Institute for Maternal and Child Health, IRCCS “Burlo Garofolo”, Trieste, Italy; 3) Medical Genetics Unit, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (FG), Italy; 4) Department of Pediatrics, University of Torino, Torino, Italy; 5) Oasi Institute for Research on Mental Retardation and Brain Aging (IRCCS), Troina, Italy; 6) Unit of Pediatrics and Medical Genetics, IRCCS Associazione Oasi Maria Santissima, 94018 Unit of Pediatrics and Medical Genetics, IRCCS Associazione Oasi Maria Santissima, 94018 Troina (EN), Italy.

The harmful effect of inbreeding is well known by geneticists and several studies have already reported cases of Intellectual Disability (ID) in consanguineous families caused by recessive variants. Nevertheless, despite the recent findings for other neurological disorders like Alzheimer disease, Parkinson disease and schizophrenia, the effect of inbreeding on ID is still poorly investigated in outbred populations. Here we present a Runs of Homozygosity (ROHs) study performed in a 612 individual cohort of ID patients.

We first investigated the ROHs distribution considering the complexity of the phenotype (presence of malformations, seizures, micro/macrocephaly, etc.), and we didn’t find a different amount of homozygosity comparing syndromic and non-syndromic ID patients. In a second stage, we focused on the effect of ROHs on the ID degree, starting from the IQ. Our data revealed the presence of significantly larger ROH stretches in severe ID cases compared to non-severe ID ones (p=7.1×10-3), together with an increased percentage of ROH coverage of ISID genes (p=1.0×10-2).

According to the recent findings on ID in autism, this study reveals that autosomal recessive variants have an important role in ID and that they mainly and specifically affect the cognitive impairment modulation, in spite of the ID cause. This is not surprising, since ID degree is a highly complex trait whose overall variability is likely not determined by a single pathogenic mutation. Other environmental and genetic factors modulate the phenotype and one of the most important seems to be accounted for by ROHs.
1116T
Association of Mitochondrial DNA levels with Frailty and All-Cause Mortality.
F.N. Ashar1, A. Moes1, A.Z. Moore2, M.L. Grove2, P.H.M. Chaves3, J. Cotre1, A.B. Newman1, A.M. Mattioni2, K. Bandeen-Roche1, E. Boerwinkle1, J.D. Walston4, D.E. Arking1. 1) Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA; 2) Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; 3) Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; 4) Benjamin Leon Center for Geriatric Research and Education and Department of Medicine, Florida International University, Miami, USA; 5) Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, USA; 6) Division of Geriatric Medicine and Gerontology, Johns Hopkins University School of Medicine, Baltimore, USA; 7) Department of Biostatistics, Johns Hopkins University, Baltimore, USA.
Age-related declines in mitochondrial function have long been hypothesized to underlie multiple biological changes that increase vulnerability to multiple disease states, functional and cognitive decline, and ultimately, mortality. Evidence to support this comes from the fact that specific variants in mitochondrial DNA (mtDNA) have been shown to modulate risk for several age-related disease states. While the association of mitochondrial variants with age-related disorders is well-established, the effect of mtDNA copy number, which reflects energy reserves and oxidative stress, on aging and mortality in the general population has not been addressed. To address this gap, we examined mtDNA copy number in two large multi-center prospective studies—the Cardiovascular Health Study (CHS) and the Atherosclerosis Risk in Communities (ARIC) study—with a total of 16,401 samples of European and African descent focusing on two primary phenotypes—prevalent frailty in CHS, and all-cause mortality in ARIC and CHS. In race-stratified meta-analyses, we demonstrate a significant inverse association of mtDNA copy number with age, with a reduction of 0.12 standard deviation units with a 10 year differences in age (P=2.78 x 10^-6), and higher mtDNA copy number in women relative to men (meta-analysis OR=1.33, 95% CI 1.14 - 1.51, P=6.91 x 10^-4). Furthermore, we show that lower mtDNA copy number is significantly associated with prevalent frailty in 4,109 self-identified white participants from CHS (OR=0.91, 95% CI, 0.85-0.97, P=0.005). Finally, a meta-analysis of mtDNA copy number was performed in an Asian and sex-adjusted, race-stratified analysis of 16,401 participants from both cohorts with a pooled hazard ratio of 1.47 (95% CI 1.33-1.62, P=4.24x10^-14) for the lowest quintile of mtDNA copy number relative to the highest quintile. Based on the association with mortality and observation of a higher mtDNA copy number in women relative to men, we tested whether a mito-protective effect may be related to the differences in life expectancy between men and women. We estimate that mtDNA copy number is associated with 3–10% of the effect of sex on mortality, suggesting that mtDNA copy number may be a contributing factor to the disparity of the phenotype. In summary, we report that a single, easily implemented measure of mtDNA copy number, isolated from whole blood decades before the event of interest (death), is predictive of all-cause mortality.

1117S
Height and intracranial volume (ICV) are both highly heritable traits as shown by twin studies. Additionally, it has been observed that the two quantities are positively correlated. We investigated the contribution of height-associated variants to the variability of ICV in a sample of ~100,000 European and African descent focusing on two primary phenotypes—prevalent epilepsy following a single seizure is about 45%, which are both much higher than the prevalence of Type 2 diabetes (6.4% of all Europeans and 6.2% of all Africans). The low prevalence of single seizures among the very large sample sizes now available. The MultiBLUP software (Multi-trait Unbiased Prediction) is a classic prediction tool widely used in plant and animal breeding. It is computationally fast, but is suited only to simple polygenic traits, and not to complex genetic architectures. We therefore introduce MultiBLUP, which generalises BLUP in a way that retains its good computational properties while increasing its flexibility so that it can perform well for a wide range of genomic architectures. In tests on many disease traits, it consistently achieves better risk prediction than rival methods, while being many times faster than the next best method and able to handle much larger datasets. For example, applied to imputed data for inflammatory bowel disease (13,000 individuals, 1.5M SNPs), MultiBLUP runs in a few hours, and achieves better prediction performance than BLUP or genetic risk scores (AUC 0.68 compared to 0.58 and 0.61, respectively). Similar improvement is achieved for Celiac Disease (AUC 0.86 compared to 0.79 for both BLUP and genetic risk scores). MultiBLUP can also be used for predicting continuous traits (we demonstrate this for cholesterol phenotypes), and for non-human data (we consider 139 mice phenotypes). The low prevalence of most diseases means that prediction in a general population remains difficult, but there are many settings in which subgroups of patients can be identified using classical risk factors. For example, the prevalence of Type 2 diabetes among obese individuals is almost 50%, while the chance of developing epilepsy following a single seizure is about 45%, which are both much higher than the population prevalences. SNP-based prediction using MultiBLUP can be very powerful for such subgroups. Moreover, with meta-analysis consortia established for many major diseases, MultiBLUP can take advantage of the very large sample sizes now available. The MultiBLUP software will be freely available at www.ldak.org after publication.
Hair e-QTLs - delineating the genetic basis of gene expression in human hair follicle and its implication for the interpretation of hair loss disorders.

S. Heilmann1,2, H. Schulz3, N. Karbalai4, A. Hofmann1,2, L.M. Hochfeld1,2, F.A. Degenhardt1,2, S. Herms1,2,5, J. Becker1,2, T. Kätzel1,2, F.F. Brockschmidt1,2, A.M. Hillmer6, P. Hoffmann1,2,5,7, B. Müller-Myhsok4, M.M. Nöthen1,2, 1) Institute of Human Genetics, University Bonn, Bonn, Germany; 2) Department of Genomics, Life & Brain Center, University of Bonn, Germany; 3) Cologne Center for Genomics (CCG), University Cologne, Cologne, Germany; 4) Max-Planck-Institute of Psychiatry, Munich, Germany; 5) Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland; 6) Cancer Therapeutics and Stratified Oncology, Genome Institute of Singapore, Singapore, Singapore; 7) Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany.

The human scalp hair plays an important role in our social and cultural life and its undesirable loss is often perceived as psychologically stressful in affected individuals resulting in a reduced quality of life. This has provoked many studies on the genetic and biological causes of hair loss. Candidate gene and genome-wide association studies have been successful in the identification of genetic risk factors that predispose to hair loss disorders such as alopecia areata and male-pattern baldness. However, the biological context in which these genetic risk factors exert their biological effect is often unknown. Although the functional annotation of the human genome is becoming increasingly detailed, further experiments are still warranted to elucidate the often tissue- and cell-specific biological contexts. Here, the analysis of genetic variants that influence gene expression (eQTLs) has gained major importance. The aim of our study therefore was to systematically map eQTLs in human hair follicle. For this purpose we performed genome-wide genotyping of blood-DNA samples and transcriptome profiling of hair follicle RNA samples from 100 healthy male donors using the Illumina OmniExpress v1.1 and HT12v4 array. After quality control, a total of 562,176 autosomal SNPs (MAF ≥ 5%, call rate ≥ 95%) and 21,696 expression probes (quantile normalized, detection P-value < 0.05 in at least one sample, no SNP with MAF>1% within probe) for 97 individuals remained for eQTL analysis. A total of 5,438 SNPs were found to regulate expression of 856 transcripts in cis or trans, corresponding to almost 1% of all SNPs analyzed. Among the top cis-regulated genes are CHURC1, KCTD10, WDR41. Top trans-regulated genes include RPS26L, PEX16 and the long-non-coding RNA MGC57346. Elaborate data analysis is currently underway and the results will be presented at the meeting. The present hair-eQTL data set will help to further elucidate the underlying biological mechanisms of hair loss disorders such as male-pattern baldness and alopecia areata and hair morphological traits such as hair thickness or curliness. Promising strategies include the interpretation of GWAS findings using the hair-eQTL data, tests for association of the limited set of functionally relevant hair-eQTLs in case-control samples for hair related traits and disorders as well as pathway- and score-based approaches.
1120S
Next generation sequencing approaches for the identification of novel genes in spinocerebellar degeneration. M. Couteller1,2,3,4,5, L. Raymondy1,2,3,4,5, C. Tesson1,2,3,4,5, M. Mairesse1,2,3,4,5, K. Konig1,2,3,4,5, M. Jaccoupy1,2,3,4, T. Esteves1,2,3,4,5, C. Goizet1, M. Gaussen1,2,3,4, F. Darios1,2,3,4, A. Durr1,2,3,4,5, I. Delalle2,5,6, G. Rouleau7,8, S. Zuchner9, A. Brice1,2,3,4, G. Stenvand1,2,3,4,5, J. Konop9, A. Sackin9,10, M. Synofzik11, M. Greggi1,2,3,4,11. We have screened our extended cohort for these genes. Functional analyses are in progress. Genotype-phenotype correlation of several families with spinocerebellar degeneration is published (Synofzik et al., Brain 2014) or submitted. In these families, we gathered the strengths of Next Generation Sequencing technologies and high-resolution chip arrays for linkage analysis. As of today, only 60% of SCD cases are genetically elucidated even though more than 150 causal genes have been described. Aiming at the identification of new causative genes, we gather the strengths of Next Generation Sequencing technologians (NGS), and the SPATAX network with its cohort of more than 10000 SCD patients. Whole Exome Sequencing (WES) was performed in 54 families (1-3 patients/family). Along with classical criteria based on database frequencies and effect on the encoded protein, variants were sorted with linkage analysis data in the most informative families, or with gene network analysis in the smaller ones. We selected the best putative causal variants. Somatic coding region variants (SNV’s) in genes: NRM, INTS4, GDPD4, GAB2, and PIK3R5. Another noteworthy family presented with both heterozygous and homozygous mutations of the same gene, leading to interesting intra-familial genotype-phenotype comparison. In yet another family, a heterozygous mutation of a gene already described in an autosomal recessive heavier neurocutaneous phenotype shed the light on the putative broadening of phenotypes that will be brought along by NGS. Two new genes in autosomal recessive CA are published (Synofzik et al., Brain 2014) or submitted. In the remaining families, we isolated 41 candidate genes in autosomal dominant CA, one of which is involved in 2 families; 7 in HSP, again with mutations of one gene in 2 separated families; and 4 in autosomal recessive CA with associated hypogonadism. We designed targeted gene panels aimed at screening our extended cohort for these genes. Functional analyses are already ongoing for 6 of these genes. In conclusion, we report results from WES in a cohort of 54 families with SCD, with clues on frequency of already known genes, novel elements on the complexity of phenotype-genotype correlations, as well as several promising candidate genes. We furthermore demonstrate how high-throughput and medium-throughput sequencing may be used efficiently in novel genes identification and validation. Further results on extended cohorts will finally be available shortly.

1121M
Homozygosity mapping and candidate gene screening in Attention Deficit/Hyperactivity Disorder (ADHD) in Highly Inbred Saudi Arabian Families. F. Alnaemi1, A. Abeber2, A. Adi3, A. Almagrashi4, A. Taher5, S. Qasem6, M. Ghazuddin7, N. AlTassan8,9. 1) King Faisal Specialist Hospital & research Center, Riyadh, Saudi Arabia; 2) Psychiatry Department, King Faisal Specialist Hospital & research Center. ADHD is not considered to have a single genetic cause but rather a complex interaction of genetic and environmental factors. Association studies have identified a number of susceptibility loci on 4q13.2, 5q33.3, 7p, 11q22,14q12 and 17p11. Moreover, meta-analysis of candidate genes has implicated genes for dopamine, norepinephrine, and serotonin neurotransmitter systems in the etiology of ADHD. The objective of this study is to identify ADHD susceptibility in highly consanguineous and inbred families of Saudi Arabia. To achieve this goal genotyping using high-resolution chip array and candidate gene screening has been performed. Regions of homozygosity on chromosomes 3,4,5,6, 7,11,10,13,17 and 18 were identified in different families. Subsequent bidirectional screening of candidate genes in these regions identified a number of novel and previously reported single nucleotide variants (SNV’s) in DDR1, GAB2, INT54, ALG8, GDPD4, NRM and BTNL2 genes. None of these variants segregated with the autosomal recessive inheritance in these families. The large number of LOH regions identified and the lack of segregation of variants identified in candidate genes reinforce the complex genetics of ADHD, and points to other possible genetic mechanisms as an explanation to the phenotypes observed.

1122T
Linkage analysis, homozygosity mapping and whole exome sequencing to identify new genes in consanguineous families with juvenile myoclonic epilepsy. B. Ouled Amar Bencheikh1,2,3, F. Lahjouji4, K. Khaldi4, H. Langhari1, D. Spiegelman1,2, A. Dionne Laporte1,2, H. Belaid1, P. Dion1, R. Ouazzani7, G. Rouleau1, P. Cossette1, 1) Centre de Recherche du Centre Hospitalier de l’Université de Montréal, Notre Dame Hospital, McGill University, Montreal, Quebec, Canada; 2) Clinical Neurophysiology Department, Specialty Hospital, CHU Ibn Sina, Rabat, Morocco; 3) Montreal Neurological Institute and Hospital, Department of Neurology and Neurosurgery, McGill University, Montréal, QC, Canada.

Background. The genetic (idiopathic) generalized (GGEs) epilepsies account for 40 to 60% of the etiologies of epilepsy. Most of these syndromes have a complex transmission, with phenotypic and genetic heterogeneity as well as incomplete penetrance of the causative genes. However, despite this complexity, many genes and loci responsible for epilepsies have been identified. The identification of these loci and genes has been possible by genetic linkage studies and recently by whole exome sequencing. To identify new genes of GGEs, we characterized clinically three consanguineous families from Morocco with juvenile myoclonic epilepsy (JME) and autosomal recessive transmission. Genotyping of each family subjects was done by Illumina Omni-Express SNP chips for the linkage mapping. The Multipoint linkage analysis was carried out with Merlin by using an autosomal-recessive model with complete penetrance. The homozygosity mapper was used to identify intervals of homozygosity. Then, the exome sequencing was performed for 8 patients from these three families. The exome capture was done by the Agilent SureSelect V4 and the sequencing on Illumina HiSeq2000. The annotation and the calling of the variants was done by using an in-house data pipeline using the GATK software. Results. We identify one locus in each family confirmed by homozygosity mapper. Then, we prioritize the variants identified by exome sequencing within these intervals of homozygosity. We filtered the variants considering the frequency of the disease at 1% and excluded all variants with frequency greater than 1% in public and in our in-house database (2000 exomes). We have identified a total of 328669, 281434 and 325231 shared variants in family 1, 2 and 3 respectively. Some of these variants segregate well with the disease. All the variants filtered were validated by Sanger sequencing. Conclusion. The use of Linkage analysis and homozygosity mapping, combined to whole exome sequencing in consanguineous families with juvenile myoclonic epilepsy have great potential to identify the causative genes of this disease and provide novel therapeutic issue for epilepsy.
1123S Genetic characterization of a homozygous 9p deletion in a patient with hyper IgE syndrome and progressive multifocal leukencephalopathy supports deficiency of DOCK8 as a causal factor for both diseases. C. Sun1, A. Day-Williams1, H. McLaughlin2, I. Jelic3, T. Harris1, R. Martin1, J. Carulli1, 1) Translational medicine, Biogen Idec, Cambridge, MA; 2) Department of Neurology, University Hospital Zurich, Frauenklinikstrasse 26, Zurich, Switzerland.

A 30 year old male Caucasian patient with eczematoid dermatitis since early childhood was found having massively elevated serum immunoglobulin E (IgE) level (26'800 KUL), and slightly increased peripheral blood eosinophiles (1.6 x 10^9/L). Therefore, he was diagnosed with hyper IgE syndrome (HIES). In fall 2008, the patient developed progressive neurological deficits caused by large intracerebral lesions as detected by MRI, suggestive of encephalitis. Further investigation showed that the patient was HIES, and predisposes to PML development. Our data provide further evidence that DOCK8 deletion is causal in this patient. DOCK8 deficiency causes a pleiotropic immune dysfunction in this patient. DOCK8 deletion causes a homozygous 9p deletion and that the breakpoint is very close to telomere repetitive sequence. Cloning and dideoxy sequencing of a long range PCR product from primers flanking the hypothetical centromeric and telomeric breakpoints defined chimerically the nucleotide level (10,047-586751). This interval is 9ptel-specific SNPs near the potential telomeric end of the break points, on which two nuclear families could be distinguished. The nine adults (6 females, 3 males) tested, six (5 females, 1 male) were found to have the ZNF804A deletion; two females were unaffected with respect to neuphysicopathologic characteristics, three had intellectual disability, personality disorder, and major depressive disorder, respectively. The one male proband with the ZNF804A deletion was diagnosed with schizophrenia and learning difficulties. A second male without the deletion was also diagnosed with schizophrenia. To our knowledge this is the first study to report on the broader neuropsychiatric phenotypes associated with rare deletions overlapping ZNF804A. It appears that the deletion increases the risk for a range of neuropsychiatric conditions in females, though not all females manifest disease. The recurrence of schizophrenia in two males (one without the deletion) in this family suggests that additional factors are likely contributing to the variable expression and incomplete penetrance. Further study, including next generation sequencing, will be critical for understanding the sex related differences in neuropsychiatric expression associated with the ZNF804A deletion.

1125T Sex differences in neuropsychiatric expression of rare deletions overlapping schizophrenia susceptibility gene ZNF804A. C. Lowther1,2, A. C. Irving3, T. C. F. Marsh3, C. Szu-Tu1,2,1, H. McLaughlin1, 1) Institute for Medical Science, University of Toronto, Toronto, Ontario, Canada; 2) Clinical Genetics Research Program, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; 3) The Centre for Applied Genomics and Program in Genomics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; 4) Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada; 5) Department of Psychiatry, University of Toronto and University Health Network, Toronto, Ontario, Canada. 

Intelligence disability (ID), autism spectrum disorder (ASD), and schizophrenia each have a strong genetic component and significant sex differences. There is a 4:1 male sex bias in ASD, an excess of males institutionalized due to severe mental retardation and IQ <50. In schizophrenia, the sex ratio is approximately 1.5:1, with females更多的 than males. X-linked disorders have long been posited to contribute in part to the observed sex bias in ASD and ID and new studies suggest males manifesting these disorders may be more susceptible to the damaging effects of copy number variations (CNVs). Few studies as yet have investigated sex bias in the expression of a rare autosomal CNV. We recently identified two distinctly related probands ascertained for schizophrenia with a very rare 790 kb deletion overlapping ZNF804A, a top candidate gene for schizophrenia. Using family history and locally available census data we were able to trace the origin of this inherited deletion. We performed in-depth phenotyping of several relatives of these probands including comprehensive medical, psychiatric, and physical assessment to identify major lifetime features and illnesses. DNA samples for ZNF804A deletion testing were collected. We report, for the first time with a homogenous deletion and the breakpoint is very close to telomere repetitive sequence. Cloning and dideoxy sequencing of a long range PCR product from primers flanking the hypothetical centromeric and telomeric breakpoints defined chimerically the nucleotide level (10,047-586751). This interval is 9ptel-specific SNPs near the potential telomeric end of the break points, on which two nuclear families could be distinguished. Of the nine adults (6 females, 3 males) tested, six (5 females, 1 male) were found to have the ZNF804A deletion; two females were unaffected with respect to neuropsychicopathologic characteristics, three had intellectual disability, personality disorder, and major depressive disorder, respectively. The one male proband with the ZNF804A deletion was diagnosed with schizophrenia and learning difficulties. A second male without the deletion was also diagnosed with schizophrenia. To our knowledge this is the first study to report on the broader neuropsychiatric phenotypes associated with rare deletions overlapping ZNF804A. It appears that the deletion increases the risk for a range of neuropsychiatric conditions in females, though not all females manifest disease. The recurrence of schizophrenia in two males (one without the deletion) in this family suggests that additional factors are likely contributing to the variable expression and incomplete penetrance. Further study, including next generation sequencing, will be critical for understanding the sex related differences in neuropsychiatric expression associated with the ZNF804A deletion.
**1127M**

Association analysis of HLA-DQB1 gene with narcolepsy without cataplexy and idiopathic hypersomnia. I. Miyagawa,1 H. Toyoda,2 H. Kojima,2 T. Futagami,2 SS. Khor,1 A. Hirakata,1 M. Yamazaki,1 H. Saji,1 Y. Honda2, M. Honda1,4 K. Tokunaga1. 1) Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; 2) HLA Foundation Laboratory, Kyoto, Japan; 3) Japan Somnology Center, Neuropsychiatric Research Institute, Tokyo, Japan; 4) Personalized Medicine Laboratory, Department of Psychiatry and Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan.

Narcolepsy without cataplexy (NA w/o CA) is a lifelong disorder characterized by excessive daytime sleepiness and rapid eye movement (REM) sleep abnormalities, but does not exhibit cataplexy. Several studies have reported that NA w/o CA is associated with human leukocyte antigen HLA-DQB1*06:02 similar to narcolepsy with cataplexy (NA-CA). The sample sizes of the studies were relatively small because NA w/o CA is an infrequent condition, which makes it difficult to find a large number of samples. We have therefore formed a collaborative research group to promote the study in Japan. In the present study, we examined HLA-DQB1 in 146 Japanese patients with NA w/o CA and 1,418 control subjects. The frequency of HLA-DQB1*06:02 in the patient group was significantly higher than that in the control group (allele frequency: 16.8% vs. 7.6%, OR: 2.40; carrier frequency: 31.5% vs. 14.7%, P = 1.5×10^{-7}, OR: 2.68). After controlling for the effect of HLA-DQB1*06:02, distributions of HLA-DQB1 alleles were compared between NA w/o CA and NA-CA to assess whether the genetic backgrounds of the two diseases have similarities. The distribution of HLA-DQB1 alleles in HLA-DQB1*06:02-negative NA w/o CA was significantly different from that in NA-CA (P = 8.4×10^{-7}). On the other hand, the patterns of the HLA-DQB1 alleles were similar between HLA-DQB1*06:02-positive NA w/o CA and NA-CA. HLA-DQB1 analysis was also performed in 171 Japanese patients with idiopathic hypersomnia (IHS). No significant associations were observed between IHS and HLA-DQB1 alleles. We found a significant difference in HLA-DQB1 allele distribution between IHS and NA-CA (P = 2.0×10^{-6}). The findings suggest that HLA-DQB1*06:02-positive NA w/o CA has an autoimmune pathogenesis in common with NA-CA, but HLA-DQB1*06:02-negative NA w/o CA and IHS might be unique from NA-CA.

**1128T**


Autistic disorder is a severe neuropsychiatric disorder with strong genetic component. Recent large scale exome sequencing studies have demonstrated the importance of de-novo mutations in Autism. For familial autism, the contribution of rare highly penetrant genetic variants is not well understood. Identification of genes responsible for familial autism is important as genetic architecture of familial autism might differ from de-novo autism. To identify the contribution of rare variants to familial autism we performed whole exome sequencing in 25 families with Autism spectrum disorder from the University of Washington autism families that include affected cousins and thereby minimizing variant sharing between affected family members. We define candidate variants as variants that are private (not present in dbSNP132, 1000 genomes or ESP6500) and functional (coding, missense, stop-gain, stop-loss, splice). For each family between 2 and 25 variants met our selection criteria. Based on overlap with de-novo variants and literature searches we have selected 25 candidate genes for further evaluation. To establish the association of the identified candidate genes with autism we have performed a large gene-based case control study. For all candidate genes whose protein coding sequence encompasses more than 50 Kbp, we have performed Molecular Inversion Probe (MIP) capture and targeted sequencing in 960 NIMH and 288 UW familial autism cases and 960 NIMH unaffected controls. The variant calling and annotation were performed in the same manner as for exome sequencing with modifications due to MIP capture and multiplexing to ensure comparable coverage representations of the genes in both case and control samples. Enrichment of rare variants in the affected subjects was evaluated using Plink tools. Our preliminary analysis has identified one novel gene for which rare functional variants are more common in cases as compared to controls (P<0.01). Detail analysis of the of the case-control study and analysis will presented.

**1129S**

Further evidence for DLGAP2 as an ASD/ID candidate gene. H. Poquet1, 2, L. Faivre1, 2, S. El Chehadeh1, 2, J. Morton2, 3, H. Goel1, 2, B. Isidro2, C. Lecaignec2, J. Andreux2, B. Delobbe2, M. Lefebvre1, 2, C. Juillet, 1, 3, A. Collinet2, 6, A. Mosca-Boidron4, 6, 1) Centre de Génétique et Centre de Référence Maladies Rares ‘Anomalies du Développement et Syndromes Malformatifs de l’Intérêrégion Ést’, Hôpital d’Enfants, CHU Dijon, France; 2) Centre de Ressources Autisme de Bourgogne, Hôpital d’Enfants, CHU Dijon, France; 3) Service de Pédiatrie, Hôpital d’Enfants, CHU, Dijon, France; 4) Laboratoire de Cytogénétique, Plateau Technique de Biologie, CHU Dijon, France; 5) EA 4271 GAD – Génétique des Anomalies du Développement –, IFP 100 - Santé STIC, Université de Bourgogne, Dijon, France; 6) Laboratoire de Cytogénétique, Centre de Génétique Chromosomique, Hôpital Saint-Vincent de Paul, Groupe Hospitalier de l’Institut Catholique de Lille; 7) Institut de Génétique Médicale, Pôle de Biologie Pathologie Génétique, Centre Hospitalier Régional Universitaire de Lille; 8) Unité de Génétique Chromosomique ou Cytogénétique, Service de génétique médicale, Institut de Biologie, CHU Nantes; 9) Unité de Génétique Clinique, Service de Génétique Médicale, Institut de Biologie, CHU Nantes; 10) Clinical Genetics Unit, Birmingham Women’s Hospital, Edgbaston, Birmingham, UK; 11) Hunter Genetics, Newcastle, Australia.

The Autism Spectrum Disorders (ASDs) comprise a range of early-onset persistent neuro developmental conditions, characterized by significant impairments in reciprocal social interaction and communication, accompanied by repetitive restricted stereotyped behaviors. The core symptom of ASD typically coexists with other medical conditions such as Intellectual Disability (ID) and seizures. The involvement of rare (<1% frequency) copy number variations (CNV) of varying expressivity and penetrance as risk factors in ASD/ID phenotypes has been highlighted previously in large series of patients. The DLGAP2 (discs, large (drosophila) homolog-associated protein 2) gene, whose glutamatergic postsynaptic density product may play a role in synaptogenesis and plasticity, has been identified as a novel candidate gene on the basis of a de novo 8p23.3 duplication intersecting DLGAP2 observed in a sporadic nonsyndromic ASD male. Another patient with sporadic de novo duplication involving DLGAP2 and ASD/ID had been previously reported in the literature. It has also been suggested that increased DLGAP2 gene expression may contribute to the pathogenesis of schizophrenia. Based on these results and after fine phenotyping of another patient with a de novo 8p24.2p32.3.2.1.8 Mb duplication involving DLGAP2 and presenting with ASD/ID with ASD features, and 3 sporadic males with ID and ASD with a duplication inherited from an asymptomatic parent. This study supports the hypothesis that rare CNV encompassing DLGAP2 gene could predispose to ASD/ID, with the principle of incomplete penetrance and variable expressivity. It also further supports the existence of common predisposing factors to ASD, ID and schizophrenia.
demonstrates that the loss at 16p12.1 microdeletion alone may give rise to and another had a 580 kb loss involving multiple exons of the NRXN1 gene region, varying from 478 kb to 909 kb. Four cases had small copy number ties, and autism. All the deletions spanned the 520 kb commonly deleted tal anomalies, microcephaly, speech delay, seizures, endocrine abnormali-
was studied by oligo-SNP array (CaytoScan HD, Affymetrix) at Quest Diag-
delictions or duplications. A cohort of 37 cases with 16p12.1 microdeletion large copy-number variants (10 of 42, 24%) as per Girirajan, S et al, 2010. Cytogenetics, Quest Diagnostics, Nichols Inst, San Juan Capistrano, CA; 2) Dept
C.M. Stein

Second thought about 16p12.1 microdeletion syndrome: is two-hit a complex genetic basis of vocalized speech, hypothesized to be a human-
limited trait, is unknown. We used extensive phenotype and genotype data from children and families with disordered speech to demonstrate that vocal-
ized speech is set in a complex cognitive network connected through many endophenotypes. Speech sound disorder (SSD) is one of the most common types of communication disorders, with prevalence rates of 16% at 3 years of age, and an estimated 3.8% of children continuing to present speech delay at 6 years of age. SSD is manifested through several correlated endophenotypes, including phonological memory, phonological awareness, vocabulary, speed of processing, oral motor skill, reading decoding, spelling, and spoken language; many of these same domains are affected in language impairment and dyslexia. Several studies have identified promising genetic associations between communication disorders and specific genes, but have not accounted for the correlation among endophenotypes. Previously, we developed a structural equation model (SEM) that depicted how endopheno-
types assessed at preschool age influenced endophenotypes at school-age (Lewis et al. 2011). However, that SEM did not explicitly model familial structure and also did not incorporate genetic effects. We have recently released ‘strum’ software, which implements the SEM framework for family data (Morris et al, 2010). Here, we have refined the speech sound SEM with strum, and identified significant polygenic influences on phonological awareness (p=0.02), phonological memory (p=0.04), and the correlation between phonological awareness and phonological awareness (p=0.02), and between phonological awareness and speeded naming (p=0.03). Next, we incorporated candidate gene association data into the model, based on single gene associations we previously observed with DRD2, AVPR1A, and ASPM (Stein et al. 2014), and found that DRD2 has an independent influence on phonological memory (p=0.004) after accounting for polygenic and resid-
ual correlations among these phenotypes. By synthesizing comprehensive endophenotype and genetic data with SEM, these findings reveal the genetic complexity of SSD and other communication and learning disorders.

Second thought about 16p12.1 microdeletion syndrome: is two-hit a valid model? F. Boyar1, L. Ross1, J. Kelly2, J.C. Wang1. 1) Dept Cytogenet-
ics, Quest Diagnostics, Nichols Inst, Salt Lake City, UT; 2) Dept Cytogenetics, Quest Diagnostics, Nichols Inst, Chantilly, VA.

The chromosome 16p12.1 deletion syndrome (520kb; OMIM #136570) was suggested to be associated with high incidence of carrying additional large copy-number variants (10 of 42, 24%) as per Girirajan, S et al, 2010. The presence of 16p12.1 microdeletion was considered to predispose patients to develop neuropsychiatric phenotypes, which were exacerbated by epigenetic, environmental or genetic abnormalities such as other large deletions or duplications. A cohort of 37 cases with 16p12.1 microdeletion was studied by oligo-SNP array (CytoScan HD, Affymetrix) at Quest Diag-
nostics Nichols Institute. The patients’ age ranges from newborn to 23 years old. A wide spectrum of phenotypic presentation was found, including intellectual disability, developmental delay, learning disability, cranio-skele-
tal anomalies, microcephaly, speech delay, seizures, endocrine abnormalities, and autism. All the deletions spanned the 520 kb commonly deleted region, varying from 478 kb to 909 kb. Four cases had small copy number variants of unclear clinical significance, such as a maternal duplication, 16p12.1 and a large or clinical significant copy number variant, one case had a 1.7 Mb gain of 22q11.2 at the DiGeorge syndrome critical region (OMIM #608363), and another had a 580 kb loss involving multiple exons of the NRXN1 gene (OMIM #600565). Our findings are unable to confirm the two-hit model as previously proposed. The absence of other large or clinically significant microdeletions or microduplications in 95% (35/37) of cases from this cohort demonstrates that the loss at 16p12.1 microdeletion alone may give rise to neuropsychiatric phenotypes as a single event.

First report on a multiplex, consanguineous, family with autism and chromosome 16p12.1 microdeletion: H. Mansour1, D. Eltabei1, V. Nimgaonkar1, D. Chennawi1, M. Sathanoori1, J.M. Sikela1, H. Mansour3, B.A. Lewis1, N.J. Morris1, S.K. Iyer2,4,6,7. 1) Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylva-

Purpose: Autism is a neurodevelopmental condition characterized behav-
iorally through social and communicative impairments and repetitive behav-
iors. DUf1220 protein domain copy number is dramatically increased specifically in the human lineage and has been implicated in brain growth and neurodevelopment. We previously demonstrated that, in individuals with autism, DUf1220 (subtype CON1) copy number is linearly associated, in a dose response manner, with increasing severity of each of the three core symp-
toms of autism. Given these findings we investigated associations of CON1 copy number and speech delay, a common marker of severity, in an addi-
tional independently assayed group of individuals with autism. Methods: DNA samples from an additional 50 non-Hispanic white unrelated individuals with autism were obtained from the Autism Genetic Resource Exchange (AGRE) and assayed by custom high density arrayCGH. Age of first mean-
ful word was obtained by the Autism Diagnostic Interview-Revised, a care-
giver interview used by clinicians to assist with autism diagnostics. Initially age of first meaningful word was dichotomized where greater than 24 months was classified as delayed first word. Unequal variance t-tests were performed to assess mean differences in DUf1220 copy number between delayed first word and not delayed groups. This was followed with a linear model examining the relationship of CON1 with month of speech acquisition adjusted for sex, age and simplex vs multiplex status. Results: Age of first word in this group ranged from 10 to 66 months with a mean of 28.7 months. On average individuals with delayed first word had a 0.065 (p=0.030) higher CON1 copy ratio (copy number) than the group who did not have a delayed first word. Similarly a linear association was identified which increased copies of CON1 were progressively associated with greater speech delay (p=0.019). Conclusions: Using ddPCR we previously impli-
cated the dosage of DUf1220 CON1 in the severity of the core symptoms of autism. Here we have extended these findings to include speech delay. This finding implicates an increased copy number of this gene in core symptoms of autism independent from the AGRE and assayed with different methods. Taken together, these findings lend further support to the view that DUf1220 CON1 dosage is an important contributor to autism severity, and that the same gene copy number is implicated in human brain expansion is also involved in the symptoms of autism.

First report on a multiplex, consanguineous, family with autism and chromosome 16p12.1 microdeletion: H. Mansour1, D. Eltabei1, V. Nimgaonkar1, D. Chennawi1, M. Sathanoori1, J.M. Sikela1, H. Mansour3, B.A. Lewis1, N.J. Morris1, S.K. Iyer2,4,6,7. 1) Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylva-

BACKGROUND: Several genetic risk factors have been detected for Autism spectrum disorders (ASD), including copy number variations (CNVs) such as duplications of the chromosomal 15q region. Maternal deletions or inactivation of genes in this region leads to Angelman syndrome, but if paternal, will cause Prader-Willi syndrome. Consanguinity is reported to be associated with the risk as well. METHOD: We report on a consanguineous, first-cousin marriage, Egyptian family having 4 siblings with ASD. We also report on 2 of these cases, an unaffected male and a maternal grandfather. Diagnoses were made following structured interviews, including the ADI-R. Array Compara-
itive Genomic Hybridization (aCGH), physical examination by clinical geneticist, IQ testing and EEG were conducted. Co-morbid seizure, intellec-
tual disability or ADHD have been reported in affected siblings. RESULTS: aCGH indicates that 3 of the 4 affected children have maternally-inherited interstitial duplication 15q11.2-q13. The mother and 3 affected siblings have intellectual disability. Several facial dysmorphic features are shared between mother and affected children, including the autistic child who does not have the duplication. CONCLUSION: To our knowledge, this is the largest multiply affected, inbred family reported to date with 15q interstitial duplication. How-
ever, the duplication does not segregate completely with ASD in this family, suggesting the presence of other risk factors, possibly related to deleterious maternal and paternal imprinting effects.

Purpose: Autism is a neurodevelopmental condition characterized behav-
iorally through social and communicative impairments and repetitive behav-
iors. DUf1220 protein domain copy number is dramatically increased specifically in the human lineage and has been implicated in brain growth and neurodevelopment. We previously demonstrated that, in individuals with autism, DUf1220 (subtype CON1) copy number is linearly associated, in a dose response manner, with increasing severity of each of the three core symp-
toms of autism. Given these findings we investigated associations of CON1 copy number and speech delay, a common marker of severity, in an addi-
tional independently assayed group of individuals with autism. Methods: DNA samples from an additional 50 non-Hispanic white unrelated individuals with autism were obtained from the Autism Genetic Resource Exchange (AGRE) and assayed by custom high density arrayCGH. Age of first mean-
ful word was obtained by the Autism Diagnostic Interview-Revised, a care-
giver interview used by clinicians to assist with autism diagnostics. Initially age of first meaningful word was dichotomized where greater than 24 months was classified as delayed first word. Unequal variance t-tests were performed to assess mean differences in DUf1220 copy number between delayed first word and not delayed groups. This was followed with a linear model examining the relationship of CON1 with month of speech acquisition adjusted for sex, age and simplex vs multiplex status. Results: Age of first word in this group ranged from 10 to 66 months with a mean of 28.7 months. On average individuals with delayed first word had a 0.065 (p=0.030) higher CON1 copy ratio (copy number) than the group who did not have a delayed first word. Similarly a linear association was identified which increased copies of CON1 were progressively associated with greater speech delay (p=0.019). Conclusions: Using ddPCR we previously impli-
cated the dosage of DUf1220 CON1 in the severity of the core symptoms of autism. Here we have extended these findings to include speech delay. This finding implicates an increased copy number of this gene in core symptoms of autism independent from the AGRE and assayed with different methods. Taken together, these findings lend further support to the view that DUf1220 CON1 dosage is an important contributor to autism severity, and that the same gene copy number is implicated in human brain expansion is also involved in the symptoms of autism.
113T

Screen for somatic mosaic mutations in unexplained Dravet syndrome patients. CT. Myers1, JM. McMahon2, IE. Scheffer2, HC. Metford1. 1) University of Washington, Department of Pediatrics, Division of Genetic Medicine, University of Washington, Department of Pediatrics, Division of Genetic Medicine, Seattle, WA, USA; 2) University of Melbourne, Department of Medicine, Florey Institute of Neurosciences and Mental Health, Austin Health, Melbourne, Australia.

Dravet syndrome is an infantile-onset epileptic encephalopathy characterized by refractory seizures, cognitive arrest or regression associated with ongoing epileptic activity, and a poor prognosis. 80% of Dravet cases are caused by de novo pathogenic mutations in the STXBP1 gene. However, a number of cases remain unexplained with molecular diagnosis is still unknown. Given the strong correlation between genotype and phenotype in solved Dravet cases, we hypothesize that somatic mutations in these known genes (SCN1A, GABRA1, and STXBP1) are causative in a subset of currently unsolved cases. Somatic mutations, well recognized as disease causing in cancer, are implicated in more than 30 monogenic disorders. We propose to screen for somatic mutations in a cohort of 51 patients with unexplained Dravet syndrome that have already undergone candidate gene screening or exome sequencing. Targeted sequencing using single molecule Molecular Inversion Probes (smMIPs) will be used to screen for low level somatic mosaicism in DNA from blood and buccal swabs. The smMIPs technology has been used to detect allele frequencies as low as 0.2% (Hiatt et al., 2013) and searching for mutations in another tissue type such as buccal swab or blood can enhance the detection sensitivity. This study will collect 51 blood samples, 10 buccal swabs, and are actively collecting more samples. We are also in the process of testing the smMIPs capture protocol and analysis pipeline for variant calling. It is likely that these studies will identify previously undetected mutations that will explain a subset of Dravet diagnoses.

113SS

Genome-wide linkage analyses of non-Hispanic White families identifies several novel loci for familial late-onset Alzheimer’s disease. J. Jaworski3, B.W. Kunkle4, C. Reitz5, A.C. Naj6, L.S. Cantwell6, A. Partch7, G.W. Beecham7, E.R. Martin8, W. Raskind9, R.P. Mayeux1, J.L. Haines2, L.A. Farrer2, G.D. Schellenberg5, M.A. Pericak-Vance1. 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 2) Taub Institute for Research on Alzheimer’s Disease, Columbia University, New York, NY, USA; 3) Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 4) School of Medicine, University of Washington, Seattle, WA, USA; 5) Department of Biostatistics & Bioinformatics, Case Western Reserve University, Cleveland, OH USA; 6) School of Medicine, Boston University, Boston University, Boston, MA, USA.

While more than two dozen loci that contribute to late-onset Alzheimer disease (LOAD) have been identified, few have been implicated in non-Hispanic white families. For LOAD, genome-wide association studies (e.g., PSEN2 mutations) that explain risk in families heavily burdened with LOAD have been found. Linkage/identity-by-descent (IBD) sharing analyses in large multiplex pedigrees represent a well-powered set of designs to isolate regions and loci of interest for sequencing studies. In order to identify highrisk families, we performed extensive parametric and non-parametric multipoint linkage analysis on 386 individuals in 42 non-Hispanic white (NHW) families (5-15 cases per pedigree). The 42 NHW families selected met three criteria: 1) exhibiting dominant inheritance of LOAD; 2) no mutations at known familial AD loci; and 3) low occurrence of the APOEε4 allele. All were among families selected for whole genome sequencing (WGS) through the genome wide neuroimaging study (GENIE) which has shown that LOAD is driven by the neurobiological systems implicated in neurodegeneration.

1136M

Microsatellites in the 5’ flanking region of AVPR1A were associated with social behavior scales of autism spectrum disorders in the Korean population. J. Park1, B.W. Kunkle2, K. Kim1, H. Gim3, Y. Jung3, H. Yoo1,2. 1) Seoul National University Bundang Hospital, Biomedical Research Institute, Gyeonggi-do, South Korea; 2) Seoul National University College of Medicine, Seoul, South Korea; 3) Department of Pharmacology, Eulji University College of Medicine, South Korea; 4) Department of Psychology, Chungbuk University, Chungju, Korea.

Background: Impairment in social interaction and communication is core features in autism spectrum disorder (ASD). The arginine vasopressin receptor 1A gene (AVPR1A) has been implicated in regulation of social behavior. In this study, we aimed to identify genetic factors that may influence social behavior in ASD patients in the Korean population.

Methods: We compared 178 patients (mean age, 15.91±3.67 years) and 452 healthy controls (mean age, 15.75±3.45 years) with AVPR1A microsatellite markers. ASSQ and AQ scores were completed by parents using a computerised module. We conducted association analyses of microsatellite markers with ASSQ and AQ scores.

Results: We observed significant association between AVPR1A microsatellites and social behavior scales. ASSQ and AQ scores were significantly lower in the AVPR1A microsatellite group compared to the control group (p=0.004).

Conclusion: Our findings suggest that genetic factors located on the X-chromosome microsatellites may influence social behavior scales in ASD patients in the Korean population.
1138S
Personality Genetics and Health in Super Seniors, J.M.T. Nelson1, 2, A.R. Brooks-Wilson1, 2, 3.
1) Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada; 2) Genome Science, Canada’s Michael Smith Genome Sciences Centre, Vancouver, British Columbia, Canada.

Background: Few seniors are free of chronic disease and the genetic factors that contribute to long term health are not well understood. We have established a study of super seniors, individuals who are over the age of 85 who have never reported being diagnosed with Alzheimer disease, cancer, diabetes, cardiovascular disease or pulmonary disease. It is important to determine why some individuals age with more comorbidities and others with none at all. Longevity and health are associated with specific personality traits. Personality consists of a person’s emotional, attitudinal and behavioural responses, which are related to neurotransmission in the brain. Genes that encode neurotransmitters, their receptors or affect processes in their metabolism are candidates for study on the effects of personality. Genetic variation in personality-related genes may affect health by influencing behaviour and lifestyle. Hypothesis: Genetic variants in genes involved in neurotransmission or that underlie personality disorders influence healthy aging. Objective: To determine if genetic variation in personality-related genes is associated with the super senior phenotype. Methods: 1) A literature search has been conducted for genes involved in the physiology of personality and personality-related disorders. Five neurotransmitter-related candidate genes have been chosen. In addition, a literature search was conducted for genome wide association studies (GWAS) of personality traits and personality-related disorders. Single nucleotide polymorphisms (SNP) have been selected and include tag SNPs, SNPs that meet genome-wide significance in GWAS of personality traits, and SNPs that were associated with personality traits in more than one study. 2) Genotyping of the SNPs will be performed using the Sequenom MassArray method. Variable number repeat random repeats will also be genotyped using PCR and fragment analysis. 3) We will test for association with the super senior phenotype using a set of 493 European-ancestry super seniors and 431 European-ancestry controls. Identifying personality-related genetic variants that influence healthy aging may lead to pharmacological opportunities to improve long term health in our aging population.

1139M
Gene-based pleiotropy across five major psychiatric disorders. D.R. Nyholt, H. Zhao. QIMR Berghofer, Brisbane, QLD, Australia.

Background: Studies have indicated genetic overlap between the five disorders in the Psychiatric Genomics Consortium (PGC): autism spectrum disorder (ASD), attention deficit-hyperactivity disorder (ADHD), bipolar disorder (BDP), major depressive disorder (MDD), and schizophrenia (SCZ). In this study, we aimed to identify specific genes overlapping the five psychiatric disorders utilizing a novel gene-based approach. Methods: Optimized gene-based tests were performed utilizing genome-wide association (GWA) results from the PGC analysis of single-nucleotide polymorphism (SNP) data for the five disorders in 33352 cases and 27886 controls of European ancestry. After accounting for correlation (i.e., non-independence) of gene-based test results due to linkage disequilibrium we examined the significance of the proportion of genes nominally associated across the five disorders. Pathway and network-based analyses were performed on the sets of genes significantly overlapping the disorders. Results: We found highly significant overlapping genes between SCZ and BDP, moderate overlap between SCZ and MDD, SCZ and ASD, and ADHD and BDP. After combining discovery sets, we found significant overlap across SCZ, BPD and MDD, across SCZ, BPD, MDD and ASD, and across BPD, MDD and ASD/ADHD. No significant overlap was observed between the individual adult-onset disorders and ADH. Pathways implicated by the genes overlapping the adult-onset disorders include MAPK signalling, calcium signalling, dopa ventral axis formation, chemokine signaling, and malaria. Pathways for BPD and ASD include glycophospholipid biosynthesis globo series, and pathogenic Escherichia coli infection. Pathways implicated by genes overlapping BPD and ADHD include glutathione metabolism, anabolic and catabolic processes. Additionally, combining gene-based association results across disorders identified numerous genes surpassing our cutoff for genome-wide significance. Discussion: Utilizing a novel approach, we identified numerous genes associated across multiple psychiatric disorders. These results extend previous findings from single SNP-based genetic overlap analyses by providing important insight into the likely genes and biological mechanisms underlying the observed genetic correlation and co-morbidity between these major psychiatric disorders.

1140T
Fine mapping of schizophrenia risk locus CSM1 (rs10503253) in Indonesian samples revealed association with BPD. D.B. Wildenauer1, A.R.A. Kusumawardhani1, D.M.B. Widenauska1, B. Benyamin2, S.G. Schwab2, Indonesian Schizophrenia Genetics Consortium. 1) University of Western Australia, Perth, Australia; 2) Department of Psychiatry, University of Indonesia, Jakarta, Indonesia; 3) Queensland Brain Institute, Brisbane, Australia; 4) Department of Psychiatry, University of Erlangen-Nuremberg, Germany.

Schizophrenia is a devastating complex genetic disorder with a highly polygenic background. Recent GWAS findings in large combined Caucasian samples suggested rs10503253, located in the intronic region of the CUB Sushi and multiple domains 1 gene (CSM1) on 8p23.2 as a risk locus for schizophrenia. We hypothesized that risk loci detected in Caucasian samples are also evident in samples from a different ethnic background. Our sample comprised 1067 individuals with schizophrenia and 1111 mono-affected controls consecutively ascertained in five Mental State Hospitals in the area of greater Jakarta, Indonesia. Structured interviews (DIP in Bahasa translation) were performed in all cases and consensus diagnosis made according to the DSMIV criteria by psychiatrists. Using a test panel with 374 evenly spaced SNPs (illuminia) all samples were tested for genotyping quality and sampling errors. Principal component analysis revealed ethnic homogeneity of cases and controls with close relationship to East Asian populations (genomic inflation factor 1.02). Association of rs10503253 at 8p23.2 was confirmed (P=0.0086, same marker, same allele). For fine mapping, we selected 22 SNPs in the region (4.17-4.225Mb) and genotyped these using Taqman technology. Lowest P-value was obtained for rs10102283 (nominal p=0.00083, p=0.02 adjusted for multiple testing), which is 2.5kb apart from rs10503253. Next we analysed the region further by imputing with variants identified in this region in Asian individuals by the 1000 genome project. We selected 134 variants with info ≥0.9. Four indels (MAF around 0.45) in the association peak (1.4kb) revealed nominal p-values of 3x10^-6 (P<0.005 after adjusting for multiple testing). Odds ratios were in the range of 1.4. CSM1 is expressed predominately in the brain. The protein plays a role in complement activation and has been found to be a binding site for micro RNAs. Even though the indels are intronic, they may affect structure and function, affecting splicing or transcription and thus quantitatively influence the function of the protein in complement activation.

1141S
Transethnic HLA comparison in narcolepsy. H.M. Ollila1, E. Mignot1, J. Faraco1, J. Ravel1, F. Harri1, L. Lin2, J. Hallmayer2, E. Mignot1.
1) Stanford University, Palo Alto, CA., Select a Country; 2) Department of Pulmonary, Critical Care Medicine, Peking University People’s Hospital, Beijing, China.

Narcolepsy is an autoimmune sleep disorder, which is characterized by severe daytime sleepiness. Recently, the onset of narcolepsy has been associated with the 2009 H1N1 influenza. In addition, it is strongly associated with Human leukocyte antigen (HLA) DQB1*06:02, and over 98 percent of narcoleptics carry this haplotype. However, the role of other HLA loci has remained unresolved. The aims of this study were to study the contribution of various HLA haplotypes to narcolepsy and clinical phenotypes related to narcolepsy. We used direct HLA genotyping and HLA imputation. The study was performed in Caucasian and Asian samples comprising altogether over 3000 cases and 10000 controls. Clinical phenotypes included age of onset, hallucinations, onset after vs. before 2009, absence of cataplexy, sleepiness and sleep paralysis. We found that DQB1*06:02 effect was strongly modulated by DQA1 genotype so that those carrying DQA1*01:02 had highest risk and those with DQA1*01 non 01:02 were protected from narcolepsy, suggesting allelic competition in narcolepsy. Analysis of clinical phenotypes revealed that cases carrying DQB1*03:01 had nearly two years earlier age of onset. DQB1*06:02 was rarer in individuals that got narcolepsy after 2009. The findings highlight the importance of HLA in the etiology of narcolepsy. Analysis of HLA-A, B, C and HLA-DR loci are ongoing.
1142M
Comparative sequencing of the PARK2/PACRG/QKI locus in Early Onset Parkinson’s disease. W.C. Macedo1,2, M.E.D. Sauer1, H.A.G. Teive2, R.P. Munhoz2, C.M. Probst2, M.T. Mira1. 1) Graduate program in health sciences, Pontificia Universidade Católica do Parana - PUCPR, Curitiba, Paraná, Brazil; 2) Department of Genetics, Universidade Federal do Parana - UFPR, Curitiba, Paraná, Brazil; 3) Department of Neurology, Hospital de Clinicas - UFPR, Curitiba, Paraná, Brazil; 4) Associação Paranaense dos Portadores de Parkinsonismo - APPP, Curitiba, Paraná, Brazil; 5) Department of Genomic and Bioinformatics, Instituto Carlos Chagas - FIOCRUZ-PR, Curitiba, Paraná, Brazil.

Early Onset Parkinson’s disease (EOPD [MIM 600116]) is a neurodegenerative disorder characterized by death of midbrain dopaminergic cells and clinical outcome before 45 years of age. EOPD is under strong genetic control; mutations in the PARK2 gene have been consistently associated with familial and non-familial autosomal recessive EOPD cases. Interestingly, an attempt to reproduce PD in PARK2 knockout mouse strain has failed. On the other hand, the Qk1 mouse, which develops vigorous limb shaking a few days after birth, are carriers of a spontaneous, recessive deletion of Quaking (Qk), Pacrg and Park2 neighboring genes located at chr17A. In humans, these three genes are also adjacent to each other and located at chr6q25-q27, homologous to the murine chr17A. We hypothesized that mutations in human QKI and/or PACRG and PARK2, together or not with PARK2 mutations, are necessary or sufficient for the development of EOPD. To investigate this hypothesis, we performed comparative sequence analysis of the coding sequences of QKI, PACRG and PARK2, seeking for potentially pathogenic variants. Study sample was composed by two DNA pools: one including 36 EOPD cases and the other, 17 individuals with age higher than 65 years of any neurodegenerative disease. Sequencing was conducted as implemented on the Ion PGM platform. Primary data was subjected to a filtering protocol designed to eliminate sequencing artifacts; variants approved by the quality filters were subjected to in silico prediction analysis, searching for changes in regulatory sites of mRNA, splicing sites and poly-A peptide sequence. A total of 82 variants were identified and 38 were validated by the quality criteria (seven new and 31 known). From this set, functional high impact variants were only found in PARK2; two of them were selected for PCR confirmation and performed in a subset of 60 PD families with confirmed consanguinity, using Agilent technologies and illumina HiSeq2000. 5) Department of Genomic and Bioinformatics, Instituto Carlos Chagas - FIOCRUZ-PR, Curitiba, Paraná, Brazil.

1143T
Strategies for identifying new genes in autosomal recessive Parkinson’s disease. S. Lesage1,2, A. Nicolas1, M. Sabahou2, V. Drouet2, C. Proukakis1, A. L. Leutenegger1, A. H. Schapira1, C. Grace2, E. Drouet2, K. Gancheva2, A. Dürr1, H. A. G. Teive1, O. Corti2, E. Kabashi1, A. L. Leutenegger1, A. Bricol1, French PDG group. 1) U1217, INSEIRM, Paris, Paris, France; 2) Inserm UMR S946, Fondation Jean Dausset, Paris, Paris, France; 3) Centre d’Investigation Clinique Pitié Salpétrieres, CNRS U1142, Paris, France; 4) Behavioural Neurology and Movement Disorders Unit, Department of Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey; 5) Faculté de Médicine, Université d’Alger, Algier, Algeria; 6) AP-HP, Pitié-Salpêtrière Hospital, Department of Genetics and Neurogenetics, Paris, France.

Early onset Parkinson’s disease (EOPD) is a neurodegenerative disorder characterized by death of midbrain dopaminergic cells and clinical outcome before 45 years of age. EOPD is under strong genetic control; mutations in the PARK2 gene have been consistently associated with familial and non-familial autosomal recessive EOPD cases. Interestingly, an attempt to reproduce PD in PARK2 knockout mouse strain has failed. On the other hand, the Qk1 mouse, which develops vigorous limb shaking a few days after birth, are carriers of a spontaneous, recessive deletion of Quaking (Qk), Pacrg and Park2 neighboring genes located at chr17A. In humans, these three genes are also adjacent to each other and located at chr6q25-q27, homologous to the murine chr17A. We hypothesized that mutations in human QKI and/or PACRG and PARK2, together or not with PARK2 mutations, are necessary or sufficient for the development of EOPD. To investigate this hypothesis, we performed comparative sequence analysis of the coding sequences of QKI, PACRG and PARK2, seeking for potentially pathogenic variants. Study sample was composed by two DNA pools: one including 36 EOPD cases and the other, 17 individuals with age higher than 65 years of any neurodegenerative disease. Sequencing was conducted as implemented on the Ion PGM platform. Primary data was subjected to a filtering protocol designed to eliminate sequencing artifacts; variants approved by the quality filters were subjected to in silico prediction analysis, searching for changes in regulatory sites of mRNA, splicing sites and poly-A peptide sequence. A total of 82 variants were identified and 38 were validated by the quality criteria (seven new and 31 known). From this set, functional high impact variants were only found in PARK2; two of them were selected for PCR confirmation and performed in a subset of 60 PD families with confirmed consanguinity, using Agilent technologies and illumina HiSeq2000. 5) Department of Genomic and Bioinformatics, Instituto Carlos Chagas - FIOCRUZ-PR, Curitiba, Paraná, Brazil.

1144S
Could somatic copy number alterations contribute to sporadic Parkinson’s disease? C. Proukakis1, K. Gancheva2, C. Grace2, D. Pease1, D. Brazma3, E. Kara1, J.W. Taanman1, H. Houliden1, A.H. Schapira1, E. Nacheva1, 1) Institute of Neurology, University College London, London, United Kingdom; 2) 2Academic Department of Haematology, University College London, United Kingdom.

Objective: to determine feasibility and sensitivity of detecting somatic copy number alterations (CNAs) in Parkinson’s disease brain. Background: Parkinson’s disease (PD) is usually sporadic, and despite heritability estimates of around 30%, the majority of the cases are unexplained. Alpha-synuclein (SNCA) protein is central to pathogenesis, but SNCA missense mutations and copy number variants (CNV) are rare. We have recently hypothesised that mosaicism due to post-zygotic somatic mutations in early embryogenesis in the neuronal cell lineage could contribute to pathology. We have already screened over 500 brain samples for somatic coding SNCA mutations, without detecting any. There is, however, recent clear evidence of somatic CNAs in brain, which are methods would not have detected, but their role in diseases remains unknown. Methods: We have designed a custom Agilent CGH array with dense spacing of probes for PD-related genes, including SNCA and PARK2. Dilution of DNA from a SNCA duplication carrier with wild-type to give various levels of increased SNCA DNA (thus creating artificial mosaics for CNAs) allows determination of the sensitivity, and comparison of different analytical methods. To determine if interphase cell FISH (fluorescent in situ hybridization) can allow reliable detection of lower level mosaicism, probes for SNCA and PARK2 will first be tested on fibroblasts with known CNVs. Results: Our array can detect 7.5-15% increase of SNCA genomic DNA, depending on the analytical settings used. FISH probes for SNCA and PARK2 have been verified in both normal and positive control fibroblasts. Reproducible data can be obtained by FISH analysis of substantia nigra. Conclusion: We have developed methods and tools to allow us to further investigate our hypothesis of somatic CNAs in sporadic PD cases.
1145M

Whole-genome sequencing to identify genes implicated in Familial Parkinsonian Tauopathy. M. Sanchez-Contreras 1, S. Fujioka 1, C. Pottier 2, A.J. Strongosky 2, P.C. Boeve 2, J.E. Parisi 3, P.M. Tack 2, N. Aoki 1, M. Baker 1, V. Sossi 2, D.W. Dickson 3, A.J. Stoessl 3, O.A. Ross 1, Z.K. Wszolek 4, R. Rademakers 5, 1) Department of Neurosciences, Mayo Clinic Jacksonville, FL; 2) Department of Neurology, Mayo Clinic Jacksonville, FL; 3) Department of Neurology, Mayo Clinic Rochester, MN; 4) Department of Physics and Astronomy, University of British Columbia, BC, Canada; 5) Pacific Parkinson’s Research Centre, Division of Neurology, University of British Columbia & Vancouver Coastal Health, BC, Canada.

Familial Parkinsonian Tauopathy (FPT) is a group of neurodegenerative disorders characterized by the pathological inclusions of phosphorylated tau protein. In some entities, as Alzheimer’s disease, genetic variants have been proposed to contribute to tau pathohlogy. However, genetic causes of other less common tauopathies, as Progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD), are poorly understood. To identify novel causal genes implicated in tauopathies, we provided a detailed study of a large family of 64 individuals with 8 affected patients with an inherited tauopathy that is likely autosomal dominant with reduced penetrance. The index case developed progressive speech and language difficulties at age 64 years. Examination 4 years later disclosed non-fluent aphasia, word-finding difficulties, circumspection, frontal release signs, and right-sided bradykinesia, rigidity, and pyramidal signs. She died 3 years later, at the symptomatic onset. Neuropathologic features included numerous ballooned neuronfilament-positive neurons, tau-positive astrocytic plaques, and oligodendroglial coiled bodies, all typical of CBD.

Two other family members were diagnosed clinically with Parkinsonism and behavioral problems, 2 with Parkinson’s disease, 1 with amyotrophic lateral sclerosis, 1 with dementia with Lewy bodies and 1 with severe dementia with frontotemporal lobar degeneration. DNA was available from the proband and one first-degree cousin, clinically diagnosed with dementia and Parkinsonism at the age of 54. After exclusion of mutations in MAPT, PGRN and LRRK2 in these affecteds, we performed whole-genome sequencing. Analysis of the genomes of these patients resulted in a list of novel and rare variants that were shared among the affecteds and follow an autosomal dominant inheritance pattern. This list of potential tauopathy variants was confirmed by Sanger sequencing and subsequently screened in a series of healthy control individuals. We identified six novel missense variants in 6 genes that are highly associated in a total of 6 confirmed-variants which were absent or very rare in controls. SCN1A, OPRL1, CAPRIN2, UBN1, NEURL4 and CCDC9. The presence of additional variants in these candidate tauopathy genes currently is being studied in additional families with a similar phenotype as in 3 additional tauopathy families. This approach provides a select list of potential tauopathy genes by whole-genome sequencing in a family with pathologically confirmed CBD that may help identify novel pathways involved in pathological tau aggregation.

1146T

Distinct genetic variants in Alzheimer’s disease, Parkinson’s disease and type 2 diabetes. S. J. Chung1, S. Y. Kim1, J. Kim1, M. -J. Kim2, Y. Kim1, J.H. Lee1, 1) Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea; 2) Psychiatry, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea; 3) Neurology, Bobath Memorial Hospital, Seongnam, South Korea.

An association of the dopaminergic system with Alzheimer’s disease (AD) and Parkinson’s disease (PD) is well established. We aimed to investigate whether genome-wide significant loci of type 2 diabetes mellitus (T2DM) are associated with AD and PD. Study subjects were 400 AD cases, 500 PD cases, and 500 controls. Genomic DNA was available from the proband and one first-degree cousin, clinically diagnosed with dementia, 1 with multiple sclerosis, 1 with dementia, and 1 with progressive gait and speech problems. She died 5 years after the symptomatic onset. Neuropathologic features include classical neurofibrillary tangles and neuritic plaques. The death was due to a complication of severe depression. A full postmortem examination was not allowed. A diagnosis of CBD was assigned postmortem.

We aimed to investigate whether genome-wide significant loci of type 2 diabetes mellitus (T2DM) are associated with AD and PD. Study subjects were 400 AD cases, 500 PD cases, and 500 controls. Genomic DNA was available from the proband and one first-degree cousin, clinically diagnosed with dementia, 1 with multiple sclerosis, 1 with dementia, and 1 with progressive gait and speech problems. She died 5 years after the symptomatic onset. Neuropathologic features include classical neurofibrillary tangles and neuritic plaques. The death was due to a complication of severe depression. A full postmortem examination was not allowed. A diagnosis of CBD was assigned postmortem.

We found that T2DM and Parkinson’s disease (PD), both age-related chronic conditions, are poorly understood. To identify novel causal genes implicated by the pathological inclusions of phosphylated tau protein. In some entities, as Alzheimer’s disease, genetic variants have been proposed to contribute to tau pathohlogy. However, genetic causes of other less common tauopathies, as Progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD), are poorly understood. To identify novel causal genes implicated in tauopathies, we provided a detailed study of a large family of 64 individuals with 8 affected patients with an inherited tauopathy that is likely autosomal dominant with reduced penetrance. The index case developed progressive speech and language difficulties at age 64 years. Examination 4 years later disclosed non-fluent aphasia, word-finding difficulties, circumspection, frontal release signs, and right-sided bradykinesia, rigidity, and pyramidal signs. She died 3 years later, at the symptomatic onset. Neuropathologic features included numerous ballooned neuronfilament-positive neurons, tau-positive astrocytic plaques, and oligodendroglial coiled bodies, all typical of CBD.

Two other family members were diagnosed clinically with Parkinsonism and behavioral problems, 2 with Parkinson’s disease, 1 with amyotrophic lateral sclerosis, 1 with dementia with Lewy bodies and 1 with severe dementia with frontotemporal lobar degeneration. DNA was available from the proband and one first-degree cousin, clinically diagnosed with dementia and Parkinsonism at the age of 54. After exclusion of mutations in MAPT, PGRN and LRRK2 in these affecteds, we performed whole-genome sequencing. Analysis of the genomes of these patients resulted in a list of novel and rare variants that were shared among the affecteds and follow an autosomal dominant inheritance pattern. This list of potential tauopathy variants was confirmed by Sanger sequencing and subsequently screened in a series of healthy control individuals. We identified six novel missense variants in 6 genes that are highly associated in a total of 6 confirmed-variants which were absent or very rare in controls. SCN1A, OPRL1, CAPRIN2, UBN1, NEURL4 and CCDC9. The presence of additional variants in these candidate tauopathy genes currently is being studied in additional families with a similar phenotype as in 3 additional tauopathy families. This approach provides a select list of potential tauopathy genes by whole-genome sequencing in a family with pathologically confirmed CBD that may help identify novel pathways involved in pathological tau aggregation.

1147S

Defects of ARHGAP36 in patients with developmental delay and autism. S. Fan1, L. Dukes-Rimskey1, R.C. Rogers1, A. Chaube1, C. Skinner1, L. Wang2, A.K. Srivastava1,2, J.C. Seif Research Institute for Human Genetics, Greenwood Genetic Center, Greenwood, SC; 2) Department of Genetics and Biochemistry, Clemson University, Clemson, SC.

Rho family GTPase activating proteins (GAPs) are known to regulate physiological functions including embryogenesis, neurodevelopment, cytogenesis and differentiation and act as negative regulators of Rho GTPases. Several genes implicated in intellectual disability (ID) and autism spectrum disorders code for proteins that function as regulators or effectors of Rho signaling. In our study, we identified 2 novel, rare variants of ARHGAP36 in 1 patient with ID and 245 male patients with developmental delay/autism and identified a novel missense alteration, c.1316G>A (p.R439H), in a 16-year-old male patient with absent speech, hyperactivity and autism. The alteration was absent in 8500 individuals in the NHLBI dataset. Defects of ARHGAP36 have been identified in at least 7 other patients, including individuals with autistic spectrum disorder, intellectual disability and dysmorphic features and severe hypothyroidism. Thus, the deletion of the IGSF1 gene is likely responsible for hypothyroidism in the proband and we speculate that his younger brother is also likely to develop hypothyroidism. To understand the physiological function of ARHGAP36, we further determined ARHGAP36 expression in several tissues including human fetal brain. We identified and analyzed genes that are co-expressed with ARHGAP36. On the basis of probe-to-probe correlation coefficient calculated from an integrated set of 2,968 microarray expressions from 7 healthy human tissues, we found that ARHGAP36 was co-expressed with genes significantly enriched with Gene Ontology terms including synapse, postsynaptic membrane, synaptic transmission and neurogenesis. Altogether, our findings suggest a potential role for the ARHGAP36 gene in developmental delay and autism.

1148M

HIV-related cognitive impairment shows association with polymorphisms within the dopaminergic system in substance dependent and independent populations. R. J. Jacobs1, M. Pizzurutto1, J. Murray1, S. Morgello1,2,3, 1) Neurology, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Neurosciences, Icahn School of Medicine at Mount Sinai, New York, NY; 3) Pathology, Icahn School of Medicine at Mount Sinai, New York, NY.

It has been postulated that drugs of abuse act synergistically with HIV, leading to increased neurotoxicity and neurocognitive impairment. The CNS impacts of HIV and drug use converge on the mesocorticolimbic dopamine (DA) system. Use of popular substances of abuse, such as alcohol, marijuana and methamphetamine, as well as opiate and cocaine dependence. To replicate and expand these studies, we have increased our sample size and added additional polymorphisms within several genes of the dopaminergic system (DRD1-5, COMT, DAT). We observed that polymorphisms of DRD2 are associated with opiate and cocaine dependence in our population. In Caucasian subjects, we observed significant association of multiple genes and dopamine dependence with polymorphisms within DRD2 and COMT genes, while a significant association within the DRD1 gene was observed in African-American individuals. Using linear regression analysis, we next examined these polymorphisms for associations with neuropsychological performance in global and cognitive domain T-scores (Motor, Processing Speed, Verbal Fluency, Learning, Memory, Executive Functioning, Working Memory) while controlling for opiate and cocaine dependence. While significant associations were observed in nearly every domain across both populations for multiple polymorphisms, the most consistent relationships were observed in the motor domain with several DRD2 polymorphisms while African-American subjects had its most significant associations in working memory (COMT and DRD3) and memory (DRD1) domains. For all of these associations, the odds of observing significant correlations of the same sign and direction were nearly 6 times greater than random, indicating that the correlations were opposite to what was seen in subjects without dependence. Future studies will focus on increasing the sample size of the population as well as increasing the number of genes within the dopaminergic circuitry. We conclude that studies to examine genetic risk for HAND must carefully account for substance dependence patterns when assaying dopaminergic systems, as the neurobiological substrates of cognition in HIV populations may vary with tonic alterations secondary to chronic substance exposures.
1149T  HLA-DRA is strongly associated with Parkinson’s disease in Iranian population. J. Jamshidi1, H. Darvish2, B. Emanuellazadeh3, A. Movafagh4, 1) Department of Biochemistry, Fasa University of Medical Sciences, Fasa, Iran; 2) Department of Medical Genetics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Background and Objective: Parkinson’s disease (PD) is a progressive neurodegenerative disorder which impairs the patient’s motor skills. The rs3129882, a noncoding variant in HLA-DR, was found to be associated with PD using several genome-wide association studies (GWAS). The aim of this replication study was to explore the relationship between this variant and PD in Iranian patients. Materials and methods: A total of 520 unrelated patients and 520 healthy controls. The mean age of patients was 59.5±14.5 years and were composed of 278 men and 242 women. The diagnostic criteria for PD were based upon the UK Parkinson’s disease Society Brain Bank Clinical Diagnostic Criteria. Genomic DNA was extracted from peripheral blood using a standard salting out method. The rs3129882 variant was genotyped through polymerase chain reaction—restriction fragment length polymorphism (PCR-RFLP). We examined the association of rs3129882 (A/G) with the risk of PD, under four models (codominant (G/G vs. G/A vs. A/A), dominant (A/A or G/A vs. G/G), recessive (G/G or G/A vs. G/A) and overdominant (G/G or A/A vs. G/A)) using SNPassoc package of R version 3.0.1. Odds ratio (OR) together with 95% confidence interval (CI) was estimated; and a p < 0.05 was considered as statistically significant for the tests. Results: Distributions of the rs3129882 (A/G) polymorphism in both PD and control group were in Hardy-Weinberg equilibrium. The alleles frequencies of rs3129882 (A/G) were distributed differently in PD group and control group (p2 = 4.641, OR1=1.209, 95% CI: 1.017-1.436, P = 0.033). Our results indicate that the overdominant, dominant and recessive association of the SNP with PD risk was significant, where the A allele was observed to be protective. Discussion: HLA is associated with numerous neurological disorders and association of HLA with Parkinson’s disease was recently revealed. Our result suggests that the rs3129882 (A/G) polymorphism may be a risk factor for PD in Iranian population. It is assumed that the intrinsic variant plays a role as a cis-acting regulatory element which correlates significantly with overexpression of DR antigens in substantia nigra. The replication in our study is not so far beyond expectation because of the previous evidences for involvement of neuroinflammation and adaptive immunity in PD pathogenesis. This association emphasizes an important biological pathway in the etiology of the disease and points to a potential target for new therapies.

1150S Ancestral haplotypes of BHLHE40 in non-24-hour rhythms and bipolar disorder. D.F. Kriple1, S.A. Ament2, C.M. Nierengert3, J.R. Kelsoe1, 1) Psychiatry, UCSD, La Jolla, CA; 2) Institute for Systems Biology, Seattle, WA.

A previous study (Kriple et al., Psychiatry Investigation, in press, 2014) demonstrated a haplotype group in the circadian gene BHLHE40 that was associated with patients displaying non-24-hour free-running circadian rhythms. This haplotype group and several of its SNVs were also nominally associated with bipolar disorder. The only exonic variant of the 16 in the haplotype group, rs908078, is a synonymous variant which may alter splicing. Also, some variants may alter transcription binding sites. With further study of the haplotypes, we now found that 7 of the 16 key minor variants were identical with the chimpanzee reference, and 6 matched the reference allele of the majority of 9 non-human primates. Each of these 7 variant alleles was also reported at least once in available early sequencing data from 3 Neanderthals and the Denisovan, so they likely have an ancestral primate origin. According to 1000 Genome data (Phase 1), 5 of the 7 minor alleles are distinctly more common among Africans than Europeans or Asians. In sum, current data suggest 7 SNVs and 7 ancestral haplotypes. Although each of these variants has been negatively selected in modern humans, particularly outside of Africa, we group have not been negatively selected in modern humans, particularly outside the tropics. Perhaps altered circadian dynamics were favored after the discovery of fire and especially after migration to more northern latitudes where more time must be spent inside shelters. This would not be surprising, since Drosophila and several small mammals have recognized genetic variation among populations related to adaptations to latitude and photoperiod. Whole-genome sequencing of BHLHE40 was now examined 325 members of multiplex bipolar disorder pedigrees, among whom 74 individuals displayed these SNVs, but the sequencing was too shallow to examine association powerfully. Recent PGC GWAS analyses have given no support to any BHLHE40 variant association with bipolar disorder, but Selvadur (et al. 2010) found a nominally-significant association with down-regulated expression of BHLHE40 by beta-cells of bipolar disorder subjects. The ancestral haplotype group has not been well represented in GWAS studies, further replication of expression studies and expanded studies with whole-gene-region sequencing will be needed to clarify the import of BHLHE40 for circadian and affective disorders.

1151M MAPT non-coding variation in neurodegenerative disorders. C. Labbé5, M. Heckman6, K. Ogaki1, O. Lorenzo-Belancor1, A. Orotola1, R. Walton5, D. Senior5, R. Uitti1, Z. Wszolek5, O. Ross1, 1) Neurosciences, Mayo Clinic, Jacksonville, FL; 2) Section of Biostatistics, Mayo Clinic, Jacksonville, FL; 3) Neurology, Mayo Clinic, Jacksonville, FL.

Genome-wide association studies have identified over 20 loci associated with increased risk to sporadic Parkinson’s disease (PD). The MAPT gene is one of the most systematically replicated locus in PD. The gene encodes protein tau which aggregate in brain inclusions and define the neurodegenerative diseases called tauopathies. A common non-recombining MAPT haplotype (MAPT H1) has been associated to several tauopathies including progressive supranuclear palsy (PSP); yet MAPT H1 has also been implicated in the risk to Parkinson’s disease (which is not a traditional tauopathy). Preliminary sub-haplotype analyses suggest that different genetic variants on the MAPT H1 haplotype associate with each of these disorders. To date it remains unclear which variant(s) at the MAPT locus is(are) responsible for the risk and what is the underlying pathomechanism of disease. We set out to identify causal variants for PD and PSP within the MAPT region using next generation sequencing technologies. We captured the entire MAPT gene and 10kb on each side, a 154kb genomic region that was sequenced in 300 patients with PD, 300 patients with PSP, and 300 controls using a pooling strategy (10 DNA samples/pool). We designed 4248 amplicons using the Haloplex platform following Illumina sequencing for a total coverage of 96.3%. We identified over 4000 variants in the coding, intronic and surrounding region of the MAPT gene. Among these variants 39% are singletons, 19% have a minor allele frequency between 1% and 5%, and 42%, over 5%. We identified 43 variants located within 25 base pairs of an intron/exon boundary, of some which map to boundaries of the well-known exons 2, 3, 4a, 6, 8 and 10. We selected SNPs for stage 2 based on association tests and are using the OpenArray platform to genotype the variants in our replication series consisting of a total of 4500 independent samples. We will present results using the latest sequencing and genotyping technologies to comprehensively define the MAPT locus associated with PD and PSP and thus identify novel targets for both neuroprotective and symptomatic therapies.

1152T 9.6% of mouse gene knockouts show abnormal neuroanatomy: a resource to identify genes related to intellectual disability in human. A. Mikhailova1, A. Baud2, V. Vancolle3, A. Edwards4, M. Kannan5, H. Whiteley5, C. Wagner5, A. Durak5, I. Henri1, K. Navarro5, A. Delay5, S. Jayaram5, J. Estabe5, D. Wright5, D. Fairley5, C. Lelliott5, J. White5, D. Adams5, D. Keay5, J. Flint5, Y. Heraldt5, A. Reymond5, B. Yalcin5, 1) Center for Integrative Genomics, University of Lausanne, Switzerland; 2) EMBL-European Bioinformatics Institute, Hinxton, Cambridge, CB10 1SD, UK; 3) Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1HH, UK; 4) Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, UK; 5) Institute of Genetics and Molecular and Cellular Biology, Illkirch, 67404, France; 6) HistologiX Ltd, Biocity, Nottingham, NG1 1GF, UK; 7) European Bioinformatics Institute, Hinxton, Cambridge, CB10 1SD, UK; 3) Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, UK; 5) Institute of Genetics and Molecular and Cellular Biology, Illkirch, 67404, France.

Although intellectual disability affects 1-3% of the population, it is one of the least understood health problems. It is estimated that genetic lesions account for half of the currently diagnosed cases. Despite recent successes in identifying some of the mutations responsible, it has been suggested that up to 1000 more genes remain to be uncovered. The large number of intellectual disability syndromes is due to many causal pathophysiological mechanisms. A recent method for generating an array of quantifiable neuroanatomical abnormalities. To identify genes related to intellectual disability, we are collaborating with the Sanger Mouse Genetics Project (MGP), allied to the International Mouse Phenotyping Consortium (IMPC). We systematically study the neuroanatomy of the MGP/IMPC knockout mouse strains using a standardized set of 78 brain parameters. So far, we have assessed brain defects in 425 knockout mouse mutants (this number will double by the end of 2014). These preliminary data yielded success with the identification of 20 known intellectual disability genes including Acpa1, Ccnd1, Chd7, Mep1, Sc4m1 and Ube3b discovery of the perturbation of our approach. We also discovered 21 other genes including Mta1, Cdc104, Caprin2 and Dusp3, which when disrupted caused modification of brain structures. Our study is the largest screen of brain morphology from 1000 different models, with 21 known intellectual disability genes identified among 10% of knockout mouse mutants, and that these translate into human pathology. This offers a complementary resource to human genetic studies.
Homozygous deletions of non-coding transcriptional control sites in autism spectrum disorder. K. Schmitz-Abe, M. Higl, G. Sanchez-Schmitz, J. Bartlewski, B. Mehta, S. Servatella, A. Lamm, E. Morrow, T. Yu, C. Walsh, K. Markianos. 1) Division of Genetics and Genomics, Department of Medicine, Boston Children's Hospital, Boston, Massachusetts, USA, 02115; 2) Manton Center for Orphan Disease Research, Boston Children's Hospital, Boston, Massachusetts, USA, 02115; 3) Harvard Medical School, Boston, Massachusetts, USA, 02115; 4) The Autism Consortium, Boston, Massachusetts, USA, 02115; 5) Howard Hughes Medical Institute, Boston Children's Hospital, Boston, Massachusetts, USA, 02115; 6) Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA, 02114; 7) Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA; 8) Department of Genetics and Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, USA 02115, USA.

Autism spectrum disorder (ASD) is characterized by extensive genetic heterogeneity. A subset of ASD has been linked to disruptions of gene coding regions by de novo copy number variants (CNVs), de novo single nucleotide variants (SNVs), or inherited biallelic SNVs but most cases remain unexplained. We analyzed CNVs in 183 consanguineous families with ASD and found that individuals affected with ASD have significantly more homozygous deletions compared to unaffected siblings [22% vs. 15%]. This excess suggests that homozygous deletions might account for up to 7% of ASDs in our cohort. In contrast to most de novo CNVs and SNVs, we found that the homozygous deletions often did not include exons, but were enriched for DNA regulatory regions, with 12 of 16 disrupting ENCODE regulatory elements, a rate much higher than expected by chance (bootstrap p<0.05).

Specifically, we cross-referenced the 16 deletions with histone modification marks from the ENCODE project. We selected H3K4Me1, H3K4Me3, and H3K27Ac as three of the best characterized histone modifications corresponding to regions of active transcription and enhancer activity, and then examined ChIP-seq data gathered from 7 cell lines (GM12878, H1-hESC, H5MM, HUVEC, K562, NHEK, NHLF). Our data suggest an important new mechanism of ASD, and highlight the importance of patterned gene activation in cognitive and social function.
Schizophrenia is one of the most debilitating disorders affecting about 1% of the world population. The etiology is complex involving a major genetic contribution as well as environmental factors interacting with the genetic susceptibility. Several lines of evidence from post-mortem and brain imaging in schizophrenia patients suggest that Gamma-amino butyric acid (GABA) deficits may contribute to the pathophysiology of schizophrenia. Changes in GABA receptors in schizophrenic brain may represent a primary pathogenic mechanism, such as modifications to GABA receptor subunit assembly and/or structure contribute to the symptoms of schizophrenia. The GABA receptors are a class of receptors that respond to the neurotransmitter gamma-amino butyric acid (GABA), the chief inhibitory neurotransmitter in the vertebrate central nervous system. Brain-derived neurotrophic factor (BDNF) is one of the most important modulators of glutamatergic and GABAergic synapses. BDNF can play a permissive role in shaping synaptic networks, making them more susceptible for the occurrence of plastic changes. Reelin (RELN), a neuregulator secreted by GABAergic interneurons regulates the synaptic plasticity. Together, these genes play an important role in regulation of neurotransmission and synaptic plasticity. The objective of the study was to investigate the association of polymorphisms in GABA (A) receptors, BDNF and RELN genes with schizophrenia in a South Indian population. DNA was isolated from 300 patients and 300 healthy controls after obtaining informed consent. The study was approved by the institutional ethics committee. Samples were genotyped using allele specific amplification followed by fluorescence detection. Genetic association analysis of single nucleotide polymorphisms: rs2279020 (GABRA1), rs3219151 (GABRA6), rs2229944 (GABRB2), and rs211037 (GABRG2), BDNF -270C/T, rs6265(BDNF), rs727531(RELN), rs7341475(RELN), rs6951875(RELN), rs362719 (RELN) with predisposition to schizophrenia. Samples were genotyped using allele specific amplification followed by fluorescence detection (KASPar), PCR-RFLP and PCR sequencing. Allele and genotypic frequencies were calculated for patients and controls and association with predisposition to schizophrenia was tested.
1159S
Common and rare genetic risk factors converge in protein interaction networks underlying schizophrenia. X. Chang1, L. Lima1, Z. Wei1, J. Lu1, P. Steinman1, H. Hakonarson1. 1) Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Department of Computer Science, New Jersey Institute of Technology, Newark, NJ, USA.
Schizophrenia is an idiopathic brain disorder with profound genetic heterogeneity. The major forms of genetic risk factors in schizophrenia are common single nucleotide polymorphisms (SNPs) and rare copy number variants (CNVs). Recent studies have shed new light on the genetic basis of schizophrenia that de novo mutations play a prominent part in sporadic form of schizophrenia. However, the underlying genetic mechanisms remain unclear. We first performed a network-based pathway analysis on our previous meta-analysis of GWAS in schizophrenia and bipolar disorder, which contains 13,394 cases and 34,676 controls from 16 cohorts. The largest connected component (LCC) identified in our PPI network is significantly enriched in many pathways relevant to synaptic plasticity, neuronal development and signaling transduction such as long-term potentiation, RAP1 signaling pathway, GmRH signaling pathway. We also observed the proteasome pathway and ubiquitin mediated proteolyis pathway to be significantly enriched. Dysfunction of the ubiquitin-proteasome pathway has been implicated in the pathology of various neurodegenerative diseases. However, growing evidence revealed that schizophrenia patients have aberrant gene expression patterns in the UPP suggesting the UPP may also contribute to the deficits in schizophrenia. In order to add more genetic pieces to the schizophrenia puzzle, we collected literature reported genes disrupted with CNVs and de novo mutations in schizophrenia patients for the network analysis. We found the size of LCC were significantly increased, bigger than 10k genes. Based on known common SNPs and de novo mutations group of interacted genes with a significant combined effect to schizophrenia, we developed an edge-based network search algorithm for detecting causal gene modules in PPI networks. The most significant module is comprised by UBP1, CTNNA3 and N-methyl-D-aspartate (NMDA) receptor associated genes. We then combined the top three modules and constructed a small sub-network of 47 nodes and 76 edges, which is highly enriched in long-term potentiation and Calcium signaling pathway. In summary, we demonstrated that gene-gene interactions and common variants are likely to influence same genes or functionally interacting genes.

1160M
Alzheimer’s disease (AD) is a complex disorder influenced by genetic and environmental factors. We previously analyzed the phenotypic and genetic variance explained by SNPs known to be associated with AD, APOE alone, and by ~2 million genotyped or imputed SNPs from the Alzheimer’s Disease Genetics Consortium (ADGC), using the Genome-wide Complex Trait Analysis (GCTA) software. We found that 33% of total phenotypic variance is explained by common SNPs (n~2 million), 6% by APOE alone (including the e2 and e4 alleles), and 2% by other known markers—meaning 25% of phenotypic variance remains unexplained by known AD markers. In this study we follow-up our previous work using an expanded version of the ADGC dataset. In this expanded version, SNPs were imputed using the 1000 Genomes dataset, resulting in more complete imputation of both common and rare variation. In the present analysis we estimate phenotypic and genetic variance explained by the SNPs. Additionally, while our previous study included only APOE and nine other loci, the present study includes all the IGAP and previously known AD SNPs (n=23, includes APOE e4 and e2). In addition we will evaluate the contribution of SNPs in genes known to harbor rare variants for AD (APP, PSEN1, PSEN2, TREM2, P3D3, ADAM10). We discuss the current state of known AD heritability and suggest where resources should be focused to identify the remaining undiscovered AD SNPs.

1161T
Probing the shared polygenic underpinnings of anorexia nervosa and five other major psychiatric disorders. L.M. Huckins1, K.S. Mitchell2, L. Thorton3, D. Coller1, P.F. Sullivan4, C.M. Buikse5, E. Zeggini1, WTCCC3 Consortium; GCAN Consortium. 1) Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Boston University, MA, USA; 3) University of North Carolina at Chapel Hill, NC, USA; 4) King’s College London, London, UK.
Anorexia nervosa (AN) is marked by extremely low body weight and intense fear of gaining weight. Comorbidity with other psychiatric disorders (PsDs) is common (up to 55.2%) We evaluated shared genetic determinants of AN and PsDs by testing whether polygenic risk scores derived from genome-wide data of other PsDs can predict AC status. We obtained allele score for major depressive disorder (MDD), bipolar disorder (BPD), autism (AUT), attention deficit hyperactivity disorder (ADHD) and schizophrenia (SCH), and a set of cross-disorder risk alleles, from the Psychiatric Genomics Consortium (PGC). We divided each of these sets into 10 Pt significance level thresholds. The test set comprised AN cases and controls from a published WTCCC3 AN GWAS study. For every sample in this test set, we produced a polygenic risk score as a weighted sum of allele risk scores. Logistic regression was used to assess whether each of the five polygenic scores predicted AN case-control status. We computed pseudo R^2 values, and compared these to the values obtained when randomly permuting case-control status to measure the proportion of variance in AN explained by each PsD risk score. Our pseudo R^2 values were ~0.5-1%, comparable to pseudo R^2 values found by the PGC in a recent cross-disorder polygenic score analysis of five PsDs [1]. We used the baseline results to compute an empirical p-value for each significance threshold and compared these to the baseline significance threshold, at the lowest Pt threshold (Pt<0.001) for AN vs AUT (p=0.0009), MDD (p=0.009), and SCZ (0.0008), and at the second lowest Pt threshold (Pt<0.01) for BPD (p=0.0004). We demonstrated for the first time a shared genetic etiology between AN and other PsDs using genome-wide data. Cross-disorder analyses may prove to be a powerful approach to identifying overlapping susceptibility loci across PsDs [1] Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) “Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis” The Lancet 381:1371-1379.

1162S
Genes regulated by epigenetic mechanisms in determining general intelligence (g) are over-represented in disorders that affect cognition. P. Cha1, K. Kobayashi1, Y. Ando1, R. Taka2, T. Miyakawa3, T. Toda1, 4. 1) Division of Molecular Brain Science, Kobe University Graduate School of Medicine, Kobe, Hyogo, Japan; 2) Center for Genetic Analysis of Behavior, Section of Behavior Patterns, National Institute for Physiological Sciences, Okazaki, Aichi 444-8585, Japan; 3) Division of Systems Medical Science, Institute for Comprehensive Medical Science, Fujita Health University, Toyoa, Aichi 470-1192, Japan; 4) Division of Neurology, Kobe University Graduate School of Medicine, 7-5-1 Kusunokicho, Chuo-ku, Kobe 650-0017, Japan.
General intelligence (g) is a common core shared by cognitive tasks. Despite the high heritability of “g” that increases with age, molecular mechanisms that underlie “g” remains poorly understood. To study epigenetic mechanisms associated with “g,” we conducted 5 mouse behavioral studies on 41 inbred mice and identified mice with high, medium, and low cognition. Through subsequent genome-wide gene expression profiling by using hippocampal RNAs of these cognitively discordant mice followed by gene-set enrichment analysis (GSEA), we identified 17 genes and several pathways that are expressed differentially between mice with high and low cognition. Since inbred mice are genetically homogeneous, we expect most of the differences that we observed are resulting from epigenetic mechanisms. In addition to memory and learning, the role of epigenetic mechanisms in cognition and neurological disorders has recently come to the lights. To investigate if genes implicated in our study provide insights into mechanisms underlying “g” and disorders that affect cognition, we evaluated enrichment of genes reported to be associated with each of the cognitive function/intelligent quotient (IQ), intellectual disability (ID), autism, Parkinson’s disease (PD), Alzheimer’s disease (AD), Rett Syndrome, and Schizophrenia in the top 100 genes, as well as their interacting proteins, identified in our study. Information on genes reported to be associated with IQ, ID and each of the aforementioned disorders was retrieved from published databases (PubMed and/or the databases (AlzDB, AlzGene etc), whereas interacting proteins were identified from Protein Interaction Network Analysis (PINA) platform. Overlap enrichment analysis indicated that the 100 genes or genes encoding their interacting proteins are significantly enriched in genes reported to be associated with IQ, ID, PD and AD (p<0.05 in 100,000 permutation test). Subsequent investigation of the commonly shared genes may shed light on common mechanisms and pathways that underlie “g” and the aforementioned disorders.
1163M
Novel Locus in 15q23 Implicated in Recovery after Severe Traumatic Brain Injury. Y.P. Conway1,2, D. Ren3,4, S. Beers5, A. Puccio6, D. Okonkwo7, D.E. Weeks2,4, 1) School of Nursing, Univ Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh PA; 3) Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, PA; 4) Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 5) Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA.

Traumatic Brain Injury (TBI) often results in chronic disability, however there is great variability in symptom development and severity of disability with little knowledge to date to explain the biological underpinnings for this variability. Genomics, though rarely used in the context of recovery following injury, provides a toolbox for assessment of the underlying biology that may explain the variability in phenotypes post-TBI. This project utilized a cohort of subjects recruited immediately after a severe TBI and followed for up to 24 months post injury using the neurobehavioral rating scale (NRS) to conduct a genomewide analysis using the Illumina Human Exome Beadchip. After data quality checks and filtration using the GWAStools R package, 386 subjects were analyzed and six SNPs within a 414kb region in 15q23 were implicated in NRS scores at six months post-TBI (p=3.8x10^-7 through p=4.8x10^-5) after controlling for age, sex, and extent of initial injury. This region has been implicated previously with rate of progression in subjects with mild cognitive impairment and contains several novel, biologically relevant candidate genes. This is the first study to conduct a non-candidate driven association study to identify genes potentially implicated in recovery after TBI.

1164T
Long non-coding RNAs associated with synapse are differentially expressed in autistic brains. W. Ju1, Y. Wang2, X.L. Zhao3, M. Piony1, E.C. Jenkins1, W.T. Brown1, N. Zhong1,2,3, 1) New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY; 2) Shanghai Jiao Tong University Children’s Hospital, Shanghai, China; 3) Peking University Center of Medical Genetics, Beijing, China.

Autism spectrum disorders (ASD) are characterized by significant impairments in reciprocal social communication and the presence of repetitive and/or restricted behaviors. The most recent surveillance data showed that ASD affects as many as 1 in 68 children within the United States. The increased incidence of ASD may reflect environmental exposures interaction(s) with genetic factor(s). Several studies have demonstrated that non-coding RNAs may involve in the development of ASD and the long non-coding RNA (lncRNA) may be differentially expressed in ASD brains and contribute to ASD risk. We have studied 25 pairs of blood samples collected from ASD vs. controls to investigate the IncRNA in the ASD brains. We identified both up-regulated and down-regulated IncRNAs from ASD brains. Among the up-regulated 2,329 IncRNAs, 1,053 were intergenic, 210 exon sense-overlapping, 69 intron sense-overlapping, 592 natural antisense, 392 intronic antisense, and 13 bidirectional. Among the 1,408 down-regulated IncRNAs, 644 were intergenic, 129 exon sense-lapping, 96 intron sense-overlapping, 443 intronic antisense, and 96 bidirectional with no natural antisense. Six pathways, constructed with three up-regulated IncRNAs involved in long-term depression, long-term potentiation, and synaptic vesicle cycling and three in down-regulated IncRNAs involved in Huntington’s disease, Alzheimer diseases, and Parkinson’s disease, showed neurological IncRNAs may be differentially expressed in peripheral bloods. Among the IncRNAs identified, 27 (11 up-regulated and 16 down-regulated) were found as the synapse associated. Functionally, they are localized at the loci of gene SYT (synaptotagmin), SYN (synapsin), STX (synlatxin), SV2C (synaptic vesicle glycoprotein 2C), SDCBP2 (syntenin), SYNM (synemin), SYNRG (synergin gamma), and SYNO2L (synaptopodin 2-like protein). Accordingly, a subset of synaptic mRNAs, transcribed from genes SYT, SYN, STX, SYNGR, SDCBP2, SYNO2L, SNAP25, SYDE1, STXBP6, SYNJ2BP, SYNCRIP, SNCA, SDC2, SYC3, SYPL2, SYBU, and SYCE1, were identified. Identification of synapse-associated IncRNAs that were differentially expressed in the ASD peripheral bloods may have opened a new avenue to investigate the epigenetic mechanisms underlying ASD and to explore the potential of applying the differentially expressed IncRNA as a biomarker for early detection of ASD clinically.

1165S
Evidence pointing to abnormal energy metabolism in two genetic animal models of epilepsy, A.H.B. Matsos1, A.S. Vieira1, V.D.B. Pascoal1,2, C.S. Rocha1, M.F.D. Moraes2, C.V. Maurer-Morelli1, D.R. Nascimento2, S. Martins2, A.S. Martins2, A.C. Valle4, A.L.B. Godard5, I. Lopes-Cendes1, 1) UNICAMP, Campinas, Brazil; 2) UFMG, Belo Horizonte, Brazil; 3) UFF, Nova Friburgo, Brazil; 4) USP, Sao Paulo, Brazil.

Background: Wistar audiogenic rat (WAR) is a genetic epilepsy model susceptible to audiogenic seizures, after high-intensity sound stimulation. Another genetic model we have recently identified is the generalized epilepsy with absence seizures (GEAS) rat. The aim of this study was to determine the molecular pathways involved in the susceptibility to seizures of these two strains using gene expression analysis. Methods: We obtained total RNA from five susceptible WAR [hippocampus and corpora quadrigemina (CQ)], five control Wistar and five WAR naive (without stimulation) as well as from hippocampus and somatosensory cortex (SC) of five GEAS rats and five control Wistar. Gene expression analysis was performed using microarray technology, and analyzed in R environment using the Affy and RankProd packages from Bioconductor, as well as the MetaCore® platform to identify molecular networks, gene ontology categories and gene interactions. Results: Genes with differential expression and a possible biological role in epileptogenesis were validated by qRT-PCR. Results: In WAR, expression profile showed a total of 1624 differentially expressed transcripts in the CQ and 1351 differentially expressed in the hippocampus compared with controls, with 616 upregulated and 1008 downregulated in CQ and 660 upregulated and 691 downregulated in the hippocampus. Enriched gene ontology categories identified in WAR were involved in oxidative phosphorylation, neurophysiological process GABA-A receptor life cycle. Genes validated by qRT-PCR were Grin1, Nedd8, Il18 and Slc1a3. The comparison of these genes between WAR naive and WAR showed that overall these genes are upregulated in WAR. In GEAS rats expression profile showed a total of 2307 differentially expressed transcripts in the hippocampus and 2282 differentially expressed in the SC compared with controls, with 1039 upregulated and 1268 downregulated in hippocampus and 991 upregulated and 1291 downregulated in the SC. The top enriched gene ontology categories included: oxidative phosphorylation, LRRK2 in neurons in Parkinson’ s disease. Conclusion: Interestingly, among the gene ontology categories abnormally expressed in both models (WAR and GEAS rats) we found oxidative phosphorylation, indicating that an abnormal energy metabolism in the central nervous system may be an important mechanism leading to epileptogenesis in these two genetic models of epilepsy. Supported by FAPESP, BRAZIL.
1166M Genetic influences on metabolite levels in Alzheimer’s Disease. P. Poity1, M. Kim2, L. Whiley2, P. Sharr2, S. Lovestone1, J. Powell1, C. Legido-Quigley2, R. Dobson1, AddNeuroMed Consortium. 1) King’s College London, Institute of Psychiatry, De Crespigny Park, London, UK; 2) King’s College London, Institute of Pharmaceutical Science, London, UK; 3) Department of Psychiatry, The University of Hong Kong, Hong Kong; 4) Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK; 5) NIHR Biomedical Research Centre for Mental Health and Biomedical Research Unit for Dementia at South London and Maudsley NHS Foundation Trust.

A better understanding of the biological mechanisms underlying Alzheimer’s Disease (AD) is required. AD is one of the major challenges for healthcare in the 21st century and with estimated longer life expectancy, the worldwide numbers of demented patients are expected to reach 81.1 million in 2040. Studies have shown that there are no demonstrable differences in the promotion of using associations with blood metabolites, the repertoire of small molecules present in cells and tissue, as functional intermediate phenotypes in biomedical and pharmaceutical research. In AD, a number of non-targeted blood metabo- lomic studies have been performed highlighting the role of lipid compounds in disease initiation and progression. In addition, recent studies investigating the genetics of the human metabolome have identified genetic variants in metabolism-related genes that lead to differently profiled metabolic phenotypes, termed ‘genetically influenced metabolomes’ (GIMs), which provide new insights into the role of inherited blood metabolic variation. The aim of this study was to investigate genetic influences on human plasma metabolites in a sample of >400 AD patients and controls, in a two stage approach (Stage 1 N=102, Stage 2 N=315), in order to survey regions of the genome that more strongly affect metabolites in (a) a phenotype driven metabolomic data quality control, linear regression analyses were run to identify genetic influences on each metabolic feature. Results were pooled together by inverse-weighted meta-analysis. We report novel and previously reported associations between genetically and metabolically sNPs and metabolites at p<10^-6 and 3 associations at p<5x10^-8 in stage 1, as well as evidence for associations between loci and two metabolites previously shown to be altered in AD (p<5x10^-8). Results of stage 2 and meta-analysis will be reported.

Additional analyses in additional epilepsy patients and healthy control cohorts are required to conclusively confirm or reject this position.

1166S Investigating polygenic contributions of common hippocampal variants to epilepsy predisposition. C.D. Whelan1, D.P. Hibar2, J. Stein3, D. Speed1, S. Bisoyi4, M. Johnson5, D. Goldstein3, N. Delanty3, S.E. Medland1, P.M. Thompson1, G.L. Cavalleri4, The International League Against Epilepsy Consortium on Genetics of Complex Epilepsies. 1) Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Leinster, Ireland; 2) Imaging Genetics Center, Keck School of Medicine, University of Southern California; 3) UCL Genetics Institute, University College London; 4) Institute of Neurology, University College London; 5) Division of Brain Sciences, Imperial College London; 6) Duke Centre for Human Genetics, Duke University Medical Center, Durham, North Carolina; 7) Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 8) Departments of Human Genetics and Psychiatry, Radboud university medical center, Nijmegen, The Netherlands.

High-dimensional analysis of the principal component of feature loading at the selected loci (PC 1) is a common feature of localisation-related epileptic sites (LREs), present in 50-75% of all surgical resections in the disorder. The underlying cause of HS is debated. Animal models and post-mortem cell counts suggest that HS can result from epileptogenesis, and some MRI investigations have highlighted a familial component to HS and concomitant neuronal loss within hippocampal regions. Recently, the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium identified genome-wide significant signals correlating with hippocampal volume, in a study of 29,037 individuals. We tested the hypothesis that variants predisposing to changes in hippocampal volume may, as a group, contribute to epilepsy predisposition. To investigate this, we summarised variation across nominally-associated ENIGMA SNPs into quantitative ‘risk’ scores, weighted for local linkage disequilibrium and effect size, and related these scores with magnetic resonance imaging (MRI) samples of epilepsy patients (n=2,502), (ii) four epilepsy ‘subtypes’, including LREs (n=1,801), lesional epilepsies (n=280), nonlesional epilepsies (n=614) and idiopathic generalised epilepsies (IGEs; n=194) and (iii) an independent sample of healthy controls (n=191). Results did not reveal a significant association between disease state and risk score; observed scores only explained a small fraction (0.0-0.2%) of total variance in our risk model. Our findings suggest that being genetically predisposed to having smaller hippocampal volumes may not be a significant risk factor for epilepsy. However, further prospective genetic and metabolomic analyses in additional epilepsy patients and healthy control cohorts are required to conclusively confirm or reject this position.

1169M Using polygenetic risk scores of psychiatric disorders to predict Neuroti- cism. L. Colodro Conde1,2, K. Verweij, E. Byrne1, R.A. Power3, N. Martin1, S. Medland1, D. Speed1, S. Sisodiya2, J. Powell*, The International League Against Epilepsy and Schizophrenia, Keck School of Medicine, University of Southern California; 2) Imaging Genetics Center, Keck School of Medicine, University of Southern California; 3) UCL Genetics Institute, University College London; 4) Institute of Neurology, University College London; 5) Division of Brain Sciences, Imperial College London; 6) Duke Centre for Human Genetics, Duke University Medical Center, Durham, North Carolina; 7) Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 8) Departments of Human Genetics and Psychiatry, Radboud university medical center, Nijmegen, The Netherlands.

Neuroticism is conceptualized as an endophenotype for major depressive disorder (MDD). Neuroticism and MDD are moderately heritable and 45-55% of genetic variance in MDD is shared with neuroticism (1). To test whether increasing Neuroticism provides an index of the risk for MDD, we examined the extent to which Neuroticism can be predicted by a polygenic risk score (PRS) derived from the Psychiatric Genomics Consortium (PGC) MDD analyses. Additional analyses were performed in a Tourette syndrome (TS) and childhood bipolar disorder (BP) and schizophrenia (SCZ) accounted for additional variation in Neuroticism. Phenotypic information was obtained from four studies undertaken at the QIMR Berghofer Medical Research Institute. Phe- notypes were available for 20702 individuals, of whom 5295 were genotyped. Personality data were collected using the Eysenck Personality Questionnaire. Scale scores in Neuroticism were transformed by taking the arcsine of the square root, corrected for age and sex effects and standardised. After correcting for multiple imputation, PRSs were available for 2549 individuals (36.2% males, mean age = 35.83, SD = 12.14). PRSs were calculated using the PLINK profile score method A multiple regression on the profile scores of Neuroticism was performed controlling for ancestry (using three principal components). Comprehensive analyses failure to detect SNP effects for Neuroticism on Neuroticism (p>0.05). PRSs with absolute correlation to the PGC MDD risk score thresholds of p<0.001, 0.01, 0.1 and 0.5 were identified. As hypothesised, the MDD PRSs predicted Neuroticism explaining 0.17, 0.31, 0.49, 0.48 and 0.45 of the variance (p=0.04, 0.05, 4.2x10^-4, 5.0x10^-4 and 7.2x10^-4 respectively). However, the PRSs for bipolar disorder (BP) and schizophrenia (SCZ) accounted for additional variation in Neuroticism. The PGC MDD PRS was the only predictor that remained significant in a multiple regression on Neuroticism. While nominally significant at a p ≤ 0.05, the PRSs for ADHD, AUT, SCZ and BP did not improve the prediction of Neuroticism. Although limited by the power of both the PRS analysis and the original PGC GWAS, the present analyses support the use of Neuroticism as a specific endophenotype for MDD.
1170T Machine Learning Derived Disease Risk Prediction for Anorexia Nervosa. Y. Guo1, Z. Wei1, B. Keating1,2, H. Hakonarson1,2,3, GCAN, WTCCC3. 1) The Center for Applied Genomics, Abramson Research Center, The Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA; 2) Department of Computer Science, New Jersey Institute of Technology, Newark, NJ 07102, USA; 3) Department of Pediatrics, School of Medicine University of Pennsylvania, Philadelphia, PA 19104, USA.

Anorexia nervosa (AN) is a serious complex psychiatric disease with a strong contribution from the patients’ genetic background. In addition to identifying signals through genome-wide association (GWA) studies, researchers have been using genetic information and machine learning methods to predict risk of diseases in which genetics play an important role.

In this report we collected whole genome genotyping data on 3,940 AN cases and 9,266 controls from the Genetic Consortium for Anorexia Nervosa (GCAN), the Wellcome Trust Case Control Consortium 3 (WTCCC3) and the Children’s Hospital of Philadelphia (CHOP), and applied machine learning methods for predicting AN disease risk. The prediction performance is measured by area under the receiver operating characteristic curve (AUC), indicating how well the model distinguishes cases versus controls. Logistic regression model with the lasso penalty technique generated an AUC of 0.693, while Support Vector Machines and Gradient Boosted Trees reached AUCs of 0.691 and 0.623, respectively. Our results of different sample sizes also suggested that larger datasets are required to optimize the machine learning models and achieve higher AUC values. To our knowledge, this is the first attempt to assess AN risk by incorporating genetic data, and it will pave the way for improved AN risk evaluation, eventually benefitting AN patients and families.

1171S Genome-wide Association Study of Quantitative Autistic Traits in the General Population. T. Nishiyama1, S. Hosono1, M. Watanabe1, I. Ozer1, H. Ito2, H. Tanaka2, H. Kishino2, T. Kawaguchi2, F. Matsuoka2, K. Matsuo1. 1) Department of Public Health, Aichi Medical University School of Medicine, Nagakute, Japan; 2) Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan; 3) Laboratory of Biometry and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo, Tokyo, Japan; 4) Center for Genomic Medicine, Kyoto University, Kyoto, Japan; 5) Department of Preventive Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan.

Copy number variations and single nucleotide variants, many of which are highly penetrant, are increasingly identified to confer susceptibility to autism spectrum disorder (ASD). In contrast, only a modest number of common variants associated with ASD have been reported through genome-wide association studies (GWAS) (Weiss LA 2009, Wang K 2009, Anney R 2010, Xia K 2013). These common variants have weak effects on ASD susceptibility, and thus sample size would be prohibitively large in practice. Recent studies both of the general population and of twins have shown that quantitative autistic traits are continuously distributed with no natural boundary between normality and autism (Van der Hulst-Jonker M 2008). A strong correlation leading to the notions that the extreme of quantitative autistic traits represents the clinical syndrome of autism. Based on this dimensional conceptualization, we conducted GWAS of quantitative autistic traits in the general population and instead of relying on case-control sample of ASD. This methodology seems promising because of easy sampling.

In this study, we used non-cancer participants in the Hospital-based Epidemiological Research Program II at Aichi Cancer Center Hospital (HERPACC-II) (Hamanaka Y et al. 2013). Quantitative autistic traits were measured using the Japanese version of the Subthreshold Autistic Traits Questionnaire (SATO) (Nishiyama T 2014) in the second wave of surveys. The sample was genotyped on a IlluminaHuman 610 Quad BeadChip with 576,736 SNP markers. The present study was approved by the Ethics Committee of Aichi Cancer Center and informed consent was obtained from all participants.

After removing SNPs that failed the quality control criteria a priori defined, 488,823 markers were used for the analysis. The SNPStats package in R was used to test the association between SNP genotypes and the SATO score, using additive models adjusted for gender. Using only 358 subjects currently available (we have a plan to increase up to around 1000 in this year), we found four SNPs (rs11577757, rs10071163, rs11862060 and rs10444399) significant at p-value < 1 x 10^-6. These four SNPs are located in NEIL1, ABAT, PDE4D, ABAT and LSM9-SLC9A6, respectively, and were reported to be associated with ASD in some way. This result suggests potential advantages of the methodology using quantitative autistic traits in the general population.

1172M Diagnostic exome sequencing of patients with Autism Spectrum Disorder overwhelmingly detects mutations in newly characterized genes, which supports a de novo paradigm and the convergence of disrupted pathways in neurodevelopmental disease. Z. Powis1, C. Mroskie1, K. Radtke1, D. Shinide1, K.F. Gonzalez1, L. Shahmirzadi1, D. El-Khechen1, B. Tippin1, E. Chao1, R. Baxter1, S. Teng1, 1) Ambry Genetics, Aliso Viejo, CA, Select a Country; 2) University of California Irvine, Irvine, CA 92697.

Diagnostic exome sequencing (DES) is an increasingly effective tool for diagnosis in intractable cases where the underlying cause is believed to be genetic. Autism Spectrum Disorders (ASDs) have traditionally been thought to be highly heritable but often the exact genetic origin has been difficult to elucidate. It is estimated that the positive rate for ASD patients is around 15% with the three first tier tests (karyotyping, CMA, and fragile X). However, the diagnostic yield for the other identifiable etiologies for ASD remains unknown. Of the first 500 unscreened cases referred for DES to Ambry Genetics (Aliso Viejo CA), 50 (10%) were reported to have a ASD. All of these individuals had additional features such as seizures, intellectual disability, dysmorphic features or other organ involvement, and 40 (80%) had previous uninformative results in at least one of the three-first tier tests. For these 50 ASD patients, trio DES analyses were performed for 41 families. Among them, mutations in characterized genes have been detected in 12 (24%) (ANK2, CDKL5, CHD8, ELP2, FHL1, IQSEC2 (2 cases), KMT2A, PANK2, SETD5, SHANK3, and one individual with mutations in both AN03 and NALCN). 10 of these cases were de novo alterations. In addition, 7 patients (14.0%) had an uncertain result (SYK, KDM6B, MTRR, PI1A, ZBTB20, ADCY5, and SHROOM4). 6 of these 7 cases were inherited (85.7%). 62 (4%) of these cases had a novel result (HERC1, MIR133A, ZBTB20, ADCY5, SHROOM4). These observations indicated that 24% of syndromic ASD patients received a definitive molecular diagnosis via DES after negative first-tier genetic testing. The vast majority of the associated genes were newly identified, non-classical ASD genes and thus might evoke traditional sequential testing approaches to be employed. The rapid expansion of the understanding of ASD genetics accentuates the great potential and continuing challenges for genotype-phenotype correlation in the clinical setting. Our clinical data also strongly support the de novo paradigm as well as the convergence of disrupted pathways in neurodevelopmental disorders.

1173T Common polygenic variation and risk for childhood-onset schizophrenia. K. Ahn1, S. An2, J. Rapoport3. 1) Child Psychiatry Branch, National Institute of Mental Health, Bethesda, MD, 2) Department of Environmental Health Sciences, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD.

Childhood-onset schizophrenia (COS) is a rare and severe form of the disorder, with more striking abnormalities in brain development and more striking uneven prepsychotic development compared to later onset disorder. We previously documented that COS patients, compared with their healthy siblings and with adult-onset patients (AOS), carry substantially more rare (0.1% in control) chromosomal copy number variations (CNVs), spanning large genomic regions (>100kb). (Ahn et al. in press). Here we interrogated the contribution of common polygenic variation to the genetic susceptibility for schizophrenia. We examined the association between a direct measure of genetic risk of schizophrenia in 126 COS probands and 98 healthy siblings. Using data from the schizophrenia and autism GWAS of the Psychiatric Genomic Consorita, we selected risk related sets of SNPs from which we conducted polygenic risk score comparisons for COS probands and their healthy siblings. COS probands had higher genetic risk scores of schizophrenia and autism than their siblings (P<0.05). These initial results show that childhood onset schizophrenia shares common risk alleles with autism. Finally, these findings suggest that COS patients may have more salient genetic risk than do adult onset patients.
1174S
A psychometric GWAS finds specificity of variants associated with level and change in immediate and delayed verbal memory after age 60. T.E. Attur1, S. Stroebel1, P. Gerretsen1,2. 1Dept Neurogenetics, Univ Toronto, Toronto, ON, Canada; 2Dept Neurology, University of Southern California, Los Angeles, CA.

Substantial individual differences exist in memory ability and the rate of age-related memory decline among older adults. Phenotypically, there is a strong association between immediate recall (IR) and delayed recall (DR), but candidate gene findings suggest that genetic influences on IR and DR may be partially distinct, with stronger genetic influence on initial level of recall than on rate of decline. However, this may be due to measurement error; with limited measurement occasions, there is greater precision for estimating level than change. Our goal was to combine a psychometric approach with GWAS to identify whether unique genetic variants are associated with level and change in IR and DR in older age. The sample comprised 12,326 participants (mean age=66.7, range of 50 to 107) in the Health and Retirement Study (HRS) who provided DNA and cognitive data (77,693 data points) and were assessed up to 8 times between 1996 through 2012. IR and DR were assessed with an auditory learning task in which individuals were read a list of 10 nouns, asked to immediately recall the list, then asked to recount them after a several-minute delay. Genetic data comprised 2.4 million SNPs from the Illumina Human Omni-2.5 Quad beadchip. We used the following analytical steps: (1) Applied a two-spline mixed effects models to estimate each individual’s level at age 60 and rate of change after 60 in both IR and DR; (2) Conducted four GWAS for IR and DR level and decline, adjusting for sex and ancestral informative markers; (3) Examined whether the same SNPs associated with level and slope of IR were similarly associated with DR, among others with the same ancestry. Overall, we identified several SNPs which were strongly associated with each outcome (P<5.0x10^-6), and the implicated regions for IR and DR level and decline were mostly non-overlapping.

1175M
GWAS analysis of Insight into illness in Schizophrenia. V. De Luca, A. Bani-Fatemi, A. Graff, P. Gerretsen. Dept Neurogenetics, Univ Toronto, Toronto, ON, Canada.

Impaired insight into illness is a core feature of schizophrenia characterized by a failure to recognize the presence of a mental illness, symptom misattribution, and unawareness of the need for treatment or negative consequences associated with the disorder. Impaired insight is associated with treatment non-adherence and poor clinical outcomes. The degree of insight impairment tends to remain stable over the illness course, improving moderately with age, but can fluctuate with illness exacerbations, i.e. psychotic episodes. Impaired insight is associated with a number of premorbid factors, such as lower functioning, education, IQ, and personality traits. In particular, impaired insight into illness is consistently associated with illness severity, global cognitive impairment, lower premorbid intelligence (i.e. IQ), executive dysfunction, and memory deficits. A number of genes are implicated in schizophrenia, including DISC1 and NR51. Yet, despite the high heritability of schizophrenia (81%), few genes are reliably replicated and together indicate that disorders with language-associated deficits share genetic etiologies. Future studies should further examine to what extent these disorders are related using human-based and model-based approaches.

1176T
Genetic influences of language development in typically developing children and children with autism spectrum disorders. J.D. Eicher1, J.R. Gruen2, 1Pediatric Imaging Neurocognition and Genetics study; 2Departments of Genetics, Yale University, New Haven, CT; 2Departments of Pediatrics and Investigative Medicine, Yale University School of Medicine, New Haven, CT.

The ability to effectively communicate is an essential milestone of child development, particularly in social interaction and academic achievement. Deficits in communication skills severely hamper development and lead to long-term negative consequences. One of these vital communication skills is language acquisition, the ability to learn and use words and write. Children with various language disorders, including dyslexia, language impairment, and autism spectrum disorders, have deficits in receptive vocabulary. As each of these disorders is a complex trait with a significant genetic component, we examined the genetics of receptive vocabulary skills in typically developing children and in children with autism spectrum disorders. First, we performed a genome wide association study (GWAS) of receptive vocabulary in 440 typically developing children in the Pediatric Imaging Neurocognition and Genetics (PING) study. There were no significant association with markers in MAGI-1 (p=3.6x10^-4) and BMP7 (p=8.25x10^-6). MAGI-1, a scaffolding protein important in cell-cell junctions, has been previously implicated with bipolar disorder, schizophrenia, associative learning, and modulation of AMPA synaptic behavior. BMP7 is important in neural development and was recently identified as a prospective candidate gene for language related disorders (Boeckx and Benitez-Burraco 2014). Next, we assessed whether genes previously associated with language— including MAGI-1, BMP7, DCDC2, KIAA0139, MRPL19-GFC2, ATP2C2, CMIF, and KIAA0943 were associated with language skills in a meta-analysis of two autism spectrum disorders cohorts: the Autism Genome Research Exchange (AGRE) and Simons Simplex Collection (SSC). We found evidence of associations with ATP2C2 (p=4.89x10^-6), KIAA0371 x1 (p=1.00x10^-1), and MRPL19 (p=9.31x10^-6), indicating that genes previously associated with other language disorders (such as dyslexia and language impairment) also influence communication skills in children with autism spectrum disorders. These results identify genetic that influence childhood communication development and indicate that disorders with language-associated deficits share genetic etiologies. Future studies should further examine to what extent these disorders are related using human-based and model-based approaches.

1177S
Genome-wide meta-analysis reveals significant association between CHRNA4 variants and nicotine dependence in cohorts of European ancestry. N.C. Gaddis1, D.B. Hancock1, W.G. Reginson2, N.L. Saccone3, S.M. Lutz2, P. Kraft4, D.B. Hancock3, J.E. Hokanson3, L.J. Bierut4, T.E. Thorgersson5, E.O. Johnson1, K. Stefansson6, 1) RTI International, Research Triangle Park, NC; 2) deCODE Genetics, Reykjavik, Iceland; 3) Washington University, St. Louis, MO; 4) University of Colorado Anschutz Medical Campus, Aurora, CO; 5) Harvard University School of Public Health, Boston, MA.

Cigarette smoking is a major contributor to cancer, vascular disease, and lung disease and is the leading cause of preventable mortality worldwide. CHRNA4 is a gene encoding a high affinity acetylcholine receptor on chromosome 15q25 (CHRNA5-CHRNA3-CHRNB4) and chromosome 8p11 (CHRNB3-CHRNA6). To identify novel genetic loci associated with smoking we conducted a genome-wide association study (GWAS) and meta-analysis of nicotine dependence (ND) in five cohorts of European ancestry totaling 16,244 subjects: the largest GWAS meta-analysis of ND to date. We defined ND using the Fagerstrom Test of Nicotine Dependence (FTND), which is a multi-dimensional measure of physiological dependence to nicotine. After completing standard quality control and 1000 Genomes imputation for each of the five cohorts, we tested approximately 10 million genetic variants for association with a three-level FTND-based phenotype (mild, moderate, and severe ND), separately by cohort, using a linear regression model adjusted for age, sex, eigenvectors to correct for population stratification, and other cohort-specific covariates as needed. We then combined the results from the five cohorts in an inverse variance-weighted meta-analysis with genomic control applied to each cohort. As expected, variants in the CHRNA5-CHRNA3-CHRNB4 region show strong association with ND, with p-values as low as 1x10^-17. We also observed genome-wide significant association for a variant in the 3’ untranslated region of CHRNA4 (p=3.60x10^-9), with odds ratios for the five cohorts ranging from 1.05-1.20 and 1.11-1.43 for moderate and severe ND, respectively. This variant is predicted to alter a microRNA binding site that may regulate CHRNA4 expression. In total, we found 22 CHRNA4 variants with p-values less than 1x10^-4, all of which had consistent directions of association. From the five cohorts, these microRNA evolutionarily (4-20%) are being associated with an increased risk of ND. CHRNA4 is known to have high affinity for nicotine, and our findings build on the evidence from several smaller studies suggesting that CHRNA4 variants contribute to smoking-related phenotypes.
Identification of novel candidate genes in canine noise phobia -a model for human panic disorder. O. Hakosalo1,2, K. Tira1,2, R. Savitamo1,2, M. Silpanpää1,2, J. Kere1,2, H. Lohi1,2, 1) Department of Veterinary Biosciences and Programs Unit, Molecular Neurology, University of Helsinki, P.O.Box 63, 00014 University of Helsinki, Helsinki, Finland; 2) The Folkhäl- san Research Center, Helsinki, Finland; 3) Department of Biology and Mathematical Sciences, University of Oulu; 4) Department of Bioscience, All and Nutrition, Karolinska Institutet, and Center for Biotechnology, Sweden; 5) Research Programs Unit, University of Helsinki.

Noise phobia (NP), a fear of loud noises, is a severe anxiety disorder in dogs, which may result in panic attacks. It provides a natural animal model for human panic disorder. NP is common (20-40%) in many breeds with high heritability (h² 0.56) estimates but its genetic background remains unknown. We aimed to discover the genetic cause by a genome wide association analysis. We focused on German Shepherds because the breed is popular and present sufficient phenotypic variation in NP. We used our validated owner-completed anxiety questionnaire to create a categorical NP phenotype and genotyped altogether 310 German Shepherds (GS) (86 cases and 224 controls) to map the NP locus. We found a genome widely significant association in a 4 Mb region of GFAS0 harboring several interesting candidate genes. Targeted capture and resequencing of the associated region from 39 dogs revealed several disease segregating variants that were enriched in three promising candidate genes. Further validation and functional characterization of variants in a larger GS cohort and other breeds should implicate the causative gene. This study will improve the understand- ing of the genetic background of noise phobia, and at best reveal a new candidate gene for human anxiety and help in developing a model for potential therapeutic approaches.

CoreExome Chip study of low-frequency variants identifies genome-wide significant hits associated with anorexia nervosa. K. Hatziitoulou1, L.M. Huckins1, L. Thornton1, L. Southam1, D. Collier1, P. Sullivan1, C.M. Bulik2, E. Zeggini1, GCAN Consortium. 1) Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom; 2) Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 3) Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King’s College London, London, UK.

Anorexia nervosa (AN) is a neuropsychiatric disorder presenting with extremely low body weight, and a marked fear of gaining weight. AN has the highest mortality of any psychiatric disorder, and affects roughly 0.9% of women. Very little is known about the biological mechanisms which underlie AN; two GWAS have been completed and have yet to identify genome-wide significant hits. No effective medications are available, and treatment outcome for AN remains unacceptably poor. Our study comprises 2,376 female AN cases and >22,000 controls, genotyped on the Core-Exome Chip. Samples derive from eight different populations; cases and controls are ancestrally matched. The CoreExome Chip enables us to study both common and low-frequency variants simultaneously; our study is the first to examine the role played by low-frequency and rare variants in AN. Analysis is currently complete across three of the eight contributing populations: Norway, 87 cases, 100 controls; Finland, 163 cases, 5,300 controls; and the UK, 181 cases and 10,034 controls. We have performed a meta-analysis across these three populations and thus far have identified four genome-wide significant signals: exm370124, exm462797, exm464785, exm2166592. These four variants are all low frequency, missense variants. We looked for the frequency of these SNPs in both cases and controls. All SNPs were extremely low frequency in the control populations, with highest MAF between 0.005 and 0.01. SNPs were also low frequency in the cases, with highest MAF between 0.01 and 0.10. Effect sizes for each SNP were high, and the same direction of effect was noted for every SNP in at least 2/3 populations. Maximum effect sizes for each SNP were between 6.6 and 74.5. One of these associated variants (exm464785) lies in RASGRF2, a gene that has previously been associated with eating disorders (Wade et al. 2013), albeit not at a genome-wide significant level. This is the first genome-wide significant variant that has been associated with AN. We hope that this will enable further studies into the functional mechanisms underlying AN, and perhaps be a first step towards establishing effective medications and treatment. Further, all four hits that have been identified are very low frequency and could not possibly have been identified in previous GWAS studies. This may be a good indication that low-frequency, CoreExome chip type studies have potential to reveal new associated variants across a range of psychiatric disorders.

Multi-ethnic meta-analysis in a cohort of 110,266 individuals identifies novel shared and sex-specific loci associated with smoking initiation. E. Jorgenson1, L. Shen1, A. Hamdovic2, H. Choquet2, T. Hoffmann2, Y. Banda3, M. Kvale3, N. Risch3, C. Schaefer1, L. Sakoda1, 1) Kaiser Perma- nente Division of Research, Oakland, CA; 2) Departments of Pharmacy and Psychiatry University of New Mexico, Albuquerque, NM; 3) Institute for Human Genetics University of California, San Francisco San Francisco, CA 94143.

Smoking is a major public health problem in which both genetic and environ- mental factors play an important role. Genome-wide association studies have identified six loci influencing smoking behavior, and additional investi- gation is needed to clarify how these and other genetic factors affect the chronology of smoking behavior, specifically smoking initiation, persistence, and cessation. Family and twin studies have suggested that significant sex differences exist in smoking behavior, with the heritability of smoking initiation being higher in women than in men. We conducted a genome-wide association meta-analysis of smoking initiation across race/ethnicity (non- Hispanic white, Latino, Asian, and African American subjects) and sex strata in the Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort. Over 6.6 million SNPs were imputed with imputation r²>0.8 and minor allele frequency > 0.05 using the 1000 Genomes reference panel. We identified four loci that exceeded a genome-wide significance level of 5×10-8. An intergenic locus near TMEM182 on chromosome 2 (p=2.65×10-8) and SNPs in SDK1 (p=2.52×10-8) were associated with smoking initiation in both men and women. Additionally, SNPs in BDNF (p=3.48×10-8) and NCAM1 (p=5.20×10-9) were associated with smoking initiation in women but not in men. While BDNF was previously reported to be associated with both smoking initiation and persistence, our other findings are novel. We examined additional evidence for the association of these loci with smoking initiation in the combined-sex association results among European populations made publicly available by the Psychiatric Genomics Consortium. Those results provided support for both BDNF (p=2.01×10-5) and NCAM1 (p=6.76×10-3), but not for the other two loci (p>0.05), indicating both a need for additional replication in large, multi-ethnic samples and the power of combining evidence across race/ethnicity groups to provide additional insights into the genetics of smoking behavior.
1181T

First GWAS in DBH confirms strong cis-acting variants and lends support for its role as an intermediate phenotype in post-traumatic stress disorder. A.X. Maihofer1, M. Mustapic1,2,4, D.G. Baker3, D.T. O'Connor2, C.M. Niievergel1,3. 1) Department of Psychiatry, University of California at San Diego, La Jolla, CA; 2) Department of Medicine, University of California at San Diego, La Jolla, CA; 3) VA Center of Excellence for Stress and Mental Health (CESAMH), VA San Diego Healthcare System, La Jolla, CA; 4) Division of Molecular Medicine, Ruder Boškovic Institute, Zagreb, Croatia.

Dopamine beta-hydroxylase (DBH) catalyzes formation of norepinephrine. DBH is expressed in noradrenergic nerve terminals of the central and peripheral nervous systems, as well as in chromaffin cells of the adrenal medulla. DBH is present in cerebral spinal fluid and plasma as stable heritable trait. Differences in DBH expression or activity might reflect a role in the pathogenesis of cardiovascular and neuropsychiatric disorders. The genetic mechanisms underlying DBH activity and its secretion have been only partially explained. Thus we conducted a genome-wide association search for loci contributing to human plasma DBH (pDBH) activity. In a population sample of European ancestry, we identified 3 common trait-determining variants (top hit rs1611115, p = 7.2 × 10^{-5}) in the proximal DBH promoter. Each of the 3 variants had an effect on transcription and acted additively on gene expression. We replicated the associations in a population sample of Native American descent (top hit rs1611115, p = 4.1 × 10^{-15}). Additionally, we identified yet another genome-wide significant SNP at the LOC385787 locus on chromosome 12 as potential trans-quantitative trait locus (QTL) (rs4255618, p = 4.62 × 10^{-8}). Analysis conditioned on 3 DBH promoter variants identified a third genomic region on chromosome 9q contributing to pDBH variation: a likely cis-QTL adjacent to DBH in SARDH (rs7040170, p = 1.31×10^{-14}).

The identification of SNPs with such strong effects on pDBH opens the possibility of utilizing Mendelian randomization (MR) approaches to test causal effects of this intermediate trait on disease. Therefore we performed an exploratory MR analysis in a sample of participants recruited from the Marine Resiliency Study (MRS), a large, prospective study of post-traumatic stress disorder (PTSD) involving active-duty United States Marines bound for deployment to Iraq or Afghanistan. The SNP rs1611115, which explained 50% of the variability in pDBH, was employed as a genetic instrument to estimate the causal effect of the association of pDBH and re-experiencing symptoms was significant (beta = 0.21, p = 0.002), indicating the possibility that pDBH is a causal component in the development of re-experiencing symptoms.

1181M

Genome-Wide Association Study in APOE ε4 Negative Subjects Identifies a Novel Locus in 17q21.31 for Alzheimer Disease. G. Jun1,2,3, C. Ibrahim-Verbaas4, C. Bellenguez1,2,9, M. Vronskaya11, J. Chung1, J.C. Bis12, J. Williams11, S. Seshadri4, G.D. Schellenberg14, K.L. Lunetta2, P. Amouyel7,8,9,10, P. Holmans51, C.M. van Duijn3, L.A. Farrer1,2,3,4,5, International Genomics of Alzheimer’s Project (IGAP) Consortium. 1) Medicine, Boston University, Boston, MA, USA; 2) Epidemiology, Boston University, Boston, MA, USA; 3) Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, USA.

APOE ε4 is the most significant genetic risk factor for Alzheimer disease (AD) and may mask effects of other loci. We re-analyzed genome-wide association study (GWAS) data from the International Genomics of Alzheimer’s Project (IGAP) Consortium in APOE ε4-positive (10,246 cases and 11,924 controls) and APOE ε4-negative (7,231 cases and 19,603 controls) subgroups as well as in the total sample allowing for interaction between a SNP and APOE ε4 dosage. Suggestive associations (p < 1 × 10^{-4}) in the discovery sample were evaluated in an independent replication sample containing 4,474 subjects (ε4-positive: 1,250 cases and 536 controls, ε4-negative: 1,158 cases and 1,329 controls). In the discovery set, we observed suggestive associations from 15 loci in either ε4-positive or ε4-negative subjects (5 novel loci in ε4-negative subgroup: SOX/CLDN18, C12q23.2/ACSL6, FFNB11/HBEFG, KANSL1/LRRC37A, and CDR2L) and two novel loci in the total sample for interaction (THSD7A/MEM10G and ABCA9/DFC14BL). Of these, five distinct loci from 4 known (CR1, BIN1, CLU, and MS4A cluster) and one novel locus (KANSL1/LRRC37A) were genome-wide significant (p < 5 × 10^{-8}) in meta-analysis of the discovery and replication sample. The ability to detect genome-wide significant associations in ε4-negative subjects is noteworthy.

Novel associations among ε4-negative subjects were observed in the MAPT region covering approximately 100 kb including KANSL1 and LRRC37A (best SNP, rs2372703, ε4-positive: OR = 0.86, CI = 0.76-0.98, p = 0.02; ε4-negative: OR = 0.73, CI = 0.65-0.81, p = 5.8 × 10^{-8}). Although 17 SNPs in this region initially showed genome-wide significant p-values, analyses conditioned on the lead SNP (rs2372703) revealed no additional independent associations. However, despite these findings, no SNPs showed statistical evidence of genome-wide significant interaction by ε4 status in the meta-analysis of discovery and replication. Association of variants in the MAPT region with AD may represent a novel genetic risk factor among individuals who do not carry the ε4 risk allele. Further examination of this region is warranted to identify functional determinants of AD risk among persons lacking the ε4 allele.
1183S

Genome-wide association study of sensory disturbances in the inferior alveolar nerve after bilateral sagittal split ramus osteotomy. D. Nishizawa 1, D. Kobayashi 1, 2, 3, Y. Takasaki 1, S. Kasa 1, Y. Aoki 1, J. Hasegawa 1, T. Kakizawa 1, K. Ikeda 1, K. Fukuda 1.
1) Psychiatry and Behavioral Science (Addictive Substance Project), Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan; 2) Department of Dental Anesthesiology, Tokyo Dental College, Tokyo, Japan; 3) Department of Dentistry and Oral surgery, Tokyo Metropolitan Tama Medical Center, Tokyo, Japan; 4) Department of Dentistry and Oral surgery, National Hospital Organization, Takasaki General Medical Center, Gunma, Japan; 5) Department of Oral Health and Clinical Science, Division of Oral and Maxillo-facial Surgery, Tokyo Dental College, Tokyo, Japan.

Background: Bilateral sagittal split ramus osteotomy (BSSRO) is a common orthognathic surgical procedure. Sensory disturbances in the inferior alveolar nerve, including hypoesthesia and dysesthesia, are frequently observed after BSSRO, even without distinct nerve injury. The mechanisms that underlie individual differences in the vulnerability to sensory disturbances have not yet been elucidated.

Methods: The present study investigated the relationships between genetic polymorphisms and the vulnerability to sensory disturbances after BSSRO in a genome-wide association study (GWAS). A total of 304 and 303 patients who underwent BSSRO were included in the analyses of hypoesthesia and dysesthesia, respectively.

Hypoesthesia was evaluated using the tactile test 1 week after surgery. Dysesthesia was evaluated by interview 4 weeks after surgery. Whole-genome genotyping was conducted using Illumina BeadChips including approximately 300,000 polymorphism markers.

Results: Hypoesthesia and dysesthesia occurred in 51 (16.8%) and 149 (49.2%) subjects, respectively. Significant associations were not observed between the clinical data (i.e., age, sex, body weight, body height, loss of blood volume, migration length of bone fragments, nerve exposure, duration of anesthesia, and duration of surgery) and the frequencies of hypoesthesia and dysesthesia.

Significant associations were found between hypoesthesia and the rs2063640 polymorphism (trend model: combined $\chi^2 = 23.07$, nominal $P = 6.633 \times 10^{-7}$), between dysesthesia and the rs2063640 polymorphism (trend model: combined $\chi^2 = 23.07$, nominal $P = 1.563 \times 10^{-7}$), and between dysesthesia and the nonsynonymous rs2677879 polymorphism (trend model: combined $\chi^2 = 16.56$, nominal $P = 4.722 \times 10^{-5}$; dominant model: combined $\chi^2 = 16.31$, nominal $P = 5.369 \times 10^{-5}$).

The rs502281 and rs2063640 polymorphisms were located in the flanking region of the ARID1B and ZPLD1 genes on chromosome 18, whose official name is “methyltransferase like 4” and “ARID domain 1B (SWI1-like)” and “zona pellucida-like domain containing 1”, respectively. The rs2677879 polymorphism is located in the METTL4 gene on chromosome 18, whose official name is “methyltransferase like 4”.

Conclusions: The GWAS of sensory disturbances after BSSRO revealed associations between genetic polymorphisms located in the flanking region of the ARID1B and ZPLD1 genes and hypoesthesia and between a nonsynonymous genetic polymorphism in the METTL4 gene and dysesthesia.

1184M


Crying, for emotional reasons, is considered unique to humans. Crying habits vary greatly–some crying easily and others rarely. Thus far it is not clear why, though emotional stability has been shown to play a role and excessive emotion appears to have a familial tendency. In this genome-wide association study, we searched for possible genetic associations with crying easily using a large sample of unrelated customers of 23andMe, Inc., with European ancestry. 60,48 cases self reported that they ‘cry easily’ and 94,871 controls said that they did not ‘cry easily’. Females are three times more likely to self-identify as ‘crying easily’ than males and individuals who cry easily have a significantly higher rate of psychiatric disorders in our data.

We identified genome-wide significant associations with 7 loci, including rs62335062 (P=2.6x10^-31, OR=1.111) upstream of IRX2, rs1796282 (P=5.0x10^-14, OR=0.928) in an intron of ZNF423, rs2178574 (P=7.2x10^-11, OR=1.066) in an intron of LRRTM4, rs2744475 (P=2.1x10^-10, OR=1.060) upstream of TAF2B, rs10838125 (P=2.5x10^-10, OR=1.058) between TTC17 and HSD17B12, rs877614 (P=5.7x10^-10, OR=1.054) in an intron of BIN3 and close to EGR3, and rs16903275 (P=3.9x10^-8, OR=0.94) in microRNA-9-2 (MIR9-2). The top two associated genes, IRX and ZNF423, have been shown to be involved in cerebellum development and brain regionalization. The cerebellum is an important brain region that has been linked to crying. LRRTM4 gene product is known to trigger the formation of excitatory synapses and shows highly selective expression in the brain. Mutations in LRRTM family genes have been associated with human handedness and schizophrenia. TAF2B product, an important factor in the development of ectodermal and neural tissues, is involved in monoaminergic regulation and has been associated with neonatal temperament, alcohol addiction, adolescent depressive symptoms, and attention deficit hyperactivity disorder.

The expression of EGR3 is rapidly regulated by neural synaptic activity and can be mediated by the 

1185T

Identification of a novel locus for human-directed fear in dogs. K. Tiira 1, 2, O. Hakosalo 1, 2, R. Sarvioaho 1, 2, M. Stållbrand 3, J. Kero 1, 2, H. Lohi 1, 2, 1) Department of Veterinary Biosciences and Research Programs Unit, Molecular Neurology, U.P.O.Box 63, 00014 University of Helsinki, Helsinki, Finland; 2) The Folkhälsoan Research Center, Helsinki, Finland; 3) Department of Biology and Mathematical Sciences, University of Oulu; 4) Department of Biosciences and Nutrition, Karolinska Institutet, and Center for Biotechnology, Sweden; 5) Research Programs Unit, University of Helsinki.

Behavioural, pharmacological, clinical and etiological studies suggest that dogs suffer from human anxiety-like disorders with shared neurobiological etiologies. Fearfulness has obtained high heritabilities in dogs ($h^2=0.46$) and inbred population structures facilitate gene mapping. We have initiated a large program to dissect the genetic and environmental correlates of different canine anxiety traits. To map the fear locus using canine HD SNP chip arrays (173K) we identified a genome-wide significant association in a 4 Mb region at CFA11 using a linear mixed model analysis with puppy socialization as a covariate. Our ongoing analyses aim to replicate the association in larger cohorts available across breeds and to identify the causative genes. This is the first report for a fear locus in dogs and provides insights to the genetic and environmental factors in canine anxiety with a translational potential to human anxiety.
1186S
Genetic determinants of survival in patients with Alzheimer’s disease. X. Wang1,2, O. Lopez3,4, RA. Sweet1,4,5, S. Becker3,4, MM. Bammada1, E. Feingold1,2,3, PY. Demirci1,2, IM. Kambhampati1,4,5, 1) HUMAN GENETICS, University of Pittsburgh, PITTSBURGH, PA; 2) Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 3) Department of Neurology, University of Pittsburgh, Pittsburgh, PA; 4) Alzheimer’s Disease Research Center, University of Pittsburgh, Pittsburgh, PA; 5) Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA.

There is a strong genetic basis for late-onset of Alzheimer’s disease (LOAD) and thus far over 20 genes/loci have been identified that affect the risk of LOAD. In addition to disease risk, genetic variation at these loci may also affect components of the natural history of AD, such as survival in AD. However, the role of these known loci in survival of AD patients has not been explored extensively. In the present study, we examined the role of 22 known LOAD genes with the time to death in 983 AD patients recruited from a referral clinic. In a secondary analysis, genome-wide association was examined to identify novel loci that may influence survival in AD. Survival analysis was conducted using Cox proportional hazards regression under an additive genetic model adjusted for the baseline MMS score, education, an additive genetic model and age was an covariate. We identified 13 single SNP associations (P<0.05) either within or adjacent to 10 LOAD genes and 3 of them remained significant in gene-based analyses (BINP, P=4.79E-04; INPP5D, P=4.49E-02 and APOE, P=3.23E-02). Genome-wide association analysis revealed 8 suggestive novel loci at P<1e-05 (ALDH4A, P=3.24E-06; IL19, P=6.62E-07; NCKAP5, P=1.37E-06; CDCDC3, P=2.25E-06; NARS2, P=3.14E-06; PKNOX2, P=8.03E-06; SDR9C7, P=2.42E-06; and SALL4, P=2.25E-06). These results indicate that in addition to some known LOAD genes, genetic variation in novel loci may also affect survival of AD patients. Additional large follow-up studies in independent samples are required to confirm our potentially novel candidate loci findings.

1187M Integrative systems approaches to deciphering the genetic landscape of late-onset Alzheimer’s disease. Y. Zhao1, V. Mäkinen1, Q. Meng1, X. Yang2, 1) Integrative Biology and Physiology, University of California, Los Angeles, Los Angeles, CA 90095; 2) EMBL Australia, South Australian Health and Medical Research Institute, Adelaide Australia 5000.

Late-onset Alzheimer’s disease (LOAD) is a complex neurodegenerative disorder, with the genetic components estimated to account for 60-80% of the disease variability. Recent genome-wide association studies (GWAS) implicated a number of susceptibility loci for LOAD. However, the identified susceptibility loci are substantially inconsistent across GWAS and the underlying genetic mechanisms are still largely unknown. In this study, we utilized an integrative systems approach that leveraged a multitude of genetic and genomic datasets, including 1) LOAD GWAS, 2) LOAD eQTL, 3) LOAD expression data, 4) genetic variation in novel loci, and 5) data-driven gene regulatory networks from multiple human and mouse cohorts. The integration of these diverse data sources enabled tissue-specific investigations on whether the genetic variants associated with LOAD GWAS were concentrated on gene subnetworks (i.e., specific parts of gene regulatory networks) and whether novel key regulators in the subnetworks could be identified based on data-driven network structures. Our results confirmed the involvement of many well-known LOAD-related processes, such as lipid metabolism, immune and inflammatory response, endocytosis/intracellular trafficking, and cell migration to be significantly enriched for LOAD risk variants across the nine LOAD GWAS datasets. More interestingly, we detected several novel subnetworks related to mitochondrial protein import, viral infection, and pyrimidine metabolism which show significant expression changes between LOAD patients and nondemented subjects. Finally, by utilizing the gene-gene relationships revealed by the network architecture, we detected key regulator genes, both known (e.g. NCKAP1L and FYB) and inferred (e.g. PTPN6), in these LOAD subnetworks. Our results shed light on the complex mechanisms underlying LOAD and highlight potential novel targets for the debilitating disease.

1188T TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9ORF72 hexanucleotide repeat expansions. M. Gallagher1,2, E. Suh3, M. Grossman4, L. Elman5, L. McCluskey5, J.Q. Trojanowski6, V. Lee6, 1) Department of Pathology & Laboratory Medicine, Center for Neurodegenerative Research, Perelman School of Medicine, University of Pennsylvania, PA; 2) Department of Neurology, Perelman School of Medicine, University of Pennsylvania, PA; 3) Department of Psychiatry, University of Pennsylvania, PA; 4) Department of Psychiatry, University of California, Los Angeles, Los Angeles, CA; 5) International Collaboration for FTD.

Late-onset Alzheimer’s disease (LOAD) is a complex neurodegenerative disorder, with the genetic components estimated to account for 60-80% of the disease variability. Recent genome-wide association studies (GWAS) implicated a number of susceptibility loci for LOAD. However, the identified susceptibility loci are substantially inconsistent across GWAS and the underlying genetic mechanisms are still largely unknown. In this study, we utilized an integrative systems approach that leveraged a multitude of genetic and genomic datasets, including 1) LOAD GWAS, 2) LOAD eQTL, 3) LOAD expression data, 4) genetic variation in novel loci, and 5) data-driven gene regulatory networks from multiple human and mouse cohorts. The integration of these diverse data sources enabled tissue-specific investigations on whether the genetic variants associated with LOAD GWAS were concentrated on gene subnetworks (i.e., specific parts of gene regulatory networks) and whether novel key regulators in the subnetworks could be identified based on data-driven network structures. Our results confirmed the involvement of many well-known LOAD-related processes, such as lipid metabolism, immune and inflammatory response, endocytosis/intracellular trafficking, and cell migration to be significantly enriched for LOAD risk variants across the nine LOAD GWAS datasets. More interestingly, we detected several novel subnetworks related to mitochondrial protein import, viral infection, and pyrimidine metabolism which show significant expression changes between LOAD patients and nondemented subjects. Finally, by utilizing the gene-gene relationships revealed by the network architecture, we detected key regulator genes, both known (e.g. NCKAP1L and FYB) and inferred (e.g. PTPN6), in these LOAD subnetworks. Our results shed light on the complex mechanisms underlying LOAD and highlight potential novel targets for the debilitating disease.


Cocaine induces paranoia (CIP) is a common consequence of cocaine use that can cause serious morbidity. We used pathway analysis to identify target genes to aid in our understanding of CIP pathophysiology, based on the principle of identifying reproducible GWAS-derived gene lists from two independent datasets. GWAS data set was generated by using the Illumina Omni1-Quad microarray, as described in our published cocaine dependence (CD) GWAS. The two populations in this study were African American (AA; 1273 cases and 824 controls, 52% male) and European American (EA; 951 cases and 749 controls, 58.9% male). All sample subjects were interviewed using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA). The CIP phenotype was ascertained via the question: “Have you ever had a paranoid experience when you were using cocaine?” The control group included only subjects who reported lifetime cocaine use. Gene sub-networks were obtained by using the dmGWAS R package on each population. The P-values from the GWAS (briefly, logistic regression on sex, age, and the first three PCs - Gelernter et al. 2013) was combined with gene-gene interaction data to identify significant gene-gene interaction modules. The AA and EA lists were evaluated for overlapping gene content (which could result from different SNPs in the two populations). Each list was filtered by both lists tested by SKAT-O. Gene lists were submitted to the Database for Annotation, Visualization and Integrated Discovery (DAVID) for check for enrichment pathways according to KEGG (Kyoto Encyclopedia of Genes and Genomes) gene impact database. Enriched pathways were considered a total of 22061 genes. From these, four genes out of the 40 identified in AAs and the 88 in EAs were identical between the two populations: GABBR2 (AA p = 0.00065, EA p = 0.0076), endothelin receptor, DL1C1 (AA p = 0.0015 , EA p = 0.0065) and ABLIM2 (AA p = 0.0023, EA p = 0.00013). Gene lists from both populations were enriched in several KEGG pathways, including Focal Adhesion (Padj 1.20E-09), ECM-Receptor Interaction (Padj 6.30E-07), etcs Separately, the AA group was most enriched in the Pathway for Lysine degradation, the EA group was most enriched in the Neuroactive Ligand-Receptor Interaction Pathway (Padj 1.40E-03), while the EA group was most enriched in Focal Adhesion Pathway (Padj 6.30E-07). Interaction pathways including GABBR2, ADCY8, DL1C1 and ABLIM2 appear the most promising for CIP risk, based on our dataset.
1190M
A genome-wide screen for fear of heights susceptibility loci in a Finnish isolate. I. Hovatta1,2, Z. Misiewicz1, T. Hiekkanen1,2, T. Paunio1,3, T. Vartiainen1,2
1) Department of Biosciences, University of Helsinki, Helsinki, Finland; 2) Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland; 3) Unit of Public Health Genomics, National Institute for Health and Welfare, Helsinki, Finland.

Objectives: It is classified as a specific phobia of naturalistic type under the DSM-V. No specific genetic variants predisposing to acrophobia have been identified. Our goal was to map genes predisposing to acrophobia. We genotyped and analyzed 575 microsatellite markers across the genome in the Finnish schizophrenia cohort, in which acrophobia segregates independently of schizophrenia.

Methods: We performed non-parametric linkage analysis using statistical software packages FASTLINK 4.1P and MERLIN. As a follow-up, we are performing joint analysis and association analyses with PSEUDOMARKER to fine-map the identified loci and to find putative founder alleles predisposing to acrophobia. Finally, we will perform haplotype analysis using MERLIN to identify potential risk haplotype(s) shared among affected family members. Our results from single point linkage analysis revealed two suggestive loci with recessive model: 13q34 (Zmax=2.88) and 5q32-q34 (Zmax=2.15). Variants in D-amino acid oxidase activator (DAOA) gene located at the 13q34 locus were previously associated with schizophrenia. Consistently, we found linkage in individuals affected both with schizophrenia and acrophobia and not those affected only by the latter.

Conclusion: This study confirmed previous findings of schizophrenia and acrophobia and identified two suggestive loci with recessive model: 13q34 and 5q32-q34 (Zmax=2.15). Variants in DAOA gene located at the 13q34 locus were previously associated with schizophrenia. Consistently, we found linkage in individuals affected both with schizophrenia and acrophobia and not those affected only by the latter. However, the 5q32-34 seems specific for acrophobia as omitting individuals with comorbid schizophrenia did not significantly affect the LOD score. The singlepoint non-parametric linkage analysis showed the strongest linkage for 5q32-34 (Zmax=2.15). Our study revealed several suggestive loci segregating independently in schizophrenia and acrophobia. The concordance between results from parametric and non-parametric linkage analysis for region 5q32-34 with individuals only affected with acrophobia indicates this region as a genetic risk factor for acrophobia.

1191T
Ancient human mtDNA variation is associated with Autism spectrum disorder in Europeans. D. Chaitkin1, O. Berdeneva2, M. Lyova3, A. Lazikov4, J. Leipziger5, D. Hadley6, H. Hakonarson6, D. Wallace1,2,3
1) Center for Mitochondrial & Epigenomic Medicine, Children’s Hospital of Philadelphia Research Institute, Philadelphia, PA; 2) Department of Psychiatry and Human Behavior, University of California at Irvine, Irvine, CA; 3) Center for Biomedical Informatics, Children’s Hospital of Philadelphia Research Institute, Philadelphia, PA; 4) Center for Applied Genomics, Division of Genetics, Children’s Hospital of Philadelphia, Philadelphia, PA; 5) Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA.

Autism spectrum disorders (ASD) are thought to result from genetic defects in synaptic biogenesis, structure, and maintenance. GWAS and CNV studies in autism have led to the identification of a number of chromosomal loci involved in the structure of the synaptic assembly pathway. However, when all of the synaptic loci are added together they still account for only a small proportion of ASD cases. Therefore, additional pathways must be relevant in the genetics of ASD. To explain the complex genetics of ASD, alternative ASD pathways would need to encompass a large number of genes, affect a quantitative trait, and be relevant to the synapse. Mitochondrial bioenergetics fulfills these criteria. Therefore, we are testing the hypothesis that a significant proportion of the risk for developing ASD is the result of inheriting partial defects in genes involved in mitochondrial bioenergetics. To test our hypothesis, first we determined if ancient human mtDNA variation is associated with ASD risk. To accomplish this, we used two independently generated and genotyped cohorts, the Autism Genetic Resource Exchange (AGRE) and the CHOP Autism Case-Control (ACC). The genome-wide mtDNA haplotype data were generated by Illumina chip analysis. We deduced the mtDNA haplogroups for all the AGRE and ACC samples first by converting the Yoruban mtDNA SNPs included in these arrays to their corresponding positions on the list and then using MitoMap to assign an mtDNA haplogroup to each sample. To limit the number of variables and thus increase the statistical power of the study, we only utilized cases and controls that harbored European haplogroups. Generalized linear modeling analysis of the AGRE cohort suggested that the mitochondrial lineage and T2 haplogroup are risk factors for ASD (J lineage: OR = 2.22, CI=1.21-4.05, p-value=0.0094; T2 haplogroup: OR=1.79, CI=1.26-2.55, p-value=0.0012). Congruently, in the ACC cohort we found that haplogroup T2 is strongly associated with ASD and a risk factor (OR=1.90, CI=1.20-3.03, p-value=0.0042). These data demonstrate that ASD risk is modified by mtDNA haplogroup thus supporting our mitochondrial bioenergetic hypothesis of ASD.
The contribution of uncommon coding variants to risk for Alzheimer’s disease, frontotemporal dementia, and progressive supranuclear palsy: an exome array study of the multi-ethnic GIFT cohort. J.A. Chen, Q. Wang, J. Davis-Turak, Y. Li, D. Chatzopoulos, H. Chur, C. Colman, C. DeCarli, T. Foroud, A. Huang, A. Kanyidas, E. Klein, W. Kukull, J. Lee, A. Levey, M. Mendez, J. Miller, W. Poon, J. Ringman, A. Rosen, R. Sears, J. Shapiro, A. Varpelainen, K. Wojta, B.L. Miller, D.H. Geschwind, G. Coppola. 1) Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA; 2) Department of Neurology, University of Southern California, Los Angeles, CA; 3) Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, CA; 4) Department of Neurology, University of California, Davis, CA; 5) Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; 6) Memory and Aging Center, University of California, San Francisco, CA; 7) National Alzheimer’s Coordinating Center, University of Washington, Seattle, WA; 8) Department of Neurology, Emory University, Atlanta, GA; 9) Department of Neurology, University of California, Los Angeles, CA; 10) Department of Nutritional Sciences, Rutgers University, New Brunswick, NJ; 11) Department of Pathology and Laboratory Medicine, University of California, Davis, CA.

Genetic contributions to neurodegenerative diseases such as Alzheimer’s disease, frontotemporal dementia, and progressive supranuclear palsy have become increasingly well characterized. However, few studies have examined the effect of low-frequency coding variants on a genome-wide level. Here, we used the Illumina HumanExome BeadChip array to genotype a large number of exonic variants and other polymorphisms in a cohort of patients with dementia and Alzheimer’s disease (224 with Alzheimer’s disease, 168 with frontotemporal dementia, and 48 with progressive supranuclear palsy), and 224 non-demented controls from the Genetic Investigation in FrontoTemporal dementia and Alzheimer’s disease (GIFT). An additional multi-ethnic replication cohort of 240 Alzheimer’s disease patients and 240 controls was used to validate suggestive findings. Association testing, on both the variant level and the gene level, was performed. No novel loci were detected with genome-wide significance in these tests, likely due to the limited power to detect an effect from rare variants of modest effect size. However, using a gene-based association (SKAT) approach that allowed for both coding and non-coding variants, we identified a number of loci that were significantly associated with Alzheimer’s disease risk. These loci included known Alzheimer’s disease genes such as PICALM, MAPT, and ABCA7, as well as novel candidate genes such as CDH1 and SPP1. This study highlights the importance of considering both coding and non-coding variants in the genetic susceptibility of Alzheimer’s disease and suggests that further study is needed to better understand the role of these novel candidate genes in the disease process.

A variant in Cadherin 1 (CDH1) achieves near genome-wide significance in African Americans using a liability model. J. Meiz, J. Chung, K.L. Lunetta, J. Haines, R.P. Mayeux, M.A. Pericak-Vance, G.D. Schellenberg, L.A. Farrer. 1) Alzheimer’s Disease Genetic Consortium, 1) BU Alzheimer’s Disease Center (Neurology), Boston University School of Medicine, Boston, MA; 2) Biomedical Genetics (Medicine), Boston University School of Medicine, Boston, MA; 3) Department of Biostatistics, Boston University School of Public Health, Boston, MA; 4) Department of Epidemiology and Biostatistics, Case Western Reserve University School of Medicine, Cleveland, OH; 5) Taub Institute on Alzheimer’s Disease and the Aging Brain (Neurology), Department of Neurology, Columbia University, NY; 6) Gertrude H. Sergievsky Center (Neurology), Columbia University, New York, NY; 7) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL; 8) John P. Hussman Institute for Human Genetics, University of Miami, Miami, FL; 9) Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; 10) Department of Ophthalmology, Boston University School of Medicine, Boston, MA; 11) Department of Epidemiology, Boston University School of Public Health, Boston, MA.

Late onset Alzheimer’s disease (LOAD) risk is influenced by multiple known genetic, clinical, and environmental factors. Including these factors as covariates in traditional genetic analyses of subjects ascertained by phenotype (case-control ascertainment) may decrease power to detect an association. However, a liability model can be used to extract risk information from known genetic and non-genetic factors without a reduction in power. A recent genome-wide association study (GWAS) by the Alzheimer’s Disease Genetic Consortium (ADGC) found that, in addition to the APOE-ε4 allele, a variant in the CDH1 gene (rs115550680) was significantly associated with LOAD at the genome-wide level in African Americans. In the current study, we conducted a GWAS in African Americans, employing a liability model that included these genetic risk factors as well as age and sex. Subjects included 1,910 well-characterized African American LOAD cases and 3,829 cognitively normal African American controls from 9 datasets from the ADGC. First, we used logistic regression (or logistic generalized estimating equations (GEE) for family-based datasets) to derive a liability score (Pearson residuals) and then conducted a genome-wide association study (GWAS) using linear regression (or linear GEE for family-based datasets) to meta-analyze across datasets. We obtained near genome-wide evidence of association (p=2.21×10^-8) between the liability score and a SNP in CDH1, a gene which encodes a calcium dependent cell-cell adhesion glycoprotein. In the ADGC’s previous African American GWAS of LOAD that used a traditional case-control model, the same CDH1 SNP reached a p-value of 1.33×10^-5, two orders of magnitude larger than using the liability model. The current study suggests a liability model can improve association signals in GWAS. We plan to add more clinical risk factors to our model with the hope of further increasing our power to detect an association.
A genome wide association study on fine motor speed. C.L. Saltzabal, J.A. Smith, J.C. Bis, L. Yu, A. Beiser, W. Zhao, J.I. Rotter, A.S. Buchman, P. Au, S.T. Turner, W.T. Longstreth, S.R. Kardia, D.M. Psaty, D.A. Bennett, A.L. Fitzpatrick, S. Seshadri, D.A. Bennett, T.H. Mosley, Neuro-CHARGE consortium. 1) Department of Neurology, Boston University School of Medicine, Boston, MA, USA; 2) Framingham Heart Study, Boston, MA, USA; 3) Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA; 4) Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; 5) Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, IL, USA; 6) Boston University Schools of Medicine and Public Health, Boston, MA, USA; 7) Institute for Translational Genomics and Population Sciences and Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; 8) Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA; 9) Division of Nephrology and Hypertension, Department of Internal Medicine, Mayo Clinic College of Medicine, Rochester, MN, USA; 10) Departments of Neurology and Epidemiology, University of Washington, Seattle, WA, USA; 11) Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA; 12) Department of Medicine-Geriatrics, University of Mississippi Medical Center, Jackson, MS, USA.

Background: The Finger Tapping Test (FTT) is one of the most extensively used tasks to evaluate fine motor speed, and lower scores are often indicative of brain damage. However, genetic variants associated with fine motor functions have not been investigated in large community samples. The aim of this study is to relate genome-wide genetic variation to fine motor speed in a large sample of middle-aged and elderly individuals using FTT scores. We used a novel approach, ExcoT, to define the average number of taps for the right and left hands over 10 or 15 seconds. We conducted two separate meta-analyses of genome-wide association studies (GWAS) of FTT scores in 7,580 individuals of European and 1,405 of African-American descent from the population-based cohorts of the Neuro-CHARGE consortium. Genotype imputation was based on the 1,000 Genomes reference panel. Results across cohorts were combined with pooled inverse variance-weighted meta-analysis methods using METAL and applying genomic control. We considered p-values of < 5 x 10^(-6) as findings of interest. Results: In the European sample, we identified suggestive intronic variants in ASTN1 (rs78553813, p=1.5 x 10^(-5)) and ALSZ2 (rs3731702, p=1.9 x 10^(-5)) even after correcting for the top hits, which may play a role in the pre-pulse/startle response, and possibly in alcohol dependence and other behavioral disorders, in this Mexican American cohort. (Supported by AA06420, DAO30976).

Large-scale genetic predictor of gene expression associated with risk of bipolar disorder. K.P. Shah, E.R. Gamazon, N.J. Cox, H.K. Im. Department of Medicine, University of Chicago, Chicago, IL.

Genome-wide association (GWA) studies have successfully identified thousands of loci associated with complex diseases, many of which have also been associated with gene expression phenotypes across a variety of tissues, known as expression quantitative trait loci (eQTLs). The enrichment of eQTLs in GWAS studies suggests that altered gene expression plays a role in risk for a number of complex diseases. Recognizing the important role of gene expression in risk for complex diseases, we developed and applied a novel approach, PrediXcan, to predict gene expression traits in GWAS studies and used these predicted gene expression traits for association with disease. We developed large-scale predictors of gene expression for all genes using gene expression data from the whole blood of control individuals as part of the GTEx Pilot Project. We then tested each of these predicted gene expression levels for association with bipolar disorder (BDP) using GWA data from the GAIN sample. We found that predicted levels of RFNG are associated with risk of BDP (p-value = 2.59 x 10^(-6)) even after correcting for multiple testing (p-value Bonferroni corrected = 4.19 x 10^(-5)). Higher predicted RFNG expression is associated with lower risk of BDP in this study. Our RFNG result replicated in the WTCCC GWA study as well (p-value = 0.00017). RFNG, RFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase, is part of the Notch signaling pathway and involved in neurogenesis. Previous studies have highlighted the importance of Notch signaling in psychiatric disease through GWA studies and linkage studies. Specifically, NOTCH4 has been associated with risk of schizophrenia and mutations in NOTCH3 have been seen in patients with CADASIL, a rare autosomal dominant disease characterized by cortical white matter death, stroke, mood disorders, and dementia. Our results suggest that RFNG gene expression may be relevant to the development of BDP further highlighting the importance of the Notch signaling pathway in neuropsychiatric disease risk. This study also shows the potential of our PrediXcan approach for identifying genes relevant to disease phenotypes.
1198S  


The most recent schizophrenia GWAS reported >100 independent associated loci, implying the disorder is highly polygenic. To better understand the pathology of neuropsychiatric disease, we formed the CommonMind Consortium (commonmind.org) to generate large-scale data (RNA-seq, ChIP-seq, DNA-seq/genotyping) from human post-mortem brain samples. Here, we focus on eQTL using RNA-seq (Illumina HiSeq2000 - paired end reads) of the dorsolateral prefrontal cortex (BA9/46) in 265 schizophrenia cases and 289 controls. Effects of known clinical (gender, age of death, medications) and technical (brain bank, post-mortem interval, RNA quality, sequencing batch) covariates, as well as hidden confounders, were removed using surrogate variable analysis (SVA). Data were normalized via voom and a linear model was applied to detect eSNPs, adjusting for genetic structure. Genotypes were assayed on the Illumina Infinium HumanOmniExpressExome8 chip and were imputed to 1000 Genomes. Preliminary differential expression analysis using linear models implemented in voom/limma identified 15.6% of all expressed genes as differentially expressed between cases and controls (FDR 5%). These genes were enriched for DNA variants associated with schizophrenia, including rare (allele frequency < 0.1%) non-synonymous variants in Swedish case-control exome sequencing (p=0.012) and common GWAS variants (p=0.045). Preliminary eQTL analysis of the assayed genotypes (~260,000 with MAF≥0.05) identified 102,461 and 85,015 proximal eSNPs (distance < 1Mb) at an FDR of 5% in controls and patients with schizophrenia, respectively; representing 53.1% and 44.3% of expressed genes. 51,823 eSNPs were common to both groups, and 24.8% of expressed genes have at least one eSNP in common. Preliminary results show that 54 of the 108 GWAS regions from the Psychiatric Genomics Consortium contain eSNPs. Complete analysis using imputed SNPs will examine condition dependent eQTL incorporating epigenetic information, and identify likely gene candidates in regions indicated by GWAS in hopes of contributing to the understanding of the biological mechanisms contributing to disease. This large dataset will be made public in early 2015, including a catalogue of brain-expressed genes and their isoforms in cases and controls, which will likely gene candidates in regions indicated by GWAS in hopes of contributing to the understanding of the biology of schizophrenia.

1199M  


Recent advances in genome- and exome-wide methods, including SNP genotyping arrays and high-throughput sequencing, have facilitated identification of genes associated with autism spectrum disorder (ASD) and schizophrenia (SCZ). This ever-expanding list of genes offers an unprecedented opportunity to elucidate the biology of ASD and SCZ. However, the large number of genes involved coupled with the many biological functions of each gene confound interpretation. Working on the hypothesis that the many genes involved in each disorder must disrupt a much smaller number of shared biological processes to result in a common phenotype, we previously identified the midfrontal prefrontal cortex and primary motor-somatosensory cortex as a point of convergence in the expression of ASD genes in the human brain. Since this initial analysis, the number of ASD-associated genes has increased more than three-fold, and additional loci continue to be implicated in SCZ. By applying this methodology in a cross disorder analysis of ASD and SCZ we aim to identify novel components of the neuropathology and to determine whether these are shared or distinct between the disorders. Using expression data from the Psychiatric Genomics Consortium, we identified 51,823 eSNPs were common to both groups, and 24.8% of expressed genes have at least one eSNP in common. Preliminary results show that 54 of the 108 GWAS regions from the Psychiatric Genomics Consortium contain eSNPs. Complete analysis using imputed SNPs will examine condition dependent eQTL incorporating epigenetic information, and identify likely gene candidates in regions indicated by GWAS in hopes of contributing to the understanding of the biological mechanisms contributing to disease. This large dataset will be made public in early 2015, including a catalogue of brain-expressed genes and their isoforms in cases and controls, which will likely gene candidates in regions indicated by GWAS in hopes of contributing to the understanding of the biology of schizophrenia.
1201S
AUTOSOMAL DOMINANT CEREBELLAR ATAXIA AND MENTAL IMPAIRMENT WITH A NOVEL NONSENSE MUTATION OF PRKCG. H. Shimazaki, J. Honda, T. Nago, M. Namekawa. Division of Neurology, Dept. of Internal Medicine, Jichi Medical University, Shimotsuke, Tochigi, Japan.

Backgrounds: Autosomal dominant cerebellar ataxias (ADCA) comprise clinically and genetically heterogeneous neurodegenerative disorders characterized by progressive ataxia with other neurological system disturbances. Machado-Joseph disease (MJD) / Spinocerebellar ataxia type 3 (SCA3), SCA6, SCA31, dentatorubral-pallidolysian atrophy (DRPLA), SCA2 and SCA1 are frequent ADDAs, but other types are rare in Japan. We encountered a family with two patients of cerebellar ataxia with mental impairment. We attempt to identify the causative gene mutation of this family with ADA. Methods: We investigated the proband with neurological examination, brain MRI, SPECT and gene analyses. Results: The neurological examination revealed cerebellar ataxia and cognitive impairment, but tremor or myoclonus were not observed. Brain MRI demonstrated cerebellar vermian atrophy. Brain SPECT showed cerebellar hypoperfusion. We could not detect CAG repeat expansions of SCA1, 2, 3.6, 7.8, 12, 17, DPRA genes, and sequencing analyses could not reveal pathologic substitutions in the 5’ UTR of the purkinjen gene. Whole-exome sequencing (WES) could identify the novel heterozygous nonsense mutation in the PRKCG gene, which is the causative gene for SCA14. This heterozygous mutation was confirmed by Sanger sequencing, and found in another patient, co-segregating within the family members and not found in controls. Conclusion: We could identify PRKCG mutation in this ADCA family. SCA14 usually showed cerebellar ataxia with tremor or myoclonus. Mental impairment is rare in SCA14. To the best of our knowledge, PRKCG gene mutations were only missense mutations in the previous literatures. This family has atypical clinical features with a novel nonsense mutation in the PRKCG gene. We could consider this nonsense mutation of PRKCG gene is associated with atypical clinical symptoms compared with typical ones of SCA14.

1202M
SPG7 mutations in a French-Canadian family affected by a recessive spastic ataxia. M. Tetreault1,2, S. Yang3, K. Choquette3, K. Boycott3, J. Majewski1,2, B. Bras3, C4R Consortium. 1) McGill University and Genome Quebec Innovation Center, Montreal, QC, Canada; 2) Department of Human Genetics, McGill University, Montreal, QC, Canada; 3) Neurogenetics of motion laboratory, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada; 4) Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa. Ottawa, ON, Canada.

The current implementation of whole exome (WES) and whole genome (WGS) sequencing, in research has greatly accelerated the identification of disease-causing genes in Mendelian disorders causing rare diseases that were poorly diagnosable. The diagnosis of ataxias is difficult due to rarity, clinical and genetic heterogeneity, and still today a large number of patients remain without a molecular diagnostic. We recruited a French-Canadian (FC) family (10 individuals: 4 affected cases) with an adult onset autosomal recessive spastic ataxia with dysarthria. Due to the heterogeneity and the difficulty to identify a disease gene by relying only on clinical features, we decided to send one affected individual for sequencing. WES was performed using Agilent SureSelect capture kit and Illumina HiSeq. A list of variants were not observed. Brain MRI demonstrated cerebellar vermian atrophy. Brain SPECT showed cerebellar hypoperfusion. We could not detect CAG repeat expansions of SCA1, 2, 3.6, 7.8, 12, 17, DPRA genes, and sequencing analyses could not reveal pathologic substitutions in the 5’ UTR of the purkinjen gene. Whole-exome sequencing (WES) could identify the novel heterozygous nonsense mutation in the PRKCG gene, which is the causative gene for SCA14. This heterozygous mutation was confirmed by Sanger sequencing, and found in another patient, co-segregating within the family members and not found in controls. Conclusion: We could identify PRKCG mutation in this ADCA family. SCA14 usually showed cerebellar ataxia with tremor or myoclonus. Mental impairment is rare in SCA14. To the best of our knowledge, PRKCG gene mutations were only missense mutations in the previous literatures. This family has atypical clinical features with a novel nonsense mutation in the PRKCG gene. We could consider this nonsense mutation of PRKCG gene is associated with atypical clinical symptoms compared with typical ones of SCA14.

1203T
Discovery, validation and genotyping of CNVs by analysis of genome sequence and microarray. D. Antaki1, M. Gujrati2, W. Brandt1, J. Majewski3, C. Consolo3, J. Sabatier3, I. Tabei3, K. Schmitz3, A. Cussac2, C. Deloukas3, C. Corsello2, 1) Institute for Genomic Medicine, La Jolla, CA; 2) Department of Psychiatry, UCSD, La Jolla, CA; 3) Department of Cellular and Molecular Medicine, UCSD, La Jolla, CA; 4) Autism Discovery Institute, Rady Children’s Hospital, San Diego, CA.

Advances in Next Generation Sequencing technology have enabled detection of an unprecedented array of structural genetic variants including Copy Number Variants (CNVs). CNVs confer risk for a variety of psychiatric and neurodevelopmental disorders, including Autism Spectrum Disorder (ASD). As we begin to apply whole genome sequencing approaches in our genetic studies of ASD, sensitive and accurate methods for CNV detection are needed. We applied a combination of DNA analysis techniques to the investigation of a cohort of 161 samples (28 ASD trios and 19 discordant sibling pair quad families), including deep 40X whole genome sequencing (325 bp libraries with 100 bp paired end reads), and two independent genotyping array platforms (Affy 6.0 and Illumina 2M). Genotyping was executed via a clustering algorithm, partitioning around medoids, resulting in significant noise reduction commonly found with SNP chip hybridization assays while distinctly segregating CNVs into clusters of heterozygous or homozygous duplications or deletions. Our robust multifaceted clustering method further enabled the detection of de novo CNVs. We demonstrate the advantage of a multidimensional approach in clustering and genotyping CNVs as opposed to the standard unidimensional approach.

1204S
Meta Analysis of Case-Control Autism Exome Sequencing Data. J.A. Kosovich1,2, K. Roeder3, B. Devlin2, J.D. Buxbaum4,5,6,7,8,9,10,11, M. Daly1,2,12, Autism Sequencing Consortium. 1) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA; 2) Stanley Center for Psychiatric Research and Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA; 3) 3Ray and Stephanie Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; 4) Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; 5) Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; 6) Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 7) Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 8) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 9) Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 10) Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 11) The MINDChild Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA; 12) Harvard Medical School, Boston, Massachusetts, USA.

Autism spectrum disorders (ASDs) currently affect 1 in 66 individuals and are highly heterogeneous in their phenotypic outcome ranging from severely affected individuals with speech impairment, intellectual disability, and severe stereotypic behaviors to mildly affected individuals. As we begin to apply whole genome sequencing approaches in our genetic studies of ASD, sensitive and accurate methods for CNV detection are needed. We applied a combination of DNA analysis techniques to the investigation of a cohort of 161 samples (28 ASD trios and 19 discordant sibling pair quad families), including deep 40X whole genome sequencing (325 bp libraries with 100 bp paired end reads), and two independent genotyping array platforms (Affy 6.0 and Illumina 2M). Genotyping was executed via a clustering algorithm, partitioning around medoids, resulting in significant noise reduction commonly found with SNP chip hybridization assays while distinctly segregating CNVs into clusters of heterozygous or homozygous duplications or deletions. Our robust multifaceted clustering method further enabled the detection of de novo CNVs. We demonstrate the advantage of a multidimensional approach in clustering and genotyping CNVs as opposed to the standard unidimensional approach.
1205M Interstitial duplication Xp11.4 and triplication of Yq11.22 leading to disruption of TSPAN7 and NLGN4Y in a child with autism. W.S. Baek, Neurology, Parkside Medical Group, Upland, CA. 1) Introduction TSPAN7 (tetraspanin 7) is a gene of the tetraspanin superfamily located at Xp11.4, which is highly expressed in the prefrontal cortex. TSPAN7 mutations have been reported to be associated with non-syndromic X-linked mild to moderate mental retardation. NLGN4Y (neuroligin 4, Y-linked) is a protein-coding gene located at Yq11.22, of which missense mutations leading to loss-of-function have been reported in autism. We report an unprecedented case with a combination of both Xp11.4 duplication and Yq11.22 triplication leading to partial duplication of the TSPAN7 gene presenting as autism. Keywords: tetraspanin 7(TSPAN7), neuroligin 4, Y-linked (NLGN4Y), copy number variant (CNV), autism Materials and Methods: Case Report A 5 year-old Hispanic boy presented with autism. Birth and family history were unremarkable. He had speech delay and behavioral problems. He was receiving special education and speech therapy. He was not dysmorphic. Results: SNP microarray analysis detected a 1.40MB interstitial duplication of Xp11.4 and a 5.67 MB interstitial triplication of Yq11.22, which was found to be a familial inheritance. Conclusions: The interstitial duplication of Xp11.4 led to disruption of the TSPAN7 gene. As the duplication disrupts the gene, this may result in a dominant-negative effect. Inheritance of an X chromosome copy number variant (CNV) from a clinically normal mother may yield clinical consequences in male offspring due to differences in gene dosage. The interstitial triplication of the Yq11.22 led to disruption of the NLGN4Y (neuroligin 4) gene. Missense mutations of NLGN4Y leading to loss-of-function have been reported in autism. Although direct inheritance of a Y chromosome CNV from a phenotypically normal father cannot be excluded, we hypothesize that this too might have had an additive effect on the proband.

1206T TRPM1, the transient receptor potential cation channel M1, harbors rare putatively damaging missense variants disproportionately transmitted to affected sibs in schizophrenia quads. S. Gulsun,1* T. Walsh1*, M.-C. King1*, J.M. McClellan2*. 1) Department of Medicine (Medical Genetics), University of Washington, Seattle, WA; 2) Department of Psychiatry, University of Washington, Seattle WA. Schizophrenia can be caused by de novo damaging mutations, either point mutations or CNVs. We extended this model to search for rare inherited damaging mutations in schizophrenia. We screened exome sequence data of the parents of 81 quads, each including a proband with schizophrenia and a healthy sibling. For each gene, all rare putatively damaging variants were identified and aggregated gene by gene. Applying the rare variant transmission disequilibrium test with discordant siblings (RV-TDT-DS), we calculated chi-square values for transmission disequilibrium of the variants for every gene. Genome-wide, the gene TRPM1 was the most extreme outlier for disequilibrium in transmission of rare damaging variants. At TRPM1, 14 parents carried putatively damaging variants, which were transmitted to 11 of 14 affected children and to 2 of 14 unaffected siblings (X2 = 11.63, 1 df; P=0.006). All rare putatively damaging variants were missense mutations at highly conserved residues. TRPM1 encodes the Transient Receptor Potential Cation Channel M1, which is highly expressed in retina, melanocytes, and brain. In rod bipolar cells, TRPM1 mediates synaptic transmission, specifically by depolarizing bipolar cells in response to glutamate release from photoreceptors. Recessive loss-of-function mutations in TRPM1 lead to congenital stationary night blindness. Recurrent gene-disrupting CNVs at TRPM1 have been previously identified in both schizophrenia and autism patients. TRPM1 lies in the 1.5 MB hotspot for such CNVs on 15q13.3. We suggest screening other cohorts of quads or trios with schizophrenia or autism for transmission disequilibrium at TRPM1, and conversely, evaluating the frequency of mental illness phenotypes in parents of children with congenital stationary night blindness. We also suggest evaluating the possible role of TRPM1 in synaptic transmission in brain structures.

1207S Screening for Mutations in Non-Syndromic Autosomal Recessive Intellectual Disability Genes in Non-Consanguineous Intellectual Disability and Autism Spectrum Disorder Patients. K. Lin,1* C. Mittal1*, J. Davidson1*, N. Vasi1*, A. McNaughton1*, A. Mikhailov1, K. Roetter1, M. Hudson1, C. Windpassinger2, P. Magee3, L. Al Ayadhi4, W. Kaschnitz5, E. Petek6, D.J. Stavropoulos7, M.J. Carter8, M. Ayub9, J.B. Vincent1. 1) Dept Psychiatry, Queen’s Univ, Kingston, ON, Canada; 2) MIND LAB, CAMH, Toronto, ON, Canada; 3) Human Molecular Genetics Institute, Medical University of Graz, Graz, Austria; 4) King Saud University, Riyadh, Saudi Arabia; 5) Department of Paediatric Laboratory Medicine, Hospital for Sick Children, Toronto, ON, Canada; 6) Department of Paediatrics, Division of Developmental and Behavioral Pediatrics, Sick Children, University of Toronto, Toronto ON Canada. Intellectual disability (ID) and autism spectrum disorder (ASD) are believed to occur each with a prevalence of ~1%. In addition, ID is present in up to 70% of individuals with ASD, and in ID populations as many as 28% may meet criteria for ASD. Both ASD and ID are frequently the result of genetic aberrations, however only a fraction of patients currently receive a genetic diagnosis. Genes identified for ID have frequently been implicated in ASD and vice versa. Autosomal recessive (AR) mutations are the main cause of ID in populations where consanguinity rates are high, but even in outbred populations the rate may be 13-24%. Here we use a targeted next generation sequencing approach using Amplicase primer pools and the Ion Proton platform to assess the involvement of known recessive ID genes in outbred ID and ASD populations, and comparison to some known dominant and linked genes. We combined this with a pooling strategy to enable mass screening. Our primer pools targeted known ID and ASD genes, including 67 genes reported as non-syndromic ARID genes, 7 reported as syndromic ARID genes, 3 reported as genes encoding X-linked ASD or ID genes, and 6 reported as X-linked ID or ASD genes. To date, with pools of 20, we have screened ~1,200 ASD and 740 ID individuals from outbred populations. We included as positive controls two samples, with known homozygous base substitution or deletion. Our work flow and pipeline aimed to identify nonsense and damaging missense mutations and indels, homozygous or compound heterozygous for AR, heterozygous for dominant and X-linked recessive and hemizygous for X-dominant. Alignment and analysis was performed using Ion Reporter. The control base substitution was present in ~10% of reads for that pool, indicating good coverage and identification, however the Sbp deletion was present in only 8 out of 1300 (0.6%), suggesting the alignment for indels is not as strong. In addition, we have identified numerous variants of interest in NS-ARID genes such as MAN1B1, but very few in others. We have also identified variants of interest in X-linked genes PTCHD1 and MECP2. A pooling strategy for targeted gene sequencing is an efficient means of screening large numbers of patients for multiple candidate genes, however improved alignment algorithms are needed to identify indels. Some AR genes may be relatively common causes of ID and ASD in outbred populations.

1208M Association study of TREM2 exon 2 variants with late-onset Alzheimer’s disease in Iranian elderly population. Z. Mehrjoo1, A. Najmabadi2, S.S. Abbedini1, K. Kamali3, H. Najmabadi1, H.R. Khorram Khoshid1. 1) Department of Research Center for Neurological and Genetic Sciences, Tehran, Iran; 2) Division of Biological Sciences, University of California, San Diego, California, USA; 3) Reproductive Biotechnology Research Center, Avicenna Research Institute, Academic Center for Education, Culture and Research (ACECR), Tehran, Iran. Dementia, a major elderly disability, is increasing all over the world, especially in developing countries. The most common form of dementia is Alzheimer’s disease (AD [MIM 104300]), which is defined as a neurodegenerative disorder that affects memory and thinking ability; it then impairs basic body movement, and eventually leads to death. Recent studies on AD disclosed significant association between the rs-75932628-T variant of the TREM2 [MIM 605086] gene and AD in European and North American populations, but not in Han Chinese individuals. However, the rs-75932628-T allele frequency in this cohort of Iranian elderly population was higher than other reported populations (0.8%) but did not show a statistically significant association with AD (odds ratio [OR]: 4.8; 95% confidence interval [CI]: 0.54 to 43.6; P = 0.270). Although we identified more rare variants in AD patients compared to controls, the abundance of TREM2 rare variants do not reach a statistically significant association with AD, which could be a result of low sample size. Therefore, to continue this study and to obtain meaningful results, we are collecting and genotyping more samples.
1209T
Increased Genome-wide Burden of Rare Coding Variants in Schizophrenia, L.M. Olde Loohuis1, J.A.S Vorstman2,4, A.F. Ort1, K.A. Staats1, T. Wang1, J. DeYoung1, R.S. Kahn2, R.M. Cantor1,3, R.A. Ophoff1,2,3,1 Center for Neurobehavioral Genetics, University California Los Angeles, Los Angeles, CA; 2) Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands; 3) Department of Human Genetics, University of California Los Angeles, Los Angeles, USA.

Schizophrenia is a highly heritable polygenic disorder with suspected neuropsychological and immunological components. While GWAS continue to uncover common single nucleotide variants (SNVs) contributing to this phenotype, rare variants of the order of 1-10% are often understudied. The current study aimed to uncover additional genes contributing to schizophrenia by analyzing the cumulative effects of coding variants occurring in controls or cases only. Our analysis is based on exome SNP array genotyping of 1,022 schizophrenia cases and 251 controls from a homogeneous population from the Netherlands. Using CONDEL, a scoring tool for estimating the functional impact of non-synonymous SNVs, we assigned a burden score to each individual by summing the scores of every variant carried. In cases versus controls, we observe a significantly increased individual genome-wide rare variant burden for deleterious non-synonymous variants (empirical P = 0.024). In addition, an increased number of genes with double hits (empirical P = 0.035), as well as splice site variants (empirical P = 0.003) were seen. The genes implicated by the increased burden of rare coding variants were significantly enriched with the well-established schizophrenia susceptibility genes (empirical P = 0.033), and are enriched for those that are expressed in fetal brain and spleen. We demonstrate that non-synonymous rare variants are important in the etiology of schizophrenia. Previous reports suggest the enrichment of schizophrenia-specific non-synonymous variants in schizophrenia in a set of pre-selected genes. Our results now extend these and implicate non-synonymous SNVs at a genome-wide level. This signal is enriched for genes implicated through GWAS studies of schizophrenia, implying the contribution of common and rare variants to schizophrenia from identical loci. The genes containing these rare coding variants overlap significantly with genes expressed in the fetal brain and spleen, highlighting the potential involvement of neurodevelopment and immune system in disease etiology.

1211M
Identification of Molecular Markers in Parkinson’s Disease Using Next Generation Sequencing, S.M. Sperber1, Y. Shih2,3, R. O’Rourke2, K. Kornets5, E. Berry-Kravis3, E. Spector2,3, D. Hait1. 1) Genetic Testing Laboratory, Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mt. Sinai, New York, NY; 2) Children’s Hospital of Colorado, Aurora, CO; 3) University of Colorado AMC, Aurora, CO; 4) Rush University, Chicago, IL.

Parkinson’s disease (PD) is a heterogeneous neurodegenerative disorder with insidious onset and devastating progressive decline currently affecting over 4.5 million older Americans. Prevalence is projected to double by 2030 causing considerable economic burden. High impact genes that contribute to PD have been identified in individual families and by GWAS studies. How different genetic variants influence the PD phenotype has yet to be fully explored. A targeted test could rapidly identify variants in multiple genes simultaneously that provide risk information and may be informative for potential phenotypic categorization necessary for future therapies.

Here we identified genetic variants contributing to PD in a cohort of clinically defined PD patient samples compared to genome population data using a targeted panel of genes associated with PD and next generation sequencing techniques. Ninety-nine patients meeting Gelb criteria for PD were interviewed, examined, and had blood collected for genotyping. The mean age was 68.9±10.3 years, 65% male, and age of onset of disease was 62.1±10 years. Mean UPDRS motor score was 25.6±11 and mean UPDRS total score was 40±16. A set of high impact genes associated with PD was exon sequenced in a targeted manner, using a PCR-based enriched panel. Sequencing the cohort’s genes identified on average, greater than 300 variants per patient. One hundred and nine variants of potential clinical interest were identified. Of the 109 variants, 26 were previously identified variants. The majority of variants (n=51) were found to be associated with a reduced risk of PD and 47 are likely benign. The patients exhibited variant profiles that highlight the utility of genetic testing to further define the gene pathways and their associated phenotypic characteristics.

The target next panel is sensitive, cost effective and efficient. The genotypic data pointing to particular pathways for further investigation. Acknowledgements: Parkinson’s Disease Foundation, Anti-Aging Foundation, Denver Genetic Laboratories.
1213S
Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. M.G. Butler1, S.K. Rathi1, H. Wang2, D. Stephan2, A.M. Manzardo1. 1) Psychiatry & Behavioral Sciences and Pediatrics, University of Kansas Medical Center, Kansas City, KS, USA; 2) Silicon Valley Biosystems, Foster City, CA, USA; 3) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA.

Classical autism or autistic spectrum disorder belongs to a group of genetically heterogeneous conditions known as Autism Spectrum Disorders (ASD) with a predominant male to female (4:1) ratio. The overall occurrence rate is higher than epilepsy or Down syndrome. Although understood in females, heritability estimates are as high as 90%. A recent review of genetic linkage data, candidate genes and genome-wide association studies along with high resolution DNA microarray and next generation sequencing have led to a compilation of clinically relevant candidate and known genes for ASD for a total of 629 genes (Butler et al., 2014). We chose to undertake whole exome sequencing of 30 well-characterized Pakistani females with autism (average age, 7.7 ± 2.6y; age range, 5 to 16y) selected from the Autism Genetic Research Exchange (AGRE) repository (www.agre.autismspeaks.org) from multiplex families having a high probability of causation due to gene disturbances. Genomic DNA (Surg) samples were used for whole exome sequencing via paired-end next generation sequencing approach using the Illumina HiSeq2000 platform with the Agilent SureSelect Human All Exon v4 - 51Mb. Primary sequence data files were aligned to the reference. variants called and functional significance of each variant was calculated to generate a rank-ordered list. The list of disease-causing genes were developed with the primary selection criteria using a machine learning-derived Classification (G2M) score and adding cutoff levels for other predictive parameters (GERP, PolyPhen2, and SIFT). Prior candidate genes and genomic variants were then subjected to further screening for biological significance. We identified between 100 and 300 genes showing genomic variants of novel or candidate genes for autism per subject using the Classification (G2M) score, increasing >0.7 further narrowed the list to 10 to 20 genes. Seventy-eight genes were identified as meeting our selection criteria, range of 1 to 9 genes per female. Five subjects presented with disturbances of X-linked genes (SYTL4, GPRAS2, PIR, IL1RAPL1, GABRP). The cadherin, protocadherin and ankyrin repeat gene families were most commonly disturbed (e.g., CDH6, FAT2, PCDH8, CTNNA3, ANKRD11) along with genes related to neurogenesis and neuronal migration (e.g., SEMA3F, MIDN) indicating the usefulness of whole exome sequencing in females with ASD from multiplex families to enable identification of known and novel gene mutations.

1214M
Novel compound heterozygous PIGT mutations caused multiple congenital anomalies-hypotonia-seizures syndrome 3. M. Nakashima1, H. Kashiri2, Y. Murakami2, M. Kato3, Y. Tsurusaki1, N. Miyake1, M. Kubota2, T. Kinoshita3, H. Saitoi1, N. Matsumoto1. 1) Yokohama City University Graduate School of Medicine, Yokohama, Japan; 2) Division of Neurology, National Center for Child Health and Development, Tokyo, Japan; 3) Research Institute for Microbial Diseases and World Premier International Immunology Frontier Research Center, Osaka University, Osaka, Japan; 4) Department of Pediatrics, Yamagata University Faculty of Medicine, Yamagata, Japan.

Recessive mutations in genes of the glycosphingolipidoses (GPI-anchor synthesis pathway have been demonstrated as causative of GPI deficiency disorders associated with intellectual disability, seizures, and diverse congenital anomalies. In this study, we performed whole exome sequencing in a Japanese patient with progressive encephalopathies and multiple dysmorphisms with hypophosphatasia and identified novel compound heterozygous mutations, c.250G>T (p. Glu84) and c.1342C>T (p. Arg448Tyr) in PIGT encoding a subunit of the GPI transamidase complex. The surface expression of GPI-anchored proteins (GPI-APs) on patient granulocytes was lower than that of healthy controls. Transfection of the Arg488Tyr mutant PIGT construct, but not the Glu84* mutant, into PIGT-deficient cells partially restored the expression of GPI-APs DAF and CD59. These results indicate that PIGT mutations caused neurological impairment and multiple congenital anomalies in this patient.

1215T
Common, low frequency, and rare coding variants in CHRNA5 contribute to nicotine dependence in European and African Americans. E. Olsson1, N.L. Saccone2, E.O. Johnson3, N. Breslau4, D. Haltzmann5, K. Doheny6, L. Fox7, S.M. Gogarten8, K. Hetrick9, C.C. Laurie10, B. Marosy1, J. Stites1, J. Rice1, A. Goate1, L.J. Bierut4. 1) Department of Psychiatry, Washington University School of Medicine, St Louis, MO; 2) Department of Genetics, Washington University School of Medicine, St Louis, MO; 3) Behavioral Health Epidemiology Program, RTI International, Research Triangle Park, NC; 4) Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI; 5) Department of Psychiatry, University of Minnesota, Minneapolis, MN; 6) Center for Inherited Disease Research, Johns Hopkins University, Baltimore, MD; 7) Department of Biostatistics, University of Washington, Seattle, WA; 8) Institute for Behavioral Genetics, University of Colorado, Boulder, CO.

Purpose: The functional nonsynonymous variant rs16969968 in the 5′ nicotinic receptor subunit gene (CHRNA5) is the single strongest genetic risk factor for nicotine dependence in European Americans (MAF=0.35), and evidence suggests a similar contribution to risk in African Americans (MAF=0.06). Additional variants in CHRNA5 likely increase liability for nicotine dependence. This study examines targeted sequence data from approximately 3,000 unrelated cases and controls to assess the influence of CHRNA5 coding variation on nicotine dependence. Methods: Community-based recruitment enrolled subjects aged 25-45. Cases had a Fagerstrom Test for Nicotine Dependence score of ≥7 and controls had a score of 0 or 1. Custom next-generation sequencing with mean on-target coverage of 180X was performed on regions associated with smoking behaviors. Logistic regression was used to model case-control status with the variables sex, age, ethnicity, individual common variants (MAF≥1%), and collapsed rare variants (MAF<0.5%). Results: Sequencing identified 24 nonsynonymous variants and 2 frameshift deletions of high quality in CHRNA5. The well-studied rs16969968 was the only common variant among these, and the minor allele was associated with increased risk of nicotine dependence (OR=1.2, p=0.0007). Three low frequency nonsynonymous variants were identified, and all individually trended in the risk direction. The collapsed low frequency variant carried a minor allele was associated with nicotine dependence (OR=1.45, p=0.01). The remaining 22 coding variants were rare (each occurred in 1-4 individuals), and similarly, the collapsed term revealed a risk effect (OR=2.39, p=0.03). Nagelkerke’s adjusted R2 was used to assess the proportion of nicotine dependence variation explained by CHRNA5 variants after adjustment for sex and age. In European Americans, the well-studied rs16969968 variant gave an R2 of 0.9%, and the addition of all other CHRNA5 coding variants increased the R2 to 2.4%. In African Americans, rs16969968 gave an R2 of 0.2%, and adding all coding variants increased the R2 to 0.8%. Conclusions: Our findings suggest that common, low frequency, and rare coding variants in CHRNA5 are associated with increased risk of nicotine dependence. Coding variation in CHRNA5 accounts for over 2% of the nicotine dependent variation in European Americans and close to 1% in African Americans.
1216S
Trio-based exome sequencing indicates the ion homeostasis is relevant to bipolar disorder. N. Matoba1,2, M. Kataoka1,2, K. Fujii3, T. Kato4.
1) Laboratory for Molecular Dynamics of Mental Disorders, RIKEN BSI, Wako, Saitama, Japan; 2) Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan; 3) Department of Child Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan; 4) Department of Psychiatry, Dokkyo Medical University School of Medicine, Mibu, Tochigi, Japan.

Background: Bipolar disorder (BD) is a mental disorder characterized by recurrent manic and depressive episodes. Heritability of BD is reported around 59-87% (Smoller et al., 2003). Though recent genome wide association studies (GWAS) showed associated with several SNPs, each candidate SNP has a weak effect and can collectively explain only 38% of heritability (Lee et al., 2011). Therefore, we hypothesize that multiple rare variants may also contribute to BD.

Methods: We recruited participants mainly through E-mail newsletter of Bipolar Disorder Research Network Japan. All probands and their parents were interviewed by a trained psychiatrist using M.I.N.I. In total 59 trios including 35 original trios and independent 24 trios were analyzed. From the exome sequencing data of these trios, we selected only rare variants referring to public databases such as dbSNP137. These variants were classified as "transmitted" or "un-transmitted" based on the inheritance pattern. Then we investigated the characteristics of the set of transmitted variants compared to those in un-transmitted.

Results: There was no significant difference of numbers of variants between two groups. Gene ontology analysis showed significant enrichment of 57 GO terms in transmitted CNVs compared to those in independent 24 trios. These GOs included calcium ion homeostasis. In addition, this finding was not observed in trios of healthy siblings of probands with autism spectrum disorder (ASD), which showed slightly similar trends in some of those GOs.

Discussion: Our finding highlights the role of multiple rare damaging variants related to ion homeostasis in the pathophysiology of BD, which has been supported by genetic association with SNPs of calcium channel genes by GWAS and elevated calcium levels in blood cells among others. The findings need to be replicated in a larger set of families including multiple populations. Since we directly compared "transmitted" and "un-transmitted" variants, the results of gene enrichment analysis are robust to artifacts due to gene contents and so on. This simple but robust approach could also be utilized to other complex genetic diseases.

1218T

Background: Autism Spectrum Disorders (ASDs) are a group of neurodevelopmental conditions with a demonstrated genetic etiology. Rare (<1% frequency) copy number variations (CNVs) account for a proportion of the genetic events involved, but the contribution of these events in non-European ASD populations remains relatively unexplored. As clinical microarray testing becomes the adopted standard of care across medical genetics labs worldwide, studies describing the genetic architecture in these other ASD populations will become even more important. Here, we report on rare CNVs detected in a cohort of individuals with ASD of Han Chinese background.

Methods: DNA samples were obtained from 104 ASD probands and their parents who were recruited from Harbin, China. Probands were diagnosed using the Autism Behavior Checklist (ABC) and Childhood Autism Rating Scale (CARS). The DNA samples were genotyped on the Affymetrix CytoScan HD platform, a high-resolution array comprised of nearly 2.7 million probes. CNVs were confirmed in multiple CNV-calling algorithms. Rare CNVs were identified by comparing data with 873 technology-matched controls from Ontario and 1,235 additional population controls of Han Chinese ethnicity. Results: 8.6% of the probands had at least 1 de novo CNV (overlapping the GIOYP2, SPRY1, 16p13.3, 16p11.2, 17p13.3, 17p13.2, DMD, and NAP1L6 genes/loci). Rare inherited CNVs affected other plausible neurodevelopmental candidate genes including GRID2, LINGO2, and SLC9A12. An Asian-specific copy number polymorphism (duplication) was also identified in a proband with synteny block deletions described as etiologic in ASD and other developmental disorders.

Conclusions: Our findings help define genomic features relevant to ASD in the Han Chinese and emphasize the importance of using ancestry-matched controls in medical genetic interpretations.
1219S
Next Generation Sequencing for the study of ALS and other Motor
Neuron Diseases. C. Gellera1, B. Castellotti1, V. Pensato1, S. Magni1,
D. Di Bella1, G. Lauria2, N. Ticzozi2, F. Taroni2. 1) Unit of Genetics of
Neurodegenerative and Metabolic Diseases, Fondazione IRCCS - Istituto
Neurologico Carlo Besta, Milan, Italy; 2) Unit of Neurology, Fondazione
IRCCS - Istituto Neurologico Carlo Besta, Milan, Italy; 3) Department of
Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Ita-
liano, Milan, Italy.

ALS is an adult-onset, progressive, neurodegenerative disease caused by
the selective loss of motor neurons with a large prevalence of sporadic
cases (SALS). The pathogenesis is largely unknown, but genetic factors
likely play a major role in the disease. Although mutations in several genes
have been described, the underlying cause is still unknown in about 50%
of familial ALS (FALS) and in the large majority of SALS (80%). Genetic
determinants are unlikely to be explained by traditional strategies (linkage
analysis or candidate gene approach). In order to better define the genetic
contribution in ALS and to fill the genetic gap in FALS, we set up an
integrated NGS approach as follow: 1) customized gene panel screening,
2) wide exome sequencing (WES), 3) variant filtering and validation. The
panel has been designed by TruSeq Custom Amplicon Studio Design Illum-
ina and includes 51 ALS causative genes. Validation process indicates
that these genes are analyzed with a coverage of > 90%. Though still in a limited
number of SALS (n=26) (selected to be negative for SOD1, TARDBP
and C9ORF72 gene mutations), data analysis allowed the identification of
several variants of possible pathogenic significance. After a first filtering
step which excludes intronic variants and known polymorphisms with MAF
>1%, we get a report showing on average 50 variations per sample in exonic
or splicing regions. In 10 cases we identified at least one variant of possible
pathogenetic significance: 2 cases with a missense variant in GARS gene, 1
case with a missense mutation in VCP gene already described in association
with the disease, 1 case with a missense mutation in OPTN gene, 1 case with
a stop mutation in NEFH gene, 1 case with a missense mutation in DCTN1
gene, 1 case with 2 missense variants in SETX gene, 1 case with a missense
variant in SQSTM1 gene, 1 case with a splicing mutation in ASAH1 gene.
The genetic determinants validated in the laboratory have a high specific
definition process not only for FALS but also for SALS or other
overlapping motor neuron phenotypes. WES is currently in progress for
85 FALS of our cohort in an international collaborative study (John Landers,
Massachusetts, Vincenzo Silani, University of Milano, Cinzia Gellera, Istituto Neurologico Besta, Milano) supported by ARISLA foundation.

1220M
Exome sequencing of familial temporal lobe epilepsy with hippocampal
scarring in parent-offspring trios. S.S. Cherry1, J.K.L. Wong1, P.C.
Sham1, L. Baum2, P. Kwan2. 1) Department of Psychiatry, University of Hong
Kong; Pokfulam, Hong Kong; 2) School of Pharmacy, Chinese University of
Hong Kong, Shatin, Hong Kong; 3) Department of Medicine, University of
Melbourne, Australia.

Mesial temporal lobe epilepsy related to hippocampal sclerosis (MTLE-
HS) is recognized to be the most drug resistant type of epilepsy, and HS is
the most common pathological substrate identified in pharmacoresistant
temporal lobe epilepsy. The condition is characterized by relatively well-
defined clinical, electrographic, radiologic, and pathologic changes. Despite
attempts in finding genetic variants associated with epilepsies by genome-
wide association studies on European and Chinese subjects by our team and
collaborators, limited common variants reached genome-wide significance.
Meanwhile, the causes of the pathologic changes in MTLE-HS remain
unknown, and no preventive intervention has been identified. Discovery of
novel genetic markers could allow understanding patients with age-related
epilepsy subtypes for early surgical or pharmaceutical intervention, and ul-
mately lead to genetic therapies. To further explore the contribution of new,
rare variants in MTLE-HS, we plan to sequence exomes of 30 trios of
proband with their unaffected parents. So far we have sequenced 9 trios,
and 10 de novo mutations were identified. Among these, two de novo
mutations are within genes involved in neurological functions (NLGN3 and
NBEAL1) and might be involved in disease development. NLGN3 was
reported to be associated with X-linked autism in a Swedish family and
even the high comorbidity of autism spectrum disorders. This suggests that
epilepsy, this de novo mutation might contribute to MTLE-HS development.
1223M

Identification of rare variants from exome sequencing in a large family with dyslexia. A. Carrion-Castillo1, C. Franks1,2, B. Franke3, S.E. Fisher4,5.
1) Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands; 2) Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands; 3) Radboud university medical center, Radboud University, Nijmegen, Netherlands.

Dyslexia is one of the most common human neurodevelopmental disorders (5-10% prevalence), and it has a complex genetic aetiology. So far, the common disease - common variant approach has not identified many risk alleles. Most of the heritability for dyslexia therefore remains unexplained. With the goal of identifying a rare variant with an individually substantial effect on this complex trait, we analyzed a 30 member multigenerational pedigree with dyslexia, in which evidence for a susceptibility locus in Xq27.3 has previously been identified by parametric linkage analysis based on microsatellite genotyping (de Kovel et al. 2004). In the current study, we sequenced the exomes of 10 affected members of the family. After quality- and frequency-based filtering (MAF ≤ 0.01) we identified 177 exonic, non-synonymous, rare variants present in at least 7 affected members and not in a set of unrelated people. Microsatellite genotypes and common SNPs called from the exome data were combined to perform non-parametric linkage analysis, which defined three major regions of interest (NPL > 2) on chromosomes 2q, 20q and Xq, and suggestive regions (NPL > 1) on 13 other chromosomes. The 177 variants identified by sequencing were then filtered in the rest of the family members by sampling inheritance vectors from a Markov chain Monte Carlo analysis of the multi-locus marker data. We used the imputed allelic dosages for exomic variants to carry out family-based association analysis and to determine which sequence variants reside on the haplotypes that defined the three NPL regions in the family. Linkage results were used to prioritize a subset candidate variants for validation by Sanger sequencing in all the family members.

1224T

Analysis of major amyotrophic lateral sclerosis genes in Japan. R. Nakamura1, J. Sone1, N. Alsutra2, H. Watanabe4, D. Yokoi1, H. Watanabe1, M. Ito1, J. Senda1, F. Tanaka2, G. Sobue3,1, the Japanese Consortium for Amyotrophic Lateral Sclerosis research (JaCALS).
1) Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan; 2) Department of Neurology, Yokohama City University Graduate School of Medicine, Yokohama, Japan.

Background
Amyotrophic lateral sclerosis (ALS) is an adult-onset neurodegenerative disease characterised by progressive upper and lower motor neuron loss. Five to ten percent of ALS patients have a family history of the disease, whereas most patients are sporadic. Mutations in over 20 genes have been identified as causative or susceptibility genes of ALS in Japan. A recent study of over 200 Japanese sporadic ALS patients found an allelic association of the APOE epsilon4 allele with sporadic ALS.

Methods
We conducted a genome-wide association analysis of ALS patients in Japan. We used DNA from 219 exomic variants from sporadic ALS patients in Japan. The frequency of each allele was compared between cases and controls. Analyses of additional families were performed.

Results
We observed differences between cases and controls. Analyses of additional families confirmed this finding. Examining the different classes of coding variation, our multiplex families with dyslexia may account for the origin and manifestation. To this end, we have proposed that the discordance for this disease in monozygotic twins is only 50%, suggesting that traditional genetic elements alone may not account for their origin and manifestation. To this end, we have proposed that the discordance for this disease in monozygotic twins may be due to ontogenic de novo mutations. Further, the de novo events may include copy number variations, structural variants, indels and single nucleotide variants. In order to test this hypothesis, we have analyzed DNA extracted from blood samples of two pairs of monozygotic twins discordant for schizophrenia and their parents using Complete Genome Sequencing.

The analysis of the data has revealed that the genomic sequence of monozygotic twins harbor unique variations at the DNA level. The best explanation for some of these differences includes de novo events. Also, a number of these de novo variants present in the affected twins are not shared with their co-twin and may contribute to their disease discordance. We have confirmed a number of such using Real Time PCR technology. These include a CNV loss at 7q11.21, and CNV gains at 15q11.2 and 12p24, respectively. In addition, a large tandem spanning spanning from 1p36.31 to 1p36.33 which overlaps many genes including the GABRD that has been previously associated with schizophrenia in humans. Further, a number of single nucleotide variants and small insertions and deletions have been identified to be present only in the affected twin and not found in parental samples. The results are compatible with the suggestion that the genetic differences between monozygotic twins may arise during ontogeny. More importantly, depending on the gene(s) affected these differences may play a role in the discordance of twins for diseases, including schizophrenia.
1227T
Association analysis of MAPT with cerebrospinal fluid tau using targeted sequencing data in older adults with mild cognitive impairment or Alzheimer’s disease. K. Delers1, K. Nho1, S. Kim1, M.W. Weiner2, T. Forour1, J.Q. Trojanowski3, L.M. Shaw2, R.C. Green2, A.W. Toga2, A.J. Saykin1. 1) Indiana University School of Medicine, Indianapolis, IN; 2) University of Pennsylvania School of Medicine, Philadelphia, PA; 3) University of California, San Francisco, San Francisco, CA; 4) University of Southern California, Los Angeles, CA; 5) Brigham and Women’s Hospital and Harvard Medical School, Boston, MA.

The microtubule associated protein tau (MAPT) gene encodes the protein tau, one of the main contributors to the pathogenesis of Alzheimer’s disease (AD). SNPs of MAPT have previously been associated with AD and cerebrospinal fluid (CSF) levels of tau. However, most studies focused on only six haplotype-tagging SNPs of MAPT. While SNP genotyping covers only limited portions of the whole genome, whole-genome sequencing (WGS) has been used to obtain the most comprehensive genetic variation of an individual. Thus, we identified complete common variants (minor allele frequency (MAF) > 5%) of the MAPT gene from WGS data to investigate a comprehensive genetic association to CSF tau. WGS of 782 non-Hispanic Caucasian participants from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) was used to identify common variants within MAPT resulting in 870 SNPs after standard quality control steps. Linear regression models were then used to investigate the association of MAPT common variants with CSF tau covering for age, sex, and APOE status. Permutation testing was used to adjust for multiple-testing comparisons. Gene-based analysis of MAPT was significantly associated to CSF tau (p=0.03, corrected). MAPT SNPs rs117199550 (p=0.018, corrected) and rs63750072 (p=0.039, corrected) were significantly associated with CSF tau and were in linkage disequilibrium (r²=0.616, D'=0.861). Rs117199550 is intronic and rs63750072 is located in an exon excluded from the major isoforms expressed in the brain but has been implicated in Frontotemporal dementia. Our results suggest that variance within MAPT is associated with CSF tau levels. To our knowledge, this is the first study to perform a comprehensive association analysis of MAPT variants with CSF tau using WGS. Our findings also suggest that WGS might enable the discovery of new variants associated with AD-related biomarkers.

1229M
Exome sequencing in extended families with age-related macular degeneration reveals enrichment of genes involved in extracellular matrix pathway. R. Priya1, S. Perez1, M. Mutuddi2, K. Branham2, M. Othman2, J. Heckel2, A. Swaroop3. 1) Neurobiology-Neurodegeneration & Repair Laboratory (N-NRL), National Eye Institute, Bethesda, MD; 2) Department of Molecular and Human Genetics, Banaras Hindu University, Varanasi, India; 3) Department of Ophthalmology, University of Michigan, Ann Arbor, MI, USA.

Age-related macular degeneration (AMD) is a major cause of visual impairment in elderly population. Genome wide association studies have identified 20 common susceptibility loci, which explain up to half of the AMD heritability. More recently targeted sequencing has helped identify penetrant, high-risk low-frequency coding variants at some of the known AMD loci. To further understand the genetic contributions of rare exonic variants, we undertook whole exome sequencing (WES) in 19 AMD families with 2-5 affected members. Exome of 69 samples from 19 families were captured using Agilent SureSelect Human All Exon kit and sequencing data was generated on Illumina GAIIx. Reads were aligned to the human reference genome (NCBI build 37.3) using Burrows-Wheeler Aligner, variants were called using GATK and annotations were done using ANNOVAR. We identified 343 rare segregating variants (MAF ≤ 0.05%) in these families. Subjecting this list to the pathway analysis using the database for annotation, visualization and integrated discovery (DAVID) recognized genes in extracellular matrix (ECM) as top functional clustering. ECM pathway along with complement and lipid metabolisms pathways have been associated with AMD based on GWAS analysis. Thus it was interesting to see this association in exome sequencing analysis. Many of these candidates were highly expressed in human fetal retinal transcriptome that are relevant in AMD pathophysiology. Taken together, our data shows that exome sequencing in extended AMD families can provide important insights into the disease etiology.

1228S
Deep whole genome sequencing reveals multiple hits in non-coding sequence of autism risk genes. F. Hormozdiari1, M. Duyzend1, B. Coe1, J. Huddleston1, D. Hanna1, J. Smith1, P. Sudmant1, D. Nickerson1, E. Eichler1-2. 1) Genome Sciences, University of Washington, Seattle, WA, Select a Country; 2) Howard Hughes Medical Institute, Seattle, WA.

Exome sequencing of thousands of children with autism spectrum disorder (ASD) has led to the discovery of hundreds of candidate genes based on the excess of disruptive mutations in when compared to unaffected siblings. The combination of de novo coding mutations and known pathogenic copy number variations (CNVs), however, still only accounts for a fraction of simplex ASD cases (~35% of the affected probands). In order to assess the importance of non-coding variation, we performed deep genome sequencing (>50 fold sequence coverage) on 16 autism trios where no de novo truncating mutation had been identified based on exome sequencing. We generated a comprehensive assessment of all types of genetic variation (SNVs, indels, CNVs and mobile element insertions) using a combination of tools (GATK, FreeBayes, Pindel, CommonLaw and GenomeStrip). As expected, we observed a linear correlation between paternal age and number of de novo SNVs per proband (r = 0.49). An average of 45 de novo SNVs and 1760 private deletions (>200bp) of which 128 mapped within or near genes. 25% (4/16) of trios showed a pattern of multiple de novo mutations and private deletions in the non-coding region of genes previously implicated in autism and neurodevelopmental disease. One proband, for example, carried four de novo mutations in the intronic regions of ARID1B, ROBO1 (x2) and STXBP4. Another proband inherited a 35 kbp and 150 kbp deletion of MBDS and DISC1, respectively, in addition to a de novo mutation within ARID1B. A third patient had a de novo mutation in the noncoding sequence of a metabotropic glutamate receptor, GRM3, in addition to three inherited deletions ranging in size from 12 to 120 kbp in autism risk genes CACNA2D4, ARID1B and SCN2A. The convergence of multiple rare non-coding mutations of a potentially deleterious nature in probands provides support for an oligogenic model of autism where the disruption of gene regulation in multiple targets contributes to disease etiology.
1230T

1) Department of Medicine, University of Pennsylvania, Philadelphia, PA; 2) Division of Statistical Genetics, University of Michigan, Ann Arbor, MI; 3) HudsonAlpha Institute for Biotechnology, Huntsville, AL; 4) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 4) Department of Psychiatry and the Behavioral Sciences, Zilkha Neurogenetics Institute, University of Southern California, Los Angeles, CA; 5) Medical Research Council Social, Genetic and Developmental Psychiatry Center, Institute for Psychiatry, King’s College London, London, UK; 6) National Institute for Health Research Biomedical Research Centre for Mental Health, The Maudsley Hospital, South London School of Medicine, and the Behavioural Sciences, King’s College London, London, UK; 7) Department of Psychology, University of Minnesota, Minneapolis, MN; 8) Department of Psychiatry, University of Michigan, Ann Arbor, MI; 9) Molecular and Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI; 10) University of Michigan Depression Center, Ann Arbor, MI; 11) Center for Addiction and Mental Health. Department of Psychiatry, University of Toronto, ON, Canada.

Bipolar disorder 1 (BP I) is characterized by alternating periods of mania and depression and affects >1% of individuals worldwide. Bipolar disorder is highly heritable and a substantial fraction of the heritability appears attributable to common variants. Genetic contributions from rare (<5% minor allele frequency (MAF)) and low (5 ≤ MAF < 5%) frequency variants are less well defined.

Exome sequencing studies are exploring the contributions of coding variation, but in non-coding regions, rare and low frequency variants have been less well queried. Our aim is to find rare and low frequency non-coding or coding variants with strong effects on BP I risk. We are performing an exome wide-genome-exome sequencing (covegence of exome废旧), blood-based DNA from >3,770 European ancestry BP I cases and unrelated controls with no history of mental illness (BRIDGES study). Our current analysis contains 2,793 individuals: 1,382 BP I cases and 1,411 controls. We identified 43.4 million SNPs using the GoolCloud pipeline of which 3.7 million are low frequency, 12.2M rare, non-singleton and 21.4M singletons. We have 80% power (additive model) at an alpha of 5 × 10−8 to detect a single variant with 5, 1 or 5% MAF with odds ratios (OR) > 2, 3.6 or 5.3. Likewise, we have 80% power to detect at least one out of 10 variants with MAF of 5, 1 or 0.5% MAF for variants with OR > 1.7, 2.6 or 3.6, respectively. Using logistic analysis of single variants with adjustment for 10 principle components, we did not observe genome-wide significance association for any variant. We also did not observe multiple testing corrected significant gene-level associations of low-frequency (MAF < 1%) protein truncating and/or protein altering variation using SKAT-O or the CMC test, or from burden testing of deleterious coding singleton variants. Our current analyses suggest that, similar to other common, complex diseases, there are no or few moderately rare or low-frequency variants of large effects on risk for BP I. Larger sample sizes and sequencing-based studies will be necessary more fully query the contributions of BP I risk variants from across the allelic spectrum.

1232M

1) Div of Neurol/ Mol Brain Sci, Kobe Univ, Japan; 2) Med Informatics, Cen for Genome Med, RIKEN, Japan; 3) Dept of Med Genom Sci, the Univ of Tokyo, Japan; 4) Kurume Univ, Japan; 5) Dept of Neurol, Juntendo Univ, Japan; 6) Takamatsu Neurol Clinic, Japan; 7) Dept of Neurol. Natl Cen Hsp of Neurol and Psy, Japan; 8) Genotyping Development, Cen for Genome Med, RIKEN, Japan; 9) Dept of Neurol, The Univ of Tokyo, Japan; 10) The Univ of Chicago, USA. The Univ of Tokyo, Japan.

Parkinson’s disease (PD) is a complex disorder caused by multiple genetic variants, and one of the most common neurodegenerative disorders worldwide. We previously reported a Japanese SNP-GWAS which detected 4 PD-risk loci (PARK16, BST1, α-synucleine, and LRRK2) reported by our previous SNP-GWAS. Genetic variants with strong PD-risk do not exist within these 4 PD-loci, indicating that these 4 PD-loci will contribute to this disease as common variants. We will subsequently test association between whole exonic SNVs and PD to identify novel PD-genes harboring rare-variant risks. Moreover, in parallel, to identify further common variant PD-traits, we performed Japa-

1233T
Identifying Genetic Variants Associated with Anorexia Nervosa via Exome Sequencing, Q. Wei1, E. Prueitt2, A. Adam3, R. Cone1, B. Li1.

1) Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN; 2) Department of Neuroscience and Pharmacology, Rudolf Magnus Institute of Neuroscience, University Medical Centre, Utrecht, The Netherlands.

Anorexia Nervosa (AN) is a psychiatric illness characterized by inability to maintain a minimal normal weight and persistent fear of gaining weight. It has the highest rate of mortality among psychiatric disorders and is estimated to affect 0.5%-1% of the population. Family and twin studies indicated that AN is highly heritable, and the estimated genetic contribution ranges between 50% and 75%. However, the genetic etiology of AN is largely unknown. Previous genome-wide association studies assessing common variants failed to reveal genome-level significance. In this study, we carried out exome sequencing to identify rare variants associated with AN. We sequenced 24 AN patients and 47 of their relatives from pedigrees with high load of AN as well as 65 unrelated cases, and obtained 178 controls from National Database for Autism Research. In our analysis we focused on genes that are expressed in the prefrontal cortex of brain and functional variants (nonsynonymous, stop and splicing) that are deleterious predicted by Polyphen-2 and SIFT, LRT, MutationTaster, or CADD. Four genes (XIRP2, EXOG, MUSK, GRM6) harboring rare deleterious variants are enriched in AN patients with p < 0.001 in the single variant test. Furthermore, we carried out gene-based burden tests by collapsing multiple rare variants in a gene, and identified KIAA0317 with p < 0.001. We are exploring different grouping and weighting schemes for gene-based burden tests as well as pathway and network-based analyses to potentially increase power.

1231S
Tri-based Whole Genome Sequence Analysis of a Cousin Pair with Refractory Anorexia Nervosa, P. Shih1, A. Van Zeeland2, A. Bengen3, T. Carland3, V. Bansa4, P. Magistretti4, W. Berrettini5, W. Kaye5, N. Schork1, J. Carlson7,10

1) Dep of Psychiatry, Univ California, San Diego, La Jolla, CA; 2) Cypher Genomics, La Jolla, CA; 3) SRI, Palo Alto, CA; 4) The Scripps Research Institute, La Jolla, CA; 5) J. Craig Venter Institute, La Jolla, CA; 6) Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; 7) University of Pennsylvania, Philadelphia, PA.

Anorexia Nervosa (AN) has an onset during adolescence and is characterized by emaciation, fear of gaining weight despite being underweight, and has the highest mortality rate of all psychiatric illnesses. Despite the serious health and psychosocial consequences of this illness, very few treatments are effective at reversing the core symptoms of AN. AN is highly heritable and show a homogeneous clinical presentation of persistent food refusal and high anxiety traits. However, AN etiology is believed to be heterogeneous as no major susceptibility gene has been consistently replicated in multiple populations. AN symptoms and personality traits tend to be present in unaffected family members of the patients, suggesting that certain shared genetic factors within each family may contribute to unique phenotype risk of the affected. To gain insights into the role “private variants” may play in AN and to test genetic information from family members of AN, here we leveraged a family-based study design combined with whole genome sequencing to search for genetic variants that may influence AN risk in an affected cousin pair together with their parents. By capitalizing on the homogeneous presentation of the two cousins, who both have a diagnosis of refractory AN, we report methods by which we interro-
A Population-based Approach for Detecting Rare Recessive Variation Implicates the Cholesterol Biosynthesis Gene DHCR24 in Autism Spectrum Disorder and Intellectual Disability. E.T. Lim1,2, Y. Chan1,2, S. Goetz1, D. Spatt2, K. Kratz1, M.B. Johnson1,2, M. Chahrour1,2, J.N. Hirschhorn1,2, S. Raychaudhuri2,4, J.M. Silverman5, A. Kolevzon6, J. Buxbaum5, F. Winston5, R.I. Kelley2, M.J. Daly4,5, C.A. Walsh1,2, T.W. Yu1, 1) Boston Children’s Hospital, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) Kennedy Krieger Institute, Baltimore, MD; 4) Brigham and Women’s Hospital, Boston, MA; 5) Mount Sinai School of Medicine, New York, NY; 6) Massachusetts General Hospital, Boston, MA.

We developed and applied a statistical method (RAFT for Recessive Allele Frequency-based Test) for detecting rare recessive variants to a whole-exome genotyping dataset for ~1,000 families and discovered a rare missense variant (allele frequency = 0.5%) in DHCR24, a cholesterol synthesis gene, that is inherited recessively in 3 children from a single European American family affected by autism spectrum disorder (ASD). Subsequently, we discovered another 2 families of Middle Eastern ancestry with 6 children affected by ASD and intellectual disability with 2 other rare recessive missense variants in DHCR24 (RAFT P = 7.64×10−10).

Recessive deleterious mutations in DHCR24 have been associated with the cholesterol biosynthesis disorder desmosterolosis. But other than neurodevelopmental impairment, two of the three families did not exhibit additional distinctive features to suggest that diagnosis. Cholesterol is known to be important for the formation of myelin sheaths and membrane lipid rafts, and is a signaling molecule involved in critical developmental pathways that regulates skeletal and limb development. Recessive mutations in several genes that affect the precursors for cholesterol synthesis have been implicated in various neuro-developmental diseases, such as Smith-Lemli-Opitz syndrome, desmosterolosis and lathosterolosis. DHCR24 is also named Seladin-1 (Selective Alzheimer’s Disease Indicator-1) based on initial discoveries that the mRNA expression of the gene was found to be down-regulated in individuals with Alzheimer’s Disease. To evaluate the cholesterol synthesis functionality of the missense variants identified in ASD patients, we adapted a protocol to express human DHCR24 mutant proteins under the control of a GAL1 promoter in S. cerevisiae (Gilk et al., 2010) and optimized a biochemical assay for measuring desmosterol to cholesterol conversion (Waterham et al., 2001) to evaluate the pathogenicity of the missense variants detected in these families with autism and intellectual disability. Having validated the approach, we are now applying it to the mutations we have observed to test if the missense variants result in reduced cholesterol synthesis.

Mutational and transcriptional analysis in Autism Spectrum Disorders support their oligogenic model disturbing common functional pathways. M. Codina-Solà1,2,3, B. Rodríguez-Santiago1,4, J. Santoyo-López5,6, A. Homs1,2,3, M. Rigau5, G. Aznar Lain5, M. del Campo6, B. Gner6, E. Gabau10, MP. Botella1,2, A. Gutiérrez-Arumí1,2,3, G. Antiholo12, LA. Pérez-Jurado1,3, I. Cusco1,2,6, 1) Universitat Pompeu Fabra, Barcelona, Barcelona, Spain; 2) Hospital del Mar Research Institute (IMIM), Barcelona, Spain; 3) Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Spain; 4) qGenomics, Barcelona, Spain; 5) Andalusian Human Genome Sequencing Centre (CASEGH); 6) Medical Genome Project (MGP) Sevilla, Spain; 7) Pediatric Neurology. Hospital del Mar. Barcelona; 8) Hospital Universitari Vall Hebron, Barcelona, Spain; 9) Hospital Universitario Cruces. Osakidetza, Spain; 10) Hospital Universitari Parc Taulí Sabadell, Barcelona, Spain; 11) Neuropediatrics. HUA (Hospital Universitario de Alava), Alava, Spain; 12) Department of Genetics, Reproduction and Fetal Medicine, Institute of Biomedicine of Seville, University Hospital Virgen del Rocío/CSIC/University of Seville, Seville, Spain.

Background: Autism spectrum disorders (ASD) are a group of neurodevelopmental disorders with high heritability. There is high genetic heterogeneity and recent genetic findings support a double or multiple-hit model. Methods: We have performed Whole Exome Sequencing (WES) in 36 ASD patients to identify ASD causative mutations and rearrangements along with RNAseq to study the transcriptomic consequences. Pathway enrichment studies were performed by ConsensusPathDB. Results: We have detected likely causative mutations with monogenic models in 6 cases; 3 de novo loss-of-function mutations (genes SCN2A, MED13L and KCNV1), one de novo missense (CUL9) and two X-linked inherited mutations (genes MOA, CDKLS). Inherited missense mutations in candidate genes, rarely mutated in controls, including SCN2A and MED13L were also found in the ASD cohort, suggesting the presence of ASD risk variants, contributing in an additive manner to the phenotype. Transcriptomic studies validated 90% of WES variants, identified deregulation (overexpression of SEM6AB, MECP2), Allele Specific Expression (FUS, EP300) and nonsense mediated decay (RTT1, ALG9) as a consequence of rare mutations. Pathway enrichment studies of inherited rare variants reveal the presence of recurrent alterations, including the previously implicated pathways PI3K-Akt signaling and Axon Guidance. Conclusions: WES has proven to be a highly efficient technology to identify the molecular defect in a proportion of probably monogenic ASD forms in our cohort (16.6%), as well as to detect potential risk variants contributing to an ASD disorder in an additive oligogenic manner. Concomitant whole blood transcriptomic data allowed us to evaluate in part the functional consequences of genetic variants. Deregulation of common functional pathways by rare inherited variants such as PI3-Akt signaling pathway and Axon Guidance could guide the identification of pharmacological treatments.
1236T
Whole-exome sequencing of multiplex families identifies several rare coding variants in known and novel Late-Onset Alzheimer genes. B. W. Knukel1, M. A. Kohli2, K. L. Hamilton1, W. R. Perry1, R. M. Carney1, P. L. Whiteside5, J. R. Gilbert1, E. E. Martin1, G. W. Beecham3, J. L. Haines3, M. A. Pericak-Vance1. 1) Hussman Institute for Human Genetics, University of Miami, Miami, FL; 2) Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH.

We performed whole-exome sequencing (WES) on 28 multi-generational, late-onset Alzheimer disease (LOAD) families to identify rare LOAD risk variants. The pedigrees chosen for WES display a dominant inheritance pattern, have an average of eight Alzheimer-affected subjects, are free of known Alzheimer disease (AD) mutations, and do not cluster for the APOE4 allele. 1-4 affected subjects per pedigree were sequenced in 22 families. Results for these families were combined with six previously reported pedigrees (4-9 affected subjects sequenced). Variants were filtered for segregating, non-synonymous or splice-site rare variants (MAF<0.005) assuming autosomal or X-linked dominant models. Filtered variants were examined for implication as LOAD candidate genes by comparing results across families, and by comparison to GWAS-confirmed LOAD susceptibility genes and biologically relevant KEGG AD pathway genes. Six new rare missense variants in the genes PLEKHG5, ADAMTS4, ORL13, ZBTB11, CDDC39, and SEMA4B segregated with disease status in two or more different families. An additional family segregated a separate rare missense variant in PLEKHG5, and another family segregated a different rare missense mutation in ADAMTS4. Three candidate genes (ABC2, CD163L1, DMD) from our previously reported WES received additional support, and a new variant in the previously reported TTC3 gene segregated in two different pedigrees. Several of these genes have been functionally implicated in dementia-related processes and diseases (i.e. SEMA4B regulates axonal extension and synapse development; ABCG2 is involved in amyloid-β clearance across the blood brain-barrier; TTC3 is a Down syndrome critical region protein crucial for neuron survival), supporting their potential role in LOAD pathogenesis. In addition to these candidates, filtered variants also segregated with disease in the two LOAD GWAS genes (SORL1 and EPHA7), and several KEGG AD pathway genes (LPL, CACNA1D, RYR3, GRIN2A and NCSTN), making them important. This work confirms the hypothesis that rare missense variants, found in LOAD-affected individuals, are enriched in the LOAD risk loci identified by GWAS. We’ll discuss the results of our search for LOAD candidate genes, and conclude that our analyses suggest that the genetic contribution of non-coding functional DNA elements to autism.

1237S
Targeted resequencing of non-coding functional DNA elements in autism. D. Mahotra1, A. Watts2, T. Chapman3, N. Pongthongkum4, A. Gore5, A. Fung6, K. Zhang7, J. Sebat8. 1) Department of Pediatrics, University of California San Diego, La Jolla, CA, USA; 2) Department of Bioengineering, University of California San Diego, La Jolla, CA; 3) Institute of Genomic Medicine, University of California San Diego, La Jolla, CA; 4) Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA; 5) Byester Center for Genomics of Neuropsychiatric Diseases, University of California San Diego, La Jolla, CA; 6) Neuroscience Discovery and Translational Area, Pharma Research & Early Development (pRED), F. Hoffmann-La Roche Ltd, Basel, Switzerland.

While the role of rare de novo loss of function single nucleotide variants (SNVs) and copy number variants (CNVs) in autism is now firmly established, the genetic contribution of non-coding functional DNA elements to autism risk is unknown. We developed an inexpensive and high-throughput padlock probe based assay to resequence 5 megabase (Mb) of non-coding functional genome in 500 autism quad families from Simons Simplex Collection (SSC). We designed 77,277 padlock probes to capture the target region which mainly includes a) conserved mutation hotspots (CMH) in the human genome and b) putative enhancer and promoter elements of 95 high-risk autism genes. We performed a series of optimization experiments to develop an assay that a) captured 99.7% of target bases, b) had high sensitivity and specificity to detect SNVs confirmed empirically by testing the assay in two hapmap trios that were whole-genome sequenced at deep coverage by the 1000 genomes project and c) had greater than 40X coverage per base per sample in a 384 sample multiplexed Illumina HiSeq sequencing run. Our experimental approach and study design will provide insights into the role of rare variation in the non-coding functional genome in the risk for autism.

1238M
Identification of Rare Variants for Bipolar Disorder by Exome Sequencing in Multiplex Families. S. Ramdas1, J. Li2, AB. Ozel3. 1) Bioinformatics, University of Michigan, Ann Arbor, MI; 2) Department of Human Genetics, University of Michigan, Ann Arbor, MI.

Bipolar Disorder is a heritable psychiatric disease characterized by alternating manic and depressive episodes. The phenotypic and genetic complexity in this disorder has made the identification of causal genes or biological pathways a considerable challenge. In this study, we focus on 34 multi-generation multiplex BP families from the NIMH repository. Owing to the high occurrence of BP in these families, we hypothesize that these families represent a nearly Mendelian subset of bipolar cases and transmit one or a few high-impact coding variants that result in the disorder. We genotyped and performed exome sequencing for 90 affected individuals in 34 families representing first cousin pairs or more distant relatives, aiming to detect functionally damaging rare variants in regions shared by affected relatives. Variant filtering were applied to identify those that are bi-allelic, have a minor allele frequency of less than 1% in European samples from the Exome Sequencing Project and 1000 Genomes, and are damaging missense, nonsense, or splicing site variants, resulting in a median of 508 variants per family. As exome chip genotype data allow inference of genomic segments shared among family members identically by descent (IBD), we used Beagle to identify IBD regions shared by BP3 cases within each family. This led to a further reduction in the number of candidate variants, with a median of 147 variants. We found that HLA genes HLA-DRB1 and HLA-DRB5 showed damaging mutations in 8 and 7 families respectively. These genes have been previously reported to be associated with schizophrenia. Also found in our final list of genes are glutamate receptors GRK3 and GRIN3B, and calcium channel receptor CACNA1C, all of which are known risk loci. Validation and segregation analyses using 254 additional members of the same families are underway to further prioritize the candidate genes and pathways (This study is supported by the IMHRO - Johnson & Johnson Rising Star Translational Research Award.).

1239T
A Cohort for Researching Autism Genetics in New Zealand. B. Swan1, J.C. Jacobsen1, R. Hill2, B. Tsang3, M. Taylor4, D. Lover5, J. Taylor6, R.G. Snell7, K. Lehner8. 1) Centre for Brain Research and School of Biological Sciences, The University of Auckland, Auckland, New Zealand; 2) Department of Neurology, Auckland City Hospital, Auckland, New Zealand; 3) Paediatrics & Newborn Services, Waiwamata District Health Board, Auckland, New Zealand; 4) Diagnostic Genetics, LabPLUS, Auckland City Hospital, Auckland, New Zealand; 5) Genetic Health Service New Zealand, Auckland City Hospital, Auckland, New Zealand.

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that presents with impaired social communications and restricted behavioural traits. It has a strong genetic aetiology and is observed in communities throughout the world. We have recently established a study investigating the genetic contributions to ASD in New Zealand, and to this end have developed a unique research cohort consisting of diverse populations living in New Zealand.

Recruitment of subjects was performed through collaboration with clinicians and an online project registry (www.mindsforminds.org.nz). Since August 2013, 899 participants have either registered themselves or been registered by their care-giver (232 self-registered, 667 registered by carer). Upon registration, participants can elect to enter information detailing age, sex, ethnicity, diagnosis, clinical specialists, prior DNA testing, and other relevant diagnoses for the participant or their family. Within this database we currently have an ASD research cohort that includes 684 diagnosed with ASD or Asperger’s (Asp) (492 and 192, respectively). The male to female ratio of participants diagnosed with ASD is 5:1 (2.3:1 for Asp). Self-reported ethnicity background reflects New Zealand’s population: 757 European (84.2%), 122 Maori (13.6%), 25 Polynesian (2.8%), and 35 Asian (3.9%) (individuals can register as more than one ethnicity). The 2006 census national population frequencies of these ethnicities in New Zealand is Euro- pean 76.8%, Maori 14.9%, Polynesian 7.2%, and Asian 9.7%. The ASD/ Asp cohort includes 119 multiplex families, 320 individuals with co-occurring neurological disorders [ASD:196; Asp:124], and 78 individuals with co-occurring gastrointestinal disorders [ASD:52; Asp:26]. This database is subse- quently being used to identify participants for exome and genome sequenc- ing. We are currently coordinating exome sequencing on our first group of subjects with the goal of eventually sequenc- ing the entire cohort.
1240S
Mutations in adaptor protein AP-5 subunits lead to peripheral neuropathy, spastic paraplegia and parkinsonism with aberrant endolysosomes. M. Madoé1, J. Edgar2, F. Darios3, T.N. Jepperson4, J. Liles3, C. Blackstone5, M.S. Robinson2, J. Hirst6, M.C. Krue1. 1) Sanford Children’s; 2) Cambridge University; 3) INSERM; 4) Vanderbilt; 5) NIH NINDS.

AP-5 is one of five adaptor protein complexes that play important roles in intracellular trafficking. We identified 7 patients with mutations in AP-5 subunits whose clinical phenotype ranged from isolated peripheral neuropathy to spastic paraplegia with parkinsonism. Primary cultured fibroblasts from affected patients showed accumulation of enlarged endolysosomes. Patient cells also displayed autofluorescence and accumulation of lysosomal storage material by electron microscopy. These findings indicate that AP-5 mutations may lead to diverse neurological phenotypes and further support links between endosomal-lysosomal dysfunction and hereditary spastic paraplegia and parkinsonism.

1241M
Targeted sequencing of African American autism spectrum disorder patients reveals loss of function variants in novel autism genes. P. Whitehead1, A.J. Griswold2, D. Van Booven3, N. Duerk3, J.A. Rantus1, J.M. Jaworski1, S.H. Stiller1, M.A. Schmidt1,2, W. Hults1, J. Konidari1, M.L. Cuccaro1,2, E.R. Martin1,2, J.L. Haines3, J.R. Gilbert1,2, J.P. Hussman1,2, M.A. Pericak-Vance3. 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL; 3) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 4) Hussman Institute for Autism, Baltimore, MD, USA.

Whole exome sequencing studies in autism spectrum disorder (ASD) have identified no single factor accounting for the genetic risk, though several candidate genes and molecular pathways have been implicated. These studies have been performed primarily in individuals of white European ancestry. However, the incidence of ASD is similar across ethnicities (1 in 68 children in whites, 1 in 81 in blacks) begging the question whether the risk genes identified are generalizable across races and ethnicities. We are currently investigating rare ASD risk variants by integrating GWAS with targeted massively parallel sequencing of candidate regions from GWAS Noise Reduction analyses of European ASD datasets (Hussman et al., 2011). We targeted 689 GWAS-NR associated genes and evolutionarily conserved intronic and intergenic regions. We sequenced 140 ASD cases and 180 controls of African ancestry determined by principal component analysis using Eigenstrat. We identified 10,859 coding single nucleotide variants (SNVs) in exons of targeted genes. Among these exonic mutations are 5,544 nonsynonymous variants and 93 LOF alterations (46 stop gains, 6 stop losses, and 41 splice sites). LOF variation has been previously implicated as ASD risk factors in whites so we examined the genes affected by such alterations in our black cohort. There was no overall enrichment for LOF mutations in black ASD cases (p = 0.553) and only a single previously identified ASD candidate gene was affected by a LOF (MYO16). A total of 29 other genes had at least one LOF in a black ASD case and not in a control. These include novel candidate genes within pathways previously implicated in ASD. A stop mutation was found in a single case in KCNIP1 (voltage-gated potassium channel interacting protein 1) which modulates GABAergic signaling in mouse models. A stop in STRAP3 (SLIT-ROBO Rho GTPase activating protein 3) was identified in another case and the Rho-GTP signaling gene has been implicated by copy number screens in ASD cases, but in implicated ASD pathways. This suggests that novel genes associated genes reveals loss of function variants in novel genes in black ASD cases, but in implicated ASD pathways. This suggests that novel genes in similar underlying pathways might contribute to the genetic risk of ASD in other ethnicities.

1242T
Exome sequencing of 43 sporadic cases with an autism spectrum disorder in a local cohort of families identifies severe de novo variants and implicates additional genes in ASD pathogenesis. W. Banks1, D. Cunningham1, E. Hansen1, K. Ratliff-Schaub2, E. Butter3, D. Schulteis4, C. Boreman1, B. Kelly2, P. White5, G. Herman1,2. 1) The Research Institute at Nationwide Children’s Hospital, Columbus, OH; 2) Center for Microbial Pathogenesis, The Research Institute at Nationwide Children’s Hospital, Columbus, OH; 3) Department of Pediatrics, Nationwide Children’s Hospital, Columbus, OH; 4) Department of Pediatrics, Wright-Patterson Air Force Base, OH.

Autism spectrum disorders (ASD) are a common group of complex neurodevelopmental disorders, with over 500 CNVs and genes implicated in ASD pathogenesis to date. To identify additional loci involved in ASD, we undertook whole exome sequencing (WES) of 43 simplex trios from the Central Ohio Registry for Autism (CORPA), a local cohort of over 200 families. Psychometric and implicates additional genes in ASD pathogenesis. These include novel candidate genes within pathways previously implicated. These candidate genes and molecular pathways have been implicated. These findings indicate that AP-5 mutations may lead to diverse neurological phenotypes and further support links between endosomal-lysosomal dysfunction and hereditary spastic paraplegia and parkinsonism.

1243S
De novo and rare inherited mutations implicate the transcriptional coregulator TCF20/SPBPB in autism spectrum disorder. A.O.M. Wilkie1,2, J.C. Inns2, E.R. Martin2, J.A. Griswold1,3, S.H. Slifer1,2, M.C. Kruer2,3, F. Darios1,3, M.S. Robinson1,2, S.H. Slifer1,2, M.C. Kruer2,3, F. Darios1,3, M.S. Robinson1,2, S.H. Slifer1,2, M.C. Kruer2,3, F. Darios1,3, M.S. Robinson1,2, S.H. Slifer1,2, M.C. Kruer2,3, F. Darios1,3.

Autism spectrum disorders (ASDs) are common and have a strong genetic basis, yet the cause of ~70-80% ASDs remains unknown. By clinical cytogenetic testing we identified a family in which two brothers had ASD, mild intellectual disability and a chromosome 22 pericentric inversion, not detected in either parent, indicating de novo mutation with parental germlinal mosaicism. We hypothesised that the rearrangement was causal of their ASD and localised the chromosome 22 breakpoints. Methods: The rearrangement was characterised using fluorescence in situ hybridisation (FISH) and FISH in situ hybridisation, inverse PCR and 454 sequencing. Open reading frames and intron/exon boundaries of the two physically disrupted genes identified, TCF20 and TNR6B, were sequenced in 342 families (260 multiplex and 82 simplex) ascertained by the International Molecular Genetic Study of Autism Consortium (IMGSAC). Results: IMGSAC family screening identified a de novo missense mutation of TCF20 in a single case and significant association of a different missense mutation of TCF20 with ASD in 3 further families. Through exome sequencing in another project we independently identified a de novo frameshifting mutation of TCF20 in a white family with ASD and mild intellectual disability identifying a significant association of TNR6B mutations with ASD. Conclusions: TCF20 encodes a transcriptional coregulator (also termed SPBP) that is structurally and functionally related to a single paralogous gene, RAF1. Notably TCF20 encodes the coiled-coil forming domain found in the behavioural phenotypes of the Smith-Magenis and Potocki-Lupski 17p11.2 deletion/duplication syndromes, in which ASD is frequently diagnosed. Our study provides the first evidence that mutations in TCF20 are also associated with ASD and should encourage further investigation for the functions of TCF20/SPBP in neurodevelopment and the role of mutations in ASD.
1244M

MEF2C haploinsufficiency is a frequent finding in patients with autism spectrum disorders. A. Ziegler1,2, R. Delorme3-5, P. Bonneau1,6, A. Guilmotet3,7, T. Bourgeron3,5,6, D. Bonneau1,6, 1 Department of Biochemistry and Genetics, CHU Angers, Angers, France; 2 Human Genetics and Cognitive Functions, Institut Pasteur, Paris, France; 3 CNRS URA 2182 "Genes, synapses and cognition," Institut Pasteur, Paris, France; 4 Assistance Publique-Hôpitaux de Paris, Robert Debré Hospital, Department of Child and Adolescent Psychiatry, Paris, France; 5 University Denis Diderot Paris 7, Paris, France; 6 UMR CNRS 6214 INSERM 1083, Angers, France.

Since the first description of MEF2C haploinsufficiency syndrome in 2009, 42 patients have been reported either with deletion or point mutation of this gene. Intellectual disability (ID) is the core disorder of this syndrome but autistic features such as stereotypic movements and lack of social communications are commonly reported. To further assess the role of MEF2C in autism spectrum disorder (ASD), we looked for MEF2C point mutations by Sanger sequencing and for MEF2C copy number variations by SNP-array in a cohort of 195 patients with ASD and a mild to severe ID. We identified a de novo frameshift mutation leading to a premature stop codon in 2 siblings with ASD and severe intellectual disability. A maternal germinal mosaicism was confirmed based on the haplotype. A de novo MEF2C deletion was also found in a patient with ASD and a mild intellectual disability. In this study, we observed MEF2C haploinsufficiency in 1.5% (95% confidence interval: 0-3.2%) of cases with ASD. This is the first study to look specifically at MEF2C in ASD. Of note, a mutation in MEF2C was also found once by whole-exome sequencing in another cohort of 175 trios with ASD. According to this notable MEF2C haploinsufficiency frequency in ASD, we strongly recommend to pay specific attention to this gene when performing non-targeted genetic screening in ASD with a comorbid ID even for patients without severe ID.

1245T

Assessing the role of methylation in autism brains. S.E. Ellis1, S. Gupta1, A. Moses2, A.B. West3, D.E. King1, 1) Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD; 2) Department of Neurology, University of Alabama School of Medicine, Birmingham, AL.

The high prevalence of autism spectrum disorder (ASD) and the limited understanding of relevant risk factors have sparked an intense focus on elucidating the genetic basis of ASD. GWAS, exome-sequencing, and RNA-Seq have all helped the field move forward, but the vast majority of the risk remains unexplained. Given that prenatal brain development is both critical to the progression of ASD and a period in which development is sensitive to alterations in epigenetic pathways, many groups have turned to investigating the role of methylation in ASD. To date, microarray studies have looked for methylation differences between ASD brains and controls; however, these have been limited by the digital nature of methylation arrays and small sample sizes. We report results from the largest study to date, relying on next-generation methylation sequencing (reduced representation bisulfite sequencing) in 63 cortical brain samples (BA19), comprising 34 controls and 29 individuals with autism, a subset of 63, all of whom were ascertained using ADOS and ADI-R (among other exclusions). After extensive quality control, we performed a single site analysis across 739,980 CpGs looking for methylation differences between cases and controls. This analysis identified 28 differentially methylated CpGs (q<0.05), eight of which demonstrate a mean methylation difference greater than 10 percent (5 hyper- and 3 hypo-methylated in ASD). While none of these 28 sites fall in regions demonstrating differential gene expression (as measured by RNA-Seq), six are exonic, nine fall within promoters, 16 are within CpG islands (CGI), and five are within CGI shores, suggesting a potential functional role for these individual sites. In particular, one exonic CpG site in SHANK3 demonstrates a 6.4% decrease in methylation in ASD cases. While the need to validate this finding by an independent method remains, this finding is of particular interest for future investigation as SHANK3 has frequently been reported to have a role in ASD. Finally, ongoing studies will leverage these data, along with RNA-seq carried out on these same individuals, to gain a more complete understanding of the relationship between methylation and gene expression in the prefrontal cortex with risk for ASD.

1246S


We analyzed data from exome sequencing in 9 probands with autism. In each of the autistic probands our analyses revealed the presence of rare and low frequency potentially functionally damaging variants in genes that encode enzymes involved in synthesis of methyl residues and/or genes that encode enzymes involved in the transfer of methyl residues to or from chromatin, (histones and DNA). A rare allele in lysine specific methyltransferase KMT2C variant rs10453200 occurred in 3 of the 9 autistic probands; a rare allele at KMT2C variant rs78352660 occurred in 2 autistic probands. A rare allele KMT2C variant at rs200804156 occurred in 1 autistic proband. It is interesting to note that one proband had 3 rare variants in KMT2C and one rare variant in KMT2D lysine specific methyltransferase rs201114196. Low frequency potentially damaging variants in the MTRR gene 5 methyltetrahydrofolate homocysteine methyltransferase reduclease rs2287780 occurred in 2 autistic probands. An autistic twin pair had a rare potentially damaging variant in MTHFS S-10 methylenetetrahydrofolate synthetase. The variants described were potentially damaging and were inherited from one or other of the parents where testing of family was possible (5 out of 7 families) and were present in heterozygous state in the probands. We identified rare potentially damaging variants in genes that encode proteins involved in chromatin modification. Rare potentially damaging variants, each identified in 1 proband occurred in KDM5A lysine specific demethylase at rs200804533, in KAT2B lysine acetyl transferase at rs148960024, in CHD8 chromatin helicase at rs200465274 in CHD6 chromatin helicase at rs61752057. Studies by Melnyk, James and coworkers have implicated methylation dysregulation in autism PMID: 21519954. There is increasing evidence that abnormalities of chromatin modification are frequent cause of cognitive disorders in children, (Ronan and Crabtree PMID: 23568486). It is important to consider the possibility that rare or unusual variants and pathogenic copy number variants in a small number of genes may play roles in the etiology of autism and that de novo and inherited variants may have impact.

1247M

Mutation Screening in Saudi Parkinson’s Disease Patients. B.R. Al-Mubarak1, A. Magrashi1, M. AlTurki2, S. Boholega2, T. Akhairallah2, B. AlTawfi1, H. Abou Al-Shaar1, L. Al-Joma1a, N. Al Tassan1, 1) Genetics, King Faisal Specialist Hospital, Riyadh, Saudi Arabia; 2) Neurosciences, King Faisal Specialist Hospital, Riyadh, Saudi Arabia.

Parkinson’s disease (PD) is classified as the second most prevalent aging-related neurodegenerative disease after Alzheimer’s, and is predicted to affect approximately 8-9 million individuals by 2030. The disease is characterized by the occurrence of four cardinal symptoms: bradykinesia, resting tremor, rigidity and postural imbalance. In most cases the disease occurs sporadically with complex etiopathology involving a combination of genetic propensity and environmental factors. However, familial cases also exist accounting for 15-20% of the patients. Mutations in PARKIN, SNCA, DJ-1, PINK1, and LARRK2 genes have been reported in patients with different modes of inheritance of PD. In this study we sought to determine the genetic basis underlying PD in 97 Saudi individuals affected with either sporadic or familial form of the disease through screening the entire ORF region including SNCA, PARKIN, LARRK2, DJ-1, and PINK1 genes. We used direct sequencing to identify disease-related variants. Direct sequencing of PARKIN, SNCA, DJ-1, PINK1 and LARRK2 genes in these samples failed to reveal pathogenic mutations with the exception of one familial case homozygous for c.938 C>T (p.Thr313Met) mutation in PINK1. Moreover, the mutation analysis was extended to include more recently identified PD susceptibility genes particularly, FBXO7 , GIGYF2 , VPS35 , and UCHL1 . Out of the 121 detected SNVs, only few were potential disease-causing mutations, while the majority was either present in ethnic match controls or did not segregate with the mode of inheritance. The absence of pathogenic mutations in the screened genes in our cohort doesn’t rule out the possibility of their involvement in the development of the disease, as gene expression maybe perturbed. However, it also suggests that other genetic factors may be implicated in these patients. These findings are in keeping with the long-held notion of the complexity and genetic heterogeneity of PD.
124BT
Genetics of dementias in a Turkish cohort. R. Guerreiro1,2, G. Guven2, J. Bras1, L. Darwent1, J.R. Gibbs1, N. Unaltuna2, H. Gurvit2, B. Bilic2, H. Hanagaki3, M. Emre3, A. Singleton4, J. Hardy1, E. Lohmann5,6. 1) Department of Molecular Neuroscience, UCL Institute of Neurology, University College London, Queen Square, London, UK; 2) Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA; 3) Department of Genomics, Institute for Experimental Medicine, Istanbul University, Istanbul, TR; 4) Behavioral Neurology and Movement Disorders Unit, Department of Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, TR; 5) Department Of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, GERM.

Turkey has a population of over 77 million, with a growth rate of 1.5% per year and an average life expectancy of 73.2 years. Large family units, high birth rates, and high rates of consanguineous marriages can often be found in this population, making genetic diseases a common health problem. In parallel, the improvement of living standards and health services will increase survival rates and, consequently, the incidence of Alzheimer’s disease (AD) [MIM 104300] and other dementias. The current study is the first large molecular study of dementias in Turkey. Our main goals were: 1) to establish the presence of mutations in the most common genes known to cause dementias (APP, PSEN1, PSEN2, MAPT, GRN and C9orf72); 2) to characterize the mutation spectrum in these genes by describing genotype-phenotype correlations; and 3) to study recessive forms of dementia; 4) to identify novel genes/variants involved in different forms of dementia. We have established a cohort of 150 families presenting with different forms of dementias in which we have been studying by using genome-wide genotyping techniques and next generation sequencing combined with Sanger sequencing. We established a frequency of 11.2% of mutations and variants in APP, PSEN1 and PSEN2, and of 5.2% of pathogenic mutations in MAPT, GRN and C9orf72, suggesting that mutations in these genes are not uncommon in Turkey. The application of exome sequencing to the cases without mutations in the known dementia genes revealed some unexpected results with the identification of a previously reported NOTCH3 mutation in an AD family and three homozygous TREM2 mutations in Frontotemporal dementia-like 1,2. The NOTCH3 gene is located on 22q13.3 region and is predominantly expressed in the cerebral cortex and cerebellum. The exon 2 of NOTCH3 gene has been studied in a Brazilian individual with idiopathic ASD and this exome sequencing is a valid, rapid and cost-effective tool to identify genetic mutations and further investigations should be carried out to clarify the presence of mutations in this exon to investigate a possible association with the behavioral phenotype. We sequenced 190 Brazilian individuals and found a heterozygous alteration of A of the position 22: 50675139 resulting in a substitution of the amino acid threonine by alanine at codon 42. The mutation was found in a 21 year old male patient with diagnosis of autism without phenotypic changes. Karyotypic investigation and molecular testing for Fragile-X syndrome were normal. We searched for this mutation in the 1000 Genomes-Deep Catalog of Human Genetic Variation and in “NHBLI Exome Sequencing Project (ESP) Exome Variant Server” and nothing was found. Genetic studies have demonstrated the strong association of mutations in SHANK3 with susceptibility to psychiatric disorders, particularly autism, however, little is known about the neural consequences of these mutations and further investigations should be carried out to clarify the causes of this disease.

1249S
Mutation screening in exon 2 of synaptic gene SHANK3 in Brazilian individuals with Autism Spectrum Disorder. D.B. Rosan1, P.P. Nascimento2, A.C. Fett-Conte1, S. Divyakolu3, V.R. Om Sai4, V. Nagarath5, Q. Hasan1,4, Y.R. Ahuja1, 1) Department of Genetics and Molecular Medicine, Vasavi Medical and Research Centre, Hyderabad, Andhra Pradesh, India; 2) Department of Neurology, Apollo Hospitals, Jubilee Hills, Hyderabad, Andhra Pradesh, India; 3) Department of Medical Sciences, National Institute of Mentally Handicapped, Bowenpally, Secunderabad, Andhra Pradesh, India; 4) Department of Genetics and Molecular Medicine, Kamineni Hospital, LB Nagar, Hyderabad, Andhra Pradesh, India.

Background: Non-synonymous mutations/ polymorphism in amyloid precursor protein (APP) gene causes overproduction of Aβ proteins or affects its split into Aβ40 and Aβ42 peptides. Aβ42 has been considered to be a toxic peptide playing a major role in the pathogenesis of Alzheimers (AD). Similar APP plaques were observed in the brains of Down syndrome (DS) patients and high level of plasma APP was observed in patients with severe Autism spectrum Disorder (ASD). The aim of this study was to evaluate exon 16 and 17, the hotspot regions of APP gene in patients with neurobehavioral disorders like AD, DS and ASD.

Methods: A total of 75 cases were recruited in the study which included AD (n=25), DS (n=25), and ASD (n=25). Polymerase chain reaction (PCR) analysis and sequencing was carried out using exon-intron encompassing primers for the selected APP gene regions. In-silico analysis was also carried out to identify the impact of sequence variants on the protein structure.

Results: Three exonic variants, two in exon 16: V683F, H684Y and one in exon 17, H733Q were identified in sporadic AD cases. Apart from these, two in-silico analyses were also observed. In-silico analysis showed that H733Q mutation may affect the structure and function of APP, whereas H684Y mutation is neutral. In an ASD case, our analysis showed an intronic variation i.e., an A insertion at c.1964-13_1964-12insA. In-silico analysis predicted that this variation affects the elongation features of the transcript. None of the DS cases had any variation in this hotspot region.

Conclusion: Our data indicate that variations in the selected hotspot region of APP may play an important role in the aetiology of neurobehavioural disorders.

1250M
Variations in hotspot region of β-amyloid precursor protein (APP) gene in various neurodegenerological disorders from Hyderabad, a major urban city of South India. W. Thomas1, S. Divyakolu1, V.R. Sreekanth1, V. Om Sai2, V. Nagarathana3, Q. Hasan1,4, Y.R. Ahuja1, 1) Department of Genetics and Molecular Medicine, Vasavi Medical and Research Centre, Hyderabad, Andhra Pradesh, India; 2) Department of Neurology, Apollo Hospitals, Jubilee Hills, Hyderabad, Andhra Pradesh, India; 3) Department of Medical Sciences, National Institute of Mentally Handicapped, Bowenpally, Secunderabad, Andhra Pradesh, India; 4) Department of Genetics and Molecular Medicine, Kamineni Hospital, LB Nagar, Hyderabad, Andhra Pradesh, India.

Background: Amyloid precursor protein (APP) gene causes overproduction of Aβ proteins or affects its split into Aβ40 and Aβ42 peptides. Aβ42 has been considered to be a toxic peptide playing a major role in the pathogenesis of Alzheimer’s (AD). Similar APP plaques were observed in the brains of Down syndrome (DS) patients and high level of plasma APP was observed in patients with severe Autism spectrum Disorder (ASD). The aim of this study was to evaluate exon 16 and 17, the hotspot regions of APP gene in patients with neurobehavioral disorders like AD, DS and ASD.

Methods: A total of 75 cases were recruited in the study which included AD (n=25), DS (n=25), and ASD (n=25). Polymerase chain reaction (PCR) analysis and sequencing was carried out using exon-intron encompassing primers for the selected APP gene regions. In-silico analysis was also carried out to identify the impact of sequence variants on the protein structure.

Results: Three exonic variants, two in exon 16: V683F, H684Y and one in exon 17, H733Q were identified in sporadic AD cases. Apart from these, two in-silico analyses were also observed. In-silico analysis showed that H733Q mutation may affect the structure and function of APP, whereas H684Y mutation is neutral. In an ASD case, our analysis showed an intronic variation i.e., an A insertion at c.1964-13_1964-12insA. In-silico analysis predicted that this variation affects the elongation features of the transcript. None of the DS cases had any variation in this hotspot region.

Conclusion: Our data indicate that variations in the selected hotspot region of APP may play an important role in the aetiology of neurobehavioural disorders.
DYT16 revisited: exome sequencing identifies PRKRA mutations in a European dystonia family. M. Zach1, T. Castrop1, B. Schormair1,2, A. Jochim1, T. Wieland2, N. Gross1, P. Lichten1, A. Peters3, C. Gieger1, T. Strom1, 2, T. Strom1, 2, K. Oexle1, B. Haslinger1, J. Winkelmann1, 2, 3, 1) Technical University of Munich, Munich, Germany, Munich, Germany; 2) Helmholtz Zentrum München, Munich, Germany; 3) Stanford University School of Medicine, Palo Alto, CA, USA.

Objective: Recessive dystonia involves mutations in PRKRA has until now been reported only in seven Brazilian patients. The aim of this study was to elucidate the genetic cause underlying dystonia in a Polish family with autosomal-recessive, early-onset generalized dystonia and slight parkinsonism, and to explore further the role of PRKRA in a dystonia series of European ancestry. Methods: We employed whole-exome sequencing in two affected siblings of the Polish family and filtered for rare homozygous and compound heterozygous variants shared by both exomes. Validation of the identified variants as well as homozygosity screening and copy number variation analysis was carried out in the two affected individuals and their healthy siblings. PRKRA was analyzed in 339 German patients with various forms of dystonia and 376 population-based controls by direct sequencing or high-resolution melting. Results: The previously described homozygous p.Pro222Leu mutation in PRKRA was found to segregate with the disease in the studied family, contained in a 1.2 Mb homozygous region identical by state with all Brazilian patients in chromosome 2q31.2. The clinical presentation with young-onset, progressive generalization of dystonia and mild parkinsonism resembled the phenotype of the original DYT16 cases. PRKRA mutational screening in additional dystonia samples revealed three novel heterozygous changes (p.Thr34Ser, p.Asn102Ser, c.-14A>G), each in a single subject with a dystonia phenotype consistent with early-onset dystonia. Our study provides the first independent replication of the DYT16 locus at 2q31.2 and strongly confirms the causal contribution of the PRKRA gene to DYT16. Our data suggest worldwide involvement of PRKRA in dystonia.

1252S

Rapid multiplex sequencing of genes associated with progressive neurodegenerative disorders. M.O. Dorschner1, 2, 3, M.A. Weaver2, 3, 4, G.L. Carvill4, H.C. Melford5, I.P. Mata6, 7, C.P. Zabetian1, 2, 3, M. Rumbaugh1, 2, T.D. Bird3, 2, D.W. Tsuang3, 2, 1) Department of Pathology, University of Washington, Seattle, WA; 2) Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, WA; 3) Geriatric Research, Education, and Clinical Center, VA Puget Sound Health Care System, Seattle; 4) Department of Neurology, University of Washington, Seattle, WA; 5) Department of Pediatrics, University of Washington, Seattle, WA.

Alzheimer’s disease (AD) is an insidious neurodegenerative disorder that accounts for significant morbidity and mortality. While there are multiple causes of dementia, including dementia with Lewy bodies (DLB), frontotemporal dementia (FTD) and Parkinson’s disease dementia (PDD), greater than half of cases are due to Alzheimer’s disease. Dementia manifests clinically with subtle and poorly recognized cognitive deficits and slowly becomes more severe and eventually incapacitating. An accurate diagnosis early in the disease process would greatly improve the ability of physicians to predict disease progression and manifestation of specific symptoms. We developed a multigene sequencing panel for several purposes: 1) to establish an algorithm for clinical assessment of dementia, 2) for discovery-based efforts prior to exome or genome sequencing, and 3) to identify rare disease-causing variants and study the type and distribution of variants with respect to phenotypic characteristics. The coding regions of these genes were captured with gap-fill molecular inversion probes (MIPs) and sequenced using Illumina technology. Sequence data were aligned with BWA and variants called with GATK. Several genes proved difficult to capture due to GC content and sequence context. To improve the capture of these genes, we made several modifications to the protocol and separated the poorly performing MIPs from those that were working efficiently. Two captures were performed for each sample and combined after amplification. With 928 MIPs we were able to obtain adequate coverage for > 95% of the target when sequenced in pools of 192 subjects. To test the ability of our panel to detect disease-causing variants we assayed a set of positive controls with 11 known mutations. 9 of the 11 known mutations were detected. The two failures were located in the PRNP gene, in a difficult to capture segment. The assay was tested with 53 individuals clinically diagnosed with a variety of neurodegenerative disorders. MRI showed clinical atrophy but the etiology was unknown. Six individuals carried predicted pathogenic, likely pathogenic, or risk variants. The combination of molecular inversion probe enrichment and next generation sequencing provides a rapid, low cost procedure for screening large numbers of individuals. We are currently expanding and deploying this assay to screen additional populations with dementia and Parkinson’s disease.
1255S
Rare disease allele penetrance and loss-of-function tolerance in a dominant disease gene: analysis of variation in >60,000 exomes. E.V. Mini-kev1,2,3, S.M. Vallabh1,2, M. Lek1,2, D.G. MacArthur1,2, Exome Aggregation Consortium (ExAC). 1) Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 2) Prion Alliance, Boston, MA; 3) Broad Institute, Cambridge, MA.

Sequencing of control individuals at an ever-larger scale has long promised to shed light on disease biology, but sample sizes are only now approaching the numbers required to afford insights into ultra-rare diseases and into genes with short coding sequences, where fewer variants are expected a priori. Here we demonstrate the power of massive exome data sets to address two critical questions about the biology of a dominant disease gene: the penetrance of previously reported disease mutations, and the discovery of loss-of-function (LoF) variants that demonstrate the tolerance of this gene to heterozygous inactivation. As part of the Exome Aggregation Consortium (ExAC) we aggregated exome sequence data from >60,000 individuals from >60 studies and performed joint variant calling of SNPs and indels across all samples. To demonstrate this dataset’s power to illuminate rare disease biology, we analyzed rare (<0.1%) variation in the prion protein gene (PRNP). Prion diseases affect 1 person per 1 million population per year, and the PRNP coding sequence is short (253 codons). As no samples were ascertained on neurodegenerative disease phenotypes, they should offer an unbiased view of rare PRNP variation. We therefore used this dataset to answer two longstanding questions in prion disease biology: (1) are pathogenic point mutations in PRNP fully penetrant, and (2) is prion protein LoF tolerated in humans? We find that reported disease-associated missense mutations in PRNP are found in our dataset at >20 times the frequency expected based on genetic prion disease incidence, suggesting that a subset of mutations are either benign or incompletely penetrant. The pathogenic variants most common in our controls are those least frequently reported in prion disease patients, and vice versa, suggesting that these variants occupy a spectrum from mild risk factor to fully penetrant Mendelian mutation. Prion protein (PrP) knockout confers total resistance to prion disease and is associated with relatively mild phenotypes in mice, cows and goats. Reducing PrP expression is therefore considered to be a therapeutic strategy in prion disease, yet no LoF allele has ever been observed in humans. We find three heterozygous LoF alleles in our dataset, consistent with expectation, and phenotypic follow-up suggests that these individuals are healthy. This supports reduction of PrP expression as a viable therapeutic strategy in this class of diseases.

1256M
A novel insertion mutation of MAPT causes FTDP-17 with distinct pathology. H. Morino1, Y. Matsuda1, R. Ohsawa1, K. Hiraki1, T. Kurashige2, Y. Izumi2, Y. Yamazaki2, T. Takahashi2, A. Takashima2, Y. Soeda2, T. Miyasaka2, M. Higuchi2, N. Sahara2, T. Suhara2, H. Shimada2, H. Maruyama2, H. Ito2, H. Kawakami1. 1) Department of Epidemiology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan; 2) Department of Clinical Neuroscience & Therapeutics, Institute of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan; 3) Department of Clinical Neuroscience, Graduate School of Medicine, University of Tokushima Graduate School, Tokushima, Japan; 4) Department of Aging Neurobiology, National Center for Geriatrics and Gerontology, Oshu, Japan; 5) Neuropathology Department of Medical Life Systems, Faculty of Life and Medical Sciences, Doshisha University, Kyoto, Japan; 6) Molecular Neuroimaging Program, National Institute of Radiological Sciences, Chiba, Japan; 7) Department of Neurology, Graduate School of Medicine, Wakayama Medical University, Wakayama, Japan.

PURPOSE: MAPT, LRRK2, GRN, and C9orf72 are known as responsible genes for several neurological disorders, including progressive supranuclear palsy (PSP), Parkinson’s disease (PD), and frontotemporal dementia (FTD). We conducted genetic analysis of a family with PSP, PD, and FTD over two generations in order to identify a causative gene for these neurological disorders.

METHODS: We performed high-density SNP typing on three affected and two unaffected persons, followed by exome sequencing on the three affected persons. Candidate regions were selected by linkage analysis and Homozygosity Haplotyping based on the results of the high-density SNP analysis. The variants were obtained using BWA, Samtools, Picard, and GATK from the exome sequencing data. The reduced the candidates by referring to the public variant databases, and confirmed the remaining mutations by Sanger sequencing. In addition, we screened for the mutations in other patients with similar symptoms.

RESULTS: Relatively long segments of high LOD score exist on chromosome 6, 9, 14, 17, and 20. Homozygosity Haplotyping indicated common haplotype regions of chromosomes 13 and 17 in the patients, but not in the unaffected. From the result of exome sequencing two novel variants remained finally. One was located in the microtubule-associated protein tau (MAPT) gene that is known as the causative gene of frontotemporal dementia with parkinsonism on chromosome 17 (FTDP-17). The same variant was detected in two PD patients of another family, and the haplotype around the variant was identical to that of the index family. Biochemical experiment indicated that the MAPT mutation reduced the ability of microtubule polymerization, and accelerated tau aggregation. Pathological findings from autopsy of the FTD patient demonstrated notable observations.

CONCLUSION: The novel MAPT variant we identified presumably causes FTDP-17. The variant is located in the region in which several mutations were reported previously, but it is more unique that the variant is nonframeshift single amino acid residue insertion. In the further investigation, we will aim at revealing the pathogenic mechanism by evaluating biochemical changes by the mutation.
1257T
Strategy to discover new ALS causative genetic variant in Japanese ALS patients. J. Sone1, R. Nakamura2, M. Nakatohi2, H. Watanabe1, N. Atsuta1, F. Tanaka2, G. Sobue2, 1) Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan; 2) Center of advanced medicine and clinical research, Nagoya University Hospital, Nagoya, Aichi, Japan; 3) Department of Neurology and Stroke Medicine, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa, Japan.

Amyotrophic lateral sclerosis (ALS) is a progressive and fatal neurodegenerative disease characterized by progressive muscle atrophy, paralysis, and death within a few years of diagnosis. A number of hypotheses for the pathogenesis of ALS have been proposed, however, the pathogenic mechanism has not been clarified, and no effective therapy has been developed. Approximately 90% of ALS cases are sporadic (SALS) and 10% are familial ALS (FALS). Recently, many FALS causative genes have been identified with next-generation sequencing methods and reported. To clarify the pathogenesis of ALS, we are now doing a study on the genomic basis of FALS and SALS with a next generation sequencer and about 1,000 cases of DNA samples from the Japanese Consortium of Amyotrophic Lateral Sclerosis (JaCALS). At first, we established a generic screening system of already discovered ALS causative genes with an Ion Torrent sequencer. We selected 28 ALS genes as follows, SOD1, ALS2, SETX, SPG11, FUS, VAPB, ANG, TARDBP, FIG4, OPTN, VCP, UBQLN2, SIGMAR1, NEFH, DCTN1, TARDBP, FIG4, OPTN, VCP, UBQLN2, SIGMAR1, NEFH, DCTN1, TAF15, EWSR1, PRPH, GRN, CHMP2B, ZNF512B, PFN1, ATXN2, TFG, C9orf72, RNF19A and SOSTM1. We designed a primer set for multiplex PCR by Ion Ampliseq Designer. The total coverage of the designed primer set was 97.7% of all exons of the 28 ALS genes. We made a fragment library of each sample using Ion Ampliseq Express Kit 2.0, then sequenced the exomes of ALS patients with the Ion Torrent sequencing system. We subsequently sequenced the causative variant of 28 ALS genes, then divided the ALS patients into two groups, the causative positive ALS group and the causative variant negative ALS group. Next, we engaged in exome sequencing of the causative variant negative ALS group. We are now collecting exome data and analyzing it with the dbSNP database, the 1000 genome database and the HGVD database (Human Genetic Variation Database in Japan). We will present the progress of this study.

1259M
PRKAR1B mutation associated with a new neurodegenerative disorder with unique pathology. T.H. Wong1, W.Z. Chiu1, G.J. Breedveld2, K.W. L3, J.M.H. Verkerk3, D. Hondius3,4, R.K. Hukema2, H. Seelaar1, P. Fick1, L.A. Severijnen1, G.J. Lammers2, J.H.G. Lebbink5, S.G. van Duinen6, W. Kamphorst2, J.M. Rozemuller5, E.B. Bakker1, M. Neumann2,7, R. Willemse1, V. Bonifati8, A.B. Smit9, J.C. van Swieten10,12,13, Netherlands Brain Bank, The International Parkinsonism Genetic Networks. 1) Neurology, Erasmus Medical Center, Rotterdam, Netherlands; 2) Clinical Genetics, Erasmus Medical Center, The Netherlands; 3) Molecular and Cellular Neurobiology, Centre for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands; 4) Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; 5) Pathology, VU University Medical Centre, Amsterdam, The Netherlands; 6) DZNE, German Centre for Neurodegenerative Disease, Tübingen, Germany; 7) Neurology, Leiden University Medical Centre, Leiden, The Netherlands; 8) Cell Biology and Genetics and Radiation Oncology, Erasmus Medical Center, Rotterdam, The Netherlands; 9) Pathology, Leiden University Medical Centre, Leiden, The Netherlands; 10) Clinical Genetics, Leiden University Medical Centre, Leiden, The Netherlands; 11) Neuropathology, University of Tübingen, Tübingen, Germany; 12) Alzheimer Centre, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands; 13) Neurology, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands.

Pathological accumulation of intermediateofilaments can be observed in neurodegenerative disorders, such as Alzheimer's disease [MIM 104300], frontotemporal dementia [MIM 600274] and Parkinson's disease [MIM 168600], and is also characteristic of neuronal intermediate filament inclusion disease. Intermediate filaments type IV include three neurofilament proteins (light, medium and heavy molecular weight neurofilament subunits) and a-interemin. The phosphorylation of intermediate filament proteins contributes to axonal growth, and is regulated by protein kinase A. Here we describe a family with a novel late-onset neurodegenerative disorder presenting with dementia and/or parkinsonism in 12 affected individuals. The disorder is characterized by a unique neuropathological phenotype displaying abundant neuronal inclusions by haematoxylin and eosin staining throughout the brain with immunoreactivity for intermediate filaments. Combining genetic analysis, exome analysis and proteomics analysis, we identified a heterozygous c.149T>G (p.Leu50Arg) missense mutation in the gene encoding the protein kinase A type I-beta regulatory subunit (PRKAR1B [OMIM 176911]). The pathogenicity of the mutation is supported by segregation in the family, absence in variant databases, and the specific accumulation of PRKAR1B in the inclusions in our cases associated with a specific biochemical pattern of PRKAR1B. Screening of PRKAR1B in 138 patients with Parkinson's disease and 56 patients with frontotemporal dementia did not identify additional novel pathogenic mutations. Our findings link a pathogenic PRKAR1B mutation to a novel hereditary neurodegenerative disorder and suggest an altered protein kinase A function through a reduced binding of the regulatory subunit to the A-kinase anchoring protein and the catalytic subunit of protein kinase A, which might result in subcellular dislocation of the catalytic subunit and hyperphosphorylation of intermediate filaments.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1260T
Transcriptome sequencing in bipolar disorder identifies a global downregulation in the anterior cingulate and dysregulation of G protein-coupled receptors. C. Cruceanu1,2,4, P. Pavlidis3, S. Rojic1, P.P. Tan1, J.P. Lopez1,2, S.G. Torres-Platas1,2, G.A. Rouleau1,2, G. Tureck1,2. 1) Human Genetics, McGill Univ - Douglas Hosp Res Inst, Montreal, QC, Quebec, Canada; 2) McGill Group for Suicide Studies & Douglas Research Institute, Montreal, QC, Canada; 3) Centre for High-Throughput Biology and Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada; 4) Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada.

Objective: Bipolar disorder (BD) is a complex mental illness characterized by episodes of mania and depression. Gene expression changes and alterations in regulatory mechanisms have been associated with BD, mainly through candidate gene and a limited number of global microarray expression studies in post-mortem brain, but questions remain about isoform-specific dysregulation of known genes as well as the plethora of non-coding transcripts whose importance has been demonstrated recently in the brain but not characterized for bipolar disorder. Method: RNA sequencing (RNA-seq) is a powerful technique that captures the complexity of gene expression, and greatly improves upon previous approaches in both accuracy and level of information. We used RNAseq in tissue from the anterior cingulate gyrus from 13 BD cases and 13 matched controls. We sequenced an average of 180M paired-end reads per subject, and obtained information from ~23,000 transcribed coding and non-coding RNA species. Results: We computed differential expression between cases and controls using the HTSeq pipeline and detected 10 differentially expressed transcripts at a false discovery rate of <5%, replicated these findings using Cufflinks, and validated them by quantitative real-time PCR. Among the most significant results, we observed genes coding for Class A G protein-coupled receptors (GPCRs), SSTR2 (somatostatin receptor 2), CHRM2 (cholinergic receptor, muscarinic 2) and RXFP1 (relaxin/insulin-like family peptide receptor 1). Interestingly, a gene ontology analysis of the entire set of differentially expressed genes pointed to an overrepresentation of genes involved in GPCR regulation. We further followed-up the top genes by querying the effect of treatment with mood stabilizers commonly prescribed in BD through an in vitro study, and found that most of these drugs affect expression of some of these genes. Conclusion: By using total transcriptome analysis with RNAseq in the post-mortem BD brain, we identified a global downregulation in the ACC, and interesting profile of GPCR dysregulation, and pointed to several new BD genes. Furthermore, we characterized the coding transcriptome in BD and possibly identified the first long intergenic non-coding RNAs in BD. Our findings have important implications in regards to fine-tuning our understanding of the dysregulated BD brain as well as for identifying potential new drug target genes or pathways.

1261S
Targeted-resequencing gene panels for the genetic diagnosis of spinocerebellar ataxia and spastic paraplegia in Italian patients. D. Di Bella1, S. Magri1, E. Sarto1, S.M. Caldarazzo2, M. Pluman3, S. Baratta1, L. Nanetti1, C. Mariotti1, C. Gelleria1, P. Bauer2, F. Taroni3. 1) Unit of Genetics of Neurodeg and Metabolic disease, IRCCS Istituto Neurologico C Basta, Milan, Italy; 2) Institute of Medical Genetics, University of Tübingen, Tübingen, Germany.

Spinocerebellar ataxias and hereditary spastic paraplegias (HSP) are genetically highly heterogeneous group of neurological disorders involving both central and peripheral nervous system. More than 50 genes have already been identified for each disease group, but >50% of the patients remain undiagnosed. With the advent of next-gen sequencing, virtually all known disease genes can be tested at once, hugely increasing the expected diagnostic yield. We defined a comprehensive disease gene list and developed an HSP differential TruSeq Custom Panel (TSCA/SCAR) exorn-sequencing for 54 HSP genes and 76 genes for both dominant (SCA) and recessive (SCAR) spinocerebellar ataxias. Aim of the study was to analyze all the patients referred to our laboratory and negative for the common disease genes. Our data indicate that we could sequence 90% of the coding regions of these genes at more than 20X coverage. Mean coverage was 400-600X. A standard bioinformatic pipeline for mapping and annotation yielded a total of 200-400 variants in our disease genes per subject, which can be reduced to 5-10 with a filtration strategy. We included in the study 60 patients (9 AD, 15 AR, 36 S) with ataxia and 37 patients with sporadic spastic paraplegia. Overall, pathogenic mutations were identified in ~20% of patients negative for the common disease genes. In particular, in patients with ataxia, we identified mutations in 2 challenging genes (SYNE1 and SACS), which are extremely difficult to be identified even by whole exome sequencing because of their length. Moreover, we identified mutations in extremely rare SCA genes (SCAS and SCA14) in 2 large families. As respect to HSP, we could identify mutations in the large SPG11 gene and in rarer AR genes (GBA2, GTPBP3). Different missense variants of uncertain significance were identified in genes responsible for dominant ataxia and several heterozygous mutations were identified in recessive genes. All high-quality variants were confirmed by Sanger sequencing indicating reliability of this approach. Furthermore, the in-house database is required for the validation of uncertain variants and the assessment of the presence of large in/del mutations in recessive genes in which heterozygous mutations have been identified. In conclusion, our data confirm the need for multiplexed gene panel approach for the detection of the molecular causes of neurodegenerative diseases with high genetic heterogeneity. (E-Rare grant to PB and FT; Italian Ministry of Health grant to FT).

1262M
Deconstructing obsessive-compulsive disorder (OCD) by whole exome sequencing and integration of clinical endpoints and cognitive domains. L. Domenech1, K. Rabionet1, G. Escaramis1, D. Trujillano1, S. Ossowski1, A. Carracedo2, P. Alonso3, X. Estivill1. 1) Bioinformatics and Genomics Program, Centre for Genomic Regulation, Barcelona, Spain; 2) Hospital de la Santa Creu i Sant Pau, IDIBELL, Barcelona, Spain; 3) Hospital Universitario Bellvitge, IDIBELL, L’Hospitalet de Llobregat, Barcelona, Spain.

Obsessive-compulsive disorder (OCD) is a neuropsychiatric condition that affects 1-3% of the population worldwide. Genome-wide association studies for elucidating the genetic basis of OCD have given extremely limited results. We are taking a genome-centric approach to exhaustively explore the genome of selected OCD cases and to evaluate new aspects of the disease biology. As a first step in an integrative approach to dissect the biology of OCD we have performed whole-exome sequencing (WES) of 200 cases of OCD that underwent exhaustive clinical, neuro-psychological, and neuro-imaging assessment, and received standardized treatment. Control samples for WES have been the 1000 genome project European subset of WES data, the European-American data of the exome variant server (EVS), and the in-house data from over 600 WES of Spanish non-psychiatric subjects. We have implemented a Unified Mixed-Effect Models for Rare Variant Association to perform the analysis of this set of WES data. This approach can evaluate the association to disease of both rare and common variants detected by WES, implementing a combination of a Burden Test and a Multidimensional Scaling Test. In comparison to rare exome variant association studies on pathways we expect this new approach to more specifically identify protein interactions, regulatory mechanisms or protein complexes involved in OCD development. After pathway analysis and protein-protein interaction (PPI) approaches we have found an accumulation of mutations in epigenetic and neurodegenerative-related genes. Strong evidence of association of several genes in these pathways will be obtained throughout replication in characterized cohorts of OCD in the framework of the International OCD consortium and functional analysis of the identified variants. Supported by the Spanish Ministry of Economy and Competitiveness. LD is supported by a Severo Ochoa fellowship.
1263T
Targeted sequencing of a visual migraine aura locus on chromosome 9q22. M. E. Hiekkanen 1, M. A. Kaunisto 1, 2, V. Artto 3, S. Paavonen 1, E. Hämaäinen 1, M. Färkkilä 1, A. Palotie 1, 2, M. Kaljukari 1, M. Pesonen 1. 1. Folkvård Institute of Genetics, Folkhålsan Research Center, Helsinki, Finland; 2. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland; 3. Dept. of Neurology, Helsinki University Central Hospital, Helsinki, Finland; 4. The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

Migraine is a common and disabling neurological disorder characterized by recurrent multiphase attacks. The two main phases of a migraine attack are the aura phase and the headache phase. Often, patients report visual auras which typically last 5–60 minutes. This short period of time creates a technical challenge for genomic sequencing as only approximately 1/10th of patients experience aura during the genomic sequencing session. Moreover, only a subset of patients experience auras in the first family members. It is also known as choreoathetosis, due to its movement-related nature. Migraine auras can be occurring in a single individual or in families. We hypothesize that there is an isolated ACC. Altogether, our findings unravel an unprecedented role of recessive mutations in patients with CDK5RAP2.

1265M
Rare Alleles Altering Schizophrenia Risk Occur in Exons and Noncoding Functional Sequences. E. K. Lokken 1, 2, 4, S. A. Bacanu 3, 5, D. Walsh 6, F. A. O’Neill 7, K. S. Kendler 7, 8, B. P. Riley 1, 2, 4, Virginia Institute for Psychiatric Genetics, Virginia Commonwealth University, Richmond, VA; 2. Dept. of Psychiatry, Virginia Commonwealth University, Richmond, VA; 3. Clinical and Translational Research Center, Virginia Commonwealth University, Richmond, VA; 4. Health Research Board, Dublin, Ireland; 5. Queens University Belfast, Belfast, Northern Ireland.

Background: Schizophrenia demonstrates high heritability, in part accounted for by common simple nucleotide variants (cSNV), rare copy number variants (CNV) and epistatic interactions. Although heritability explained by cSNV and CNV is small compared to that explained by rSNV, rSNV in functional sequences may identify specific disease mechanisms. However, current exome methods do not capture a large proportion of potentially functional bases where rare variation may impact disease risk; as much as two-thirds of conserved sequences lie outside the exome in noncoding regions of cross-species evolutionary constraint. Methods: We reasoned that the candidate loci from the Psychiatric Genomics Consortium Phase 1 (PGC1) schizophrenia study represent good target loci to test for the impact of rare SNVs in non-coding constrained regions. We developed custom reagents to capture mammalian constrained non-coding regions, exons and 5’ and 3’ untranslated regions (UTRs) in the 12 PGC1 loci for pooled sequencing in 912 cases and 936 controls. Results: Our design contains substantially more highly conserved bases (46,412 vs. 31,609) and variants (390 vs. 193) in noncoding as compared to coding targets. Using C-alpha to detect excess variance due to aggregate risk increasing or decreasing rSNV effects, we identified signals attributable to allele-specific expression and CNVs in the non-coding regions of exonic rSNV in constrained non-coding sequences, including variants within ENCODE transcription factor binding sites, DNase hypersensitive regions, and histone modification sites in neuronal cell lines. We also observed significant excess risk-altering variation in the exonic domain of CSMD1, a gene expressed in the developing central nervous system. Discussion: These results support the hypothesis that common and rare variants in the same loci contribute to schizophrenia risk, but highlight the need to expand capture strategies in order to detect trait-relevant sequence variation in a broader set of functional sequences.

1266T
A targeted-resequencing approach for the genetic diagnosis of inherited peripheral neuropathies in Italian patients. S. Magni 1, D. Di Bella 1, M. Milani 2, 3, M. Färkkilä 1, B. Pareyson 4, 5, M. E. Hiekkala 1, 5. 1) Unit of Genetics of Neurodegenerative and Metabolic Disease, IRCCS Istituto Neurologico Carlo Besta, Milan, MI, Italy; 2) Neurology Unit 8, IRCCS Istituto Neurologico Carlo Besta, Milan, MI, Italy; 3) Child Neurology Unit, IRCCS Istituto Neurologico Carlo Besta, Milan, MI, Italy; 4) Neurology Unit, IRCCS Istituto Neurologico Carlo Besta, Milan, MI, Italy; 5) Virginia Institute for Psychiatric Genetics, Virginia Commonwealth University, Richmond, VA.

Inherited peripheral neuropathies are a complex group of genetically heterogeneous disorders caused by mutations in more than 70 genes that lead to the distal degeneration of peripheral nerves. They are divided into 3 groups according to the type of the neuropathological changes: 1) those characterized by Charcot-Marie-Tooth (CMT) disease characterized by motor and sensory neuropathy, subdivided into demyelinating (CMT1-4) and axonal (CMT2) forms; 2) distal hereditary motor neuropathies (dHMN), which have only motor involvement; 3) hereditary sensory neuropathy (HSN). These disorders not only represent a phenotypic continuum, but also show a genetic overlap and inheritance pattern variability. For these reasons, efficient genetic diagnosis would require a comprehensive and validated approach to be developed. We developed and validated two customized gene panels using TruSeqCustomAmpliC (Illunina) technique: one included 53 genes for CMT2 and dHMN, while the second one included 54 genes for CMT1-4 and HSN. We analyzed 78 patients previously screened for the CMT1A deletion/duplication and for −4–5 genes selected according to the phenotype. We identified ~300 variants per patient, which are reduced to 2-8 by filtering. All variants were subsequently imported into a local database created specifically for the purpose of: 1) filtering identified variants according to quality criteria, functional consequence and frequency of polymorphic variants in the Italian population; 3) collecting data to allow an easy and quick query also by other laboratories. Coverage analysis revealed that more than 90% of the target region has coverage greater than 40X. The overall concordance of the disease was identified in 24/78 patients (30.8%). Moreover, likely pathogenic/uncertain variants were identified in 18% of patients and one heterozygous mutation in reciprocal translocation 1;2 was identified in a patient with CMT2.

In conclusion, this approach allowed us to increased by 10-fold the number of genes analyzed, significantly expanding the diagnostic yield. Our data confirm the usefulness and time/cost effectiveness of NGS-based gene panels for the genetic diagnosis of hereditary neuropathies. (Supported by Italian Ministry of Health grant to FT).
Molecular Studies of mTOR and Tau pathways in Focal Cortical Dysplasia
M.G. Mazzetti1, F.R. Torres4, P.A.O. Ribeiro1, S.H. Avansini3, R. Secolin1, B. Carvalho1, M.G. Borges1, F. Rogério2, L.S. Queiroz2, A.C. Coan2, H. Tedeschi2, E.P.L. Oliveira2, F. Cendes3, I. Lopes-Cendes2, The Brazilian Institute of Neuroscience and Neurotechnology (BRAIN), 1) Department of Medical Genetics; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 2) Department of Anatomical Pathology; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 3) Department of Neurology; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 4) Department of Pediatrics, University of California San Diego, La Jolla, CA; 4) Department of Anatomical Pathology; School of Medicne, University of Campinas, Campinas, Brazil.

Focal cortical dysplasia (FCD) is a subtype of malformation of the cerebral cortex which is a frequent cause of intractable focal seizures, requiring surgical treatment. FCD presents cortical architecture abnormalities also observed in tuberous sclerosis (TS) and hemimegalencephaly (HME). In addition, these show abnormal expression of genes linked to the mTOR signaling pathway. Potential involvement of Tau pathway was also reported in FCD. Therefore, the similarity in histological features as well as abnormal gene expression pattern suggests that pathogenic mechanisms could be common to these three disorders. Recently, somatic mosaic mutations have been identified in patients with TS and HME. Therefore, the objective of this work is to investigate whether somatic mosaic mutations in genes belonging to the mTOR and Tau pathways are present in the central nervous system of patients with FCD. NGS was performed in genomic DNA extracted from brain tissue resected by surgery (BTRS) and peripheral blood of patients with FCD. We performed exome capture with Nextera® Expanded Kit (Illumina®) and NGS on a HiSeq 2500 bench top sequencing machine. A bioinformatics pipeline was applied, using filters to variants present in public databases. To date, BTRS samples have included 71 patients with FCD. Forty patients with FCD were sequenced and a total of 749 and 91 variants were identified in genes belonging to the mTOR and Tau pathways, respectively. Among these variants we found 107 and 12 mutations in a mosaic state (present in BTRS samples), respectively, including 77 variants not described in databases of human mutations. Genes disrupted by mutations code for proteins involved in regulation of cell growth and cellular processes such as proliferation, differentiation and development, as well as genes already implicated in other cerebral cortical malformations. Our preliminary results confirm the presence of mosaic mutations in mTOR and Tau pathways in FCD. However, it is still not clear whether these mosaic mutations are a causative factor in FCD or a consequence of mutations in another gene whose disruption causes a cumulative mutational burden by dysregulation of mitotic checkpoints. Additional experiments, including high-depth NGS, will be carried out in order to answer these questions. Supported by CEPID-FAPESP.

Evidence for association of CDH26 with Autism Spectrum Disorders.
F. Mentch1, R. Golhar1, K. Wang1,2, J. Bradfield1, S. Murray1, H. Hakonarson1-4. 1) Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Zilka Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) Department of Pathology, University of California San Diego, La Jolla, CA; 4) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA.

Background. Autism spectrum disorders (ASDs) have been associated with genetic variation near neuronal cell-adhesion molecules (NCAMs).

Evidence for association of CDH26 with Autism Spectrum Disorders. M.G. Mazzetti1, F.R. Torres4, P.A.O. Ribeiro1, S.H. Avansini3, R. Secolin1, B. Carvalho1, M.G. Borges1, F. Rogério2, L.S. Queiroz2, A.C. Coan2, H. Tedeschi2, E.P.L. Oliveira2, F. Cendes3, I. Lopes-Cendes2, The Brazilian Institute of Neuroscience and Neurotechnology (BRAIN), 1) Department of Medical Genetics; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 2) Department of Anatomical Pathology; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 3) Department of Neurology; School of Medical Sciences, University of Campinas UNICAMP, Campinas, Brazil; 4) Department of Pediatrics, University of California San Diego, La Jolla, CA; 4) Department of Anatomical Pathology; School of Medicne, University of Campinas, Campinas, Brazil.

Micro arrays for identification of mosaic mutations in the mTOR and Tau pathways. F. Mentch1, R. Golhar1, K. Wang1,2, J. Bradfield1, S. Murray1, H. Hakonarson1-4. 1) Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Zilka Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) Department of Pathology, University of California San Diego, La Jolla, CA; 4) Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, PA.

Background. Autism spectrum disorders (ASDs) have been associated with genetic variation near neuronal cell-adhesion molecules (NCAMs). Apart from association through common GWAS variants, this category of molecules has been shown to be disrupted due to rare copy number variation (CNVs) in a subset of ASD cases. Here, we investigated if rare single nucleotide variants (SNVs) in neuronal cell adhesion genes are associated with ASD. Methods: We performed exome capture and targeted sequencing of 99 cell adhesion genes, including neurexins, neureligins, semaphorins, cadherins, protocadherins, and other ASD candidate genes. A total of 225 ASD cases were sequenced on Illumina HiSeq. Variants were annotated using ANNOVAR then filtered for potential pathogenic impact. Candidate variants were subsequently tested for association in 3,100 ASD cases and unrelated 2,992 controls on Veracode and 2,134 cases, 2,216 controls on Goldengate) previously registered in the de-identified biorepository at the Center for Applied Genomics (CAG) at The Children’s Hospital of Philadelphia (CHOP). EIGENSTRAT and R:KNN1 were used to control for population stratification and PLINK for Chi-Squared test for association. Results: Of 292 annotated SNPs from the targeted sequencing studies and typed in 3,100 ASD cases and 2,992 controls, three missense mutations in CDH26, not reported previously, were found to be associated with ASD (P=0.001 P<0.007), with the top SNP surviving Bonferroni correction for multiple testing. Of those, (1) rs11086690 (MAF 0.007, OR 2.06, SIFT 0.18) is located at the end of the Ca2+ binding site, which mediates cell-cell contact, with (2) rs28409250 (MAF 0.007, OR 1.94, SIFT 0.72) residing close to a Ca2+ binding site, whereas (3) rs41310187 (MAF 0.003, OR 1.83, SIFT 0.3) is not part of any protein domain. rs11086690 is in strong linkage disequilibrium (LD) with rs28409250 (r2=0.99), while rs41310187 (r2=0.57), and rs28409250 and rs41310187 (r2=0.56) is similar. Conclusions: These associations demonstrate that both common and rare variants in CDH26 associate with ASDs, suggesting they may play a role in the pathogenesis of the disease.
Whole exome sequencing identifies MEOX2 as a candidate genetic factor in posterior cortical atrophy. E.C. Schulte1, 7, R. Pemeczky2, A. Kurz2, J. Diehl-Schmid2, D. Rujescu3, M. Hütter4, A. Petri5, C. Gieger6, T. Meitinger1, 7, H. Klünemann7, J. Winkelmann1, 2, 3, 4, 5, 7. 1) Neurologische Klinik und Poliklinik, Technische Universität München, Munich, Germany; 2) Klinik für Psychiatrie, Technische Universität München, Munich, Germany; 3) Klinik für Psychiatrie, Friedrich Schiller Universität Jena, Jena, Germany; 4) Universitätsklinik für Psychiatrie und Psychotherapie, Albert Ludwigs Universität Freiburg, Freiburg, Germany; 5) Institute for Epidemiology II, Helmholtz Zentrum München, Munich, Germany; 6) Institut für Humangenetik, Helmholtz Zentrum München, Munich, Germany; 7) Institut für Humangenetik, Technische Universität München, Munich, Germany; 8) Psychiatrie Klinik, Universität Regensburg, Regensburg, Germany; 9) Psychiatrie Klinik, Technische Universität München, Munich, Germany; 10) Department of Neurology and Neurosciences, Stanford University, Palo Alto, CA, USA.

Background: Posterior cortical atrophy (PCA) is a rare neurodegenerative condition presenting with an impairment in visual processing skills which shares clinical characteristics with Alzheimer’s disease (AD). The causes of PCA are not fully understood. Here, we performed whole exome sequencing in nine individuals with PCA in search for rare genetic variants that contribute to this phenotype. Methods: Whole exomes of 8 individuals with PCA and one trio were sequenced. Identified variants were filtered under autosomal dominant (auto dom), autosomal recessive (auto rec) and de novo modes. Burden tests (CAST) were performed for all candidate genes using data from 3891 in-house exomes and 4300 NHLBI ESP exomes as controls. Subsequently, the coding regions of the most plausible biological candidate, MEOX2, were screened for variants by high-resolution melting curve analysis in 1271 individuals with AD and 3426 general population controls. Results: We identified de novo missense variants in KIRREL3, MEOX2 and CACNA1S. Joint analysis of all 9 PCA cases yielded 44 and 37 genes harboring rare variants (MAF ≤ 0.1% for AD and ≤ 2.5% for rec alleles) in at least 30% of cases under an auto dom and rec model, respectively. None of these overlapped with any of 13 genes in which the affected son of the trio held homozygous or compound heterozygous variants. None of the variants occurred in more than a single case. Burden testing identified 7 auto dom—ANKR6/MLRY2/MLRY2, KEL, NAA16, NYX, PPP1R3E, SIGIRR—and 3 auto rec—AFAP1L2, ALPK2, IFNA10—candidate genes in which rare variants were associated with the PCA phenotype at exome-wide significance. As MEOX2 has been linked to AD, we screened the coding regions of this gene for additional rare variants in a large AD sample. Rare non-synonymous variants in MEOX2 showed a trend towards being more common in individuals with AD compared to controls (38 of 1271 AD vs. 72 of 3426 controls; p=0.07, χ² test). The de novo variant was not found again. Conclusions: We were unable to causally link rare coding variants in any gene to the PCA phenotype but created the first bias-free short-list for potential candidate genes for PCA. A de novo variant in the homeobox gene MEOX2 seemed especially interesting because individuals with extreme AD phenotypes have been shown to harbor copy number variants overlapping MEOX2 and Meox2-/- mice develop an AD-like phenotype.
Subcortical band heterotopia (double cortex syndrome) not associated with DCX or LIS1 gene mutations. E. Andermann1, 2, 5, 6, 10, F. Adrom1, 5, 6, 10, F. Dubreau1, 2, 5, 6, 10, D. Melançon5, 7, 10, D. Tampieri5, 7, 10, F. Andermann4, 5, 6, 10, B. Dobyna1, 11, 1 Neurogenetics Unit; 2 Seizure Service; 3 Department of Pediatrics and Neurology, University of Montreal Hospital and Institute; 4 Department of Neurology & Neurosurgery; 5 Department of Radiology; 6 Department of Human Genetics; 7 Department of Pediatrics; 10 McGill University, Montreal, Quebec, Canada; 11 Departments of Pediatrics and Neurology, University of Washington; Center for Integrative Brain Research, Seattle Children’s Research Institute, Seattle, WA.

Subcortical band heterotopia (SBH), also called double cortex syndrome, is a malformation of brain cortical development due to deficient neuronal migration. It is classically due to either DCX or LIS1 gene mutations, resulting in a predominantly anterior or posterior distribution of the SBH respectively. However, a number of patients do not harbor point mutations or deletions/duplications of either of these genes, suggesting even greater genetic heterogeneity. We reviewed our database of adult epilepsy patients with SBH in whom no DCX or LIS1 mutation was found. Fifteen patients presenting with epilepsy, with or without associated cognitive delay, had SBH detected on brain MRI for which no DCX or LIS1 mutation was found. They had no associated pachygryia. All patients were sporadic. They were evaluated with a karyotype and DCX and/or LIS1 mutation analysis, depending on the pattern of distribution of their SBH. A few patients had both DCX/LIS1 sequencing and deletion/duplication analysis when their SBH was diffuse. In one woman, the karyotype showed a rearrangement of chromosome 9 suggesting a paracentric inversion of the q13->q23.3 region. Karyotype in both parents was normal, indicating a de novo chromosomal rearrangement in the proband. This patient had a history of infantile spasms, morning myoclonic seizures and pervasive developmental delays. She presented a thick and diffuse SBH. DCX/LIS1 sequencing and deletion/duplication analysis were negative. A CGH microarray with 135000 oligonucleotides was added and did not reveal any deletion or duplication. A higher resolution CGH micro array performed in order to detect a possible small deletion or duplication was inconclusive for gene identification. Genetic evaluation of SBH deserves DCX and/or LIS1 sequencing and deletion/duplication analysis, but does not always reveal the genetic cause of this brain malformation. Karyotype analysis and/or CGH array may be of interest in these unsolved cases. Whole exome sequencing should be carried out to elucidate the genetic cause and provide genetic counseling for these families.
1275T
The astrocytic transporter Slc7a10 (Asc-1) is required for glycerol inhibitory function. J.T. Ehmsen1, Y. Lu2, Y. Wang3, J.D. Rothstein1, S.H. Snyder1, M.P. Mattson1, A. Hoke1. 1) Department of Neurology, Johns Hopkins, Baltimore, MD; 2) Laboratory of Neurosciences, National Institute of Aging Intramural Research Program, Baltimore, MD. Slc7a10 (Asc-1) is a sodium-independent neutral amino acid transporter known to be the primary mediator of D-serine transport in the brain. Slc7a10 transports a number of additional amino acids including glycine, L-alanine, L-serine, and L-cysteine, as well as their D-enantiomers. We find that Slc7a10 is enriched in cerebellar Bergmann glia and within a subset of astrocytes of the caudal brain and spinal cord, in a distribution corresponding to high densities of glycerol inhibitory synapses. Accordingly, we find that spinal cord glycine levels are significantly reduced in Slc7a10-null mice and that spontaneous glycerol postsynaptic currents in motor neurons of mice lacking Slc7a10 show substantially diminished amplitude, identifying the likely etiology of sustained myoclonus and early postnatal lethality previously described for these animals. These observations establish a critical role for astrocytic Slc7a10 in glycerol inhibitory function in the central nervous system, and implicate Slc7a10 as a candidate gene in human hyperekplexia and stiff person syndrome.

1276S

Introduction: Although zebrafish have become a model for seizure studies, there are no studies about neuroinflammatory response after seizure in this model; therefore, we investigated the temporal mRNA expression profile of cyclooxygenase-2 genes after pentylentetrazole (PTZ)-induced seizures in adult and immature zebrafish brain. Methods: All experiments were approved by Animal Ethical Committee/UNICAMP (#3098-1). Seven days post-fertilization (dpf) larvae and adult zebrafish were separated in Seizure (SG) and Control (CG) groups (n=5 each group). Larvae sample was composed by pooling 20 heads, and adult sample by pooling two brains. Animals from SG were exposed to PTZ 15mM, and animals from CG were handled in PTZ-free water. At 0.05h, 1h and 6h after seizure, animals were anesthetized and their heads/brains collected for RT-qPCR that was carried out in triplicates with ef1α as endogenous controls using TaqMan System. The relative quantification was calculated by the equation RQ=2-ΔΔCT. Statistical analyses were performed by Mann-Whitney test, using the GraphPad Prism with p < 0.05. Results: for seven dpf larvae, only the cox-2b mRNA was increased at 0.05h and 1h after PTZ-induced seizure when compared to CG. The mean ± SEM data were: (i) cox2a: CG0.05h 1.2±0.06 vs SG0.05h 1.3±0.12 (p = 0.42); CG1h 1.32±0.5 vs SG1h 2.33±0.52 (p = 0.055); CG6h 1.08±0.11 vs SG6h 1.04±0.06 (p = 0.34); (ii) cox-2b: CG0.05h 0.93±0.02 vs SG0.05h 1.73±0.18 (p = 0.004); CG1h 1.43±0.4 vs SG1h 2.58±0.23 (p = 0.047); CG6h 1.08±0.11 VS SG6h 1.05±0.13 (p = 2.7). For adult zebrafish, the cox-2b mRNA levels were increased only at 6h after PTZ-induced seizure compared to CG. The mean ± SEM data were: (v) cox-2a: CG0.05h 0.74±0.08 vs SG0.05h 0.75±0.08 (p = 0.42); CG1h 0.66±0.12 vs SG1h 0.62±0.06 (p = 0.34); CG6h 0.98±0.2 vs SG6h 0.87±0.22 (p = 0.07); (vi) cox-2b: CG0.05h 1.04±0.09 vs SG0.05h 1.3±0.12 (p = 0.2); CG1h 1.03±0.06 vs SG1h 1.21±0.13 (p = 0.2); CG6h 0.98±0.03 vs SG6h 1.21±0.08 (p = 0.04). Conclusion: This is the first study investigating the temporal mRNA expression profile of cox2a and cox2b genes in adult and larvae zebrafish brain after seizures. Our results showed that both genes have different mRNA expression response after PTZ-induced seizure, suggesting that zebrafish cox-2b is more similar than cox-2a from mammalian Cox-2. This study provides evidence that this little fish is a valuable model for further studies of inflammation and seizures. Support: FAPESP and CNPq.

1277M
Profiling gene expression in CFW mouse brains to refine our understanding of the genetic architecture of behavioral traits. S. Gopalan-Krishnan, F. Carbonetto1, N.M. Gonzales1, E.H. Leung1, J. Park1, E. Ayene2, J. Davis2, C.C. Parker2, A.A. Palmer1. 1) Human Genetics, University of Chicago, Chicago, Illinois; 2) Dept. of Psychology and Program in Neuroscience, Middlebury College, Middlebury, VT; 3) Stanford School of Medicine, Department of Genetics, Stanford, California.

Translating results from genome wide association studies to elucidate the underlying biology of complex human diseases and traits remains a massive challenge. It is now widely appreciated that variation in gene expression plays an important role in the genetic architecture of most complex traits. However, it is not always feasible to access the relevant tissues in human subjects to measure gene expression. Here we demonstrate the use of an outbred (CFW) mouse population to model neurosyaptic traits for methamphetamine response sensitivity, conditioned fear behavior and pre-pulse inhibition. We complement this with RNA-Seq assays to quantify gene expression in three brain regions that are relevant to these traits, the hippocampus (n = 80), striatum (n = 55) and pre-frontal cortex (n = 54). The CFW mouse stock offers several important advantages over other mouse populations for mapping, especially the resolution of QTL regions. We mapped QTLs and eQTLs in these animals using genotypes obtained from a modified Genotyping-by-Sequencing protocol and a linear mixed model. Each QTL region contains many fewer genes as compared with traditional mouse populations; however, in many cases the QTLs still implicate dozens of genes. In these cases, gene expression data can be useful to prioritize candidate genes within QTL regions. In each tissue, we identified expression QTLs (eQTL) using two approaches: a linear mixed model approach to assess support for associations, and a two-sample paired t-test to identify allele-specific expression. Using eqtlhma (Flutre et al.) to model shared eQTLs across tissues, we find that a high proportion of eQTLs (approximately 70%) are common to the three brain tissues, with a small set of tissue specific eQTLs. Within each behavioral QTL, we prioritize genes for which there is an eQTL under this QTL peak, i.e. genes with expression affected by a variant associated with behavior are ranked higher than genes without such an eQTL. Preliminary results suggest that gene expression is an important mediator of behavior, with more than half of the QTL regions containing at least one variant that is also an eQTL.

1278T
Contiguous deletion of CADPS2 and GRM8 associates with severe autism spectrum disorder. C. Hatano1, Y. Yoko1, K. Waku1, K. Enomoto1, Y. Kuroda1, I. Ohashi2, R. Kosaki3, K. Kurosawa3. 1) Division of Medical Genetics, Kanagawa Children’s Medical Center, Yokohama, Japan; 2) Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan; 3) Division of Medical Genetics, National Center for Child Health and Development, Tokyo, Japan.

Autism spectrum disorders (ASDs) are complex neurodevelopmental conditions known to have multi-hit etiology models. Some copy number variations (CNVs) are shared common events in both human families (CFDs) and de novo events, contribute to genetic weakness in ASD patients. The identification of those genes seems to support proving the clinical states. We report on a female patient with severe autism, associated with contiguous deletion of CADPS2 and GRM8 at 7q31. She was born to nonconsanguineous parents at 41 weeks of normal pregnancy. The parents were healthy and their family history was unremarkable. Her birth weight was 3328 g (+0.82 SD), length 49.4 cm (+0.47 SD), and the head circumference 34.5 cm (+1.14 SD). She is the first case of contiguous deletion of CADPS2 and GRM8 in neonate, which is associated with developmental delay (DD). She needed gavage feeding for the first three months. The standard karyotyping revealed an apparently balanced translocation, 46,XX,t(7;10)(q31.3;q23.2). Her height was 95 cm (-3.3 SD), weight 13.6kg (-1.8 SD) at 5/7 years of age. She crawled at 3 years old and was able to walk with help at 5 years of age. She threatened self-injury. Together with severe intellectual disability (ID) and behavioral abnormalities as ASD, we performed cytogenetic microarray analysis, which revealed a de novo 4.4 Mb deletion at 7q31, extending from cM 122,292,068-126,610,549 (hg19). The deleted interval contained CADPS2 (cytoplasmic activator of secretion) and GRM8 (glutamate receptor, metabotropic 8). Haploinsufficiency of CADPS2 is known to cause mild intellectual delay. While GRM8 is one of the CGH-specific CNVs that can drive attention deficit hyperactivity disorder (ADHD) in humans, we found intellectual disability and autism phenotype observed in our patient is inconsistent with a dosage effects of the two genes. This combinatorial suppression of the 2 genes attribute to the etiology. These results suggest that synergy of haploinsufficiency from contiguous genes causes neurodevelopmental and functional network of neural activity has a crucial role in the severe phenotype of ASD and ID/ID. Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1279S  
Genome-wide gene expression analysis of identical twins discordant for autism spectrum disorder. A. Safiari1, C.C.Y. Wong2, M. Amo1, A. Ronald2, L.C. Schalkwyk2, J. Milne2, R. Plomin3, F. Dudbridge2, E.L. Meaburn1,2  
1) Department of Psychological Sciences, Birkbeck, University of London, London, UK; 2) King's College London, MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, De Crespigny Park, London, UK; 3) Genome Centre, School of Biomedical and Health Sciences, King's College London, London, UK; 4) University of Exeter Medical School, Exeter University, St Luke's Campus, Exeter, UK; 5) Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, University of London, London, UK.

Autism spectrum disorder (ASD) defines a collection of complex neurodevelopmental disorders, typified by impairments in social interaction and language. Multiple lines of evidence suggest a strong genetic basis to ASD. However, identical twins do not show complete concordance indicating that epigenetic and environmental (i.e. non-genetic) factors may also contribute to ASD risk. This study aims to investigate the biological basis of non-genetic contributions to ASD risk by characterizing gene expression differences in ASD-discordant identical twins. Our sample consists of 53 sets of identical twins discordant (N = 6) and discordant (N = 36) for ASD and ASD traits, and discordant controls (N = 11) that were derived from the longitudinal, population-based Twins Early Development Study (Howarth, Davis, & Plomin, 2012). Genome-wide expression profiling of whole blood samples was initially performed using the Affymetrix Human Gene array, which was subsequently followed up with RNA-Sequencing (RNA-Seq) using the Illumina TruSeq kit, performed in order to gain a more detailed, higher-resolution profile of transcription. Whilst the array-based analysis failed to identify differentially expressed genes associated with the disorder that reached genome-wide significance, pathway-based analysis identified several pathways previously associated with ASD, these included nerve growth factor (NGF) signaling and glutathione-mediated detoxification pathways. Our current efforts are focused on analysis of the RNA-Seq dataset, investigating both differential expression and differential splicing.

To our knowledge, this study represents the largest of its kind to systematically interrogate the biological pathways of ASD in discordant twins, the identification of which will help to shed light on the biological pathways dysregulated in ASD. DNA methylation (Illumina HumanMethylation27k Human Methylation array) and DNA variation (Illumina Human OmniExpressExome) data are also available for the same sample, and future work will be focused on the integration of these diverse and multi-dimensional datasets to help tease apart the relationship between genotype, epigenotype and phenotype.

1280M  
A mouse model of 2q23.1 deletion syndrome implicates MBD5 in neuronal development. J. Young1, V. Camarena1, L. Cao1, C. Abadi1, Y. Toledo1, M. Araki1, K. Araki1, K. Waltz1 1) Hussman Inst Human Genomics, Univ Miami, Miami, FL; 2) Kumamoto University, Kumamoto, Japan.

Individuals with 2q23.1 deletion syndrome have a range of well-defined behavioral abnormalities that include intellectual disability, motor delay, behavioral abnormalities and a distinctive craniofacial phenotype. Common to these patients is a partial or total deletion of a single gene, methyl-CpG binding domain protein 5 (MBD5), suggesting that haploinsufficiency of this gene is responsible for the phenotype. To confirm this hypothesis and to examine the role of MBD5 in vivo we have generated and characterized an Mbd5 mouse model. Our analysis of heterozygous mutant mice revealed a striking similarity in the behavioral manifestations of mice compared with human 2q23.1 microdeletion carriers including abnormal behavior, cognitive impairment and motor and craniofacial abnormalities. In addition, cortical neuronal cultures and cerebellar explant cultures allowed the detection of a deficiency in neurite outgrowth. Our study indicates that the Mbd5 mutant mouse model recapitulates most of the hallmark phenotypes observed in 2q23.1 deletion carriers. These findings support the causal role of MBD5 in 2q31.2 microdeletion syndrome and suggest a role for MBD5 in neuronal processes. The Mbd5 mutant mouse model will allow us to understand the abnormal brain development underlying the emergence of 2q23.1 deletion-associated behavioral and cognitive symptoms.

1281T  
A key role for TDP2 in neuronal development and maintenance. J.H.M. Schuurs-Hoeijmakers1,2,3, F. Gómez-Herreros1,2,3, M. McCormack4,5, M.T. Greally1,2, S. Rullen3, R. Rojo-Granados4, T.J. Counihan4, E. Chaillet4, J. Conroy4, S. Ennis5, N. Delanty4,5, F. Cortés-Ledesma1, A.P.M. de Brouwer1,2, G.L. Cavalleri4, S.F. El-Khamisy1,10,11, B.B.A. de Vries1,2, K.W. Cadlecott1 1) Department of Human Genetics, Radboud University Medical Centre, Nijmegen, Gelderland, Netherlands; 2) Department of Cognitive Neurosciences, Donders Institute for Brain Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, The Netherlands; 3) Genome Damage and Stability Centre, School of Biological Sciences, University of Sussex, Sussex, UK; 4) Molecular and Cellular Therapeutics, The Royal College of Surgeons in Ireland, Dublin, Ireland; 5) National Centre for Autism Genetics, Our Lady's Children's Hospital, Crumlin, Dublin, Ireland; 6) Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER), Departamento de Genética, CSIC (Centro de Investigaciones Científicas)-Universidad de Sevilla, Sevilla, Spain; 7) Department of Neurology, University Hospital Galway, Galway, Ireland; 8) Division of Neurology, Beaumont Hospital, Dublin, Ireland; 9) School of Medicine and Medical Science, University College Dublin, Dublin, Ireland; 10) Kreb's Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK; 11) Center of Genomics, Helmy Institute, Zewail City of Science and Technology, Giza, Egypt; 12) These authors contributed equally to this work.

Patients with heritable DNA repair deficiency syndromes often have increased susceptibility to cancer and pronounced neurological dysfunction. By using exome sequencing, we identified a homozygous truncating mutation, c.425+1GT>A, in the DNA repair gene TDP2 in three brothers from a consanguineous family, as well as a second homozygous truncating mutation, c.413_414delinsAA, in an additional unrelated individual. The affected individuals were characterized by early appearance of seizures and were all intellectually disabled. They also developed progressive ataxia that eventually resulted in confinement to a wheelchair. This clinical picture indicates an important neurodevelopmental role for Tyrosyl DNA phosphodiesterase 2 (TDP2) and suggests that this enzyme is continuously required for neuronal maintenance.

TDP2 is an enzyme that repairs ‘abortive’ Topoisomerase II (TOP2)-induced DSBs. TOP2 removes torsional stress from DNA and facilitates gene transcription by introducing transient DNA double-strand breaks (DSBs). Such DSBs are normally rejoined by TOP2 but on occasion can become abortive and remain unsealed. Lymphoblastoid cells from affected individuals that lack TDP2 activity are indeed hypersensitive to TOP2-induced DSBs. Furthermore, loss of TDP2 inhibits TOP2-dependent gene transcription in cultured human cells and in mouse post-mitotic neurons. We also show that TDP2 is required for normal levels of many gene transcripts in developing mouse brain, including numerous gene transcripts associated with neurological function and/or disease, and for normal interneuron density in mouse cerebellum.

In summary, we have identified inactivating mutations in TDP2 in individuals with intellectual disability, seizures and progressive ataxia. Moreover, we show that TDP2 protects gene transcription from endogenous abortive TOP2 activity, including the transcription of many genes involved in neurolog-ical development or function. Collectively, these data identify TDP2-dependent DNA break repair as a critical guardian of gene expression and highlight abortive TOP2 activity as a threat to normal neuronal development and maintenance.
1282S
Diagnostic assessment using next generation sequencing in extremely heterogeneous neurodegenerative disorders, hereditary ataxia and spastic paraplegia.
Z. Iqbal1, L. Pihlstrøm1, A. H. Rengmark1, S. Pilar-Henriksen1, S. L. K. Rydhina1, A. J. Weddigen1, I. M. Ryback2, C. Talilaksen1, 2, M. Tot1, 1 Department of Neurology, Oslo University Hospital, Oslo, Norway; 2 Faculty of Medicine, University of Oslo, Oslo, Norway.

Hereditary ataxias (HA) and hereditary spastic paraplegias (HSP) are very heterogeneous and rare neurodegenerative disorders. Because of the extreme clinical and genetic heterogeneity of these conditions, the establishment of molecular diagnosis has been challenging, both regarding the cost and labor-intensive experimental designs. Recently, after the advent of next generation sequencing technologies, there are several opportunities to develop a single inexpensive test and easy workflows in different diagnostic settings. Here, we have ascertainment 199 clinically well-characterized patients from all-over Norway, with HA (n=61), HSP (n=44), and a mixed phenotype of HA/HSP (n=14) patients. Polyglutamine expansions had been excluded in the affected probands with ataxia phenotypes. We have followed a cost-effective, fast, and robust procedure by using a Haloplex target enrichment kit (Agilent technologies, Santa Clara, CA), in order to screen coding sequences of 91 genes implicated in these conditions, of which 89 well known and two candidate genes were included in the design. DNA from ten individuals was pooled and a single barcode was assigned, and subsequently all 12 pools were sequenced by using a single lane of Illumina HiSeq, which resulted in high quality data with an average 98.95% bases covered >80x in the targeted intervals. Our bioinformatics pipeline and criteria to assign plausible pathogenically selected about 50 candidate variants for further validation and confirmation were developed. After confirmation, the others were discarded. Our results contribute to the identification of novel potential variants in several genes, including ANO10, BEAN1, BSCL2, CYP2U1, GRID2, KCNDO3, KIF1A, KIF5A, PRKCG, PREE1, SPG7, SPG11, and SYNE1. After the ongoing validation analysis, we will be able to establish the diagnosis in a certain subset of patients, in both HA and HSP. With this approach, it will be possible to estimate mutation frequencies in the involved genes. Moreover, in a heterogeneous group of disorders, such diagnostic testing will reduce the total duration of correct diagnosis, therefore saving time for the patient and the healthcare system. In conclusion, our data will have extended implications in the field of clinical genetics and development of diagnostic procedures.

1283M
Identifying biomarkers in chronic neuropathic pain.
P. C. McHugh1, D. P. Finn1, R. Airley1, D. A. Buckley1.

Chronic pain, including neuropathic pain, is poorly understood. It is estimated that almost 1 in 5 adults experience moderate-severe pain and often receive inadequate treatment. With a growing ageing population, this is fast becoming a significant healthcare issue that requires urgent attention. Blood biomarkers can provide us with a tool for early diagnosis and in turn facilitate improved treatment strategies. We wish to identify pain-related molecules, systems or pathways (biomarkers) in human plasma with the aim of improving our understanding, diagnosis and treatment of neuropathic pain. To identify novel biomarkers that can discern neuropathic pain patients from healthy controls we are using a neuropathic pain-control pilot sample (n=20). For this we are employing a combination of quantitative PCR, PCR arrays and Affymetrix axis systems to analyse gene expression (RNA), and HPLC/Mass Spectrometry for plasma analyses. Our preliminary data identifies several molecules that can differentiate between neuropathic pain and healthy controls. To validate our findings we are currently developing two independent neuropathic pain cohorts (case-control); one at NUJ Galway and the other through the Leeds Pain Rehabilitation group. These molecules could be developed into a diagnostic test that can differentiate neuropathic pain subtypes, allowing clinicians to better treat this condition.

1284T

Genetic testing of neurological diseases is often not straightforward, mainly due to the presence of nonspecific overlapping clinical symptoms and of genetic heterogeneity. A global approach using high throughput technologies such as NGS could be of great assistance to improve their diagnosis. We have experience in the genetic assessment of neurological conditions using a 200- gene NGS targeted resequencing panel. Samples from 48 patients were submitted to our laboratory for genetic testing. Clinical indications included Spastic Paraplegia (17), Motor/Sensitive Hereditary Neuropathy (12), Ataxia (6), Parkinson (4), Miastenia (3), Joubert Syndrome (2), Frontotemporal Dementia (2), Hyperkalemic Periodic Paralysis (1) and SMA (1). Coding exons and splice-site regions of the genes associated with each pathology were analysed. Enrichment and sequencing were carried out using SureSelect Enrichment System (Agilent) and SOLID 5000/MiSeq (LifeTechnologies/ Illumina). Mean depth was established at 200x. In 7 out of 48 samples (14%), a genetic cause that confirmed the clinical diagnosis was found. Taking into account all the samples, 63 variants (confirmed by Sanger sequencing) were identified and classified into pathogenic (7), probably pathogenic (2), unknown (53) and probably benign (1). According to their effect, missense (29), synonymous (21), splice-site (8), frameshift (3) and nonsense (2) mutations were found. In 21 out of 48 samples, no variants (pathogenic/unknown) were identified. Despite being moderate, 14% of genotypes diagnosed pathologically meant that NGS was a powerful diagnostic tool for the genetic diagnosis of neurological diseases. A better understanding of unknown significance variants, that would allow classifying them into SNPs or pathogenic mutations, could significantly increase the diagnosis rate.

1285S
T. Morithara1, M. Sato2, M. Silverman1, 2, H. Akatsu1, Y. Yamaguchi2, K. Kamiino1, M. Takeda1, 1 Psychiatry, Osaka University, Suita, Osaka, Japan; 2) Chou Medical Institute, Fukushima Medical University, Toyohashi-shi, Aichi, Japan; 3) Department of Biological Sciences, Simon Fraser University, Burnaby, Canada; 4) Center for Genomic Medicine, RIKEN, Yokohama, Japan; 5) Department of Integrative Genomics, Tohoku Medical Megabank Organization, Sendai, Japan.

Amyloid beta (Aβ) deposition is the central pathology of Alzheimer’s disease (AD). The genes controlling Aβ accumulation in the brain of sporadic AD are largely unknown. Human genetic studies require huge resources and do not reveal the role of the identified genes. To overcome this difficulty, we combine mouse models of AD with transcriptomics to identify Aβ modifier genes. [Methods and Results] We have observed that APP Tg mice with DBA/2 genetic backgrounds have significantly lower levels (-74.7 to 57.7%, p<0.0001, R2=0.626) than the athymic nude Tg mice (Aβ levels not significantly different). To avoid detecting secondary affected genes by Abeta, we used non-Tg mice in the absence of Aβ accumulation and performed the first transcriptome analysis to select candidate genes differentially expressed in DBA/2 mice. The 2nd transcriptome analysis using APP-Tg mice with mixed genetic backgrounds from these three strains revealed kinesin light chain-1 (Klc1) as an Aβ modifier. The following QPCR measurements confirmed our findings. Further analysis using neuroblastoma cells suggested that Klc1 splice variant E (Klc1E) control the accumulation levels of Aβ in mouse brain. The expression levels of Klc1E were significantly correlated with the levels of Aβ (R2= 0.21 to 0.39, p<0.0001-0.0002) in these mice. We expanded mouse strain analysis by adding two more strains. Three strains with higher Aβ accumulation shared common Klc1 allele and had higher expression levels of Klc1E in contrast two strains with lower Aβ shared another common Klc1 allele and had lower levels of Klc1E. In human, KLC1E levels in brain (AD: n=10, Cont: n=14) and lymphocyte (AD: n=47, Cont: n=17) were significantly higher in AD patients compared to controls. Functional analysis using neuroblastoma cells showed that knockdown of KLC1E decrease the production of Aβ (-39.3 to -44.7%, p<0.0001). Taken together, mouse transcript, mouse genetic, human transcript and functional analysis strongly suggested that KLC1E control Aβ pathology in AD. [Conclusion] The expression levels of KLC1E control Aβ pathology. The unique combination of distinct mouse strains and model mice with transcriptomics is useful for the study of genetic mechanisms of complex diseases.
1286M

Circadian Network and Autism: Role of the JARID1 Genes. Z. Talebizadeh1, A. Shah1, D. Kalinowska2, L. DiTacchio2. 1) Pediatrics, Children’s Mercy Hospital, and University of Missouri-Kansas City School of Medicine, Kansas City, MO; 2) University of Kansas Medical Center, Kansas City, KS.

The circadian clock coordinates diverse cellular processes and functional outcomes, including behavior and cognition. Abnormalities in the clock genes may have a role in autism, but the underlying mechanism remains unknown. Members of the JARID1 gene family (JARID1a-d), histone demethylases, have been shown to be involved in the circadian molecular machinery in a recent work of our collaborator. JARID1a activates CLOCK-BMAL1, whereas JARID1b, c, and d (as represented in RGD) inhibit CLOCK. This suggests that the optimal ratio among JARID1 isoforms is vital in maintaining the proper function of the circadian system. Imbalance in these isoform ratios may contribute to the etiology of diseases. Furthermore, miR132 orchestrates translational control of the circadian clock by targeting chromatin remodeling genes, including MECP2 and JARID1a. The splicing profiles of JARID1, and factors that regulate the expression of such isoforms, are not fully known. Our hypothesis is that JARID1 mis-splicing may present in at least a subset of subjects with autism. Recent findings have shown cross links between DNA methylation and gene regulatory processes, including alternative splicing and miRNA. Therefore, we used data from a DNA methylation marker for subject stratifications. We evaluated the expression level of multiple JARID1 alternative splicing transcripts in autistic subjects stratified based on the absence or presence of a given DNA methylation-related marker (DM) in lymphoblastoid cell line-derived RNAs, using Exon array profiling, TaqMan assays, followed by DNA sequencing. A distinct pattern was detected in the expression level of alternatively spliced JARID1 isoforms for autism subjects compared with DM compared to those without DM and controls. Additional experiments, including miRNA mimics, are underway to further characterize the role of miR132 in regulating JARID1a by finding which isoforms(s) show miR132-dependent expression. This is the first study to evaluate a clock gene in autism, at the alternative splicing level in conjunction with DM markers. Our study indicates that (1) JARID1 undergoes complex splicing resulting in multiple splice variants, and (2) subject stratification using DM may assist with finding a more homogeneous subset to better understand the underlying mechanisms involved in autism. Funding: This study was supported by Patton Trust-KCALSI grant.

1287T

RNA-sequencing and gene co-expression analysis identifies novel genes and pathways in bipolar disorder. N. Akula, J.R Wendland1, K.H Choi2, B.K Lipska2, J.E Kleinman2, F.J McMahon1. 1) Human Genetics Branch, National Institute of Mental Health (NIMH), National Institutes of Health (NIH), Bethesda, MD 20892; 2) Dept of Psychiatry and Program in Neuroscience, Uniformed Services University of the Health Sciences, Center for the Study of Traumatic Stress, Bethesda, MD 20814; 3) Human Brain Collection Core, NIMH/NIH, Bethesda, MD 20892; 4) Lieber Institute for Brain Development, Baltimore, MD 21205.

Bipolar disorder (BP) is a heritable psychiatric illness affecting ~2% of the population. Genome-wide association studies (GWAS) have consistently identified several common variants associated with BP. Less consistent results have emerged from gene expression studies, partly due to limited sample size. Weighted gene co-expression network analysis (WGCNA) may be a more powerful approach, since it exploits patterns of co-expression among genes, but the approach has not yet been applied to RNA-seq data, where expression can be measured precisely over a wide dynamic range. We generated deep RNA-seq data from post-mortem dorsolateral prefrontal cortex obtained from a total of 11 BP cases and 11 age and sex-matched controls, sequenced in two batches (Akula et al, 2014). WGCNA identified 40 coexpression modules, of which 9 were preserved between the batches (Zsummary > 10). Six of the preserved modules were significantly associated with BP (meta-p < value < 0.05). Functional analysis in DAVID of genes within these 6 preserved modules revealed significant enrichment for several GO terms, including synapse, neuron projection, postsynaptic density, ion transport, and ATP binding, among others. Five of these modules were also significantly enriched for differentially expressed genes identified in Akula et al, 2014 (hypergeometric pvalue < 0.05), and 3 modules were significantly enriched (hypergeometric p-value < 0.05) for genes associated with BP in a previous GWAS (Chen et al. 2013). Thus two modules were preserved across both batches, associated with BP, enriched for genes known to be differentially expressed in BP, and enriched for prior BP GWAS hits. Some of the co-expressing hub genes in these modules include DCC, CAMK2D, KCNC3 and Npas3, which have been previously implicated in BP and related disorders. These results demonstrate that WGCNA can be applied to RNA-seq data with good results. Applied to BP, WGCNA produced results that were broadly consistent with prior findings and implicated additional genes and pathways that may contribute to risk for BP. This approach may help unify GWAS and gene expression results to suggest biological hypotheses deserving of additional study.

1288S

GluD1 is over-expressed in iPS-derived FOXG1 neurons: a potential common therapeutic target for Rett syndrome. S. Amabile1, T. Patriarchi1, A. Bartolini1, M.G. Loll2, D. Yasaki3, C. Lo Rizzo1, F. Ariani1,2, F. Mari1,2, M.A. Mancarella3,4, J.W. Helf3, I. Meloni1, A. Renieri1,2. 1) Medical Genetics, University of Siena, Siena, IT, SI, Italy; 2) Department of Pharmacology, School of Medicine, University of California Davis, Davis, CA, United States; 3) Department of Medical Microbiology and Immunology, University of California Davis, Davis, CA, United States; 4) Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy.

Rett syndrome is due to de novo mutations in MECP2, CDKL5 or FOXG1 genes. In spite of their involvement in the same disease, a functional interaction between the three genes has not been proven and disease mechanisms remain elusive. MeCP2 and FoxG1 are transcriptional regulators; CDKL5 encodes for a kinase protein involved in multiple cellular processes including gene expression. We hypothesized that mutations in the three different genes may lead to similar phenotypes by deregulating expression of common genes. To verify this hypothesis, we have used a human model based on induced pluripotent stem (iPS) cells-derived neurons from MECP2, CDKL5 and FOXG1 patients. We previously demonstrated by quantitative RT-PCR and protein analysis an over-expression of GRID1 in both iPS-cell-derived MECP2 and CDKL5 neuronal precursors and neurons (Livide Jr J Hum Genet 2014). We demonstrate here that GRID1 is overexpressed also in iPS-derived FOXG1 neuronal precursors and neurons. GRID1 encodes GluD1, a member of the delta family of ionotropic glutamate receptors that is located on the post-synaptic site and induces inhibitory or excitatory presynaptic differentiation depending on the specific brain region. GluD1 overexpression may thus favor either glutamatergic or GABAergic maturational imbalance. We therefore explored the expression of excitatory/inhibitory markers. We indeed found the expected imbalance but in opposite directions. Quantitative RT-PCR on iPSCs-derived neurons mutated in CDKL5 demonstrated that the glutamatergic markers VGLUT1 and VGLUT2 are over-expressed while the GABAergic markers GAD1 and GAD2 are down-regulated. On the contrary, iPSCs-derived FOXG1 neurons demonstrated a down-regulations of the excitatory markers and an up-regulation inhibitory markers. The different maturational imbalance found is consistent with the specific endophenotype of the patients from whom the iPSCs are derived: the FOXG1 patient does not manifest epilepsy while the CDKL5 patient has a severe epilepsy not controlled by therapy. In conclusion, at present GluD1 overexpression is the only common altered marker during neuronal differentiation derived from patients mutated in MECP2, CDKL5 and FOXG1. Modulation of the basic defect by inhibition of GLUD1 could be thus explored as a therapeutic tool in both typical Rett and its variants.
Phenotypic, molecular, functional and structural analysis of new DCX and LIS1 mutations causing the subcortical band heterotopia/lissencephaly spectrum. D. R. Amromin, 3, 4, 5, G. Brouhauchet, 6, S. Bещpected, 8, K. Topropova, 8, D. Melancon 3, 7, 8, D. Tampieri, 3, 7, 8, S. Reck-Peterson, 4, E. Andermanner, 1, 3, 4, 5, 6. 1) Neurogenetics Unit; 2) Department of Neurogenetics; 3) Montreal Neurological Hospital and Institute; 4) Department of Neurology & Neurosurgery; 5) Department of Human Genetics; 6) Department of Biology; 7) Department of Radiology; 8) McGill University, Montreal, QC, Canada; 9) Department of Cell Biology, Harvard Medical School, Boston, MA, US.

Subcortical band heterotopia (SBH), also called double cortex syndrome, and lissencephaly (LIS) are part of a spectrum of malformations of cortical development due to deficient neuronal migration, referred to as the SBH/ LIS spectrum. Most patients have either DCX or LIS1 gene mutations, associated with predominant anterior or posterior distribution of SBH, respectively. We report the phenotypic, molecular, functional, and structural analysis of new DCX and LIS1 mutations causing SBH/LIS. Patient 1 is a 46-year-old woman of French-Canadian ancestry who had onset of epilepsy at 3 months of age. She had multiple types of seizures and a diagnosis of Lennox-Gastaut syndrome refractory to antiepileptic medication. Brain MRI showed double cortex predominating in the frontal regions. DCX sequencing showed a c.578delA variant. The parents were not available for genetic testing.

Patient 2 is a 4-year-old man of British ancestry. At 6 months of age he presented mild developmental delay and delayed speech. At 2 years of age he had a partial left temporal lobe epilepsy and a right hemiparesis, with delay in his development and severe respiratory difficulties. Brain MRI showed a predominantly posterior lissencephaly associated with partial calsal agenesis and a cavum septum pellucidum, as well as diffuse diffuse cerebellar atrophy. LIS1 sequencing revealed duplication of five nucleotides in exon 8 (c.728_732dupATCAA). The parents declined genetic testing. Bioinformatic analyses of the patient showed that the change introduces a five residue stretch of altered sequence.

De novo TBR1 mutations in sporadic autism spectrum disorder patients. P. Denziotis, 1, B.J. O’Roak, 2, S.A. Graham, 3, S. Busquets-Estruch, 4, D. Dimitropoulos, 5, A. Doernier, 5, J. Gerds, 5, J. Shendure, 6, E.E. Fischer, 7, S.E. Fisher 8, 9, 1) Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands; 2) Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA; 3) Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon, USA; 4) Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington, USA; 5) Howard Hughes Medical Institute, Seattle, Washington, USA; 6) Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands.

Next-generation sequencing recently revealed that recurrent disruptive mutations in six genes - CHD8, DYRK1A, GRIN2B, PTEN, TBR1, and TBL1XR1 - may account for 1% of sporadic autism cases. TBR1 is of particular interest because it encodes a neuron-specific transcription factor of the T-box family, with established roles in patterning of the central nervous system, including regulation of neuronal identities during cortical development. The recurrence of de novo TBR1 mutations in sporadic autism spectrum disorder patients suggests that the identified mutations are likely to be pathogenic. Next-generation sequencing experiments in model systems are essential to determine the precise effect of mutations on protein function and provide insight into the molecular mechanisms of the disorder. Here we report the first functional analyses of TBR1 variants identified in sporadic autism. De novo truncating (p.A1389fsX80 and p.S351X) and missense (p.K228E and p.N374H) mutations disrupt multiple aspects of TBR1 function, including subcellular localization, interactions with co-regulators and transcriptional repression. Missense mutations (p.Q175E, p.V356M, p.Q419R and p.P421R) induce cell death in in vitro and in vivo assays. Our findings support the hypothesis that de novo mutations in sporadic autism have severe functional consequences and highlight the power of coupling novel genetic data to empirical assays of protein function to determine the precise impact of genes involved in complex neurodevelopmental disorders such as autism.

Analysis of actin cytoskeleton dynamics in stem cells from autistic patients. K. Gries-Oliveira 1, 2, M. Suzuki 1, D.Y. Sunaga 1, A.L. Sertie, 3, M.R. Passos-Bueno 1, 2, 4. 1) Bioclinics Institute- University of Sao Paulo, Sao Paulo, SP, Brazil; 2) Instituto Israelita de Ensino e Pesquisa - Hospital Israelita Albert Einstein.

Autism spectrum disorders (ASD) are genetically heterogeneous diseases, which has made the understanding of their etiology difficult. However, several genes implicated in the etiology of ASD are part of common molecular pathways, indicating these different genetic alterations may cause similar effects during neurogenesis. One of such common mechanisms may be the signaling cascade, which is involved in the regulation of dendritic spines and axonal growth and guidance. However, the relationship between these mechanisms and the etiology of ASD has been poorly explored in the literature, specially in their functional aspects. Previous results from our group have suggested a disturbance in cytoskeleton dynamics in cells from autistic patients. Thus, here we aimed to investigate the actin cytoskeleton dynamics regulation in stem cells from human exfoliated deciduous teeth (SHED) from ASD idiopathic patients. First, we incubate the cells of 13 patients and 8 controls for 24h using Rho kinase inhibitor (ROCKi) in order to depolymerize the microfilaments. To induce actin reconstitution, ROCKi was washed out and cells were treated with drugs that activate specifically Rac2, Cdc42 or RhoA. The percentage of cells presenting actin filaments were counted at 15, 30, 45 and 60 minutes after drugs application. Our results showed that 5 out of the 13 patients presented a significant lower percentage of cells with recovered actin filaments at all time points compared to controls when treated with Rac2 or cdc42 activator, and among these, 2 patients respond abnormally to RhoA activation (unpaired t-test p<0.05). Interestingly, these are the same patients that had responded abnormally to the specific actin reconstitution treatment. The other five patients respond abnormally to ROCKi activation and show a similar response to both RhoA activation, which suggests a dysfunction in the actin reconstitution for expression of Rac2, cdc42 and RhoA and we found that one of them has lower expression of these three RhoGTPases. These results suggest that, at least for a group of ASD patients, the dynamics of actin polymerization is under-regulated. This may introduce an important factor in the etiology of ASD. We believe that this study can contribute for the understanding of the common molecular mechanisms involved in ASD etiology.

Transcription and methylation reveals microglia related and non-coding RNA networks specifically altered in Dementia with Lewy Bodies. C. Humphries 1, 2, 3, M.A. Kohli 4, K. Whitehead 5, D. C. Marsh 5, 6, M. Suzuki 1, 2, S. Reck-Peterson 1, 2, 3, 4, 5, J. Gerdts 6, 7, 8, 9, S.E. Fisher 4, 5, 6, 7, 8, 9, 1) Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands; 2) Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA; 3) Department of分子 and Medical Genetics, Oregon Health & Science University, Portland, Oregon, USA; 4) Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington, USA; 5) Howard Hughes Medical Institute, Seattle, Washington, USA; 6) Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands; 7) Department of Radiology; 8) McGill University, Montreal, QC, Canada; 9) Department of Cell Biology, Harvard Medical School, Boston, MA, US.

Dementia with Lewy Bodies (DLB) is the second most common dementia. To understand the common cellular pathways involved in DLB, we investigated the transcriptional and DNA methylation in human post-mortem brain tissue. Transcription was examined using total RNA-seq permitting the identification and quantitation of known genes and non-coding RNAs. DNA methylation was examined using bisulfite-seq. All sequences were aligned to the human genome build hg19. Experimental RNA-seq reads were run on Illumina’s HiSeq2000, generating 40-65 million reads per library. The program DESeq2 was used to examine transcriptional differences between Gencode transcripts. Transcriptional analysis between DLB and Controls using a cutoff of a 1-fold change in gene expression resulted in 30,929 genes (Gencode) to have expression differences (p<0.05). Of these, 469 were significant after controlling for multiple testing using False Discovery Rate (FDR<0.05). Subsequent comparison of these 469 DLB genes to LOAD revealed seven non-coding RNAs specifically altered. To find cellular processes disrupted in DLB, we performed network analysis using weighted gene co-expression network analysis (WGCNA). We found the 4,770 genes with expression differences between DLB and CON formed 13 networks. Eight networks were highly correlated between both DLB and LOAD. The other five networks were specific to DLB. Interestingly, the genes in the network correlated to both DLB and LOAD were primarily made up of protein coding genes. Non-coding RNAs primarily made up the networks specific to DLB. Using DAVID, we found the gene networks correlated to both DLB and LOAD are involved in acetylation. Genes previously shown to be part of the disease development and progression were enriched in these networks.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1293T Loss-of-function mutations of progranulin (PGRN) in siblings with familial FTLD, E. Vitale1, S. Napoletano1, S. Pappata2, A. Postiglione1, M.T. Gentile1, L. Colucci-D’Amato3, P. Sorrentino4, G. Milan2, 1 Institute of Protein Biochemistry (IBP), CNR-National Research Council, Pozzuoli (NA), Napoli, Italy; 2 Geriatric Center “Furlone” ASL Napoli 1 - Naples-Italy; 3 Institute of Bioimaging and Biostatistics, CNR, Naples-Italy; 4 University of Naples Federico II, Naples-Italy; 5 Experimental Neuropathology, SUN Caserta-Italy.

Mutations in the progranulin (PGRN) gene, located on chromosome 17q21.32, have been linked to familial Fronto-Temporal Lobar Degeneration (FTLD) or Motor Neuron Disease (MND). A homozygous mutation in the PGRN gene was identified in a familial case of FTLD. In their study, the authors present data on the clinical and pathological features of the patient and the genetic analysis performed. They discuss the role of PGRN in neurodegeneration and highlight the importance of genetic testing in the management of patients with FTLD.

1294S Significant Enrichment of Disease-specific Polymorphisms surrounding microRNAs suggests further involvement in Schizophrenia and Bipolar Disorder, V. Williamson, M. Mamedani, G. McMichael, S. Bacanu, V. Vladimirov, Psychiatry, Virginia InstPsychiatric & Behavioral Genetics, Richmond, VA.

Schizophrenia (SZ) and bipolar disorder (BD) are debilitating neuropsychiatric disorders with substantial impact on the quality of human life. Since 2006, numerous genome-wide association studies (GWAS) have been performed to identify loci associated with SZ and BD disease risk. In one of these studies, the authors utilized a transgenic mouse model that contains multiple copies of a YAC-fragilis and a novel microduplication on chromosome 7q36.3, centered on the PGRN gene. They found significant evidence for an association between the disease and the deletion. A subsequent CGH assay performed on the DNA from the two patients demonstrated no additional evident genetic alterations. Differential gene expression measurements revealed a significant downregulation of PGRN in the two individuals carrying the deletion, when compared to controls. In addition, the correlation between PGRN mRNA expression levels and cognitive impairment was statistically significant in PBD patients. Further, they identified a PGRN protein deficiency, corroborating the defective PGRN gene expression. These results strongly suggest that PGRN could be the causative disease gene in this familial FTLD.

1295M The transcriptome of 16p11.2 syndrome patients uncovers a link between autism and ciliopathies. A. Raymond1, E. Migliavacca2, M. Geyer2,7, I. Blumenthal7, J.S. Beckmann1,2,7, M.J. Daly3,4, M.M. Loviglio4, L. Hippolyte5, A.M. Maillard5, M.M. van Haastert1, J. Andreux2, J.F. Guisella1,10, M.J. DaSilva2, J.S. Beckmann7,2,11, S. Jacquemont5, M.E. Taftkowski1,12, N. Katsanis3,12. The 16p11.2 European Consortium. 1) Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland; 2) Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland; 3) Center for Human Disease Modeling and Department of Cell biology, Duke University, Durham, North Carolina 27710, USA; 4) Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA; 5) Analytic and Translational Genetics Unit. Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, 02114; 6) Program in Medical and Population Genetics and Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA; 7) Service of Medical Genetics, Lausanne University Hospital (CHUV), Lausanne, Switzerland; 8) Department of Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands; 9) Institut de Génétique Médicale, CHRU de Lille - Hôpital Jeanne de Flandre, Lille, France; 10) Departments of Genetics and Neurology, Harvard Medical School; 11) Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland.

The transcriptome of the 16p11.2 - ~600 kb BP4-BP5 region (29.5 to 30.1Mb) is one of the most frequent known genetic etiology of autism spectrum disorder (ASD). It is also associated with a highly penetrant form of obesity and a significant increase in head circumference. Mirror phenotypes are observed in carriers of the reciprocal duplication, who present a high risk of being underweight, microcephalic and/or developing schizophrenia. We performed a comprehensive analysis of the transcriptome of individuals carrying reciprocal CNVs in 16p11.2 and analyzed the data using a gene dosage model. The genome-wide transcript perturbations correlated with clinical endophenotypes of this CNV and were enriched for genes associated with ASD. We also uncovered a significant correlation between 16p11.2 copy number changes and expression levels of genes mutated in ciliopathies. This result was replicated in orthologous mouse models, raising the possibility that ciliary dysfunction might underlie 16p11.2 pathologies. In support of this correlation, we found structural ciliary defects specifically in the CA1 hippocampal region of a mouse model duplicated for the 16p11.2 orthologous region. Moreover, using an established zebrafish model we show strong genetic interaction between KCTD13 and 16p11.2 deletions, a key driver of these phenotypes in human. In conclusion, we identified CNVs with tissue-specific gene expression in brain, with low to moderate expression in frontal cortex (FCX).
1297S
Gene expression and neuronal morphology in differentiating human induced pluripotent stem cells (iPSCs) from individuals with chromosome 15q11.2 deletions. D.K. Das1,2, K. Chowdari1, C. Celik1, L. D’Auto1, W. Joel1, A. Ghosh-Bhattacharjey1, V.L. Nimgaonkar1,2. 1) Department of Psychiatry, WPIC, University of Pittsburgh, Pittsburgh, PA, United States; 2) Department of Psychiatry, University of California San Diego, La Jolla, CA, USA.

Background: In patients with Angelman syndrome (AS), there is anhidrosis, seizures, and an increased risk of upper airway obstruction. No mutations of the CNVs have been identified yet. Here we investigated the effects of CNVs in AS on gene expression in the developing human brain.

Methods: We performed RNAseq on fibroblasts from patients with AS with CNVs, using gene expression analysis. Results: The CNVs were found to be associated with significant expression changes in the developing human brain.

Conclusions: Our findings suggest that CNVs may contribute to the pathogenesis of AS.

1299T
Transcriptome signature of schizophrenia-associated rare copy number variants (CNVs) in lymphoblastoid cell lines (LCLs). W. Moy1, J. Duan1,2, E.I. Dragalenko3, J. Freda4, M.G.S.2, H.H. Goring5, A.R. Sanders1,2, P.V. Gejman1,2,3. 1) Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem, Evanston, IL; 2) Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL; 3) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 4) Molecular Genetics of Schizophrenia (MGS) Collaboration.

Schizophrenia (SZ) has been shown to be significantly associated with multiple rare CNVs that often encompass more than one gene, making uncertain which CNV-genes and downstream pathways are relevant to pathogenesis. We carried out RNAseq (>9M depth, 1×50bp) transcriptome analyses to identify expression signature profiles associated with each SZ-associated CNV in LCLs from 27 subjects of European ancestry carrying SZ-associated CNVs (with at least 5 subjects for each CNV), namely 22q11.21, del (n=13), 16p11.2, dup (n=9), and 15q13.3, del (n=5), and 657 subjects (265 SZ cases and 392 controls) without these CNVs. A total of 10,813 genes had detectable expression with a mean log2RPKM>1. We controlled analyses for sequencing batch, sex, age, affection status, Epstein-Barr viral load, growth rate, energy status and genotypic ancestral PCs. Within each CNV region, we found >90% genes showing expression associated (FDR<5%) with CNV status, consistent with expected gene dosages. Outside the CNV regions, we found variable numbers of genes with expression associated (FDR<5%) with CNVs (N): 22q1del (516), 16p1dup (1), 15q13del (2), representing the downstream genes affected by the CNV. Gene set enrichment analysis (GSEA) of the 516 genes with expression associated with 22q1del showed an enrichment (FDR<5%) of gene ontology (GO) terms related to chromosome segregation, cell cycle, and homologous recombination. Given the small number of genes associated with other CNVs, we further explored the GSEA for a larger set of genes associated with CNV status at nominal p<0.05, but with a fold difference of expression >1.2 or <0.8 (~2SD). We observed enrichment (P<0.05) of pathways and GO terms relevant to brain disorders, such as cognition, neuron development and differentiation, and regulation of neurotransmitter levels for genes with expression associated with 22q1del and 16dup. To identify the primary CNV gene(s) possibly causing changes in the schizophrenia-related gene expression pathway, we further examined the connectivity between CNV genes and the downstream genes of each CNV-induced brain-related gene sets by DAPPLE, and identified sialophorin (SPH) and mediator complex subunit 15 (MED15) as possible driver genes at 16p1dup and 22q1del, respectively. Our transcriptome analyses in LCLs of SZ-CN V carriers thus suggested plausible gene pathways relevant to SZ for each CNV region.
Epileptic encephalopathies are severe brain disorders characterized by seizures and abundant epileptiform activity which contribute to cognitive and behavioural impairments. Landau-Kleffner syndrome (LKS) and continuous spikes and waves during slow-wave sleep (CSWS) are closely related encephalopathies with regression in language and global cognitive skills respectively. They show electroclinical overlap with Rolandic epilepsy (RE), the most frequent childhood focal epilepsy, forming a clinical spectrum of epileptic, cognitive, language and behavioural disorders. Recently it was discovered that around 20% of cases in this spectrum are caused by mutations in the NMDA glutamate receptor GRIN2A. Here we set out to determine the disease mechanism of five missense GRIN2A mutations recently described in these patients. Mutations spread across the gene were chosen: P79R, C231Y, C436R, E714K and N976S, and were inserted into GRIN2A cDNA. HEK cells were transiently transfected with the mutant GRIN2A, and standard GRIN1 constructs allowing formation of heteromers. 1) Western blotting of total protein lysates revealed that all mutations caused a decrease in GRIN2A protein levels, with levels reduced to around 65% for P79R, 40% of non-mutant levels for C231Y, E714K and N976S and to 10% for C436R. 2) Single-cell calcium imaging was used to assay glutamate potency compared to non-mutated construct by increasing the half maximally effective concentration of agonist (EC50) by 3 and 5 times respectively, meaning the mutant receptors can only be activated by higher glutamate potency compared to non-mutated receptors. Mutations P79R and C231Y decreased the destabilising. 2) Single-cell calcium imaging was used to assay glutamate potency compared to non-mutated construct by increasing the half maximally effective concentration of agonist (EC50) by 3 and 5 times respectively, meaning the mutant receptors can only be activated by higher glutamate potency compared to non-mutated receptors. Mutations P79R and C231Y decreased the destabilising.

Abstract

1301M

Transcriptional regulation at the TREM gene cluster in AD brains. M.M. Carrasquillo1, A. Allen1, J. Burgess1, M.L. Kachadoorian1, S. Aryal1, F. Zou1, H.S. Chai2, C.S. Younkin3, J.E. Crook2, V.S. Porcher2, A.A. Naidoo2, S. Middha2, S. Maharanjan3, T. Nguyen1, L. Ma4, S.J. Lincoln5, K.G. Malphrus1, G.D. Bisceglio5, R.C. Petersen4, N.R. Graff-Radford6, D.W. Dickson7, S.G. Younkin1, N. Ertekin-Taner1-3, 1) Department of Neurosciences, Mayo Clinic, Jacksonville, FL; 2) Department of Biostatistics, Mayo Clinic, Jacksonville, FL; 4) Department of Neurology, Mayo Clinic, Rochester, MN; 5) Department of Neurology, Mayo Clinic, Jacksonville, FL.

TREM2, TREML1, TREML2, and TREML4 are genetic risk factors for sporadic late-onset Alzheimer’s disease (LOAD). Recently, several single nucleotide polymorphisms (SNPs) in each of these four genes were found to be associated with LOAD. These genes are part of a larger gene cluster on chromosome 6 (TREM1, TREML1, TREML2, TREML4) with variants at this locus and their effect on risk of LOAD. Utilizing Illumina DASL array data from our published brain expression genome-wide association study (eGWAS), expression levels in ADs (197 cerebral cortex, 202 temporal cortex) were compared to levels in non-AD (177 cerebral cortex, 197 temporal cortex) via multivariable linear regression, adjusting for age-at-death, sex, APOE-ε4, RIN, and transcript levels of 5 CNS cell-type specific genes. RNAseq data from 96 LOAD brains has been obtained and will be employed for characterization of the mechanism of transcriptional changes of TREM genes.

Abstract

1302T

Polymorphism in the miRNA-433 binding site of FGF20 is a strong risk factor for Parkinson’s disease in Iranian population. S. Abtahi, H. Darvish, A. Movaffaghi, R. Dastmalchi, B. Emaralizadeh. Medical genetics, Shahid beheshti university of medical sciences, Tehran, Iran.

DNA variations at the fibroblast growth factor 20 gene have been reported to associate with Parkinson’s disease (PD). The rs12720208, a functional SNP located in the 3’ UTR region of the gene, was first reported as a risk factor for the PD. Some other studies tried to replicate the result in different populations through case-control studies. However, to our knowledge, no significant association between the rs12720208 and PD was reported. In this study, we genotyped the rs12720208 SNP in 520 PD patients and 520 healthy controls from Iran. Significant differences was found in allele and genotype frequencies between patients and controls (Fisher exact p<1×10-6). The results suggest that the rs12720208 (C/T) polymorphism is a strong risk factor for PD in Iranian population.
1303S Integration of miRNA-mRNA networks to elucidate the complexity of psychiatric disorders. C. Chen1, L. Cheng2, C. Zhang3, J.A. Badner4, E.S. Gershon5. 1) State Key Laboratory of Medical Genetics, Central South University, Changsha, Hunan, China; 2) Northwestern University, Chicago, IL, USA; 3) The University of Chicago, Chicago, IL, USA; 4) Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, IL, USA; 5) Psychiatry department, University of Illinois at Chicago, Chicago, IL, USA.

MicroRNAs (miRNAs) are small, non-coding, endogenous RNAs involved in regulating gene expression and protein translation. One single miRNA can target multiple mRNAs and a single mRNA can also be targeted by multiple miRNAs. We consider that miRNA-mRNA clusters with statistically significant associations can explore potentially regulatory mechanism and, therefore, of biological interest. In this study, we collected 89 parietal cortex samples from Stanley Medical Research Institute(SMRI), and after quality control, each sample has 420 miRNA. 19,984 mRNA and more than 1,000,000 SNPs screened. We first constructed scale-free networks including both miRNA and mRNA, and found one module exhibited differential expression between controls and psychiatric patients. In this module, mir-320e acted as one of hub nodes. Quantitative Trait Locus(QTL) result indicated mir-320e was regulated by genetic variants. Another hub gene, PDLIM5, was validated by five miRNA binding prediction software. To further investigate the causal relationship between PDLIM5 and mir-320e, we applied Network Edge Orienting (NEO) and found mir-320e regulates PDLIM5. In summary, we detected one classic regulation pathway: Genotype ->mir-320e -> PDLIM5 -> gene module -> psychotics, which can be partially explain the etiology of psychiatric disease.

1304M Disruptions to the miRNA regulatory pathway may cause an increased rate of schizophrenia in individuals with 22q11.2DS. W. Manley, S. Ryan, S. Siecinski, L.M. Brzustowicz. Genetics, Rutgers University, Piscataway, NJ.

Schizophrenia is a complex and poorly understood disease caused by the interplay of many environmental and genetic factors. The 22q11.2 deletion syndrome (DS) is a disorder that is caused by the microdeletion of part of chromosome 22 leading to a 25-fold greater chance of developing schizophrenia in affected individuals versus the general population. The missing 22q11.2 region contains DGCR8, which is required for the initial step of miRNA biogenesis. However, the 22q11.2 deletion itself is not directly the cause of schizophrenia, since 75% of individuals with this deletion do not develop the disease. We hypothesize that the 22q11.2 deletion increases the risk of schizophrenia through alterations in miRNA regulatory networks via deletion of several miRNAs. This may serve as a protective buffer against the accumulation of deleterious mutations at other schizophrenia risk loci. We have derived human neural stem cells (NSCs) from individuals with the 22q11.2 deletion. The NSCs were generated from iPSCs using Neural Induction Media (Life Technologies). In order to ensure the presence of the 22q11.2 deletion, a FISH probe for 22q11.2 (TUPLE) was used (Cell Line Genetics). Also, the levels of DGCR8 gene expression were quantified using the 7900HT Fast Real-Time PCR System (Applied Biosystems) along with Taqman Gene Expression Assays to ensure DGCR8 reduction in 22q11.2 DS NSCs relative to otherwise healthy control NSCs. We have begun to characterize disruptions to the miRNA regulatory network in the NSC lines using Taqman Array Human Microarray Cards Version 3.0. Here we will present the miRNAs that we have identified to be differentially expressed in otherwise healthy control NSCs versus 22q11.2 DS NSCs. Our preliminary data shows that of the 377 tested miRNAs, there are nearly 30% that have altered expression levels greater than 2 fold or less than .5 fold. We predict that these miRNAs could be involved in the elevated risk of developing schizophrenia in individuals with 22q11.2 deletion syndrome versus the general population. Future studies will focus on determining the biological targets of these miRNAs using miRNA target prediction software.

1305T RNA-seq analysis reveals potential link between mammalian mitochondrial fatty acid synthesis (mFAS II), RNA processing, and neurodegeneration. S.I. Mitchell1, A. Parr1, S.D. Turner2, D.C. Crawford3, D.G. Murdock4. 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA.

Mitochondria in A93F, the gene encoding the first enzyme of the mitochondrial fatty acid synthesis (mFAS II) pathway, cause combined malonic and methylmalonic aciduria (CMAMMA). Patients diagnosed with CMAMMA in adulthood display psychiatric and neurological symptoms similar to those found in Parkinson’s disease (PD). The broader role of mFAS II result in mitochondrial RNA processing defects in yeast. We describe the first evidence from mammalian cells linking the mFAS II and RNA processing pathways. These data suggest nuclear RNA processing may also be affected by a loss of mFAS II function. Defects in RNA processing have been described as a possible mechanism underlying neurodegeneration in diseases including ALS, Huntington, and Prion diseases. These three diseases were also identified in pathway analysis of our gene expression data. Results from this study point to a possible link between mFAS II dysfunction, RNA processing, and neurodegeneration.


The genetic basis and molecular mechanisms that lead to Lewy Body (LB) pathology in 15-20% of Alzheimer disease cases (LB/AD) are largely unknown. Parkinson’s disease (PD), the prototype of LB spectrum disorders, has been studied most extensively. Genetic studies including large GWAS implicated Alpha-synuclein (SNCA) and Leucine-rich repeat kinase 2 (LRRK2) in the pathogenesis of PD. In the broader context, aggregates of insoluble alpha-synuclein protein are the major component of LBs and strong evidence demonstrated the importance of SNCA over expression in PD etiology. Mitochondrial dysfunction has been implicated in PD and Bender et al have recently showed that the translocase of the outer mitochondrial membrane (Tom40) encoded by TOMM40 gene mediates mitochondrial dysfunction induced by accumulation of alpha-synuclein protein in PD. The broader role of SNCA, LRRK2 and TOMM40 genes in other LB disorders is yet to be studied. We investigated the expression regulation of SNCA, LRRK2 and TOMM40 genes in A93F mouse model for PD, as well as the mouse model for Prion disease that predispose LB pathology. Using 107 LBV/AD cases and over 400 AD control subjects we tested the associations of variants in these three loci with LB pathology in AD and identified variants, SNPs and short sequence repeat (SSR), within SNCA and TOMM40 that were significantly associated with increased risk of LB pathology. When the analyses were stratified by LRRK2 genotype, the associations with SNCA variants became stronger. Next, we investigated the expression regulation of SNCA, LRRK2 and TOMM40 in relation to LB pathology. We detected significant differences in transcript levels of each of these genes between autistic, non-autistic and AD autism cases. The results of this study provide new insights into the role of genetic and epigenetic factors in LB pathology.
1307M

Neurodegeneration with brain iron accumulation (NBIA) encompasses a group of rare neurodegenerative disorders with different clinical, brain MRI and molecular features, underlined by progressive extrapyramidal dysfunction and iron overload in the brain. To date, over 30 genes are known to be associated with NBIA. Nevertheless a large number of individuals have "idiopathic NBIA", with unknown etiology. Exome sequencing of two sisters with NBIA identified two heterozygous mutations in the REPS1 gene. REPS1 is involved in endocytosis and the two mutations (p.Ala113Glu and p.Val87Leu) affect its EH1 domain that interacts with RAB11FIP2. Western blot analysis detected a low level of REPS1 in patient’s fibroblasts. The function of REPS1 in iron metabolism is unknown, but it was shown that RAB11FIP2 functions in transferrin recycling. We investigated the iron metabolism and oxidative stress in patient’s fibroblasts. Patient fibroblasts exhibited a dramatic iron overload, measured by a colorimetric ferrozine-based assay. Consistently steady-state levels of ferritin, iron responsive protein (IRP1) and SOD2 were increased whereas aconitase activity was decreased. This indicates that REPS1 mutations induce deregulation of iron metabolism. Over-expression of the wild-type REPS1 cDNA in patient’s cells reduces the iron overload of these cells that display almost normal iron content. Our experiments demonstrate that REPS1 is a new gene of NBIA. Improvement in our understanding of the biochemistry and pathophysiology of this form of NBIA will help develop novel therapeutics for this neurological condition.

1308T
Impaired Function is a Common Feature of Neuropathy-Associated GARS Mutations, L.B. Griffin1,2, R. Sakaguchi3, D. McQuiggan3, M.A. Gonzalez4, C. Searby3, S. Züchner4, Y.M. Hou1, A. Antonellis1,5,6,7, L. B. Griffin

1) Cellular and Molecular Biology Program, University of Michigan Medical School, Ann Arbor, MI; 2) Medical Scientist Training Program, University of Michigan Medical School, Ann Arbor, MI; 3) Department of Biochemistry and Molecular Biophysics, Thomas Jefferson University, Philadelphia, PA; 4) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 5) Department of Pediatrics, Division of Medical Genetics, University of Iowa, Iowa City, IA; 6) Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI; 7) Department of Neurology, University of Michigan Medical School, Ann Arbor, MI.

Charcot-Marie-Tooth (CMT) disease is the most common inherited peripheral neuropathy, affecting one in 2,500 individuals worldwide. Six aminoacyl-tRNA synthetase (ARS) genes have been implicated in CMT disease with an axonal pathology. ARS genes encode essential enzymes responsible for charging tRNA with cognate amino acids. Thirteen mutations in the glycyrrhizin-ARS synthetase (GARS) gene have been identified in patients with autosomal dominant CMT type 2D (CMT2D) characterized by impaired motor and sensory function in the upper extremities. While studies have revealed loss-of-function characteristics of GARS mutations, only a small subset of the mutations has been rigorously tested. We evaluated nine CMT-associated GARS mutations to assess if impaired function is a common feature. Our results demonstrate a strong correlation between impaired GARS function and CMT disease. Thirteen mutations previously associated with CMT disease (SS81L) did not demonstrate impaired function in our assays, failed to segregate with disease in two families with CMT, and was identified in the general population, indicating that SS81L is not a disease-associated mutation. Supporting the notion that impaired GARS function correlates with disease pathogenesis, Wild-type endogenous GARS forms puncta in neurons in vitro and in vivo. Seven GARS mutations impair granule formation in cultured neurons, suggesting that impaired localization may result in CMT2D disease pathogenesis. It is critical to identify proteins interacting with GARS in these structures to understand granule function and how disruption of these structures may lead to CMT disease. Impaired granule formation observed for mutant GARS proteins may be due to reduced binding to protein partners required for granule formation. Preliminary mass spectrometry data reveal that GARS associates with ribosomal complex proteins, including initiation factors and RPLP0. Previous studies demonstrated that archaeal ARS enzymes associate with ribosomes via RPLP0. This interaction may be critical in neurons for tRNA recycling for translational competency in the soma and axons. Current studies to identify differential protein partners of wild-type and mutant GARS will provide insight into the function of GARS-associated granules and may help explain the dominant toxicity seen in ARS-mediated CMT disease.

1309S
Novel Cytoplasmic Roles for the RNA-binding Protein, TDP-43. R. Smith, N. Alami, J.P. Taylor. Cell & Molecular Biology, St. Jude Children’s Research Hospital, Memphis, TN.

TAR DNA Binding Protein of 43 kDa (TDP-43) is a DNA/ RNA binding protein that is found in ubiquitinated cytoplasmic inclusions in a collection of neurodegenerative diseases. While these inclusions contain ubiquitin and TDP-43, they are tau and alpha-synuclein negative, characterizing a unique “TDP-43 pathology.” Furthermore, mutations in TDP-43 have been found to play a causative role in Amyotrophic Lateral Sclerosis (ALS) and Fronto-temporal Lobar Degeneration (FTLD). The normal function of TDP-43 is unclear and how mutations in TDP-43 cause disease is still undetermined. TDP-43 contains two RNA recognition motifs and the protein is shuttled between the nucleus and cytoplasm. Although TDP-43 has been described to be predominately localized to the nucleus, cytoplasmic localization is seen in normal conditions within wild type tissue. This localization could highlight a critical but yet undefined role for TDP-43 in the cytoplasm. To address this hypothesis we utilized the Drosophila model to investigate several potential roles for TDP-43 in RNA processing. By over-expressing either human TDP-43 or the Drosophila ortholog, TBP in various larval tissues, we have elucidated potential cytoplasmic functions, such as long-range transport of RNA for local translation at the synaptic terminal. This axonal transport and formation of cytoplasmic granules is dependent on the RNA-binding domain of TDP-43. Additionally, we found that ALS mutant TDP-43 granules display disrupted anterograde movement and abnormal TDP-43 distribution in the neuromuscular junction and cell body. Additionally, in muscle cells we observed cytoplasmic TDP-43 aggregates and cell toxicity that is reversed by inhibition of RNA binding. Our data identifies a novel cytoplasmic role for TDP-43 in RNA processing and transport that is disrupted by the introduction of disease-causing mutations, leading to cell death. Our results have furthered our understanding of the normal functions of TDP-43 and will advance future studies of TDP-43 associated disease.

1310M

Monoallelic expression, including genomic imprinting, X-chromosome inactivation and random monoallelic expression of autosomal genes are epigenetic phenomena. Genes that are expressed in a monoallelic way may be more vulnerable to genetic or epigenetic mutations. Thus, comprehensive exploration of monoallelic expression in human brains may shed light on complex brain disorders. Autism-related disorders are known to be associated with imprinted genes on chromosome 15. However, it is not clear whether other imprinting regions or other types of monoallelic expression are associated with autism spectrum disorder (ASD). Here, we performed a genome-wide survey of allele expression imbalance (AEI) in the human brain using single-nucleotide polymorphisms (SNPs) arrays, in 18 individuals with ASD and 15 controls. We identified an individual with a large (~100) number of monoallelic expressed SNPs. This widespread monoallelic expression was limited to the prefrontal cortex, and was not found in the cerebellum of the same individual. Another individual was found to have a highly skewed pattern of X chromosome inactivation. This skewed inactivation was stronger in the neural cell types (neurons, oligodendrocytes and astrocytes) than in microglia. Using our data, we were also able to define the allelic expression status of known imprinted genes in the human brain. We found that many of the known imprinting genes are biallelic expressed, and other genes showed isoform-specific imprinting patterns. We were also able to discover an abnormal imprinting event in an individual with ASD, in the imprinting locus on chromosome 15q11-13. Lastly, we developed an analysis of individual-level expression, focussing on the difference of each individual from the mean. We found that individuals with ASD had more genes that were up- or down-regulated in an individual-specific manner. We also identified pathways perturbed in specific individuals. These results underline the heterogeneity in gene regulation in ASD, at the level of both allelic and total expression.
GRIP2-mediated AMPA Signaling Defects Contribute to Autism Social Behavioral Deficits. T. Niranjan1,2, A. Adamczyk1, M. Han1, R. Mejias1, R. Rose1, H.C. Bravo1,2, M. Taub1, C. Schwartz2, D. Valle2, R. Huganir2, T. Wang2,3,4,5
1. The Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Predoctoral Training Program in Human Genetics, Johns Hopkins University, Baltimore, MD; 3) Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD; 4) Department of Biostatistics, Johns Hopkins University School of Public Health, Baltimore, MD; 5) Greenwood Genetics Center, Greenwood, SC; 6) Department of Pediatrics, Johns Hopkins University, Baltimore, MD; 7) Department of Neuroscience, Johns Hopkins University, Baltimore, MD.

Autism spectrum disorders are clinically and genetically heterogeneous, and likely involve several hundred risk genes. Identification of disrupted brain signaling pathways that underlie Autism endophenotype is crucial for discovery of novel drug targets. Glutamate mediates the majority of excitatory neurotransmission in the CNS via its family of receptors including AMPA-, mGluR-, and mGlur-type receptors. Glutamate Receptor Interacting Proteins 1 and 2 (GRIP1/2) are neuron-enriched scaffolding proteins with seven PDZ domains. PDZ4-6 bind directly to the c-termini of AMPA receptor 2/3 (GluA2/3), Liprin-α, and EphrinB1/2 to form a postsynaptic protein complex, which plays a crucial role in AMPA receptor trafficking and synaptic function.

mGlUR and NMDA signaling defects are implicated in Autism-associated conditions, such as Fragile X Syndrome, Rett Syndrome, and Shank2/3 deficiencies. However, AMPA signaling defects in Autism are poorly understood. We recently identified functional GRIP1 mutations that alter GluA2 synaptic trafficking and contribute to Autism social deficits. To determine if GRIP2 may mediate AMPA signaling defects in Autism, we sequenced exons of GRIP2 in a cohort of Autism patients (n=480) and ethnically matched controls (n=480), and identified a significantly increased mutation load at PDZ2-4 (p=0.022) in Autism patients. In proband families, affected sibs who carry GRIP2-PDZ4-6 mutations show more severe deficits in reciprocal social interactions defined by ADIR’s social scores as compared to affected sibs who do not carry the mutations. These mutations have functional effect; two result in significantly reduced binding with GluA2/3, while four others significantly alter interactions with EphrinB1/2 and Liprin-α. Glutamate-mediated AMPA signaling defects and deficiencies in social novelty (p<0.001) in a modified three-chamber test, and exhibit reduced social interaction (p<0.001) in a male dyadic social interaction test. Furthermore, enhancing AMPA signaling using a receptor desensitization inhibitor improves sociability in BTBR mice (p<0.001). These results support GRIP2-mediated AMPA signaling defects as a novel mechanism for social behavioral deficits in Autism. Characterization of GRIP2 deficiency in Autism shall provide valuable insights into the pathogenesis and identify novel drug targets for correction of social behavioral deficits in Autism.

A polymorphic di-nucleotide repeat (DNR) variant in the 5’UTR of DPYSL2 gene affects its regulation via mTOR signaling. X. Pham, Y. Liu, R. Wang, A. Pulver, D. Valle, D. Avramopoulos1, Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

Schizophrenia (SZ) is a common, disabling neuropsychiatric disorder with a complex etiology. It is estimated that as much as 80% can be attributed to genetic factors. Previous linkage and association studies implicated DPYSL2 on chr8p21 as a candidate gene for SZ. DPYSL2 encodes CRMP2 which is important in axonal growth, and its dysfunction may result in neuro-developmental abnormalities. We previously identified a polymorphic dinucleotide repeat (DNR) located in the 5’-untranslated region of DPYSL2, with a characteristic 5’-terminal oligopyrimidine (5’-TOP) tract, a target for mTOR mediated regulation pathway. The 13 CT repeat allele (risk) of the DNR was found to be associated with SZ compared to the 11 CT repeat common allele (WT). We performed dual lucerase assays in HEK293 cells and mouse primary cortical neurons and the risk allele showed ~3-fold decrease in luciferase activity as compared to the WT allele in both cell types. Further, polyomse profiling of the constructs showed the fraction of luceropase mRNA in polysemus associated with the risk allele was reduced by ~3-fold compared to the WT allele. Here we show that increasing concentrations of Rapamycin, an allosteric mTOR inhibitor, reduced lucferase expression in constructs from both alleles, the risk allele remaining lower at 0-30 nM. At concentrations higher than 30 nM, both alleles reached a plateau at the same levels. The same trend was recapitulated in both HEK293 cells and mouse primary cortical neurons. Our results suggest that the difference we observe between the two DNR alleles is mediated by mTOR signaling. Using arrays of > 4,000 human transcription factors and proteins, we screened for those that bind differently to the two DNR alleles to reveal the distinct regulatory footprint of DPYSL2 and might produce the decreased gene expression observed in the risk DNR allele. We identified a number of such proteins, one of which is a ribosomal binding protein (RBP) HuD/ELAVL4 that is also involved in mTOR signaling. Further HuD/ELAVL4 has been shown to play a crucial role in neuronal differentiation. How this RBP interacts with DPYSL2 in the WT and risk form to mediate gene expression in neurons is the subject of our current studies. In conclusion, we show that DPYSL2 is regulated by the mTOR signaling pathway and a SZ associated DNR variant in the 5’-UTR affects this regulation.
1314T New mutations of CYP2U1 in patients with spastic paraplegia and exploration of mitochondria functions. C. Tessonnier1,2, KH. El Hachem1,2,3,4,5,6, N. Eckart1,2,3,4,5,6, C. Goizet1,2,3,4,5,6, D. Lacombe5,6, C. Zeledón1,2, C. Schrafft1,2, J. Yin1,2,3,4,5,6,1,2, D. Lacombe5,6, C. Zeledón1,2, C. Brice1,4, A. Durr1,4, F. Darios1, G. Stevanin1,2, 1) Institut du Cerveau et de la Moelle épinière (ICM, INSERM U1127, UPMC UMR_S1127 Sorbonne Universités, CNRS 7225, NEB), Pitié-Salpêtrière Hospital, Paris, France; 2) Ecole Pratique des Hautes Etudes, Hésam Université, Laboratoire de Neurogénétique, Paris, France; 3) Dpt de Génétique Médicale, CHUV, Lausanne, Suisse; 4) APHP, Dpt de Génétique, Paris, France; 5) Université de Bordeaux, France; 6) National Research Institute, Cairo, Egypt.

Hereditary spastic paraplegia (HSP) is considered one of the most heterogeneous group of neurological disorders, both clinically and genetically. It comprises purely and complex forms that clinically include slowly progressing lower-limb spasticity resulting from degeneration of the corticospinal tracts. At least 73 loci accounting for these diseases have been mapped to date, and mutations have been identified in 53 genes, most of which playing a role in intracellular trafficking. We recently reported that loss-of-function mutations in CYP2U1 are responsible for an autosomal recessive form of early onset HSP. This gene codes for an enzyme involved in fatty-acid metabolism. We performed the screening of CYP2U1 in a new cohort of 259 index patients with HSP or sporadic spastic paraplegia. We have now identified 2 novel causative mutations in this gene: i) c.1468G>A/ p.C490Y, at the homozygous state in a North African sporadic case with early onset and thinning of the corpus callosum, ii) c.1A>C/ p.M1L at the homozygous state in 2 siblings with early onset HSP in a consanguineous family originating from Switzerland. Interestingly, this last mutation is also present in an asymptomatic sister aged 50 years suggesting an incomplete penetrance.

In addition, previous experiments we demonstrated in human cells (fibroblasts and lymphoblast) of 2 patients with the p.D316V mutation that the pathophysiology of this clinic-genetic entity includes alteration of mitochondrial architecture and bioenergetics with increased oxidative stress. We are extending these observations to 3 additional cases using 3D electronic microscopy and preliminary results also show an alteration of the endoplasmic reticulum (ER) in fibroblasts. Moreover, overexpression experiments in COS7 cells suggest that part of the physiopathology is explained by a partial loss of colocalisation of mutated forms with the mitochondria compared to controls. Our combined results focus attention on lipid metabolism or a metabolic compensation in this individual. Exome sequencing on the 3 siblings identified some candidate variants involved in lipid metabolism or in the mitochondria that potentially explain this phenomenon.

1315S Behavioral phenotyping of mice deficient in CHRNA7. B. J. Kelleher1, P. L. K. F. P. N. J. M. B. D. L. R. E. N. S. H. E. T. E. C. C. T. E. N. S. N. E. T. S. 1,2, J. Yang1,2,3,4,5,6, D. Wang1,2,3,4,5,6, M. Maley1, S. Morais1, E. Obre6, C. Durand6, M. Zaki6, J. Lavie6, D. Lacombe6, C. Goizet6, A. Brice1,4, A. Durr1,4, F. Darios1, G. Stevanin1,2, 1) Institut du Cerveau et de la Moelle épinière (ICM, INSERM U1127, UPMC UMR_S1127 Sorbonne Universités, CNRS 7225, NEB), Pitié-Salpêtrière Hospital, Paris, France; 2) Ecole Pratique des Hautes Etudes, Hésam Université, Laboratoire de Neurogénétique, Paris, France; 3) Dpt de Génétique Médicale, CHUV, Lausanne, Suisse; 4) APHP, Dpt de Génétique, Paris, France; 5) Université de Bordeaux, France; 6) National Research Institute, Cairo, Egypt.

The mouse cholinergic receptor alpha 7 (CHRNA7) is the only known nicotinic acetylcholine receptor expressed in the adult rodent brain. It plays a role in the regulation of learning and memory, social interaction and repetitive behavior, as well as sensorimotor gating. However, deficits in social interaction and repetitive behavior have not been assessed in this mouse model. We tested heterozygous and homozygous CHRNA7 mutant mice and their wildtype littermates for repetitive behaviors in self-grooming, holeboard exploration, and marble burying test, and for social interaction behaviors in the three-chamber paradigm, partition test, and social interaction video scoring. A detailed assessment of the aforementioned behaviors will be presented and be discussed in the context of human 15q13.3 microdeletion phenotypes.

1316M Regulatory function of CACNA1C schizophrenia-associated variants. N. Eckart1,2, H. Wang1, M. Zeledón1, M. Szymanski-Pierce1, D. Vallette1, D. Auroymontautou1,2,1, Joanne Nunele, University of Plymouth (SHP), University of California, San Diego, MD; 2) Johns Hopkins University, Department of Psychiatry, Baltimore, MD.

Schizophrenia (SZ) and bipolar disorder (BP) are complex psychiatric disorders, together affecting over 3.5% of the US population. They have overlapping clinical presentations, and onset in the second or third decade of life. Association and family studies indicate a shared genetic risk. One variant that has been independently and repeatedly associated with both disease phenotypes is rs1006737, a single nucleotide polymorphism (SNP) in the third intron of the CACNA1C gene. Variants identified by association studies are often in non-coding regions of the genome and enriched in expression quantitative trait loci (eQTLs). We previously reported that the risk allele of rs1006737 is correlated with decreased expression of CACNA1C (p=0.001) in a study of 195 post-mortem tissue samples from the superior temporal gyrus (STG). In 100 post-mortem tissue samples from the dorsolateral prefrontal cortex (DLPFC) we found that the risk allele trends toward increased expression of CACNA1C, in agreement with published data from Bigos et al. in the DLPFC (p=0.002). The variant rs1006737 tags a haplotype with several other SNPs, all located in the third intron. We are testing these for allele-specific regulatory potential using dual luciferase reporter assays. Building on previous results, now using 4 biological replicates for each construct, we report that for two constructs, one containing rs2159110 and a second containing both rs1077306 and rs10744560, the risk alleles show statistically significant increases in luciferase expression compared to the common alleles when transiently transfected in HEK293 cells (p=4.8*10^-4 and p=5.9*10^-4, respectively) on their own, suggesting potential for a direct regulatory function or a metabolic compensation in this individual. Exome sequencing on the 3 siblings identified some candidate variants involved in lipid metabolism or in the mitochondria that potentially explain this phenomenon.
1318S Transcriptome analysis of Lphn3 null mutant mouse brain and implications for ADHD and Addiction. D. Walls1, S. Galaviz2, B. Baker1, M. Tucker1, T. Loeger2 1) Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX; 2) Department of Computer Science and Engineering, Texas A&M University College Station, TX 77843-3474. Lratrophin 3 (LPHN3), a GPCR, has been identified by linkage and association studies performed by separate research groups in separate cohorts as a prime candidate gene for ADHD and Addiction. Lack of readily identifiable functional polymorphisms suggests that non-coding variations determining the quantity and/or quality of LPHN3 isoforms are likely contributors to these behavioral/impulsive phenotypes. We hypothesize that this mechanism of susceptibility and gene expression profiles in neurons are essentially unknown. We generated the Lphn3 null mutant mouse in order to evaluate gene function. We have already reported behavioral characterization of these mice indicating that they are both hyperactive and display increased reward seeking behavior. Here we describe the results after comparative transcriptome analysis of 3 brain regions over time between male wild type and null mutant Lphn3 littermates. Brains were harvested and dissected at 3 time points to reflect different stages of brain development (4 days, 4 weeks, and 6 months). Three brain regions known to play roles in ADHD: the hippocampus (important for learning and memory), the prefrontal cortex (the brain area essential for most executive functions), and the striatum (the location of the mesolimbic reward processing circuitry key to ADHD) were also independently sampled. We performed 100 bp single end reads on a HiSeq5000 and got over 10 million reads per sample. Preliminary analysis has focused on changes in transcription between time points and regions based on genotype. Some of the more surprising results indicate that Serpinb3 is involved in the transcriptional regulation of all 3 time points in the hippocampus at 6 and 1 month, but is relatively unchanged at 4 days. While we fully anticipate differential gene expression over time, we are startled by the lack of correlation of genes identified as showing changes between nulls and mutants at Day 4 in comparison to the other time points. This suggests that analysis of even earlier time points might be beneficial and supports the concept that ADHD is a developmental disorder. Further data analysis will emphasize neurotransmitters and their receptors, transporters, and metabolites as well as neuronal development and survival genes. Enrichments for GO pathways will be determined.

1319M Dual-marker lineage specific sorting in heterogeneous Parkinson’s disease patient-specific iPSC-derived dopaminergic neuronal cultures. K. Belle1,2,3, B.A. DeRosa1,2,3, J.M. Van Baaren1,3, J.M. Vance1,3, D.M. Dykxhoorn1,3,1) John P. Hussman Institute for Human Genomics, Miami, FL 33136, United States; 2) Dr. John T. MacDonald Foundation Department of Human Genetics, Miami, FL 33136, United States; 3) University of Miami Miller School of Medicine, Miami, FL 33136, United States.

Parkinson’s disease (PD) is a neurodegenerative disease characterized by the loss of dopaminergic neurons (DA neurons). Model systems and cell lines have helped improve our understanding of PD neuropathology. In addition to failing to faithfully recapitulate the full spectrum of PD phenotypes, animal models lack the genetic architecture of human disease. PD studies in human patient fibroblast and immortalized cell lines are hampered by the fact that they are not the relevant tissue type affected in PD. To better understand PD genetics, we have used reprogramming and derivation techniques to derive stem cell lines genetically identical to patients diagnosed with PD, which, in turn, have been differentiated to relevant neuronal cell types, midbrain dopamine neurons (DA neurons), allowing us to perform knockdown, overexpression and no treatment conditions in cell cycle synchronized neurons to make it a powerful tool for modeling genetic diseases by functional and transcriptomic studies. This is especially important for the study of progressive neurological diseases such as PD. While studies of autopsy brains provide endpoint data, longitudinal in vitro studies to measure disease-related dysfunction and gene expression changes can be performed with iPSC derived cultures. One challenge associated with gene expression studies from iPSC culture is the heterogeneity inherent to the differentiation to specific cell types. This is due to both heterogeneity in both the cellular compositional content and to the fact that any given cell line can only be studied to a limited extent due to time limitations. In order to overcome this limitation, we have designed and tested lentiviral-based fluorescent reporter constructs that allow for the identification and isolation of specific cell types and developmental stages. Utilizing promoter reporter constructs for the transcription factor (TF) and the β-globin (β) constructs as both markers for DA cells, we were able to enrich for DA neurons mature enough to produce dopamine. We have created a dual-marker system by incorporating a previously designed promotervector for Synapsin1 (SYN1) to further increase the efficiency of DA neuron isolation. The β-globin mRNA expression and the β-globin protein expression show increase expression of mature DA markers and related genes. This dual-marker system will be further utilized in gene expression studies of PD patient derived neuronal cultures to measure subtle gene expression differences between the DANs of patients and controls.

1320T Investigating the role of RBFOX1 in human stem cell-derived glutamatergic neurons. H.N. Cukier1, B.A. DeRosa1, S. G. K. Bell1, B.A. DeRosa2,3, A.R. Dykxhoorn1,2,3, B.A. DeRosa1,2,3 1) Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX; 2) Dr. John T. MacDonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 3) University of Miami, College Station, TX; 2) Department of Computer Science and Engineering, Texas A&M University College Station, TX 77843-3474. Autism spectrum disorders (ASDs) are a group of devastating neurodevelopmental conditions that currently afflict about 1 in every 68 children. Since most genetic causes are quite rare, investigations of ASD loci using genome-wide association studies (GWAS) have identified relatively consistent signals of interest. Our group performed a joint analysis of three independent, family-based GWAS datasets for a total of 2.963 families. The top hit with a near genome-wide significance (rs30095508, joint p=6.0 x 10^-6) was in the RNA binding protein, fox 1 homolog 1 (RBFOX1) gene. RBFOX1 had a total of 9 SNPs with p<1x10^-6. RBFOX1 is a neural splicing cofactor that regulates alternative splicing. Both our group and others have identified copy number variants and point mutations in RBFOX1 in ASD individuals. In addition to ASD, RBFOX1 has been connected to bipolar disorder, epilepsy, and schizophrenia. To better understand the impact of RBFOX1 on neuronal functionality, we are evaluating how RBFOX1 expression modulates neuronal function in iPSC-derived glutamatergic neuronal cells. The viability and functionality of the glutamatergic neurons in which RBFOX1 has either been overexpressed or silenced by RNA interference are assessed at multiple time points during in vitro neurogenesis. Overexpression experiments are performed with lentiviral transduction and genomic integration the wild type form of RBFOX1 variant 4 driven under a CMV promoter, as well as variants that are RNAseq analysis of the iPSC-derived neurons under three conditions (RBFOX1 knockdown, overexpression and no treatment) will be used to complement the functional analysis to identify key networks regulated by RBFOX1. Collectively, it appears that RBFOX1 may play a broader role in ASD pathogenesis and other neurological disorders than has been previously suspected.

1321S Mutant dystrophin Dp71Δβ78-79 stimulates cellular proliferation in the inducible system PC12-Tet-On. A. Herrera-Salazar1, J. Aragón1, J. Romo-Yáñez2, A. Sánchez-Trujillo1, V. Ceja1, R. García-Villegas2, C. Montaño1, 1) Genética y Biología Molecular, CINVESTAV-IPN, México, Distrito Federal, México; 2) Fisiología, Biofísica y Neurociencias, CINVESTAV-IPN, México, Distrito Federal, México. Duchenne muscular dystrophy (DMD), a X-linked disease, is a degenerative muscle disorder caused by mutations in the DMD gene resulting in the absence of cytoskeletal dystrophin or the presence of truncated protein. DMD gene has a complex transcriptional regulation that give rise to several tissue-specific proteins named according their molecular weight (Dp427, Dp71, Dp78-79). The family-based GWAS datasets for a total of 2,963 families. The top hit with a near genome-wide significance (rs30095508, joint p=6.0 x 10^-6) was in the RNA binding protein, fox 1 homolog 1 (RBFOX1) gene. RBFOX1 had a total of 9 SNPs with p<1x10^-6. RBFOX1 is a neural splicing cofactor that regulates alternative splicing. Both our group and others have identified copy number variants and point mutations in RBFOX1 in ASD individuals. In addition to ASD, RBFOX1 has been connected to bipolar disorder, epilepsy, and schizophrenia. To better understand the impact of RBFOX1 on neuronal functionality, we are evaluating how RBFOX1 expression modulates neuronal function in iPSC-derived glutamatergic neuronal cells. The viability and functionality of the glutamatergic neurons in which RBFOX1 has either been overexpressed or silenced by RNA interference are assessed at multiple time points during in vitro neurogenesis. Overexpression experiments are performed with lentiviral transduction and genomic integration the wild type form of RBFOX1 variant 4 driven under a CMV promoter, as well as variants that are RNAseq analysis of the iPSC-derived neurons under three conditions (RBFOX1 knockdown, overexpression and no treatment) will be used to complement the functional analysis to identify key networks regulated by RBFOX1. Collectively, it appears that RBFOX1 may play a broader role in ASD pathogenesis and other neurological disorders than has been previously suspected.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1322M
Prenatal malnutrition reprogrammed postnatal gene expression in mammals’ brain. JW. Xu1,2,3, XY. Zhou4,2, YQ. Xiang5,2, T. Wang6,2, L. Ho1,2, XZH. Zhao1,2, 1 Reproductive Medicine Centre, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China; 2 Children’s Hospital and Institutes of Biomedical Sciences, Fudan University, Shanghai, China; 3 Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai, China.

Background: Converging evidence suggests that schizophrenia is a neurodevelopmental disease with environmental influences during early brain development. Prenatal nutritional deficiencies have long been implicated in the etiology of schizophrenia. The birth cohorts affected by the Dutch Hunger Winter of 1944-1945 and 1959-1961 Chinese famine who suffered a nutrition deficiency showed a two fold increase in the incidence of schizophrenia in adulthood. Material and Methods: We established a prenatal “famine” rat model, names RLP50 that induced by maternal exposure to a diet restricted with highly significant (p<0.01) expression differences, the differentially expressed genes were related to synaptic function and transcription regulation. DNA methylome profiling of the hippocampus revealed a systematic epigenetic changes of which majority showed hypermethylation (86.9%). Remarkably, plasma membrane was significantly enriched in both profiling and screens (P = 2.37×10^-9). expression and DNA methylation differential affected the same GO function and hypermethylation in the hippocampus of the RLP50 offspring. The gene associated with cognitive impairment, showed significant down-regulation.

Results: We show prenatal malnutrition reprograms adult prefrontal cortex gene expression in the offspring of RLP50, 415 genes showing modest significant expression differences (P<0.05), of which 48 genes showed highly significance (P<0.01), related to neurotransmitter and olfactory function. In the hippocampus, we identified 2987 genes with significant (p<0.05) expression differences, and 841 genes with highly significant (p<0.01) expression differences, the differentially expressed genes were related to synaptic function and transcription regulation. DNA methylome profiling of the hippocampus revealed a systematic epigenetic changes of which majority showed hypermethylation (86.9%). Remarkably, plasma membrane was significantly enriched in both profiling and screens (P = 2.37×10^-9). expression and DNA methylation differential affected the same GO function and hypermethylation in the hippocampus of the RLP50 offspring. The gene expression and DNA methylation differential affected the same GO function and hypermethylation in the hippocampus of the RLP50 offspring.

Conclusion: We show that the gene expression reprogramming was less significant in the PFC of RLP50 offspring compared with that of the hippocampus, as well as the genome-wide DNA methylation. Our results suggest that the environmental exposure may enhance postnatal brain learning and memory, and therefore contribute to predisposition to schizophrenia.

1323T
Gene expression profiling of human astrocytes treated with bexarotene and related compounds shows an increase in the neuroprotective cytokine GMCSF, R.F. Richholt1,2,3, I.S. Piras4, A.M. Persico5, M.J. Huettelmann1,2, 1) Translational Genomics Research Institute, Phoenix, AZ; 2) Arizona Alzheimer’s Consortium, Phoenix, AZ; 3) Evelyn F McKnight Brain Institute at the University of Arizona; 4) Lab. of Mol. Psychiatry and Neurogenetics, Univ. Campus Bio-Medico, Rome, Italy.

Characteristic neuropathology of Alzheimer’s disease (AD) includes the accumulation of extracellular amyloid plaques in the brain. These plaques are thought to be formed by an imbalance between beta-amyloid (Aβ) production and clearance. Recent studies and in multiple AD mouse models show that treatment with the RXR agonist bexarotene (BEX) restores cognitive functions and in some models results in reduced soluble and oligomeric Aβ. These observations position BEX as a potential agent for AD prevention therapy. RXR and LXR activation has been shown to increase expression of the cholesterol transporters ABCA1 and ABCG1, as well as APOE. These increases were attributed to the benefits of the BEX treatment on Aβ, but they also caused concern regarding its potential use in patients carrying the epsilon 4 allele variant of APOE. How these molecules facilitate Aβ clearance is not fully understood; therefore, we utilized gene expression profiling to investigate BEX and related RXR/LXR agonists in human cells. Human primary astrocytes (Lonza) were treated for 48 hours with 100nM concentrations of the following RXR/LXR agonists - BEX, honokiol, and 9-cis retinoic acid (RA). Gene expression analysis was performed with Illumina HumanHT12 v4 BeadChips and differential expression analysis was performed with the R package Limma. Hierarchical clustering and gene ontology analysis was also conducted. BEX and RA significantly upregulated ABCA1 and ABCG1, and reduced Aβ (p<0.01), but only BEX led to a decrease in the percentage of amyloidogenic Aβ. This suggests that BEX treatment may reverse cognitive impairment and amyloidosis. BEX likely represents a novel approach to upregulate GMCSF across the central nervous system.

1324S
Persistent neurocognitive decline is associated with vascular and epithelial damage to the choroid plexus and β-amyloid plaques in an outbred rat model. A.J. Wyrobek1, S. Bhatnagar1, B. Rabin1, 1 Lawrence Berkeley National Laboratory, Berkeley, CA; 2 University of Maryland, Baltimore Campus, Baltimore, MD.

CNS damage in youth may accelerate neurocognitive decline later in life and hasten the onset of neurological diseases. Understanding the molecular mechanisms of CNS damage after radiation exposures is of special importance for patients receiving cranial radiotherapy, astronauts returning from extended space missions, and has relevance to those experiencing traumatic head injuries. We employed an outbred rat model (Sprague Dawley) to investigate the time-course of molecular and cellular CNS damage after HZE irradiation with 56Fe (1 GeV/n; 10 or 100 cGy), using age-matched sham and young animals as reference. Irradiated rats showed persistent neurocognitive deficits on novel object recognition and bar-press assays. CNS transcriptomic findings pointed to persistent molecular changes in the choroid plexus (CP), the structure that produces cerebral spinal fluid (CSF). Groups of rats were exposed at 2 or 6 months (m) of age and CNS tissue was sampled at 4, 9 or 21 m later. Beginning at 4 m after exposure, there was increased CP endothelial fibrosis especially in the small fenestrated capillary vessels, and the CP epithelium produced less transhyretin (TTR) protein, a major component of CSF. TTR is a β-amyloid-binding protein that facilitates β-amyloid clearance. There was also a progressive increase in β-amyloid plaques beginning at 9 m after exposure. Unirradiated animals did not show age-related changes in CP fibrosis or TTR expression, but showed a small increase in plaque frequency with age. Our findings are consistent with the hypothesis that CNS neurotoxic exposures or trauma early in life may cause persistent CNS dysfunctions that lead to persistent neurocognitive decline associated with the diminished production of molecular factors required for β-amyloid clearance and prevention of amyloid plaque build-up in advanced aging. [Supported by NASA NNX14AC86G (AJW) and NNJ06HD93G (BR) at LBBL under DE-AC02-05CH11231].
Polymorphisms in the TCF4 gene interact with body mass index to influence lithium response among patients with bipolar disorder. E. Ryu1, A. Cuellar-Barboza2, M. Prie0, J. Geske1, C. Colby1, J. Biernacka1,2, M. Frey1. 1) Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Psychiatry and Psychology, Mayo Clinic Rochester, Rochester, MN, USA.

Background: Although lithium is considered as the first line of treatment for bipolar disorder (BD), half of BD patients do not respond to it. Recent studies have indirectly suggested a potential connection between lithium’s protective cellular effects and TCF4 on chromosome 17. TCF4 codes for a transcription factor that is involved in cell survival and neuronal differentiation, together with neuropsychiatric phenotypes including schizophrenia. Therefore, we assessed the association between genetic variants in TCF4 with lithium response to treatment of BD, allowing the genetic effect to be modified by body mass index (BMI), which has also been associated with lithium response.

Methods: Data from 341 BD patients previously treated with lithium was used to assess the association between lithium response (no response, intermediate response, or response) and 20 TCF4 SNPs previously reported to be associated with schizophrenia. Adjusting for age and sex, cumulative logit models including BMI-SNP interactions were applied to jointly test the main effect of each SNP and its interaction with BMI.

Results: Among 341 patients, 52% responded to lithium treatment, 32% showed intermediate response, and 16% had no response. The joint test showed nominally significant evidence of association between two intronic SNPs in TCF4 (rs17512836 and rs17597926) with lithium response (P=0.009). These two SNPs were in almost complete in linkage disequilibrium (r2=0.999) and showed an interaction effect with BMI (P=0.003), indicating that the effect of these SNPs on lithium response depends on BMI.

Conclusion: TCF4 is a member of the basic helix-loop-helix (bHLH) family of transcription factors that have an important role in a number of developmental processes. Our results suggest that the SNPs in this gene may interact with BMI for lithium response. Future work is needed to replicate this finding and to investigate the biological mechanism behind this association.
1328M

Whole genome analysis of high-dimensional phenotypic data: Multiple testing in the context of genome-wide analysis. S.E. Medland, B.M. Neale, C.A. Visscher, N. Jiang, M.J. Wright, P.M. Thompson, P.M. de Ruig, T. Sus代办, S. Lee, N.G. Martin, M.J. Glaun, M. Cur, M. Carless, L. Almasy, D. Glahn, J. Neary, J. Curran, J. Peralta, D. Reese McKay, R. Olvera, R. Kocian, J. Duggirala, J. Blangero, D. Ghina, M. Carless. 1) Dept. Genetics, Texas Biomedical Research Inst, San Antonio, TX; 2) Dept. Psychiatry, University of Texas Health Science Center San Antonio, San Antonio, TX; 3) School of Medicine, Yale University, New Haven, CT; 4) Research Inst, San Antonio, TX; 5) School of Medicine, University of Texas Health Science Center San Antonio, San Antonio, TX; 6) School of Medicine, University of Maryland, Baltimore, MD.

MicroRNAs associated with declarative memory phenotypes. J. Neary1, H. Kulkami2, J. Peralta1, D. Cruz2, D. Reese McKay3, E. Knowles4, P. Fox5, J. Gusrain6, H. Storring1, L. Almasy7, R. Olvera7, R. Kocian8, P. McQuaid9, J. Duggirala1, J. Blangero1, D. Ghina3, M. Carless1. 1) Dept. Genetics, Texas Biomedical Research Inst, San Antonio, TX; 2) Dept. Psychiatry, University of Texas Health Science Center San Antonio, San Antonio, TX; 3) School of Medicine, Yale University, New Haven, CT; 4) Research Inst, San Antonio, TX; 5) School of Medicine, University of Texas Health Science Center San Antonio, San Antonio, TX; 6) School of Medicine, University of Maryland, Baltimore, MD. 

1329T

Posters: Psychiatric Genetics, Neurogenetics, and Neurodegeneration

287
1331M
Expanding the phenotypic spectrum of TUBB4A-associated hypomyelinating leukoencephalopathies. S. Miyatake1, H. Osaka2,3, M. Shinya4, M. Sasaki2,4, J. Tsuchimori1, K. Higashino1, T. Yokota2, M. Morishita1, N. Ando1,2, Y. Ikuta1,2, M. Nakashima1, Y. Tsurusaki1, N. Miyake1, N. Matsuyama1, H. Saito4, 1) Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 2) Division of Neurology, Research Institute, Keio University, Tokyo, Japan; 3) Department of Pediatrics, Jichi Medical School, Tochigi, Japan; 4) Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 5) Department of Child Neurology, National Center Hospital, Tokyo, Japan; 6) Department of Pediatrics, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan; 7) Department of Pediatric Neurology, Takedo Rehabilitation Center for Children, Sendai, Japan; 8) Genetic Counselling and Clinical Research Unit, Kyoto University School of Public Health, Kyoto, Japan; 9) Department of Pediatrics, Graduate School of Medical Science, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; 10) Department of Neurology, Tokyo Metropolitan Children’s Medical Center, Tokyo, Japan.

Leukoencephalopathies are a heterogeneous group of disorders affecting the white matter of the brain. Hypomyelination with atrophy of the basal ganglia is one of those characterized by early-onset motor regression and/or delay followed by extrapyramidal symptoms. A recurrent de novo TUBB4A mutation was recently reported in 11 patients with H-ABC. Interestingly TUBB4A mutations also cause DYT4 dystonia, suggesting that TUBB4A mutations may have a broader phenotypic spectrum. Here we performed whole-exome sequencing analysis of patients with genetically unsolved hypomyelinating leukoencephalopathies, identifying eight patients with TUBB4A mutations, and allowing the phenotypic spectrum of TUBB4A mutations to be investigated. Fourteen patients with hypomyelinating leukoencephalopathies, 7 clinically diagnosed with H-ABC, and 7 with unclassified hypomyelinating leukoencephalopathy, were included. Six heterozygous missense mutations in TUBB4A, 5 of which are novel, were identified in 8 patients (6/7 H-ABC patients (the remaining patient is an atypical case), and 27 unclassified hypomyelinating leukoencephalopathy patients). In 4 cases with parental samples available, the mutations occurred de novo. The effect of the mutations on microtubule assembly was examined by mapping altered amino acids onto 3D models of the ιβ tubulin heterodimer. The p.Glu410Lys mutation, identified in patients with unclassified hypomyelinating leukoencephalopathy, directly impairs motor protein and/or microtubule-associated protein interactions with microtubules, while the other mutations affect longitudinal interactions for maintaining ιβ tubulin structure, suggesting different mechanisms in tubulin function impairment. In patients with the p.Glu410Lys mutation, the clinical course appeared generally milder compared to the typical H-ABC patients. Their basal ganglia atrophy was unobserved or minimal although extrapyramidal features were detected, suggesting its functional impairment. We conclude that TUBB4A mutations cause typical H-ABC. Furthermore TUBB4A mutations associate classes of unclassified hypomyelinating leukoencephalopathies with morphologically retained but functionally impaired basal ganglia, suggesting TUBB4A-related hypomyelinating leukoencephalopathies encompass a broader clinical spectrum than previously expected. Extrapyramidal findings may be a key for consideration of TUBB4A mutations in hypomyelinating leukoencephalopathies.

1333S
Microbiome profiling in whole blood using RNA-seq reveals disease-specific patterns. S. Mangu1, A.P. Ori2, K.A. Staats2, L.M. Olde Loohuis3, L. Junger1, H. Saitsu1, G. Josipin4, J.A. Eisen5, E. Eskin5, R.A. Ophoff6, 1) Department of Computer Science, University of California, Los Angeles, Los Angeles, California, 90095; 2) Center for Neurobehavioral Genetics, University of California, Los Angeles, Los Angeles, California, 90095; 3) Department of Evolution and Ecology, College of Biological Sciences, University of California, Davis, 95616.

Microbial communities in and on the human body represent a complex mixture of bacteria, viruses, archaea and microbial eukaryotes containing in total more than ten times the number of genes found in the human genome. Advances in high-throughput sequencing offer a powerful culture-independent approach to study the underlying diversity of microbial communities of human tissues in health and disease. Availability of comprehensive compendiums of reference microbial genomes and rRNA genes provides ample possibilities to use sequencing data to profile microbial communities present across different human tissues. Reads mapped to microbial reference databases can be used to assign taxonomy to microbial species and identify genes expressed across different samples and subjects. Here, we analyzed RNA-seq datasets from 192 humans from four disease groups (i.e. schizophrenia, bipolar disorder, amyotrophic lateral sclerosis and unaffected controls). Non-mapped RNA-seq reads (failed to map to the human genome) are mapped to a compendium of reference marker genes and microbial genomes and subsequently used to study underlying diversity of microbes in blood. We perform analysis of microbial communities using Phyllosift, a phylogeny-driven method that uses Bayesian methods to assign taxonomic IDs and to compare community diversity between samples. More specifically, using identical reference databases of rRNA and protein coding “marker” genes to profile individual microbial communities and then carries out edge principal components analysis to compare communities. Additionally we validate individual microbial communities using genomic sequence data of the same individual. Using identical procedures we are able to use non-mapped reads to profile diversity of microbial communities. We are able to confirm the microbial genomes at different abundance levels. Using sequencing data from 192 individuals we were able to predict the presence or absence of microbial species in whole blood and the observed differences in disease groups underscores the involvement of the microbiome in human health and disease.
COFS syndrome due to ERCC1 mutation without Nucleotide Excision Repair defect. Y. Capri, N. Calmels, M. Gerard, V. Belavoine, I. Doroz, D. Rodriguez, V. Laugel, O. Boespflug-Tanguy, L. Burglen, 1) Medical Genetics, Robert Debré University Hospital, Paris 19ème, Paris, France; 2) Department of Molecular Genetics, University Hospital Strasbourg, 67091 Strasbourg; 3) Department of Genetics, University Hospital Caen, 14000 Caen; 4) Department of Pediatric Neurology and metabolic diseases, APHP, Robert Debré University Hospital, 75019 Paris; 5) INSERM U1141, Paris; 6) Department of Pediatric neurology, APHP, Trousseau University Hospital, 75019 Paris, France; 7) Centre de détection des malformations et maladies congénitales du cervelet, APHP-Trousseau University Hospital, 75012 Paris; 8) Department of Pediatric Neurology, University Hospital Strasbourg, 67098 Strasbourg; 9) Department of Genetics, APHP, Trousseau University Hospital, 75012 Paris.

Cerebro-oculo-facio-skeletal syndrome (COFS syndrome) is a heterogeneous disease characterized by the association of microcephaly, congenital cataract or microphthalmia, arthrogryposis, severe psychomotor delay, growth retardation and facial dysmorphism. COFS syndrome is due to Nucleotide Excision Repair (NER) defect. ERCC6, ERCC5, ERCC2 and ERCC1 mutations have been involved in COFS syndrome. ERCC1 mutations have been reported in 1 patient with COFS syndrome and in 2 other patients; one patient developed a type II Cockayne syndrome and the other had a severe form of Xeroderma pigmentosum with neurodegeneration symptoms (both syndromes belong to the same spectrum of DNA repair disease as COFS). The patient is the 1st girl of consanguineous Algerian parents. Pregnancy was uneventful but at birth, she displayed distal arthropathy, microcephaly, and severe psychomotor delay, with cerebral dysplasia and cerebellar hypoplasia, and myelination delay. She could not walk, sit nor speak. She died at 5 years of age due to infection. The 2nd pregnancy of the couple was interrupted because of the recurrence of the disease with arthrogryposis observed at 28 weeks of gestation (WG) and cerebellar hypoplasia at 32WG, confirmed by foetal brain MRI. The pregnancy was terminated at 34WG. We report the 4th familial case of ERCC1 mutations. In this family, homozygous p.Phe-231Leu mutation was associated with COFS syndrome but no NER defect was observed. This family confirmed ERCC1 involvement in COFS syndrome but NER defect seemed inconsistent compared to what can be observed when ERCC6 and ERCC5 is mutated. The p.Phe231Leu mutation was already reported in 2 out of the 3 published patients with ERCC1 mutations at the homozygous state, and the other in combination with a nonsense mutation). The pathogenicity had already been confirmed by complementation assay and this mutation was not found in 6,000 control exomes. The absence or low consequence of ERCC1 mutations on NER function in homozygous individuals could suggest a predominance of transcriptional anomalies in COFS/Cockayne syndrome. For clinical relevant patients, functional study is probably not sufficient to decide if ERCC1 is involved. But among the 3 reported patients with the p.Phe213Leu ERCC1 mutation, clinical findings are highly variable and more patients are needed to precise the phenotype and eventually the genotype-phenotype correlation of ERCC1 mutations.
1336S Phenotypic spectrum associated with PTCHD1 deletions and truncating mutations. J.Vincent1, A. Chaudhry2, A. Noor1, B. Degagne3, K. Baker4, J. Field5, A. Bok2, F.R. Parkash6, B. Chen7, R. Bicknell8, S. Yang9, Y. Yin2, C. Cytrynbaum2, D. Dymen10, I. Filges11, B. Helm12, D.J. Stavropoulos12,13, L.J.B. Jeng14, F. Laumonnier15, C.R. Marshall16, S. Parkash17, F.L. Raymond14, A.L. Rideout18, W. Roberts19, R. Rupps20, I. Schanz21, C.T.P.M. Schubben19, J.C. Stevens19, E.P. Thomas22, A. Troutain23, S.W. Scherer24, M.J. Carter25. 1) MIND LAB, CAMH, Toronto, ON, Canada; 2) Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, University of Toronto, Toronto ON Canada; 3) Department of Pathology and Laboratory Medicine, The Hospital for Sick Children, Toronto ON Canada; 4) Department of Medical Genetics, University of Cambridge, Cambridge, United Kingdom; 5) Department of Medical Genetics, Children’s and Women's Health Centre, University of British Columbia, Vancouver BC Canada; 6) Department of Clinical Genetics, Unit of Cytogenetics, Maastricht University Medical Center, Maastricht, the Netherlands; 7) North West Thames Regional Genetics Service, Northwick Park Hospital, Harrow, United Kingdom; 8) The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada; 9) Department of Pediatrics and Adolescent Medicine, Department of Obstetrics and Gynaecology, Centre for Reproduction, Development and Growth, King’s College Hospital, London, United Kingdom; 10) Department of Genetics, Children’s Hospital of Eastern Ontario, Ottawa ON Canada; 11) Division of Medical Genetics, Department of Biomedicine, University Hospital Basel, Basel, Switzerland; 12) Division of Medical Genetics and Metabolism, Children’s Hospital of The King’s Daughters’ Eastern Virginia Medical School, Norfolk, VA, USA; 13) DPLM, The Hospital for Sick Children, University of Toronto, Toronto ON Canada; 14) Department of Laboratory Medicine, University of California, San Francisco, USA; 15) UMR_INSERM U930 Faculté de Médecine, Université François Rabelais, Tours, France; 16) The Center for Applied Genomics, The Hospital for Sick Children, Toronto ON Canada; 17) Maritime Medical Genetics Service, IWK Health Centre, Halifax, NS Canada; 18) Institute for Medical Research Wellcome Trust, University of Cambridge, Cambridge, United Kingdom; 19) Autism Pacific Research Group, The Hospital for Sick Children, Toronto ON Canada; 20) Institute of Human Genetics, University Hospital Magedenburg, Germany; 21) Department of Clinical Genetics and School for Oncology & Developmental Biology (GROW), Maastricht UMC+, Maastricht, the Netherlands; 22) Clinical Genetics Department, Guy’s and St Thomas’ NHS Foundation Trust, London, United Kingdom; 23) McLaughlin Centre for Molecular Medicine, University of Toronto, Toronto ON Canada.

Studies of submicroscopic genomic copy number variants (CNVs) have identified candidate genes associated with autism spectrum disorders (ASD) such as PTCHD1, NRXN1, SHANK2 and SHANK3. Four families have been reported in literature with exonic deletions of PTCHD1 in individuals with ASD and/or intellectual disability (ID), with minimal description of phenotypic features. Here, we describe the largest cohort of individuals with PTCHD1 deletions or truncating mutations. We report phenotype information for 21 males and one female from 15 unrelated families. Eighteen of these individuals have deletions involving PTCHD1 gene while four patients have truncating mutations within the PTCHD1 coding region. Our data suggests that individuals with disruption of PTCHD1may have subtle dysmorphic features including a long hypotonic face, prominent forehead, puffy eyelids and thin upper lip. They do not have a recognizable pattern of associated congenital abnormalities or growth abnormalities. They have mild to moderate global developmental delays, and many have prominent behavioural issues. 45.5% of subjects (10/22) have a diagnosis of ASD or have ASD-like behaviours. The only consistent neurological finding in our cohort is orfacial hypotonia and mild motor incoordination. While detailed neuropsychological studies of individuals with PTCHD1 disruptions are needed to better define the cognitive and behavioural phenotype, our findings suggest that hemizygous PTCHD1 loss of function causes an X-linked neurodevelopmental disorder with a strong propensity to autistic behaviour. Because there is limited information so far on the functionality of PTCHD1, we have not included missense mutations in the current analysis, however it would be of interest to see whether patients with missense changes predicted to be damaging and not present in controls individuals also share some of the subtle features reported here. Likewise, patients with deletions of the region upstream of PTCHD1 were not included, as it is currently not known whether these CNVs exert an effect through PTCHD1 or through another mechanism. Again, it will be interesting to evaluate the clinical features in these patients and to compare to the PTCHD1 deletion and truncation patients reported here.


Sleepwalking and night terrors are sleep disorders classified as excitement parasomnias. These disorders generate a considerable decrease in the individual’s quality of life. There is only one study that previously reported an autosomal dominant inheritance pattern in an American family demonstrating linkage with a region between the SNP type markers rs728331 and rs286819 (Licis, et al. 2011). We report second case in a three-generation Colombian family with 18 affected individuals, with a dominant inheritance pattern and reduced penetrance. However, there are some major differences between the reported family and ours. Onset prevalence and sex-related factors are the major phenotypic distinctions. We propose that we may have a new condition that is caused by different genes.

1338T The transcriptional regulator ADNP links the nBAF (mSWI/SNF) complex with autism. F. Kooy1, G. Vandeweyer1, C. Helsoomt1, A. Van Dijck1, A.T. Vulto-van Silfhout2, B.P. Coes2, J. Gerdts1, L. Rooms1, J. van den Ende1, M. Bakshi1, M. Wilson1, A. Rupps1, L.G. Hendon1, Q.A. Abdulrahman2, C. Romano1, B.B.A. de Vries1, T. Keefstra1, E.E. Eichler3, N. Van der Auwera1. 1) Department of Medical Genetics, University of Antwerp, Antwerp, Belgium; 2) Department of Human Genetics, Nijmeegen Centre for Molecular Life Sciences, Institute for Genetic and Metabolic Disease, Radboud University Medical Center, Nijmegen, the Netherlands; 3) Department of Genome Sciences, University of Washington School of Medicine, Seattle, USA; 4) HDow Hughes Medical Institute, University of Washington, Seattle, USA; 5) Department of Psychiatry, University of Washington, Seattle, USA; 6) Department of Medical Genetics, University Hospital Antwerp, Belgium; 7) Department of Genetic Medicine, Westmead Hospital, Sydney, Australia; 8) Department of Clinical Genetics, Children’s Hospital at Westmead, Westmead, Australia; 9) Clinical Genetics Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; 10) University of Mississippi Medical Center, Jackson, MS, USA; 11) Unit of Pediatrics and Medical Genetics, I.R.C.C.S. Associazione Oasi Maria Santissima, Troina, Italy; 12) Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands.

Mutations in ADNP were recently identified as a frequent cause of syndromic autism, characterized by deficits in social communication and interaction and restricted, repetitive behavioral patterns. Based on its functional domains, ADNP is presumed transcription factor. The gene interacts closely with the SWI/SNF complex by direct and experimentally verified binding of its C terminus to three of its core components. A detailed and systematic clinical comparison of the symptoms observed in our patients allows a detailed comparison with the symptoms observed in other SWI/SNF disorders. While the mutational mechanism of the first 10 patients identified suggested a gain of function mechanism, an 11th patient reported here is predicted haploinsufficient. The latter observation may raise hope for therapy, as addition of NAP, a neuroprotective octapeptide, has been reported by others to ameliorate some of the cognitive abnormalities observed in a knockout mouse model. It is concluded that detailed clinical and molecular studies on larger cohorts of patients are necessary to establish a better insight in the genotype phenotype correlation and in the mutational mechanism.
1339S Prenatal and perinatal risk factors for autism spectrum disorders. A. Anhalt1, L.R. Simard1, X-Q. Liu2, 2 Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada; 2 Department of Obstetrics, Gynaecology, and Reproductive Sciences, University of Manitoba, Winnipeg, Manitoba, Canada. Background: Autism spectrum disorders (ASD) are complex neurodevelopmental disorders. Twin studies have shown that ASD is affected by genetic factors; though, discordancy (e.g. only one twin is ASD-affected) was observed in at least 20% of identical twins. Environmental and epigenetic factors may contribute to the discordance in identical twins. Of these, various pre-/perinatal factors (e.g. abnormal/breech presentation) have been associated with ASD. However, previous studies usually do not account for age at first diagnosis or covariate effects, such as birth year or maternal age at birth. Objectives: The objectives of this study are to investigate the relationships between: 1) ASD and twinning, 2) pre-/perinatal conditions and twinning, and 3) pre-/perinatal conditions and ASD, while incorporating age at first diagnosis and covariates. Methods: We identified families with twins and/or ASD-affected individuals using administrative data from the Population Health Research Data Repository in the province of Manitoba (HIPC#2013-2014-11). We estimated the prevalence of ASD in 5 to 9 year old twins during 5-year periods. We compared the presence of pre-/perinatal complications between twins and singletons. Using survival analysis models, we examined the role of prenatal factors in ASD. Results: After applying quality control measures, we identified ~4000 families with at least one twin pair and ~4000 families with at least one ASD-affected individual; >100 of these families included twins and ASD-affected individuals. We estimated the prevalence of ASD in twins to be 0.56% in 1996-2001 and 1.15% in 2001-2006. Linkage estimates in twins are higher than published estimates for all 5 to 9 year old children during these periods in the province. In our preliminary analyses, twins were more likely than singletons to experience prenatal (OR=1.62; 95% CI=1.32-1.99) or perinatal complications (OR=2.32; 95% CI=2.12-2.54). The overall presence of prenatal conditions was not associated with ASD. Further analyses will examine specific pre-/perinatal conditions and incorporate covariate effects. Significance: Identifying potential risk factors for ASD may aid in developing effective methods for the prevention of autism, and to assist in assessing the vulnerability to ADHD. Our findings will be useful for identifying the role of genetic and environmental factors in the development of ASD.

1339M Evidence of a Genetic Basis for Developmental Topographical Disorientation. S.F. Barclay1, F. Burles2, K. Rancourt3, N.T. Becch-Hansen4, G. Iannoi5, 1 Department of Medical Genetics and Alberta Children’s Hospital Research Institute, University of Calgary, Calgary, Alberta, Canada; 2 Department of Psychology, Hotchkiss Brain Institute & Alberta Children’s Hospital Research Institute, University of Calgary, Calgary, Alberta, Canada. Background: The ability to find our way around familiar and foreign environments, known as topographical orientation, is a complex cognitive task involving many brain regions. A variety of brain lesions can result in topographical disorientation. Distinct from these cases is Developmental Topographical Disorientation (DTD), in which topographical disorientation is present in a subject with no apparent brain structural abnormalities and apparently intact general cognitive skills (Neuropsychologia 47; 30-40, 2009). The defining characteristics of DTD are that individuals (a) have not suffered any brain injury, (b) are not affected by neurological conditions, and (c) have apparently well-preserved cognitive functions; yet they (d) get lost daily in very familiar surroundings and have done so since childhood. In a study of 120 individuals with DTD (Exp Brain Res. 206: 189-96, 2010), one third of the subjects reported one or more family members with similar orientation difficulties, suggesting that DTD may be a genetic disorder. Objectives: To collect phenotypic data from families to determine whether there may be a genetic basis for DTD. Methods: We have developed an online test that includes administrative data from the Population Health Research Data Repository in the province of Manitoba (HIPC#2013-2014-11). We estimated the prevalence of ASD in 5 to 9 year old twins during 5-year periods. We compared the presence of pre-/perinatal complications between twins and singletons. Using survival analysis models, we examined the role of prenatal factors in ASD. Results: After applying quality control measures, we identified ~4000 families with at least one twin pair and ~4000 families with at least one ASD-affected individual; >100 of these families included twins and ASD-affected individuals. We estimated the prevalence of ASD in twins to be 0.56% in 1996-2001 and 1.15% in 2001-2006. Linkage estimates in twins are higher than published estimates for all 5 to 9 year old children during these periods in the province. In our preliminary analyses, twins were more likely than singletons to experience prenatal (OR=1.62; 95% CI=1.32-1.99) or perinatal complications (OR=2.32; 95% CI=2.12-2.54). The overall presence of prenatal conditions was not associated with ASD. Further analyses will examine specific pre-/perinatal conditions and incorporate covariate effects. Significance: Identifying potential risk factors for ASD may aid in developing effective methods for the prevention of autism, and to assist in assessing the vulnerability to ADHD. Our findings will be useful for identifying the role of genetic and environmental factors in the development of ASD.

1340M Linkage analysis of IQ discrepancy in autism: an attempt to replicate. A.Q. Nato1, N.H. Chapman1, H.K. Soh1, R.A. Bernier2, J.M. Viskochil2, H. Coon3, E.M. Wijsmans4, 1 Division of Medical Genetics, Department of Pediatrics, University of Washington, Seattle, WA; 2 Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA; 3 Department of Psychiatry, University of Utah, Salt Lake City, UT, USA; 4 Department of Biostatistics, University of Washington, Seattle, WA, USA. In individuals with autism spectrum disorders, performance IQ (PIQ) is often greater than verbal IQ (VIQ). Linkage analysis of IQ discrepancy (PIQ>VIQ) has identified several genomic regions with evidence of linkage, including the strongest evidence on chromosomes (chr) 10p12 and 17 [Hum Genet 2011:129:59-70]. These signals were not apparent when analyses were rerun without IQ data on participants with and without autism spectrum disorder. IQ data (AGRE and AGP) were unsuccessful. Here, we attempt to replicate these two linkage signals in large multi-generational families from Utah (UT), where IQ data are available on multiple generations.

1341T Genetic Basis of Dynamic Auditory Processing with Application to Reading Ability. J.F. Flax1, M.J. Brunn1, S.L. Wloch2, C.W. Bartlett3, L.M. Brzustowicz3, 1 Department of Medicine, University of Washington, Seattle, WA; 2 Department of Psychology, University of Utah, Salt Lake City, UT; 3 Department of Biostatistics, University of Washington, Seattle, WA, USA. A strong relationship between auditory processing and reading ability has been noted in multiple studies. Theories of auditory processing have proposed that differences in auditory processing may result in developmental dyslexia. This relationship appears to be bidirectional, with auditory processing difficulties affecting reading ability, and reading performance affecting auditory processing. While many studies have focused on differences in auditory processing between those with and without dyslexia, few studies have explored the genetic basis of auditory processing. Here, we will present the results of a linkage scan of the genome focused on auditory processing and reading ability in a large pedigreed sample (248 subjects). This study provides evidence for a genetic influence on auditory processing, with linkage signals on chromosomes 10 and 17.
Identifying endophenotypes associated with Age-related Macular Degeneration in the Amish. M. Pericak-Vance1, R.J. Sardell1, J.N. Cooke Bailey2, W. Coote3, L.D. Adams3, R. Laux3, D. Fuzzell3, L. Reinhart-Mercer3, L. Caywood4, D. Dana5, A.S. Bowman5, M.G. Nittala6, S. Sadda6, J.L. Haines7, D. Stambolian1. 1) Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL; 2) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 3) Departments of Ophthalmology and Genetics, University of Pennsylvania, Philadelphia, PA; 4) Doheny Eye Institute, Los Angeles, CA.

Age-related macular degeneration (AMD), a leading cause of blindness in older adults, is a complex disease with several known genetic and environmental risk factors. However, a substantial portion of heritability in disease risk remains unexplained. In addition to reflecting unidentified rare genetic variants, this missing heritability may also be an artifact of challenges in defining the complex phenotype. Ocular Coherence Tomography (OCT), imaging quantifies fine-scale features of AMD, potentially allowing the identification of endophenotypes associated with the disease. We ascertained both genetic samples and detailed phenotype data via OCT from both affected and unaffected related individuals from Amish populations across Pennsylvania, Ohio and Indiana. The Amish provide an excellent opportunity to analyze the heritability of complex traits given their large nuclear families, while their relatively uniform genetic and environmental background increases power to detect rare genetic variants. We sampled 51 individuals (22 males, 29 females) from 19 Amish families with mean age of 73 (range 49-99). Approximately 8% of individuals had no AMD, 51% had early AMD, 25% intermediate AMD, 9% geographic atrophy and 6% advanced neovascular AMD. In preliminary analyses, we assessed the extent to which the AMD phenotype may be independent of genotype by measuring the correlation of features between left and right eyes. We defined the phenotype using both traditional presence/absence of drusen and quantitative OCT parameters (drusen area and volume). Across all 51 individuals, presence of drusen was strongly correlated between left and right eyes; 94% of individuals were concordant between eyes for presence of small drusen, 86% for medium drusen, and 90% for large drusen. Drusen area in both a 3mm circle (Spearmans r = 0.48) and a 5mm circle (r = 0.72), and drusen volume in a 3mm circle (r = 0.57) were also correlated between eyes. These results suggested individual repeatability of quantitative phenotypic traits. OCT parameters were also correlated with the traditional Age-Related Eye Disease Study (AREDS) scale AMD grade (r = 0.42-0.57), confirming previous studies. These findings confirm the relationship, confirming that OCT parameters may help to define the disease process. Further analyses on the heritability of quantitative AMD traits may therefore enable the identification of endophenotypes associated with this disease.

Gene expression analysis of methamphetamine addicted and schizophrenia patients in correlation with their psychiatric symptoms. A. Haghighatfard1, M. Amini faskhodi2. 1) Department of biology, Science and Research Branch, Islamic Azad University, Tehran, Iran; 2) Department of biology, Tehran medical Branch, Islamic Azad University, Tehran, Iran.

Methamphetamine is a neurotoxin and potent psycho-stimulant of the phenethylamine and amphetamine classes that is used to treat attention deficit hyperactivity disorder (ADHD) and obesity. Schizophrenia is a chronic multifactorial and multigenic, disabling psychiatric disorder that affects about 1% of the population worldwide. Symptoms of schizophrenia including positive, negative and cognitive deficits. majority of methamphetamine addicted have the same psychiatric symptoms like hallucination and bizarre behavior. Researches suggesting that some parts of brain in schizophrenic patients and methamphetamine addicts are neurodegenerated but the mechanisms of this degeneration is not clear. gene expression analysis could help to understand the molecular mechanisms of this neurodegeneration and explain the cause of same psychiatric symptoms in the two groups. this researches also can help to understand the pathogenesis of schizophrenia and methamphetamine effect on brain. We have studied Methamphetamine affects on gene expression in individuals who are addicted (several time in one year), also 50 schizophrenia patients with same age, blood samples collected from 50 addicted ,50 schizophrenia patients and 50 normal person in the same range of age(23-30 years old). RNA of whole blood extracted by column method kit . cDNA synthesized and expression of 3000 gene investigated with DNA microarray technique. expression of 12 gene had been affected in patients and addicts . Over expression of NDUFS1 and NDUFS2 , and decreasing of expression in 5HT2a,DRD2,DRD3,BDNF,DISC1,NRG1,MAOA,bcl2,NDUFW1 and NDUFW2 . Over expressions were significant in schizophrenia & addicts in compare with normal but Over expressions were significant in addicts in compare with schizophrenic patients. Microarray results for two gene NDUFS1 and NDUFS2 confirmed by Real Time PCR syber green method and results analyzed by SPSS and Bonferroni statistical software. In our samples a significant affects of Methamphetamine and schizophrenia on expression of several genes that affects in mitochondria complex 1 subunits genes have been detected.It seems that the mechanism of neurodegeneration in both group is the same but rate of degeneration is higher in methamphetamine addicts . May be that’s why in PANSS psychiatric test the negative symptoms were in same scores for patients. Finally they further argue that altered circadian clock function may contribute to the pathogenesis of intellectual disability.
Early-onset Behr syndrome due to compound heterozygous mutations in OPA1


1) Unité Mixte de Recherche CNRS 6214 - INSERM 1083, Département de Biochimie et Génétique, Centre Hospitalier Universitaire, Angers, France; 2) Département de Neuropédiatrie, Centre Hospitalier Universitaire, Angers, France; 3) Clinique Sourdille, Nantes, France; 4) AP-HP, Service de Génétique Médicale, Hôpital Necker, Paris, France; 5) AP-HP, Service de Neuropédiatrie, Hôpital Necker, Paris, France; 6) AP-HP, Service d'Ophthalmologie, Hôpital Armand Trousseau, HUEP, Paris France; 7) AP-HP, Service de Neuropédiatrie & Centre de Référence de Neurogénétique, Hôpital Armand Trousseau, HUEP, UMPC Univ Paris 06, Inserm U1141, Paris, France; 8) Laboratoire de Biochimie, AP-HP Hôpital Kremlin-Bicêtre, Paris, France; 9) Institut des Neurosciences de Montpellier, INSERM U1051, Hôpital Saint Eloi, Montpellier, France.

The Behr syndrome (MIM#210000) is characterized by the association of early-onset optic atrophy with spinocerebellar degeneration resulting in ataxia, pyramidal signs, peripheral neuropathy and developmental delay. Although the disorder is believed to be inherited in an autosomal recessive manner, it may be clinically heterogeneous, encompassing several genetic aetiologies and patterns of inheritance. Recently, an adult-onset Behr-like syndrome, including optic atrophy and ataxia, was reported in two brothers carrying a heterozygous mutation in the optic atrophy type 1 (OPA1) gene. Heterozygous mutations in OPA1, a gene encoding for a dynamin-related GTPase involved in mitochondrial dynamics and mtDNA maintenance, are the main causes of autosomal dominant optic atrophy (DOA). In DOA, the optic neuropathy occurs insidiously in the first decade of life leading to various levels of visual impairment. As many as 20% of patients with DOA exhibit extra-ocular neuromuscular signs including deafness, chronic progressive external ophthalmoplegia, ataxia, peripheral neuropathy and mitochondrial myopathy with multiple mtDNA deletions, also called the “DOA plus” phenotype. Apart from these autosomal dominant forms, only a few syndromic cases have so far been reported with compound heterozygous OPA1 mutations. Here, we report four cases of children affected by the Behr syndrome associated with compound heterozygous OPA1 mutations. The four unrelated children are affected with a strikingly similar early-onset neurological syndrome associating severe visual impairment due to optic atrophy (4/4), cerebellar ataxia with cerebellar atrophy evidenced by brain MRI (4/4), peripheral neuropathy (4/4), digestive involvement (2/4) and deafness (1/4). These results confirm the importance of searching a compound heterozygosity for OPA1 in severe paediatric cases of complicated optic neuropathy.
1349M
Premorbid psychiatric diagnosis in young persons with 22q11.2 deletion syndrome who later developed schizophrenia. E. Chow1,2,3, A.S. Bassett4,5, S. Wang6, P. Barbalho2, H.M. Gormide, D.N. Nakata, I.L. Cendes, C.V. Maurer-Morelli. Department of Medical Genetics, State University of Campinas (UNICAMP), Campinas, Sao Paulo, Brazil.
Introduction: Neuroinflammatory response may play a role in the pathophysiology of epilepsy. It has been demonstrated that cycloxygenase-2 (COX-2) increases in response to an insult, as seizure. The aim of this study was to investigate the effects of indomethacin (a COX-2 non-selective inhibitor) prior to pentylenetetrazole (PTZ)-induced seizure on mRNA expression of both cox-2 genes in zebrafish larvae brain. Methods: This study was approved by Animal Ethical Committee/UNICAMP (# 3098-1). At six dpf, zebrafish larvae were separated into Seizure+Indomethacin (SG=indo; n=5) and Control+indomethacin (CG=indo; n=5) groups and incubated in indomethacin solution (110µM) for 24 hours. After the incubation, animals from SG=indo were exposed to 15mM PTZ during 20 minutes and the number of seizure-like behavior and latency to seizure onset (Stage 3 of seizure) were analyzed. Animals from CG=indo were handled in PTZ-free water. Other groups: SG - only exposed to PTZ (n=5) and CG-control (n=5) were also investigated. Each sample was composed by pooling 20 heads. Immediately after seizure, animals were anesthetized and their heads collected for reverse transcriptase quantitative-PCR amplifications that were carried out in triplicates with ef1α as an endogenous controls using TaqManTM System. The relative quantification (RQ) was calculated by the equation 

\[ \text{RQ} = 2^{-\Delta\Delta CT} \]

Statistical analyses were performed by Mann-Whitney test with p<0.05. Results: The mearsSEM obtained were: (i) coxa2a: SG=indo 0.5±0.06 vs SG 1.3±0.12 (p=0.094); (ii) cox2b: SG=indo 0.7±0.06 vs SG 1.7±0.18 (p=0.004). Interestingly, when exposed to PTZ, animals pre-treated with indomethacin showed lower latency to reach seizure: SG=indo: 4.6±0.33 vs SG: 2.92±0.17 (p=0.0004) and presented less number of seizure-like behavior response when compared to SG (SG=indo: 11.2±1.5 vs SG: 38.16±4.5; p=0.003). Conclusion: Indomethacin treatment prior to PTZ-induced seizure reduced the cox2a and cox2b mRNA expression levels compared with non-treated animals (SG). Besides, the pharmacological cox-2 inhibition increased the latency to seizure onset and significantly decreased the number of seizures during the 25min exposure. These findings support evidence that zebrafish is a valuable model for further investigations of the main role of inflammation in seizure, as well as a valuable model for anti-inflammatory screening of compounds that are potentially therapeutic for seizures. Support: FAPESP and CNPq.

1350T
Background: Advanced paternal age is associated with increased risk of psychiatric disorders. Attention deficit hyperactivity disorder (ADHD), major depression, schizophrenia, anxiety disorders, autism spectrum disorder and intellectual disability. Other studies have shown that risk for schizophrenia among sporadic cases. Methods: A total of 2450 patients with schizophrenia and 2300 healthy controls were recruited from the Schizophrenia Trio Genomic Research in Taiwan (S-TOGET) project. Early onset was defined as those below 18 years of age of the first psychotic episode. The association between paternal age and onset of schizophrenia was analyzed. Results: Although the sample size is small, results from this study would suggest that paternal age is associated with early onset in schizophrenia in sporadic cases. The association between paternal age and early onset of schizophrenia is consistent with the increase of seizures in spontaneous mutations in sperm as men age.

1351S
Cycloxygenase-2 non-selective inhibitor prior to pentylenetetrazole-induced seizures increases the latency to seizure onset and decreased the number of seizures in zebrafish. P. Barbalho2, H.M. Gormide, D.N. Nakata, I.L. Cendes, C.V. Maurer-Morelli. Department of Medical Genetics, State University of Campinas (UNICAMP), Campinas, Sao Paulo, Brazil.
Introduction: Neuroinflammatory response may play a role in the pathophysiology of epilepsy. It has been demonstrated that cycloxygenase-2 (COX-2) increases in response to an insult, as seizure. The aim of this study was to investigate the effects of indomethacin (a COX-2 non-selective inhibitor) prior to pentylenetetrazole (PTZ)-induced seizure on mRNA expression of both cox-2 genes in zebrafish larvae brain. Methods: This study was approved by Animal Ethical Committee/UNICAMP (# 3098-1). At six dpf, zebrafish larvae were separated into Seizure+Indomethacin (SG=indo; n=5) and Control+indomethacin (CG=indo; n=5) groups and incubated in indomethacin solution (110µM) for 24 hours. After the incubation, animals from SG=indo were exposed to 15mM PTZ during 20 minutes and the number of seizure-like behavior and latency to seizure onset (Stage 3 of seizure) were analyzed. Animals from CG=indo were handled in PTZ-free water. Other groups: SG - only exposed to PTZ (n=5) and CG-control (n=5) were also investigated. Each sample was composed by pooling 20 heads. Immediately after seizure, animals were anesthetized and their heads collected for reverse transcriptase quantitative-PCR amplifications that were carried out in triplicates with ef1α as an endogenous controls using TaqManTM System. The relative quantification (RQ) was calculated by the equation 

\[ \text{RQ} = 2^{-\Delta\Delta CT} \]

Statistical analyses were performed by Mann-Whitney test with p<0.05. Results: The mearsSEM obtained were: (i) coxa2a: SG=indo 0.5±0.06 vs SG 1.3±0.12 (p=0.094); (ii) cox2b: SG=indo 0.7±0.06 vs SG 1.7±0.18 (p=0.004). Interestingly, when exposed to PTZ, animals pre-treated with indomethacin showed lower latency to reach seizure: SG=indo: 4.6±0.33 vs SG: 2.92±0.17 (p=0.0004) and presented less number of seizure-like behavior response when compared to SG (SG=indo: 11.2±1.5 vs SG: 38.16±4.5; p=0.003). Conclusion: Indomethacin treatment prior to PTZ-induced seizure reduced the cox2a and cox2b mRNA expression levels compared with non-treated animals (SG). Besides, the pharmacological cox-2 inhibition increased the latency to seizure onset and significantly decreased the number of seizures during the 25min exposure. These findings support evidence that zebrafish is a valuable model for further investigations of the main role of inflammation in seizure, as well as a valuable model for anti-inflammatory screening of compounds that are potentially therapeutic for seizures. Support: FAPESP and CNPq.

1352M
miR-1202: A Primate Specific and Brain Enriched miRNA Involved in Major Depression and Antidepressant Treatment. J.P. Lopes1,2,3, R. Luna2,3, L. Crapper2, S. Wang1, M. Faraone1,2,3, L. Faggionato2, F. Casano2, B. Labonte2, G. Maussion2, J.P. Yang2, V. Verko2, E. Vigneault3, S. El Mestikawy4, N. Mechawar2, P. Pavlidis2, G. Turecki1,2,3,1) Human Genetics, McGill University, Montreal, Quebec, Canada; 2) McGill Group for Suicide Studies, Douglas Mental Health Institute, Montreal, Quebec, Canada; 3) Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada; 4) Department of Psychiatry, University of British Columbia, Vancouver, B.C., Canada.
Major depressive disorder (MDD), is a prevalent mood disorder that associates with different pleiotrophic brain expression patterns. Treatment of MDD includes a variety of biopsychosocial approaches, but in medical practice, antidepressant drugs are the most common treatment for depressive episodes, and not surprisingly, they are among the most prescribed medications in North America. While they are clearly effective, particularly for severe depressive episodes, there is important variability in how individuals respond to antidepressant treatment. Failure to respond has important implications for a depressed patient and their families. Several lines of evidence demonstrate that genes are regulated through the activity of microRNAs (miRNAs), which act as fine-tuners and on-off switches in gene expression patterns. Here we report on complementary studies using postmortem human brain samples, cellular assays and samples from a zebrafish model. miR-1202, a miRNA specific to primates and enriched in the human brain, is differentially expressed in depressed individuals. Additionally, miR-1202 regulates the expression of the Metabotropic Glutamate Receptor 4 (GRM4) gene and decreases its expression in zebrafish, an effect that is reversed by treatment with the antidepressant fluoxetine. miR-1202 is associated with the pathophysiology of depression and is a potential target for novel antidepressant treatments.
1353T
The NINDS Repository Biomarker Discovery Collection is a Public Resource for Neurodegenerative Disorders. G. Balaburski1, S. Heil1, A. Green1, C. Kopil1, M. Fraser1, M. Sutherland2, K. Gwinnt2, R. Corneva2, C. Tarn1. 1) Coriell Institute for Medical Research, Camden, NJ; 2) National Institute for Neurological Disorders and Stroke-NIH, Bethesda, MD; 3) Michael J Fox Foundation for Parkinson’s Research, NY, NY.
Neurological diseases present numerous challenges: they are devastating to patients and their families, present unique research, diagnostic and clinical difficulties and are an increasing burden on health care systems. Biomarkers for diagnosis, disease onset, disease progression and therapeutic response are urgently needed to improve research, as well as clinical and diagnostic outcomes of patients with neurodegenerative diseases such as Parkinson’s and Huntington’s disease. The mission of the NINDS Repository is to provide genetic support for scientists investigating pathogenesis of the central and peripheral nervous systems through submission and distribution of biomaterials and de-identified clinical data. Moreover, the NINDS Repository facilitates identification and validation of biomarkers for neurological diseases by virtue of being the centralized facility for storage, processing and distribution of biofluids (cerebrospinal fluid, plasma, serum, whole blood, urine) and nucleic acids (DNA and RNA) extracted from whole blood. The Repository minimizes pre-analytical variables by establishing unified standards for sample collection and shipping, securely stores samples with real time monitoring and recording systems, performs standardized laboratory processing with validated operating procedures and maintains the highest levels of quality control and quality assurance. To further incite the investigation and discovery of novel biomarkers the NINDS Repository is establishing large, long term longitudinal collections of biological samples obtained from affected neurologically healthy individuals, as well as from individuals with known genetic mutations. Currently, the NINDS Repository collects samples under either NINDS sponsored biomarker initiatives as well as in collaboration with partners such as the Michael J. Fox Foundation. Current biomarker discovery studies include: Parkinson’s Disease Biomarkers Program (PDBP), Neurobiological Predictors of Huntington’s Disease study (PREDICT-HD), Frontotemporal Dementia MAPT Carrier Study and the Fox Investigation for New Discovery of Biomarkers (BioFIND). Thus, the NINDS Repository’s mission is to provide genetic support for scientists investigating pathogenesis of the central and peripheral nervous systems through submission and distribution of biomaterials and de-identified clinical data. Moreover, the NINDS Repository facilitates identification and validation of biomarkers for neurological disorders. Samples are available upon request from the NINDS Repository web catalog (http://ccr.coriell.org/NINDS), or via NIH-sponsored resources with links to the online catalog.

1354S
SP1 inhibitors as modulators of APP and BACE1 levels in human cells: A novel drug target in Alzheimer’s disease. B.L. Bayon1*, J.A. Bailey2, B. Bay2, K. Sambamurti2, N.H. Greig3, O.K. Lahiji2. 1) Medical & Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN; 2) Laboratory of Molecular Neurogenetics, Department of Psychiatry, Neuroscience Research Building, Indiana University School of Medicine, 320 W 15th St. Indianapolis, IN, USA; 3) Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina, USA. A) According to the amyloid hypothesis, abnormalities in AD are believed to result from the over-production of amyloid-β peptide (Aβ), a product of Aβ precursor protein (APP). Dysregulation of proteins involved in the production of Aβ such as APP and β-site APP-cleaving enzyme (BACE1), may contribute to excessive Aβ deposition. The rate-limiting step in the production of Aβ is the processing of APP by BACE1. Understanding how expression of these proteins is regulated will eventually expose new drug targets. The transcription factor specificity protein 1 (SP1) coactivates the expression of the APP and BACE1 gene. We tested SP1-mediated regulation of APP with Mithramycin A, a selective inhibitor of SP1, and Tolfenamic acid, an inducer of SP1 degradation in human glioblastoma cells U873 and in human neurosphere (NSP) cultures. NSPs were cultured in Neurocult basal media plus differentiation supplements (Stem Cell Technologies). U873 (ATCC) cells were cultured and transfected, and Western blot analysis was performed as previously described (Long et al., JBC-2014). Mithramycin A (Santa Cruz) and Tolfenamic acid (Sigma Aldrich) were prepared in 1µM and 5µM doses. After 72-hour treatment or transfection, cell viability was assessed using CCK Assay (Promega), and protein lysates were analyzed by Western blot. Western blot analysis reveals a significant decrease in the expression of APP in U873 and NSP treated with Mithramycin A. NSP treated with Mithramycin A also exhibit a decrease in BACE1 expression. Treatment with Tolfenamic acid, however, does not significantly decrease APP or BACE1 expression in either cell model. APP siRNA effectively knocks down APP expression in U873 and NSP cultures. BACE1 siRNA and SP1 siRNA did not significantly affect APP levels. CTG showed no significant changes in cell viability among treatments. SP1 and APP expression was decreased after treatment with the SP1 inhibitor Mithramycin A in both U873 and human neurospheres cells. However, APP expression is not affected by treatment with Tolfenamic acid, perhaps due to the differences in the mechanisms between these SP1-inhibiting drugs. We propose that transfection with siRNAs can effectively change the expression of APP and BACE1 in both the human cells. It is essential to ascertain whether drugs or small RNAs targeting this transcription factor could be used to effectively decrease amyloid load and possibly the symptoms of AD in patients.

1355M
Association of Serotonin 2C Receptor Polymorphisms with Antipsychotic Drug Response In Schizophrenia. J. Li1, H. Hashimoto2, H.Y. Mieltzer1, J. Heydari3. 1) Psychiatry, Northwestern University Feinberg School of Medicine, Chicago, IL; 2) Graduate School of Pharmaceutical Sciences, Osaka University, Japan. The serotonin (5-HT)2C receptor (HTR2C) has been implicated in schizophrenia and response to antipsychotic drugs (APDs) through its regulatory effect on dopamine release, interaction with scaffolding proteins at the synapse, and other unknown mechanisms. Evidence from genetic association studies also implicates HTR2C in a variety of neuropsychiatric diseases. We tested the association between HTR2C polymorphisms, Cys23Ser, -795C/T, and -697G/C, and treatment response in 171 schizophrenic patients after treatment with APDs, mostly clozapine, for 6 months. We confirmed that -795C/T, but not Cys23Ser, was a cis-eQTL for HTR2C according to Braincloud data, an integrated database of genome-wide gene expression and genetic control in human postmortem dorsolateral prefrontal cortex (DLPFC) of normal subjects. Ser23 was significantly associated with treatment response at 6 months (positive symptoms, X2 = 7.540, p = 0.01; negative symptoms, X2 = 4.796, p = 0.09) in male, but not in female patients. Haplotype analysis showed that -795C-Ser maintained the same level of significant association with positive symptom improvement (X2 = 6.648, p = 0.01) but additive association with negative symptom improvement (X2 = 6.702, p = 0.01). Logistic regression after controlling for covariates showed these haplotypic associations remained significant with the same direction. Finally, a meta-analysis was performed on six studies with accessible genotyping data for rs6318 and treatment outcome. The overall odds ratio under fixed effect model is 2.00 (95%CI, 1.38-2.91, p = 0.0024) and under random effect model is 1.94 (95%CI, 1.27-2.95, p = 0.0024). In conclusion, HTR2C polymorphisms were associated with response to clozapine in male schizophrenic patients. HTR2C could be a relevant to a broad range of the psychopathology which responds to clozapine in schizophrenia.
1356T
First case of Spinocerebellar Ataxia type 1 in a Mexican female. I. Cervantes1, C.M. Morán2, J. Sanchez3, G. Castañeda4, D.M. Sanchez5, M.H. Orozco1,1, N.Y. Núñez1, J.A. Alcaraz1,3, I.P. Dávalos,5, N.O. Dávalos1, S. Ramírez1, J. Corral5, H. San Nicolás1, L. de Jorge1, V. Volpini1, D. García1.
1) Instituto de Genética Humana, CUCS, Doctorado en Genética Humana Universidad de Guadalajara, Guadalajara, Jalisco, Mexico; 2) División de Medicina Molecular, CIBO, CMNO, IMSS, Guadalajara, Jalisco; 3) División de Genética, CIBO, CMNO, IMSS, Guadalajara, Jalisco; 4) División de Neurocirugía, UMAE Hospital de Especialidades, CMNO, Jalisco; 5) Psicología, ITESO, Tlaquepaque, Jalisco; 6) Universidad de la Sierra Sur, Mihuatlán de Portillo Díaz, Oaxaca; 7) Centro de Diagnóstico Genético Moleculares, IDIBELL, Barcelona, España.

Introduction. Spinocerebellar ataxia type 1 (MIM ID #164400) is an autosomal dominant disease caused by expansion of the CAG trinucleotide repeats in the ataxin-1 gene. Clinically characterized by cerebellar signs, upper motor neuron signs and extensor plantar responses, scanning speech, incoordination, slow motor-nervous conduction and choreiform movements in some instances. Material and Methods. A female aged 24 year-old was studied clinically due to SCA with 10 years of evolution which presented gait ataxia, dysemia, dysarthria, disdiadochokinesis, scanning speech, loss of deep tendon reflexes, mild hypotonia, hypermetric bilateral saccades and normal ocular fundus; the electromyography reported sensorimotor polymyopathy with segmental demyelinization. Brain MRI scans revealed olivopontocerebellar atrophy. Laboratory studies were performed by 5mL of peripheral blood were used to isolate patient’s DNA by GeneCatcher Kit (invitrogen). The molecular analysis was made by PCR. Results. Molecular analysis revealed a number of expansion repeats of 32 exp. Discussion. In previous studies done in Mexican population there was not detected any patient with SCA1, so its frequency is unknown. This is the first time that is described a patient affected by SCA1 in Mexico.

1357S
Allelic distribution of the normal ATXN10 gene in a sample of a Peruvian Amerindian population: an exploratory study. D. Veliz-Ataní1, O. Ortega1, M. Cornejo-Olivas1,2, M. Comejo-Velasquez1,2, K. Milla-Neyra1,2, K. Espinoza-Huertas1, S. Lindo-Samanamud1, M. Inca-Martinez1, V. Marca1, I. Tirado1,2, M.L. Saraiva-Pereira1,2, L. Jardim1,2, P. Mazzetti1,2, Rede Neurogenetica1) Neurogenetics Research Center, Instituto Nacional de Ciencias Neurologicas, Lima, Peru; 2) Northern Pacific Global Health Research Training Consortium, Bethesda, MD; 3) Genetics and Biotechnology School, Universidad Nacional Mayor de San Marcos, Lima, Peru; 4) Laboratorio de Identificação Genética - Centro de Pesquisa Experimental e Serviço de Genética Médica, HCPA; 5) Programa de Pós-Graduação em Genética e Biologia Molecular, UFRGS; 6) School of Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru.

Background: SCA10 is caused by an ATTCT abnormal expansion (locus 13q12.13). In this study we describe the normal allelic distribution as well as the frequency of large normal alleles in this population, suggesting that frequency of large normal alleles is not related to the high prevalence of SCA10 in Peruvian populations. Further analysis with larger sample size and complimentary approaches are required to confirm these results.

1358M
Molecular characterization of genes modifying the age at onset in Huntington’s disease in a group of patients from Uruguay. P. Esperon1, P. Porciello1, M. Vital2, C. Cunha2, E. Bidegain3, A. Gulasso3, V. Raggio1,2, R. Risso3,1.
1) Molecular Biology Unit Clinical Biochemistry Department School of Chemistry, General Flores 2124Ave, Montevideo, Uruguay; 2) Genetic Department School of Medicine, General Flores 2124 Ave, Montevideo, Uruguay.

Background: Huntington’s disease (HD, OMIM #143100) is a neurodegenerative disorder with an autosomal dominant inheritance mode, complete penetrance and variable clinical expressivity. The genetic cause of HD is an expansion of a sequence of CAG trinucleotide repeats located in the first exon of the IT15 gene (region 4p16.3). Though the number of CAG repeats can largely explain the age at onset (AAO), a percentage of variation on AAO persists and can be attributed to modifier genes. Aim: To evaluate the influence on the AAO of HD of: the CAG repeats number and del2642 in the IT15 gene; the E2, E3 and E4 APOE gene alleles, and ADORA2A rs5751876, HAP1-1 and rs57523977, PGC1-1 rs7665116 and UCHL1-1 rs5030732 gene polymorphisms. Patients and methods: 21 patients were recruited from different national institutions. The CAG repeat number determination and gene polymorphisms were done by polymerase chain reaction and High Resolution Melting assays, respectively. Informed consent was obtained from all patients. The study follows the Code of Ethics of the World Medical Association (Declaration of Helsinki). Results: A 63.3% variation in the AAO resulted from the CAG repeat number. The genotype for the aforementioned gene polymorphisms was analysed together with the CAG repeat number and the AAO. No influence of the del2642 of IT15 gene, UCHL1-1, and APOE gene polymorphisms was found. Out of seven patients with solely the ADORA2A polymorphism, four presented anticipation on the AAO. Of the seven patients evaluated in this study, one carried 29 repeats of CAG, one the T allele of HAP-1 and one the C allele of PGC-1α gene. Conclusions: For the first time, in our population, the molecular diagnostic of HD was done. For a more extensive study of molecular aspects, a determination of all CAG repeats in patients’ blood was also performed. Considering a HD disease prevalence of 5/100.000, in a country with only a 3.3 million inhabitant population, the recruitment of a large group of patients is a big challenge. Our results, albeit qualitative, showed a novel approach to the molecular analysis of HD patients.

1359T

MicroRNAs (miRNAs) are short RNA molecules of approximately 22 nucleotides that function as post-transcriptional regulators of gene expression. They are expressed in a tissue specific manner and show different expression patterns in development and disease, and hence can potentially act as disease specific biomarkers. Several miRNAs have been shown to be deregulated in plasma and skeletal muscles of myotonic dystrophy type one (DM1) patients. Here we further expand the miRNAs associated with DM1, report the differential expression of particular miRNAs in whole blood and skeletal muscle samples of DM1 patients, and discuss the potential of these miRNAs as DM1-specific biomarkers. Eleven candidate miRNAs were analysed using quantitative real-time PCR in whole blood (n=10) and muscle biopsy samples (n=9) of DM1 patients, as well as normal control samples (whole blood, n=10; muscle, n=9). In DM1 whole blood samples, miRNA 133a, 29b, and 33a were significantly up regulated, whereas miRNA 1, 133a and 29c were significantly down regulated in DM1 skeletal muscle samples compared to controls. Our findings are aligned to those shown in other studies and point towards pathways that potentially contribute towards pathogenesis in DM1. However currently available data is not sufficient for these miRNAs to be made biomarkers for DM1, as they seem to be common to many muscle pathologies, hence lacking specificity. This reinforces the need for DM1 biomarker study to be further explored as we move towards less invasive, economical, disease-specific, and timely methods of biomarker detection and analysis.
1360S
FMR1-based “Double Hit” model and genomic studies in premutation carriers. R. Lozano1, R. Hagerman1, M. Duyzend2, E. Eichler2, F. Tassone3. 1) MIND Institute, Dept. of Pediatrics, UC Davis Medical Center; 2) Department of Genome Sciences, University of Washington School of Medicine; 3) Howard Hughes Medical Institute Seattle; 4) Dept. of Biochemistry, UC Davis.

The FMR1 premutation is defined as having 55 to -200 CGG repeats in the 5’ untranslated region of the Fragile X Mental retardation 1 gene (FMR1). The clinical involvement has been well characterized for Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) and Fragile X-associated Primary ovarian Insufficiency (FXPOI). The behavior/psychiatric and other neurological manifestations remained to be specified as well as the molecular mechanisms that will explain the phenotypic variability observed in individuals with the FMR1 premutation. We describe a pilot study of copy number variants (CNVs) in 56 participants with a premutation ranging from 55 to 192 repeats. The participants were divided into 4 different clinical groups for the analysis: those with behavioral problems but no autism spectrum disorders (ASD), those with ASD but without neurological problems, those with ASD and neurological problems including seizures, and those with neurological problems without ASD. Results: We found 12 rare CNVs (8 eight duplications and 4 four deletions) in 11 cases (19.6%) that were not found in about 8,000 controls. Three of them were at 10q26 and two at Xp22.3, with small areas of overlap. The CNVs were more commonly identified in individuals with neurological involvement and ASD. The frequencies were not statistically significant across the groups. There were no significant differences in the psychometric and behavior scores among all groups. Further studies are necessary to determine the frequency of second genetic hits in individuals with the FMR1 premutation; however, these preliminary results suggest that genomic studies can be useful in understanding the molecular etiology of clinical involvement in premutation carriers with ASD and neurological involvement in an FMR1-based “Double Hit” model.

1361M
Determination of the origin of Huntington disease based on haplotypes in a Peruvian population. I. Tirado1, 2, C. Kay3, M. Cornejo-Olivas1, 3, J.A. Collins3, M.E. Ketelaar3, S. Lindo-Samanamud1, M. Inca-Martinez1, O. Ortega1, V. Marca1, D. Veliz-Otani1, K. Espinoza-Huertas1, G. Soti2, 4, P. Mazzetti1, 5, M.R. Hayden1. 1) Neurogenetics Research Center, Instituto Nacional de Ciencias Neurologicas, Lima, Peru; 2) School of Genetics and Biotechnology, Universidad Nacional Mayor de San Marcos, Lima, Peru; 3) Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, Canada; 4) Northern Pacific Global Health Research Training Consortium, Bethesda, MD; 5) Department of Genetics and Cell Biology, Universidad Nacional Mayor de San Marcos, Lima, Peru; 6) School of Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru.

Introduction. Recent studies have associated specific HTT haplotypes with the CAG expansion causing Huntington disease (HD) in Caucasian, Asian and African populations. In Europe, HD chromosomes are predominantly associated with haplogroup A while in East Asian populations and the black subpopulation of South Africa, they are associated with haplogroups C and B respectively. There is a reported hotspot of HD cases in Peru, a Latin American country of admixed population with predominance of Spanish European and Amerindian ancestry. Aims. To determine the origin of HD in the affected Peruvian population based on HTT haplogroups and the relative admixture of Amerindian, European, Asian and African genetic background. Methods. Haplotypes of 44 HD chromosomes and 90 control chromosomes from Peruvian HD pedigrees were reconstructed from genotypes at 91 SNPs across the HTT gene region. IRB approval was obtained at Instituto Nacional de Ciencias Neurologicas. Results. HD chromosomes were found mainly on haplogroup A (85%), most specifically on the variant A1 (75%). HD occurs on A1 at the highest frequency (>90%) in the regions of Lima and Central Peru. By contrast, a small proportion (14%) are associated with haplogroup C, which occurs at the highest frequency in the South of the country, where Amerindian ancestry predominates. Conclusions. HD chromosomes from the Peruvian population are predominantly associated with haplogroup A, suggesting a European Caucasoid origin. The locally high frequency of HD on A1 in Lima and Canete, exceeding the frequency of A1 in previously studied Caucasian cohorts, may indicate a founder effect of HD in these regions. The few Peruvian HD haplotypes on C, common among East Asian HD, are likely of Amerindian origin.

Common variants can lead to complex diseases by affecting RNA splicing; such splicing QTLs (sQTLs) have been shown to explain a considerable proportion of variation in isoform usage. Most current approaches to identifying sQTLs attempt to quantify differences in transcripts at the isoform level. However, since sequencing reads/fragments are much shorter than transcripts, isoform-level quantification involves either probability models of transcripts or mapping of reads to a reference transcriptome, instead of the reference genome. Both approaches introduce various challenges and pitfalls. However, most methods only detect sQTLs related to known isoforms. We therefore present a new approach that focuses on event-level quantification of splicing within a gene and apply it to the RNA-Seq data for the 1000 Genomes Project lymphoblastoid cell lines from Geuvadis Consortium (GC). We first map reads to the reference genome using our aligner GSNAp, which performs splice-junction-aware split-read alignments, using known splice sites as well as novel splice sites it detects. Using known variants in the sample it also performs SNP-tolerant alignment and enables allele-specific expression analysis by eliminating the common problem of reference bias. These alignments are then analyzed by our new method exonfindR, which detects known and novel (i.e. ones that cannot be inferred from known transcript annotations) alternate splicing events and quantifies their expression levels in the sample. Quantification is in terms of local events e.g., skipped exon, retained intron etc, and only uses reads that provide unambiguous support for a given event i.e. splice-junction reads in most cases. Association tests are then performed to identify sQTLs. In the data from GC, we detected 37,121 known events in 14,527 genes and 3,703 novel events in 2,515 genes. At FDR cutoff of 5% in the 366 samples of European descent, we found 845 genes (known events) and 81 genes (novel events) with common (MAF ≥ 5%) cis sQTLs, where GC had previously reported 620 genes. Among our sQTLs, 38 (25 previously unreported) disrupt splice sites and likely provide a direct hypothesis for a mechanism and 35 overlap with SNPs in the NHGRI GWAS catalog. Advantages of our approach include greater power to detect association when the variant leads to an event-level change in splicing and the ability to detect novel events and sQTLs associated with them. Both tools, GSNAp and exonfindR, are freely available.

1363M Detection and prediction of deleterious mutations affecting pre-mRNA processing. D.S. Hanna, J.G. Underwood, J.D. Smith, M.O. Dorschner, D.A. Nickerson. Genome Sciences, University of Washington, Seattle, WA.

Synonymous and intrinsic variation is often ignored during human genetic analyses although such variation is known to alter signals recognized by the spliceosome or its associated trans-acting factors. This variation can alter splice site choice or generate cryptic splice sites, both of which can result in aberrant splicing events that lead to nonfunctional coding sequence or nonsense-mediated decay. For the analysis of this underappreciated class of disease-causing variation, we present Spliceosaurus, a lightweight C++ codebase architected using the Standard Template Library (STL) for rapid analysis of variation on splicing signals. Spliceosaurus uses a novel scoring algorithm based on the maximum entropy principle further scaled by weighting cis-acting splicing enhancers and silencers by genomic distance and affinity to establish U1 and U2 snRNA recognition sites on pre-mRNAs. The algorithm iteratively searches not only for variation affecting existing splice sites, but also for variation contributing to the generation of cryptic splice sites. The model is trained on the non-canonical splicing variation found in DBASS that lead to aberrant mRNA isoforms in cancerous tissue samples with specificity in the identification of variation highly impacting the splicing code. Potential irregular splicing events are interrogated by predicting new exon boundaries and the effects on the aberrant mRNA. Unlike existing splicing analysis suites, the input and output for Spliceosaurus is a standard VCF file that enables seamless integration into existing data analysis pipelines. Additionally, the code is capable of processing whole genome datasets in hours, permitting splicing analysis on the scale required for next generation sequencing approaches. We applied Spliceosaurus and correctly identified previously unreported causal intrinsic and synonymous variants in several Mendelian disorders as highly disruptive to canonical splicing [i.e., Atrichocilia/my syndrome (ACS), hereditary nonpolyposis colorectal cancer (HNPC) and X-linked parkinsonism with spasticity (XPS)] The code is freely available for academic use on GitHub: https://github.com/davidhanna/Spliceosaurus.


The rise of whole-genome expression analysis has demonstrated the power of understanding genomic-level changes in gene expression across diseases such as cancer. In addition to the large changes in whole gene expression observed in these genome-wide expression experiments via massively parallel pair-ended RNA sequencing, many novel splice isoforms of genes have been identified across different forms of cancer. However, identification of proteins that cause these splice isoforms is limited to in vitro competitive binding assays or high throughput binding experiments such as cross-linking immunoprecipitation. These assays have been helpful in identifying the space of possible protein-RNA interactions, but often fail to capture condition specific binding events. Here, we propose an integrative approach that uses sequence-level knowledge of alternative splicing events together with RNA-binding profiles and protein-protein interaction networks. We use paired-end sequencing data to identify differential splicing in a model of breast cancer, then apply a hypothesis driven scoring method that identifies putative RNA binding proteins that lead to differential splicing. We then factor in evidence from upstream protein interactions to identify the RNA binding proteins and their associated pathways implicated in the observed sequencing data. Using sequencing data collected from Mena expressing epithelial breast cancer cells, we applied this method to identify and explain alternative splicing events as well as the metastatic phenotype related to the introduction of Mena. The results identified a network of protein interactions that connect Mena over-expression with key alternative splicing events unobserved in the control condition. The network results pinpoint specific splice factors likely responsible for causing splicing changes in response to activation by upstream regulators.

1365S The Exome Coverage and Identification (ExCID) Report: a gene survey tool for clinical sequencing applications. C. Buhay1, R. Sanghvi3, Q. Wang1, K. Walker1, H. Doddapaneni1, J. Hu1, M. Wang1, Y. Han1, H. Dinh3, E. Boerwinkle1, D. Muzny1, R. Gibbs1. 1) Baylor College of Medicine - Human Genome Sequencing Center, Houston, TX; 2) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX.

The Exome Coverage and Identification (ExCID) Report was developed at the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC) to represent gene transcript and exon sequence coverage for samples analyzed with the VCFrme Whole Exome sequencing (WES) reagent. Since March 2013, the report has been used to analyze more than 13,000 WES research samples for the BCM-HGSC and more than 1500 WES clinical samples at our Whole Genome Laboratory (WGL). ExCID assesses target sequencing coverage, annotates targets with gene, transcript and exon information and reports intervals below 20X. In addition, ExCID has run batch analysis features that can compare data from hundreds of samples to reveal trends in the performance of large scale sequencing projects. Results can be visualized as ‘coverage tracks’ in popular browsers such as the Integrative Genome Browser and the UCSC Genome Browser. For a pilot survey, coverage data was examined from 34 WES samples from the WGL. This showed that the VCFrme capture reagent covered (at ≥ 20X) more than 75 percent of the medically interesting genes listed in COSMIC, HGMD, GeneTests, and ACMG Incidental Findings lists. The analysis also identified gene regions that were currently below the 20X threshold. We also analyzed data from Whole Genome Shotgun (WGS) to compare sequence coverage in WGS versus WES. ExCID output was generated on 9 WGS (30X) and 9 WES (100X) to discover and aggregate coding regions below 20X in each method. The data show regions that are potentially important as a consequence of the sequencing strategy, as well as regions unique to each strategy. Ongoing analyses include the aggregation and characterization of poorly covered regions in 600 WGS and 1000 WES samples. Results will provide insight regarding regions of the human genome that require special consideration for the development of future clinical sequencing strategies. ExCID is available under public license in the GitHub repository: https://github.com/cbuhay1/ExCID.
1366M
Scaled Sparse High Dimensional Tests for Localizing Disease Susceptible Sequence Variants. S. Cao1,2, H. Qin3,4, J. Li3,5, H. Deng1,5, Y. Wang1,2,3,4. 1) Dept of Biomedical Engineering, Tulane University, New Orleans, LA; 2) Dept of Biostatistics and Bioinformatics, Tulane University, New Orleans, LA; 3) Center for Bioinformatics and Genomics, Tulane University, New Orleans, LA.

Existing association tests are powerful to identify susceptibility variants in large samples. These tests, however, have low power in high dimensional set (HDS), which is a sample containing fewer study subjects than the number of variants tested. HDS are very common among recently deep sequencing studies. Given the assumption that most of genetic variants are neutral, i.e., only a small portion of them are functional variants, sparse regression methods are powerful to select a promising sparse set of susceptible variants from numerous genetic variants. However, they are ineffective to control the Type I error rates of HDS without functional variants. In this study, we propose a novel scaled sparse regression approach for identifying disease susceptible variants set in high dimensional data. Our method focuses on constructing statistical tests for high dimensional data, based on scaled $L_p$ ($0 < p < 1$) norm regularization. Basically, we propose two significance tests: marker wise test (single variant test) and HDS-based test (whole regional test). For marker wise test, we first apply sparse regression with scaled $L_p$ norm regularization to generate a de-biased solution. Next, we utilize the asymptotical distribution of the de-biased solution to build the significance test. For HDS-based test, we integrate the marker wise statistics to identify susceptible genetic regions. In this step, the dependence among markers is incorporated to appropriately control set-based Type I error rates while maintaining certain statistical power gain. For a wide range of simulations, the proposed approaches appropriately controlled Type I error rates and had improved detection power relative to several prominent methods such as SKAT, Global test and HDI test. When applied to a real DNA sequence data of Mexican Americans for studying hypertension, our HDS-based tests identified an additional significant susceptibility genes which were not reported by SKAT. Our promising results indicate that the marker wise test can pinpoint single suspected variants with higher resolutions, and the HDS-based test yields considerably high statistical power gain for the whole regional test, especially for high dimensional data. In addition, our methods also maintain substantial power for detecting susceptibility variants in low dimensional data or large samples. Last but not least, our method can detect both rare and common variants with almost equal efficiency.

1367T
Serapis: an archival system for large-scale genetic data. I.G. Coligui, N.A. Clarke, M.O. Pollard, J.C. Randall, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SA, United Kingdom.

As the volume of genome sequencing data has rapidly increased over the last few years, the need for high-performance systems that can cope with these storage demands and handle large datasets in an efficient way has become acute. Within the genetics research community, it is desirable to be able to easily access and retrieve large datasets from previous studies of interest. The Serapis archival system has been developed and is currently in use at the Sanger Institute to satisfy these requirements for human genetics research groups. It has been designed for the purpose of archiving large genetic datasets in a long-term storage system and facilitating search of the entire archive based on various criteria. The latter feature has been implemented by attaching relevant metadata to each archived file, stored in the form of key-value pairs. This metadata describes predefined attributes of the data, such as data source (e.g. whole exome/genome sequencing), sample id, associated publication, etc. Though it is possible to use Serapis as a general purpose system, the tool has been tailored to genetics, with built-in support for genetic file formats such as BAM and VCF. It has the ability to automatically collect metadata from within these types of files, by querying external resources, or via manual entry by users. In addition to this, the system also enforces access control restrictions, which is essential in the context of non-public human genetics datasets. In the configuration in use at the Sanger Institute, Serapis uses the IRODS system[1] for backing storage and performs all the work associated with data submission in a distributed way on the Sanger cluster. It is implemented using a distributed task management system (Celery[2]), metadata is assembled temporarily in a NoSQL database before it is permanently archived, and the submission system is accessible via a RESTful web interface. This architecture allows substantial flexibility so that it can be configured to suit the needs of a specific research team or an entire department. The current throughput of our systems is about 1TB of data archived per day (140MB/s), the limiting factor being the I/O performance of the backing store. All components of the system are available as open source code[3]. [1] http://irods.org/ [2] http://celeryproject.org/ [3] https://github.com/wtsi-hgi/serapis-web.

1368S
Leveraging Genome Mapping in Nanochannel Arrays and NGS for a Better Human De Novo Sequence Assembly. H. Dai1, A. Pang1, W. Stedman2, T. Anantharaman1, A. Hastie1, P. Y. Kwok1, A. Umarji2, E. Schad2, R. Sebr2, B.A. Bashir2, H. Cao1. 1) BioNano Genomics, San Diego, CA; 2) Mount Sinai School of Medicine, New York, NY; 3) University of California, San Francisco, CA.

Irys genome mapping technology provides direct analysis of extremely long genomic DNA (up to multi-megabases) without amplification. De novo assembly of these single molecules yields high-fidelity contiguous map information across long ranges. Its advantage over all other genome assembly methods is particular dramatic in highly repetitive regions. Genome maps thus greatly complement assemblies using relatively short second- and third-generation sequencing reads. We have constructed genome maps of human NA12878 (cell line derived from the daughter in the CEU trio) which resulted in a consensus assembly measuring 2.9 Gb and with an N50 of 4.6 Mb. With Pacific Biosciences sequence from the same cell line, we also created a sequence-based assembly with N50 length of 930 kb in parallel. By combining data from these two technologies with a custom-designed hybrid scaffolding pipeline, we were able to generate an assembly having scaffold N50 length of greater than 30 Mb covering more than 2.7 Gb of the human genome. At the same time, we were able to identify potential misassemblies in the sequence assembly as well as in the genome maps by reviewing the inconsistencies between these two complementary technologies. The hybrid scaffolds also discovered additional long range structural variations not identified in sequence assembly.

1369M
Inference for high-dimensional feature selection in genetic studies. C. Ekstrom, Biostatistics, University of Copenhagen, Copenhagen, Denmark.

Feature selection is a necessary step in many genetic applications because the biotechnological platforms provide a cheap and fast means for producing high-dimensional data. This need for dimension reduction is heightened further for example when data from different omics are combined into simultaneous integrated data analysis or when higher-level interactions among the available predictors are considered (which is the case for gene-gene or gene-environment interactions or in epigenetics). Penalized regression models such as the Lasso or the elastic net have proved useful for variable selection in many genetic applications - especially for situations with high-dimensional data where the numbers of predictors far exceeds the number of observations. These methods identify and rank variables of importance but do not generally provide any inference of the selected variables. Thus, the variables selected might be the “most important” but need not be significant. We propose a significance test for evaluating the number of significant selection(s) found by the Lasso. This method leverages the null hypothesis and uses a randomization approach which ensures that the error rate is controlled even for small samples. The ability of the algorithm to compute p-values of the expected magnitude is demonstrated with simulated data and the algorithm is applied to two dataset: one on prostate cancer and a full GWAS. The proposed method is found to provide a powerful way to evaluate the set of selections found by penalized regression when the number of predictors are several orders of magnitude larger than the number of observations.
An integrated framework for sequence variant prioritization. B. Feng, D. Goldgar, Department of Dermatology & Huntsman Cancer Institute, University of Utah, Salt Lake City, UT.

Next generation sequencing has played a key role in genetic research and is now becoming integrated into clinical practice. A major problem in utilizing such information is in the interpretation of the large number of variants discovered from such sequencing, which have diverse characteristics and unknown clinical significance. Variant prioritization is the primary step leading to interpretation, based on which variants or genes or pathways can be filtered, tested or classified. Variant prioritization can be done by many criteria, including, but not limited to, sequence conservation, predicted effect on protein, frequency, co-segregation with disease, and gene/pathway relevance. However, it is not clear how to integrate all this information into a robust and quantitative system without arbitrary filtering or scoring. In this study, we propose a Bayesian framework to combine different lines of evidence for variant prioritization. The first part of this framework is a combined deleteriousness score using the weighted naive Bayes model. By testing on a set of well classified BRCA1 and BRCA2 variants, this score has a higher area under the ROC curve than any component score such as PolyPhen2, SIFT, Mutation Taster, Mutation Assessor, LRT, GERP++, PhyloP, SiPhy, etc., or other combinatorial scores such as CADD, MetaLR, and MetaSVM. Adding minor allele frequencies obtained from the NHLBI GO Exome Sequencing Project (ESP) to the model further improved its performance. The second part of the framework is co-segregation assessment using a modified version of that from Thompson et al., where such data is available. The third part, gene relevance, is inferred from guilt-by-association (GBA) in a protein-protein interaction network. The GBA score is computed by naive Bayes label propagation on an InterConnectedness transformed weight matrix, and then converted to an empirical Bayes factor. This method showed a high level of prediction accuracy among known disease genes. As these lines of evidence are combined into an overall Bayes factor, which can be thresholded for action and/or as a weight in a gene- or pathway-based burden test for association.


Next generation sequencing (NGS) has presented researchers with the opportunity to collect large amounts of sequencing data, which has accelerated the pace of genomic research with applications to personalized medicine and diagnostics. This has resulted in the development of a large number of computational tools and analysis pipelines, necessitating the creation of best practices and reproducible integrative analysis frameworks. Several automated pipelines have been developed to tie together individual software tools, although many of these tools focus on a single NGS platform (i.e. variant calling, RNA-seq), require computational expertise, and/or are poorly documented. The Nature Protocols journal has attempted to create one solution to establish best practices, by publishing step-by-step directions for well-established NGS analysis pipelines. In addition, other best practice pipelines exist, such as the Genome Analysis Toolkit (GATK). Omics pipe (https://biosonnet.org/slab/omics_pipe) is an open-source, modular computational platform that automates best practice multi-omics data analysis pipelines. It currently automates and provides summary reports for two RNA-seq analysis pipelines, variant calling from whole exome sequencing, variant calling and copy number variation analysis from whole genome sequencing, two ChIP-seq pipelines and a custom RNA-seq pipeline for personalized cancer genomic medicine reporting. It also provides automated support for interacting with the The Cancer Genome Atlas (TCGA) datasets, including automatic download and processing of all of the TCGA data. We analyzed 100 TCGA breast invasive carcinoma data sets using Omics pipe with UCSC hg19 RefSeq annotation. Omics pipe automatically downloaded and processed the desired TCGA samples on a high throughput compute cluster to produce a results report for each sample. We aggregated the results for each of the 100 clinical datasets to compare the results of our analysis with the original publications, which revealed high overlap between the analyses, as well as novel findings due to the use of updated annotations and methods. In conclusion, Omics pipe enables researchers to analyze next generation sequencing data with little development overhead to provide reproducible, well as novel findings due to the use of updated annotations and methods.

Tandem repeat sequencing error profiles and error correction models for short read sequencing data. A. Fungammasan 1,2, G. Ananda 1,2, S. Hidayat 3, C. Sun 3, R. Harris 3, P. Maddever 2, K. Eckert 2, M. Makova 2, J. Eckert 2, 1 Integrative Biosciences, Bioinformatics and Genomics Option, Pennsylvania State University, Pennsylvania, 16802 USA; 2 Department of Biology, Pennsylvania State University, Pennsylvania, 16802 USA; 3 Department of Bioinformatics and Computational Biology, The University of Pennsylvania, Pennsylvania, 16802 USA; 4 Department of Pathology, Genetics and Cancer Research Foundation, The Pennsylvania State University College of Medicine; 5 Department of Computer Science and Engineering, Pennsylvania State University, Pennsylvania, Pennsylvania.

Prophiling Tandem Repeats (TRs) from short read sequencing data is a big challenge because of sequencing and mapping errors. We developed a pipeline that can detect the full spectrum of TR lengths from short read data. We used our pipeline to estimate the TR error rates and patterns for PCR containing and PCR-free illumina sequencing libraries using both high-depth genome-wide data and ultra high-depth of plasmid data, which yielded very similar results. We found that the error rates increase with TR repeat number, AT-richness of TR motifs, and PCR step in library preparation. Contraction errors are more common than expansion errors, and the most common type of error is contraction by 1 motif. We used these error patterns to construct an error correction model for genotyping that can assign genotypes correctly for >99.100% of TR loci depending on repeat type. We extended our error models to recommend minimum sequencing depth for genotyping. We found that the minimum sequencing depth required for genotyping increases with TR length. Also, to obtain reliable genotypes, PCR-free data require lower depth than PCR-containing data. Our error correction model has the capacity to distinguish alleles even for TR with consecutive repeat numbers.

MITOCHONDRIAL DISEASE SEQUENCE DATA RESOURCE (MSeqDR) CONSORTIUM: A Centralized Genomic Resource for Analyzing Genetic Variants of Individuals with Suspected Mitochondrial Disease. X. Gai1, L. Shen1, M.A. Gonzalez2, M. Attimonelli3, A. Stassen7, D. Navarro-Gomez1, J. Leipzig2, M. Lott1,2, D.C. Wallace3,5, D. Krotoski5, P.E. Yeske7, S. Zuchner5, M.J. Falk7, E. Eye and Ear Infirmary, Boston, MA; 2 Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 3 Dipartimento di Biotecnologie, Istituto di Biotecnologia e Biofarmaceutica, Università di Bari, Bari, Italy; 4 Department of Clinical Genetics, Maastricht University, The Netherlands; 5 Center for Mitochondrial Medicine and Epigenomic Medicine, The Children’s Hospital of Philadelphia, Philadelphia, PA; 6 NICHD, NIH, Bethesda, MD; 7 The United Mitochondrial Disease Foundation, Pittsburgh, PA; 8 Division of Human Genetics, The Children’s Hospital of Philadelphia, Philadelphia, PA; 9 University of Pennsylvania Perelman School of Medicine, Philadelphia, PA.

Mitochondrial diseases are highly heterogeneous with pathogenic variants in many genes from both the nuclear and mitochondrial genomes. While whole genome sequencing extensively explores the discovery of novel disease genes and the diagnosis of mitochondrial disease patients, it also magnifies the associated bioinformatics challenges. The later include the highly fragmented and limited nature of the current mitochondrial resources, and also the lack of a convenient and standardized way to share data across the community. Since June 2012, we have established a grass-roots effort supported by the United Mitochondrial Disease Foundation (UMDF) called the “Mitochondrial Disease Sequence Data Resource (MSeqDR) Consortium,” a centralized mitochondrial database and data resource (https://mseqdr.org/) by engaging and working with more than 100 disease experts across the globe. Major components of MSeqDR include: a) A Mitochondrial Disease Locus Specific Database (MSeqDR-LSDB) hosting 3,560 potentially disease-causing genetic variants in 1,332 genes, extracted and fully integrated from multiple resources, such as ClinVar, Ensembl, POLG database, and contributions from various diagnostic laboratories; b) A genome browser (MSeqDR-GBrowse) that allows convenient searching, browsing, and visualization of genomic variant annotations of known and predicted disease-causing genes. c) A central database of all disease-causing genes again extracted from, integrated with, and contributed by all major mtDNA genomics resources, namely MitofMap, HmIDB, and PhyloTree; c) Web-based variant annotation and analysis tools, namely HBCR, MT.AT, and MSeqDR-Gtf@MSeqDR; d) Phenotype-centric search and browse of genetic variants, fully integrated with MSeqDR-LSDB and MSeqDR-GBrowse; and last but most importantly e) powerful web-based WES analysis of individual patients, families, and cohorts via GEM.app (https://genomics.med.miamidade.edu/gem/app). The results from these disease-related associations are being added. Together, we believe this resource will support reliable genetic diagnosis in individual cases, facilitate identification of additional cases of rare genetic disorders, provide evidence to exclude from consideration sequence variants of unknown significance, and improve diagnostic efficiency and disease-causing, permit identification of modifier genetic factors underlying disease variability, and even identify genetic variants that are likely to alter response to emerging therapies.
1374S

Whole-Exome sequencing is being used to study human disease primarily by cataloguing single-nucleotide variants (SNVs). Copy-number variants (CNVs) are another form of genetic alterations that have been associated with several disease phenotypes and are harder to detect in whole-exome sequencing data due to the discontinuous nature of the target regions (exons) sequenced. Several algorithms have been published to detect rare CNVs in whole-exome sequencing data, including XHMM, Conifer and Canopy. These three methods work by normalizing read depth data, followed by using a hidden Markov Model to detect CNVs. In one dataset consisting of approximately 2000 samples, we were able to positively verify CNVs of at least 5 kb in size called by XHMM in 189 probands while CNVs in 48 probands were negative resulting in an 80% verification rate. In the second dataset consisting of 650 samples, we validated CNVs in 30 samples using qPCR where 15 CNVs (50%) were positively verified, 9 (30%) were negative, and remaining 6 (20%) could not be validated due to the quality of DNA or short length of the exon involved. In order to increase sensitivity and specificity of called CNVs, we are investigating the use of RNA-Seq to provide additional supporting evidence of CNV calls made from whole-exome sequencing data.

1375M
Branch: An interactive, web-based tool for building decision tree classifiers. B.M. Good, K. Gangavarapu, V. Babji, M. Nanis, A.I. Su. Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA.

A crucial task in modern biology is the prediction of complex phenotypes, such as breast cancer prognosis, from genome-wide measurements. Machine learning algorithms can sometimes infer predictive patterns, but there is rarely enough data to train and test them effectively and the patterns that they identify are often expressed in forms (e.g. support vector machines, neural networks, random forests composed of 10s of thousands of trees) that are highly difficult to understand. In addition, it is generally unclear how to include prior knowledge in the course of their construction.

Decision trees provide an intuitive visual form that can capture complex interactions between multiple variables. Effective methods exist for inferring decision trees automatically but it has been shown that these techniques can be improved upon via the manual interventions of experts. Here, we introduce Branch, a new Web-based tool for the interactive construction of decision trees from genomic datasets. Branch offers the ability to: (1) upload and share datasets intended for classification tasks, (2) construct decision trees by manually selecting features such as genes for a gene expression dataset, (3) collaboratively edit decision trees, (4) create feature functions that aggregate content from multiple independent features into single decision nodes (e.g. pathways) and (5) evaluate decision tree classifiers in terms of precision and recall. The tool is optimized for genomic use cases through the inclusion of gene and pathway-based search functions.

Branch enables expert biologists to easily engage directly with high-throughput datasets without the need for a team of bioinformaticians. The tree building process allows researchers to rapidly test hypotheses about interactions between biological variables and phenotypes in ways that would otherwise require extensive computational sophistication. In so doing, this tool can both inform biological research and help to produce more accurate, more meaningful classifiers.

A prototype of Branch is available at http://biobranch.org/.

1376T
Computational evaluation of the pathogenicity of noncoding sequence variants in autism spectrum disorder. A.J. Griswold1, D. Van Booven2, N. Dueker3, E.R. Martin1,2, M.D. Uccaro1,2, J.R. Gilbert1,2, J.L. Haines3, J.P. Hussmann3, M.A. Pericak-Vance1,2. 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA; 2) Dr. John T. Macdonald Department of Human Genetics, University of Miami, Miami, FL, USA; 3) Institute for Personalized Medicine, Western Reserve University, Cleveland, OH, USA; 4) Hussman Institute for Autism, Baltimore, MD, USA.

Massively parallel whole exome sequencing, whole genome association, and copy number analyses in thousands of individuals have implicated hundreds of genetic loci, including dozens of genes with rare, de novo loss of function and inherited damaging missense variants, in the genetic risk for autism spectrum disorder (ASD). A largely unstudied hypothesis is that rare variants in noncoding regions of the genome functionally contribute to ASD. As part of an ongoing project, we have sequenced evolutionarily conserved regions upstream, downstream, and within introns of 689 ASD associated genes and 942 ASD associated intergenic loci in 2099 ASD cases and 812 controls of European descent. We identified 545,916 single nucleotide variants (SNVs), 519,850 of which are noncoding. To assess which SNVs might contribute to the ASD and to prioritize them for molecular studies, we utilized three bioinformatic tools to annotate and score their functional potential. 1) FunSeq assesses the SNV against the 1000 Genomes database and prioritizes on evolutionary selection and effect of the variant on protein binding motifs. 2) GWAVA (Genome Wide Annotation of Variants) is based on annotation of SNVs with ENCODE and on genome-wide properties such as evolutionary conservation and GC-content. 3) CADD (Combined Annotation Dependent Depletion) integrates multiple functional categories and contrasts variants that survived natural selection with simulated mutations and scores them relative to all possible mutations in the genome. 7,223 variants were in the top 10% of both GWAVA and CADD. 37 variants also had a score greater than 2 in FunSeq and 372 in FunSeq, the top 2% of FunSeq scores, making these our most likely functional noncoding SNVs. Among these are case specific rare variants (MAF<0.01) in highly conserved transcription factor binding sites upstream of ASD candidate genes including fragile X mental retardation 1 (FRAXA) and T-box transcription factor 4 (TBX4). These findings suggest that noncoding variation will enhance our ability to classify SNVs as functional and broaden our understanding of the underlying genetic factors in ASD. This approach can be applied to other disorders generating large scale beyond exome sequencing data.

1377S
Similarity metrics for comparing exome sequence variants. V. Heinrich1, P.N. Robinson1,2,4, S. Mundlos1,2,3, P.M. Krawitz1,2,3,4. 1) Institute for Medical Genetics and Human Genetics, Charité Universitätsmedizin Berlin, Germany; 2) Berlin Center for Regenerative Therapies (BCRT), Charité-Universitätsmedizin Berlin; 3) Max Planck Institute for Molecular Genetics Berlin, Germany; 4) Institute for Bioinformatics, Department of Mathematics and Computer Science, Freie Universität Berlin, Germany.

Removing frequently detected variants is one of the most effective approaches to reduce the number of candidate mutations in the data analysis of next-generation sequencing studies. The incidence of a rare disorder in a population serves as an upper bound for the allele frequency or genotype frequency that can be used as a filter for dominant or recessive disorders. However, the frequentist inference requires that genotypes of a single individual are represented in the database only once. With many and decentralized data submitters the probability increases that samples of the same individual are sequenced multiple times and are contributed independently under different pseudonyms. We compare two different metrics that compute the similarity between any two exome samples in a defined target region. Each similarity value can be used to assess whether a list of sequence variants has already been submitted and provide information about the relationship between samples. This allows the identification of replicates from different enrichment procedures, sequencing platforms and bioinformatic pipelines.
1378M Novel Integrative Genomics Approach for the Discovery of MicroRNA and mRNA Signatures and target Pathways in Prostate Cancer. C. Hicks1,2, R. Jiang1, R. Ramani1, T. Koganti1, S. Girid1, S. Vijayakumar1,2
1) Cancer Institute, University of Mississippi Medical Center, Jackson, MS; 2) Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi; 3) Department of Radiation Oncology, University of Mississippi Medical Center, Jackson, Mississippi.

Background: Prostate cancer is the most common solid tumor and the second most common cause of cancer death in men in the United States. Advances in high-throughput transcription profiling have made possible molecular classification of prostate cancer. However, although these primary analyses have provided valuable insights about the molecular basis of prostate cancer, they have been unsuccessful in determining which genes have causative roles as opposed to being consequences of the prostate cancer state. Advances in genotyping and reduction in genotyping costs made possible identification of single nucleotide polymorphisms (SNPs; herein called genetic variants) associated with an increased risk of developing prostate cancer using genome-wide association studies (GWAS). However, many of the identified genetic variants (>80%) map to noncoding regions with unknown functions. Recently, microRNAs (miRNAs) a class of noncoding RNAs have gained prominence in oncology research as potential clinically actionable biomarkers. However, the functional relationships between genetic variants, miRNAs and their mRNA and pathway targets are not well defined. The objective of this investigation was to leverage GWAS information with miRNA and mRNA expression data for the discovery of MicroRNA and mRNA oncogenic signatures and pathways in prostate cancer. Methods: We used a combination of bioinformatics tools to integrated disparate genetic data in our analysis. First, we developed a comprehensive catalogue of genetic variants and associated genes from GWAS and a comprehensive catalogue of miRNAs and their targets. Next, we analyzed miRNA and mRNA expression data to identify miRNA and mRNA signatures enriched for genetic variants. Subsequently we performed network and pathways analysis to identify target molecular networks and biological pathways enriched for genetic variants. Results: The analysis revealed miRNAs and mRNA oncogenic signatures. We identified molecular networks and pathways enriched for genetic variants targeted by miRNAs. Among the identified target pathways included the androgen, androgen biosynthesis, IGF-1, JAK, STAT and prostate cancer signaling pathways. Conclusions: Our investigation demonstrates that integrative genomic analysis provides a powerful and unified approach for establishing putative functional bridges between GWAS findings, miRNAs and their mRNA and pathway targets.

1379T Medical re-sequencing analysis pipeline provides one-stop solution for identifying disease-causing mutations of Mendelian disorders. H. Hu1, T.F. Wienker1, L. Musante1, V. Kalscheuer1, P.N. Robinson2, H.H. Ropers1
1) Max Planck Institute for Molecular Genetics, Berlin, Germany; 2) Universitätsklinikum Charité, Berlin, Germany.

Next-generation sequencing has greatly accelerated the search for disease-causing defects, but even for experts the data analysis can be a major challenge. To facilitate the data processing in a clinical setting for Mendelian disorders, we have developed a novel Medical Re-sequencing Analysis Pipeline (MERAP). MERAP assesses the quality of sequencing, and has optimized capacity for calling variants, including Single Nucleotide Variants, insertions and deletions, Copy Number Variation, and other structural variants. MERAP identifies polymorphic and known causal variants by filtering against public-domain databases, and flags non-synonymous and splice-sites changes. MERAP uses a logistic model to estimate the causal likelihood of a given missense variant. MERAP considers the relevant information such as phenotype and interaction with known disease-causing genes. The MERAP algorithm performs superiorly compared with GATK, one of the widely used tools, because of its higher sensitivity for detecting indels, its easy installation, and its economical use of computational resources. Upon testing more than 1,800 individuals with mutations in known and novel disease genes, MERAP proved highly reliable, as illustrated here for 5 families with disease-causing variants. We believe that the clinical implementation of MERAP will expedite the diagnostic process of many disease-causing defects for Mendelian disorders.

1380S A novel integrated analysis framework for detecting genome-wide changes in gene expression or regulation with next-generation sequencing data. W. Huang, O.M. Urbach, L. Li. Biostatistics, NIEHS, NIH, Research Triangle Park, NC.

The cost-effective next-generation sequencing has quickly become a popular assay method for studying gene expression and genetic/epi-genetic regulation. The development of corresponding data analysis tools, as demanded by these sequencing applications, has also been fast paced, particularly in the last five years. The sequencing technologies continue to evolve so quickly, however, that the development of analysis tools still lags far behind the need. In particular, sequencing is now being used for more complicated or more challenging studies, and existing tools are no longer sufficient for these new and more advanced applications. In fact, data analysis and interpretation is becoming the bottleneck that limits more widespread sequencing applications. The main inadequacies of existing tools include: 1) ineffective at detecting true biological changes, especially when number of biological replicates is small, 2) high false positive rates, typically resulting from unrealistic underlying statistical assumptions, 3) little capability for analyzing data from more advanced experimental designs, 4) limited data visualization capability, and 5) typically restricted to one type of sequencing application. We present a new computational framework, based on our previous work (EpiCenter), to address the above challenges. Instead of using a single procedure or statistical test for every application, our new framework employs different statistical tests and filters to analyze data from different kinds of study. Specifically, our framework makes the choice according to the experimental design, number of biological replicates, type of data (e.g., histone ChiP-seq or mRNA-seq gene expression), and the organism/species (genome complexity and annotation can differ dramatically among organisms) used in the study. In addition, our framework, unlike existing methods that treat individual genes/regions separately, considers individual genes/regions as an integrated part of a genome-wide study to improve overall detection accuracy. Furthermore, our framework extends EpiCenter's data visualization capability including the generation of the UCSC hub track data. Evaluation on both simulated and actual sequencing data showed the superior performance of our framework over that of some popular existing tools.

1381M Database of disease-associated genomic polymorphisms based on assessment of reproducibility between or within human populations. T. Imanishi1,2, Y. Nagai1, Y. Takahashi4, 1) Tokai University School of Medicine, Isehara, Kanagawa, Japan; 2) National Institute of Advanced Industrial Science and Technology, Tokyo, Japan.

Genome-wide association studies (GWASs) have identified numerous single nucleotide polymorphisms (SNPs) that are associated with development of common diseases. However, because GWAS uses statistical evaluation, we cannot completely eliminate false positives that may contaminate to a certain extent. On the other hand, it is becoming clearer that genetic risk factors of common diseases are not totally universal but heterogeneous among human populations. We thus developed a new database of genomic polymorphisms that are reproducibly associated with disease susceptibilities, drug responses and other traits for each human population, and released it as “VarySysDB Disease Edition (VaDE)”. Using PubMed and NHGRI GWAS catalog, we collected 1806 GWAS papers and curated them manually. We extracted information of associated SNPs, odds ratios, p-values, study design, nationality of subjects, and many others. Also, external reproducibility assessment of SNPs carried out separately for hypertension and rheumatoid arthritis. Then, we assessed the reproducibility of each association in multiple, independent studies for each human population. Finally, we could obtain 4675 and 790 reproducible associations for 427 and 151 traits in the European and East Asian populations, respectively. Furthermore, to support finding functional SNPs in the VaDE database, we integrated data of ChiP-seq, DNasel hypersensitivity experiments, regulatory motifs, RefSeq genes, H-Inv transcripts, and linkage disequilibrium data in three major human populations that have been obtained from haploreg v2, VarsysDB, and Michigan. We also added a genome browser to visualize these data in VaDE. From the VaDE database, we searched for pleiotropic SNPs that are associated with multiple traits. As a result, we could identify 802 pleiotropic SNPs, such as those affecting metabolic traits and autoimmune diseases, which provides a clue to solve hidden relationships among various phenotypes. The VaDE database is publicly available from http://bmi-toki.jp/VaDE/. We believe that our database will contribute to the future establishment of personalized medicine and understanding of genetic factors underlying diseases.
Semantic Similarity Analysis of Patient Phenotypes for Genome Wide Genetic Diagnostics. R. James, M. Bainbridge, C. Eng, C. Shaw. Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

The clinical implementation of genome wide diagnostics is increasingly important in the molecular diagnosis of human disease. However, these genome wide approaches are challenged by the large number of variants observed in individual patients. We hypothesize that efficient integration of available clinical data characterizing patient indications for genetic testing can improve the implementation of diagnostic workflows and, ultimately, prioritization of variants. We have developed scalable quantitative analysis tools for mining indication content. Fundamentally, our approach exploits semantic similarity computed on patient phenotypes to comprehensively represent patient features in simple composite scores. Our tool can help to elucidate cohort substructure by segmenting large, heterogeneous groups of subjects into subgroups based on shared multidimensional phenotypes. Identification of this structure can help to identify and establish cohorts for associative or experimental study, accelerating the investigation of genetic variation. Identification of cohorts can also lead to discovery of novel functional variants that drive complex phenotypes. We have also developed a prototype graphical interactive data viewer exploiting our computational toolkit. This viewer can aid decision support through visual query and analysis of structured input data. This tool can help simplify identification of clinically relevant variants by facilitating indirect matching of patient phenotypes to cataloged variants already known to be associated with semantically similar patients or reported OMIM diseases. We use our tools, in the context of the Human Phenotype Ontology representation of patient indications, to analyze a pilot cohort assembled from retrospective data from the Whole Genome Diversity Panel data and show that GAMMA identifies several variants that drive complex phenotypes while other methods either fail to detect the true effects or result in many false positive identifications. While none of them consider population structure and may result false positive identifications. Here, we introduce a new methodology referred to as GAMMA that could both simultaneously analyzes many phenotypes as well as corrects for population structure. In simulated study, GAMMA accurately identifies true genetic effects without false positive identifications. While other methods either fail to detect the true effects or result many false positive identifications. We further apply our method to a real Hybrid Mouse Diversity Panel data and show that GAMMA identifies several variants that are likely to have a true biological mechanism.
1386S

**DISTMIX:** Direct imputation of summary statistics for unmeasured SNPs from mixed ancestry populations. D. Lee, T. Bigdell, B. Riley, A. Fanous, S. Bacarru. Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA.

**Background:** To improve signal detection power, many meta-analyses of genome-wide association studies use mixed-ancestry cohorts. To analyze these diverse cohorts while also increasing imputation accuracy, large multi-ethnic reference panels are used. However, the ever increasing sizes of reference panels and study cohorts makes genotype imputation computationally unfeasible using moderately sized computer clusters. Moreover, due to modelling assumptions, the commonly used genotype imputation methods lack the ability to impute genotypes for family data. Besides, genotype imputation requires subject level genetic data, which unlike summary statistics provided by virtually all studies, are not publicly available.

**Method:** To overcome these limitations, we propose a novel method/software for Direct Imputing summary STAtistics for unmeasured SNPs from MIXed ancestry population (DISTMIX) without imputing first subject level genotypes. Based on the correlation matrix of allele frequencies from the study and the ethnicities in the reference panel, DISTMIX first infers the proportions of reference panel ethnicities in the study cohort. These proportions are subsequently used as weights in calculating the ancestry-weighted genotype correlation matrix and the mixture of ethnically-specific correlation matrices, as estimated from the mixed ethnicity reference panel. This mixture correlation matrix is subsequently used by DISTMIX to impute summary statistics of unmeasured SNPs using the conditional expectation formula for multivariate normal variables.

**Results:** Experiments based on our simulated null data sets with different types of ancestry show that the proposed method controls the false positive rates. 1000 Genomes based imputation of summary statistics from the ethnically diverse Psychiatric Genetic Consortium Schizophrenia Phase 2 (PGC SCZ2) suggests that, when compared to common genotype imputation methods, our method offers comparable imputation accuracy while requiring only a fraction of computational resources. We also present some new and interesting genetic findings for PGC SCZ2 cohort, which were obtained by increasing the size of the reference panel (and likely, imputation accuracy) through the combination of the 1000 Genomes and UK10K reference panels.

1387M

**Detecting complex fusion transcripts in pediatric cancer using a novel assembly algorithm:** CICERO. Y. Li1, T. Boz1, M. Rusch1, J. Easton1, K. Boggs3, B. Vadodaria3, P. Gupta1, G. Song1, J. Ma2, C.G. Mullighan2, S.J. Baker1, R.J. Gilbertson1, J.R. Downing2, D.W. Ellison2, J. Zhang1. 1) Department of Computational Biology, St Jude Children’s Research Hospital, Memphis, TN; 2) 2Department of Pathology, St Jude Children’s Research Hospital, Memphis, TN; 3) 3The Pediatric Cancer Genome Project Validation Laboratory, St Jude’s Children Research Hospital, Memphis, TN; 4) 4Department of Developmental Neurobiology, St Jude Children’s Research Hospital, Memphis, TN.

Fusion genes are important for cancer diagnosis, subtype definition and targeted therapy. Although RNASeq is useful for detecting fusion transcripts, computational methods to identify fusion transcripts arising from internal tandem duplication (ITD), that have multiple partners, low expression or non-template insertion are limited. We developed an assembly-based algorithm CICERO (CICERO Is Clipped-reads Extended for RNA Optimization) that is able to extend the read-length spanning fusion junctions for detecting complex fusions. Using test data that include RNASeq from 3 ependymoma (EPD), 39 low-grade glioma (LGG), and 128 acute lymphoblastic leukemia (ALL), we have shown that CICERO is able to detect multi-segment fusion transcripts resulting from chromosomal internal tandem duplication or rearrangement at a highly repetitive immunoglobulin (IG) locus; all of which would be missed by existing fusion analysis methods. The overall sensitivity and accuracy of CICERO are much higher compared with existing tools such as deFuse and Tophat-Fusion. Using CICERO, we analyzed >600 leukemia transcriptomes from the St Jude/Washington University Pediatric Cancer Genome Project (PGP) and detected recurrent C11orf95-REL fusion in EPD, FGFR1 ITD in LGG, NTRK fusion in high-grade glioma and activating kinase fusions with multiple partners in ALL.

CICERO shows high sensitivity when detecting fusions with low expression, like BRAF fusions in LGG, making it useful for identifying subclonal lesions and for analyzing tumor specimens with low purity. Furthermore, the power of CICERO increases with the extended read-length enabled by improvements in next-generation sequencing (NGS) technology. Using paired-end 300bp RNASeq reads, CICERO shows the ability to assemble near full-length fusion transcripts and identify complex fusions with multiple segments.

1388T

**Mixture modeling of next generation sequencing data and its application to estimating genotype frequency.** J. Li1, S. Yoon2, S.J. Finch3.

1) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 2) Stony Brook University, Stony Brook, NY.

Studying a probability of an individual having a variant, inherited or not, is of interest in genetic studies. We present a method for modeling the frequency of single nucleotide polymorphism variants in the exome capturing sequence data of an individual. Mixture modeling is used to model the proportion of alternative allele at a position under the bi-allelic single nucleotide polymorphism model. We suggested the application of this mixture modeling in this paper. We measured the proportion of alternative allele for positions in chromosome 1 exome sequencing data of two trio from pilot study 3 in the 1000 genomes project. The measurement was based on the counts of reference and alternative allele in the pileup file from SAMtools. We fit the proportions to mixture model of two point distributions and five continuous distributions. The fitted mixture model well described the properties of the distribution of the alternative allele proportions. The estimates of mixing proportions were used to estimate the fraction of each genotype in the data. Each individual had different estimates of parameters, but the estimates of genotype frequencies were similar. We found the he estimated fractions of each trio clustered. We developed an expectation-maximization algorithm to obtain the maximum likelihood estimates of the mixture model parameters. Its application to estimate the fraction of each genotype was presented in this paper.

1389S

**A high-performance database framework for fast and easy prioritization of disease related variants from Exome Sequencing data.** B. Linghu, F. Yang, R. Bruccoleri, J. Szustakowski. Novartis Institutes for Biomedical Research, Cambridge, MA.

Exome sequencing (Exome Seq) has become a promising approach to identify disease related genetic variations. Pinpointing the small subset of pathogenic mutations amongst the thousands or millions of variants generated in an Exome Seq experiment remains a conceptual and computational challenge. One common approach is to use relational database systems to conveniently organize and query variant data for prioritization. However, traditional database systems often perform poorly when applied to such “Big Data”. Recently, a number of high-performance database systems have been developed specifically to enable analysis of extremely large data sets. Here we describe applying one such system, namely Vertica, to prioritize disease variants. Our approach leverages Vertica’s high performance capabilities to efficiently model, store, and query a comprehensive landscape of information including variant calls, variant quality metrics, predicted functional consequences, allele frequencies, disease prior knowledge, inheritance patterns, and clinical phenotypes. This framework enabled the convenient and efficient identification of candidate disease variants, with significant improvements over traditional databases. To our knowledge, this is the first demonstration that high-performance databases such as Vertica provide an efficient solution to prioritize variants from exome sequencing.
1390M Analysis of Human neurodevelopmental disorders from the systems biology perspective using the Lynx Platform. N. Maltsev1, D. Sulakhe2, D. Boerregen1, B. Xie1, A. Taylor1, A. Paciorowski2, W. Dobyns1, T.C. Gilliam1. 1) Human Genetics Department, University of Chicago, Chicago, IL, USA; 2) Computation Institute, University of Chicago, Chicago, IL, USA; 3) Departments of Neurology, Pediatrics, and Biomedical Genetics, University of Rochester Medical Center, Rochester, NY, USA; 4) Center for Integrative Brain Research, Seattle Children's Research Institute and Department of Pediatrics, University of Washington, Seattle WA, USA.

Understanding of genetic mechanisms underlying common hereditary disorders (e.g. autism, schizophrenia, brain malformations) is one of the most important challenges of biology. Recent advances in biology provided new perspectives onto studies of complex heritable disorders, including: network-based view of Human disorders; emergence of "phenomics" and a notion of "interrelatedness" of disease traits. However, the progress of biomedicine facilitated by these methodologies depends on availability of the new bioinformatics platforms supporting the development of integrative models of genetic disorders and identification of their characteristic multidimensional patterns of inheritance. We present an approach and a supporting computational platform Lynx (http://lynx.ci.uchicago.edu/) for analysis of common heritable disorders from the systems biology perspective. Our approach is three-fold and includes tools for the enrichment analysis, gene prioritization and the development of a network-based models of biological processes in health and disease. These tools utilize genomic, functional, clinical and experimental information from the Lynx knowledge base integrating over 35 public and proprietary databases and information derived from the literature using advanced text mining. We will illustrate our approach using an example of the "interrelatedness" of autism and schizophrenia, autism and infantile epilepsy and brain malformations has been high, discovery of the underlying genetic mechanisms that lead to heritable brain disorders such as autism, infantile epilepsy and brain malformations has been high, while interest in the biological mechanisms that lead to heritable brain disorders such as autism, infantile epilepsy and brain malformations has been high.

Previous studies of developmental brain disorders has led to understanding that "interrelatedness" between molecular components postulated by the systems approach will be essential for understanding both individual susceptibilities and potential therapeutic interventions. The presented approach can be used for comparative analysis of a group of developmental brain disorders that are specifically tailored to RNA-Seq data. These approaches can facilitate the work-up of disorders like dystonia that present large clinical variability and clinical overlapping for the different involved genes. WES is also the technique of choice for clinical practice as it is the new bioinformatics platform supporting the development of integrative models of genetic disorders and identification of their characteristic multidimensional patterns of inheritance. We present an approach and a supporting computational platform Lynx (http://lynx.ci.uchicago.edu/) for analysis of common heritable disorders from the systems biology perspective. Our approach is three-fold and includes tools for the enrichment analysis, gene prioritization and the development of a network-based models of biological processes in health and disease. These tools utilize genomic, functional, clinical and experimental information from the Lynx knowledge base integrating over 35 public and proprietary databases and information derived from the literature using advanced text mining. We will illustrate our approach using an example of the "interrelatedness" of autism and schizophrenia, autism and infantile epilepsy and brain malformations has been high, discovery of the underlying genetic mechanisms that lead to heritable brain disorders such as autism, infantile epilepsy and brain malformations has been high, while interest in the biological mechanisms that lead to heritable brain disorders such as autism, infantile epilepsy and brain malformations has been high.

Previous studies of developmental brain disorders has led to understanding that "interrelatedness" between molecular components postulated by the systems approach will be essential for understanding both individual susceptibilities and potential therapeutic interventions. The presented approach can be used for comparative analysis of a group of developmental brain disorders that often co-occur and share at least a subset of causative genes (e.g. autism, agenesis of corpus callosum, mid-hindbrain malformations). Our analysis allowed us to identify genes involved in esophageal adenocarcinoma cancer patients. M. Matvenko1, B. Øster2, A. Joecker1, A.-M. Hein1, P. Dekker1, R. O’Neill2, A. Krejci1, A. Arends1, N. Thomson1, C. Boyssen2, S. Mansted1, R. Forberg2, B. Knudsen1, T. Knudsen1, J.C. bio, A. Qian Company, Denmark; 2) CLC bio, a QIAGEN Company, Danmark A/S; 3) Signal Transduction Group, University of Edinburgh Cancer Research Centre in the Institute of Genetics and Molecular Medicine, Crewe Road South, Edinburgh EH2 2LR, United Kingdom; 4) Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Ústí nad Labem 766 53 Brno, Czech Republic.

High throughput sequencing technologies are currently revolutionizing the cancer research area with rapid improvements in sequencing capacity and time consumption. As a result the most time consuming step has moved from being the sequencing process itself to being the bioinformatic data analysis. RNA sequencing (RNA-Seq) is used in an increasing number of transcriptomic studies. The great advantage of using RNA-Seq data is the ability to precisely quantify transcript levels and identify novel transcripts, isoforms, and splice junctions, while further providing information of the mutational landscape down to single base resolution. To ease the hurdles associated with RNA-Seq data management there is an increasing demand for tools that are specifically tailored to RNA-Seq data. Here we describe how the newly developed CLC Cancer Research Workbench can be used to analyze and visualize RNA-Seq data with ready-to-use workflows that automatically map, quantify, and annotate transcriptomes. We identify differentially expressed genes and transcripts in Illumina HiSeq data from three esophageal adenocarcinoma cancer patients, and compare the mutational patterns in the samples with the expression values of the corresponding genes. Results are visualized in a track based browser view, which provides easy navigation as well as allowing the user to simultaneously view and annotate multiple samples and different data types (e.g. genes, transcript expression levels, and detected variants).


A core tenet of biological study is that research by one scientist or organization must be reproducible -- and thereby verified -- by third parties. The NGS analyses that are now dominating genomic research currently employ open-source computing algorithms. While these algorithms are powerful, flexible and readily accessible, they have a drawback when it comes to creating research pipelines that can be independently reproduced and verified by other researchers. Specifically, each algorithm has its own options, variables and formats that govern the manner in which it is executed. Furthermore, the algorithm’s executable code itself may go through multiple variations and versions. When these algorithms are then combined into a pipeline, the same variability applies exponentially. Unless all of these factors can be held constant across datasets, locations, and users, research will not be reproducible and verifiable by third parties. Several groups have developed software systems with Graphical User Interfaces designed to ease the manipulation of open source algorithms; however, none embodies a mechanism to enable reproducible research. BioDatomatics set out to design and create a mechanism to facilitate reproducible research for its BioDT analytics platform. The solution turned out to be surprisingly simple. Once the original researcher has “locked in” an algorithm or workflow, they can assign it a version number. The version captures all relevant metadata about the algorithm or set of algorithms. When the algorithm or workflow is shared with colleagues, the originator has the option of forcing downstream executions to run as described in the version number. In this way, research can be recreated without regard for who is running the analysis, where it is being run, or even the passage of time. This poster will explain the need for reproducible research, the inherent obstacles within open source algorithms, and will present the simple solution to the problem.

1393M Clinical whole-exome and whole-genome sequencing in dystonia: a key role for UMD knowledgebases. M. Milteni1,2, A. Blanchard1,2, L. Barre1,2, C. Paladin1,2, C. Guien1,2, A. Pinard1,2, A. Roubertie1,2, C. Bérard1,2, C. G. Collod–Bèrard1,2, M. Dufour1, C. Cornil1, C. Palacin1,2,3, C. Guien1,2, A. Pinard1,2, R. Forsberg1,2, B. Knudsen1, T. Knudsen1, J.C. bio, a QIAGEN Company, Danmark A/S; 2) CLC bio, a QIAGEN Company, Silkeborgvej 2, Aarhus, 8000, Denmark; 3) 2p53 Signal Transduction Group, University of Edinburgh Cancer Research Centre in the Institute of Genetics and Molecular Medicine, Crewe Road South, Edinburgh EH2 2LR, United Kingdom; 4) Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Ústí nad Labem 766 53 Brno, Czech Republic.

High throughput sequencing technologies are currently revolutionizing the cancer research area with rapid improvements in sequencing capacity and time consumption. As a result the most time consuming step has moved from being the sequencing process itself to being the bioinformatic data analysis. RNA sequencing (RNA-Seq) is used in an increasing number of transcriptomic studies. The great advantage of using RNA-Seq data is the ability to precisely quantify transcript levels and identify novel transcripts, isoforms, and splice junctions, while further providing information of the mutational landscape down to single base resolution. To ease the hurdles associated with RNA-Seq data management there is an increasing demand for tools that are specifically tailored to RNA-Seq data. Here we describe how the newly developed CLC Cancer Research Workbench can be used to analyze and visualize RNA-Seq data with ready-to-use workflows that automatically map, quantify, and annotate transcriptomes. We identify differentially expressed genes and transcripts in Illumina HiSeq data from three esophageal adenocarcinoma cancer patients, and compare the mutational patterns in the samples with the expression values of the corresponding genes. Results are visualized in a track based browser view, which provides easy navigation as well as allowing the user to simultaneously view and annotate multiple samples and different data types (e.g. genes, transcript expression levels, and detected variants).

1391T Identification of differentially expressed genes and somatic mutations in esophageal adenocarcinoma cancer patients. M. Matvenko1, B. Øster2, A. Joecker1, A.-M. Hein1, P. Dekker1, R. O’Neill2, A. Krejci1, A. Arends1, N. Thomson1, C. Boyssen2, S. Mansted1, R. Forberg2, B. Knudsen1, T. Knudsen1, J.C. bio, a Qian Company, Denmark; 2) CLC bio, a QIAGEN Company, Silkeborgvej 2, Aarhus, 8000, Denmark; 3) 2p53 Signal Transduction Group, University of Edinburgh Cancer Research Centre in the Institute of Genetics and Molecular Medicine, Crewe Road South, Edinburgh EH2 2LR, United Kingdom; 4) Regional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Ústí nad Labem 766 53 Brno, Czech Republic.

High throughput sequencing technologies are currently revolutionizing the cancer research area with rapid improvements in sequencing capacity and time consumption. As a result the most time consuming step has moved from being the sequencing process itself to being the bioinformatic data analysis. RNA sequencing (RNA-Seq) is used in an increasing number of transcriptomic studies. The great advantage of using RNA-Seq data is the ability to precisely quantify transcript levels and identify novel transcripts, isoforms, and splice junctions, while further providing information of the mutational landscape down to single base resolution. To ease the hurdles associated with RNA-Seq data management there is an increasing demand for tools that are specifically tailored to RNA-Seq data. Here we describe how the newly developed CLC Cancer Research Workbench can be used to analyze and visualize RNA-Seq data with ready-to-use workflows that automatically map, quantify, and annotate transcriptomes. We identify differentially expressed genes and transcripts in Illumina HiSeq data from three esophageal adenocarcinoma cancer patients, and compare the mutational patterns in the samples with the expression values of the corresponding genes. Results are visualized in a track based browser view, which provides easy navigation as well as allowing the user to simultaneously view and annotate multiple samples and different data types (e.g. genes, transcript expression levels, and detected variants).
1394T
A high-fidelity simulation validation framework for high-throughput genome sequencing with cancer applications. J.C. Mu1,*, M. Mohyuddin1, J.L. Pi2, N. Bani Asadi2, M.B. Gerstein3, A. Abyzov4, W.H. Wong3,4, H.Y.K. Lam4, 1) Department of Electrical Engineering, Stanford University, Stanford, CA; 2) Department of Bioinformatics, Bina Technologies, Redwood City, CA; 3) Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT; 4) Department of Health Sciences Research, Mayo Clinic, MN; 5) Department of Statistics, Stanford University, Stanford, CA; 6) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

High precision simulation and computational validation are essential in characterizing and assessing the accuracy of NGS analysis. We present VarSim, an integrated computational framework that leverages state-of-the-art read simulation and vast annotation databases to generate realistic high-throughput sequencing reads for normal and tumor samples and report detailed accuracy statistics. It also has the ability to isolate the analysis to specific regions of interest, such as disease genes.

For each normal or disease sample, VarSim first generates a phased diploid genome using variants, including various types of structural variations, from existing annotations and novel sites. Next, reads are simulated from this diploid genome using empirical error models. After alignment and variant-calling on the simulated reads, VarSim reports detailed statistics on the accuracy of the results in an HTML document. These statistics include alignment accuracy and variant-calling accuracy for different variant types and sizes, as well as for different categories of genomic regions, e.g., genes and repeats. Since VarSim generates a diploid genome, genotyping accuracy is also reported. For tumor samples, VarSim also uses COSMIC annotations to generate the tumor variant set. To simulate the impurity and heterogeneity of tumor samples, the reads for a tumor sample are constructed by mixing the reads generated from multiple sets of tumor variants with the reads generated from normal variants in a specified proportion. After running two more rounds of variant calling on the reads, VarSim then generates a report on the accuracy of the somatic variant calls.

To demonstrate its utility and enable rapid validations, we generated three synthetic genomes and their reads at high coverage (100x): one a normal tumor sample with a rare disease and finally a tumor-normal sample pair. We compared the accuracy statistics generated by VarSim for these genomes on popular secondary analysis tools including aligners, small variant callers, structural variant (SV) callers and somatic mutation callers. One of these tools includes our multi-algorithm SV caller, MetaSV, which improves both sensitivity and specificity of SV calling. No other simulation framework offers such a comprehensive validation of secondary analysis over a variety of whole genome sample types.

1396T
1) Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA; 2) Center for Mitochondrial Medicine and Epigenomic Medicine, The Children’s Hospital of Philadelphia, Philadelphia, PA; 3) Department of Clinical Genetics, Maastricht University, The Netherlands; 4) University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; 5) Department of Forensic Molecular Biology, Erasmus MC - University Medical Center Rotterdam, The Netherlands; 6) Division of Human Genetics, The Children’s Hospital of Philadelphia, Philadelphia, PA.

Accurate mitochondrial haplogroups determination is critical for population genetics, anthropology, and forensics. It is equally if not more important for clinical genetics as mitochondrial haplogroup associations with human diseases have been identified. Despite this, there is a lack of computational tools to facilitate automated classification starting with mitochondrial sequences. A mitochondrial sequence has to be aligned to a mitochondrial reference genome in order to identify the polymorphisms, which are then compared individually and frequently manually to the haplogroup determining polymorphisms, such as those summarized in PhyloTree (http://www.phylotree.org/). What complicates more is that the choice of optimal mitochondrial reference genome remains controversial as some researchers pointed to various limitations with the commonly used Revised Cambridge Reference Sequence (rCRS) and proposed to replace rCRS with Reconstructed Sapiens Reference Sequence (RSRS) instead. There is a strong need therefore for an alignment-free mitochondrial haplogroup classifier that is hence agnostic to the reference sequence of choice. Results: Using a k-mer approach, we developed a new Python software package called Phy-Mer that determines the associated haplogroup of a mitochondrial sequence to the highest resolution currently feasible based on PhyloTree without having to align it. Comparisons with other available manually curations, and existing haplogroup tools such as HaploGrep and Mitomaster, revealed superb performance of our tool, in terms of accuracy, sensitivity, as well as the ease to use. Phy-Mer has an added functionality such that next-generation sequence data can be used directly as input. Accurate haplogroup call can be obtained in 15-30 minutes for sample with 1 Megabase of raw reads, and about 2 minutes per sample with BAM file of either whole-exome or whole-genome sequencing data sets on a desktop computer. Availability: The source code and sequence library is being made freely available on GitHub (https://github.com/SamStudio8/frontier) and the source code is available from Data Resource Consortium (http://msqdr.org) and PhyloTree.org (http://www.phylotree.org/).

1397T
Application of Machine Learning Techniques to Next Generation Sequencing Quality Control. S. Nicholls1,2, A. Clare3, J.C. Randall4
1) Welcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Aberystwyth University, Aberystwyth, Wales, United Kingdom.

Advances in next-generation sequencing have made the concept of massively parallel DNA sequencing, reducing both the time and cost involved in performing genetic analysis. However, these "next-generation" processes are complicated and open to error, thus quality control is an essential step to ensure confidence in any downstream analyses carried out. As part of the Sanger Institute’s human genetics informatics pipeline, a system called autoqc is responsible for applying quality control to samples within the pipeline. It does this by comparing a subset of quality metrics output by samtools stats to hard-coded thresholds, determining whether a particular sample has reached a level that requires a warning, or has exceeded the threshold and failed a test entirely. Whilst this catches most of the very poor quality outputs, a large number of samples are flagged for manual inspection at a swarming threshold level; a time consuming task which can invite inefficiency and error.

Using a set of 13,455 lanelets (each containing data from one indexed sample extracted from a single lane of multiplex sequencing) drawn from 4,915 DNA samples we investigate the accuracy of a range of machine learning classifiers for a series of quality control data sets and show it is possible to generate minimal yet accurate models, creating decision trees which closely resemble the behaviour of an already existing manually curated QC system -- using stratified cross-validation, a method which measures the proportion of samples that were correctly classified as indicated by the manually curated training set, our models were able to gain scores of 95% and above. To support this work we have developed Frontier, a Python package which provides users with interfaces for the reading, storage and analysis of samtools stats output data for the purpose of working with machine learning framework. Frontier is freely available open-source software hosted at https://github.com/SamStudio8/frontier.
1399M

We present three free, open source integrated computational systems designed to aid researchers who are analyzing high throughput genome, DNA chip technology and the promising tools. In order to reduce the cost and time consuming in a large cohort genetic association study. Computational genetics approach facilitates the selection of potentially functional genetic polymorphisms and the plausible biological explanation of genetic association study results. In this study, the in-silico genetic analysis approach has been used for uncovering the potential functional genetic polymorphisms of SREBP-1 gene. The genetic polymorphism mining technique, functional effect prediction, linkage disequilibrium(LD) approach of SREBP-1 gene polymorphisms were extensively studied in reference populations from HAPMAP project. Many candidate single nucleotide polymorphisms (SNP), a nucleotide substitution, localized in regulatory region have been identified as a functional SNP. However nonsynonymous SNPs also demonstrated the possible functional effect. The large haplotype block obtained from LD analysis indicated a high nonrandom association in this region. We therefore proposed the usefulness of computational genetic approach for analyzing SREBP-1 gene and demonstrated the fascinating results of high potentially functional SNPs which can be a good candidate marker for human genetic association study. The utilization of computational genetic approach is the meaningful tool in the field of life science both in pre-and post-experimental study.

1400T
Sparse structural equations for joint phenotype-genotype network analysis. M.L. Rahman, P. Wang, M. Xiong. Human Genetics Center, University of Texas Health Science center at Houston, Houston, TX.

Multivariate linear model is the most widely used statistical method for genetic studies of multiple phenotypes. However, multivariate linear model for genetic studies of multiple phenotypes has several serious limitations. First, it cannot be used to estimate the causal relationship of genetic variants among the phenotypes. In practice, the most clinical phenotypes are often correlated. Overlooking correlation information of phenotypes will compromise identification of genetic architecture of multiple phenotypes. Second, it cannot be used to estimate the causal relationship of genetic variants among the phenotypes. To overcome these limitations, we propose sparse structural equations for modeling the phenotype-genotype networks. In other words, the multiple phenotypes are presented as exogenous variables and genetic variants are modeled as exogenous variables in the structural equations. Since the current methods of parameter estimation for structural equations will lead to fully connected networks, we use regularization principal to reduce densely connected networks to sparse networks and develop alternative direction methods of multiplier to estimate parameters in the structural equations and identify the structure of phenotype-genotype networks. We formalize the causal inference in the phenotype-genotype networks using Pearl's do operation. The directed acyclic graph (DAG) or its Markov equivalence class corresponding to the structural equation model is estimated from data and total causal effects of genetic variants on the phenotypes will be estimated by the DAG and Pearl's back door operation. The proposed method was assessed by large simulations and applied to the NHLBI's Exome Sequencing data set with 12 phenotypes and 7,194 genes which were located on the human genome. We constructed phenotype-genotype network for each pathway and estimated their causal relationships. We used test statistics to formally test association of the SNPs in the networks with 12 phenotypes and identified a total of 89 genes significantly associated with at least one phenotype.

1401S
Human Splicing Finder: An invaluable system to annotate the impact of mutations on splicing signals. G. Raa1,2, D. Salgado1,2,3, J.P. Desvignes1,2,4, A. Blanchard1,2, M. Milkina1,2, Q. Bacin1,2, C. Guien1,2, A. Pinard1,2, L. Barre1,2, S. Olschwang1,2,4, G. Collod-Béroud1,2,4, C. Béraud1,2,4. 1) Aix-Marseille Université, GMGF, 13385, Marseille, France; 2) Inserm, UMR_S 910, 13385, Marseille, France; 3) EMBL Australia; Australian Regenerative Medicine Institute (ARMI), Monash University, Building 75, Clayton, Victoria 3800, Australia; 4) APHM, Hôpital TIMONE Enfants , Laboratoire de Génétique Moléculaire, 13385, Marseille, France.

The NGS technologies have dramatically changed our approach to rare human genetic diseases, leading to the identification of many disease-causing genes: the challenge is not anymore data production but data interpretation. It has been reported that many mutations (including intronic and up to 50% of missense) are not pathogenic because of their impact on the RNA rather than due to their impact on pre-mRNA splicing machinery. Thus it is necessary to integrate this level of information in NGS bioinformatics pipelines. The Human Splicing Finder (HSF) system combines 12 different algorithms to identify and predict mutations in exonic splicing enhancers or silencers. Each mutation is analyzed on splice sites, the branch point and auxiliary sequences known to either enhance or repress: Exonic Splicing Enhancers (ESE) and Exonic Splicing Silencers (ESS). If the HSF system has quickly become an international reference, the complexity of splicing signals often result in a flow of information preventing correct interpretation by non-specialists. To make this information accessible to all we have created an expert system for data interpretation. This algorithm uses a decision tree based on mutation position: intronic vs. exonic, localization or not in a well-known splice sites branch or other splice sites or branch point. Based on these criteria, context-irrelevant signals are automatically ruled out to focus on relevant ones to evaluate the potential impact of mutations. To evaluate the data-interpretation expert system, we selected two datasets: 200 mutations reported to affect splicing (150 intronic and 50 exonic); 70 mutations (intronic and exonic) reported as having no effect on splicing signals. We then compared predictions with the manually interpreted data and with the in vitro results. The first comparison gives a 100% concordance, while the second demonstrate accuracy >96% for donor and acceptor splice sites and 85% for branch site. In addition, the best result was as expected, as the prediction of such signals is still difficult because of insufficient experimental data. Furthermore, to allow implementation of the HSF system in NGS pipelines, we created a dedicated wesbite. In the future, the HSF system could be used both in research and clinical practice by non-experts thanks to its new data-interpretation expert system and its dedicated webservice. It could be combined with other systems such as the UMD-Predictor for WGS.
results highlight the similarities and differences as well as the additional \n
phenomenon involving signals between cell types and migration of cells \n
method is the Gaussian Lasso, which we use to infer an undirected bayesian \n
number of methodological tools potentiate this analysis approach. One \n
to content-based enrichment analyses. A potentially more powerful approach \n
screening tool to identify individual candidates. When more systems level \n
fully exploiting the multivariate character of the data has remained a chal- \n
Baylor College of Medicine, Houston, TX; 3) Wellcome Trust Sanger Institute, \n
Rayner N.W. \n
Plot: A tool to automatically summarise single variant analyses. \n
308 Bioinformatics and Genomic Technology \n
Even screening tool to identify individual candidates. When more systems level \n
fully exploiting the multivariate character of the data has remained a chal-

sis (CF). We used variants found in this database along with 1000 Genomes \n
Functional TRanslation of CFTR (CFTR2) database, which assembles clini-

a machine learning method that combines the predictor of protein stability \n
about the functional consequence of missense mutations. Here we develop \n
Therefore, estimated folding free energy should give valuable information \n
and regulation of biomolecules. It is measured by the folding free energy, the \n
lead to diseases by interrupting protein folding and decreasing its overall \n
findings ought to be returned and when. In this project we considered variants \n
determination of whether a novel genetic variant at a “clinically relevant \n
If the variant is predicted to be pathogenic, it is included in the list of \n
without and with sibling. This information, along with Tute Score, was then \n
were incorporated in the analysis as prior information, and variants were \n
identified with labels such as inherited, de novo, compound heterozygous \n
and shared with sibling. This information, along with Tute Score, was then \n
used to calculate the family adjusted variant scores and gene scores in a \n
probabilistic approach. Results: In an exome sequencing study on a pedigree \n
where one subject was discovered to have idiopathic hemolytic anemia (IHA), \n
we identified two heterozygous variants in PKLR gene, a gene previ- \n
cously confirmed to be associated with IHA, ranked 7th (c.1022G>C) and \n
11th (p.R569Q) out of all variants based on Tute Scores. When applying a \n
regression model where the phenotype label of ‘anemia’ was included. \n
scores. When further applying a phenotype label of ‘anemia’ in the analysis, \n
the PKLR gene ranked 3rd based on the Tute Gene Scores. When further applying a phenotype label of ‘anemia’ in the analysis, \n
the PKLR gene ranked 3rd based on the Tute Gene Scores. This analysis \n
demonstrates that our pipeline can effectively prioritize disease-causing \n
mutations and genes in pedigrees with exome sequencing data. Conclusion: Tute can rapidly interpret large volumes of sequencing data, \n
including family-based datasets, via a web-based clinical genome interpreta-

tion platform, with comprehensive annotation types and user-friendly inter-

face for conducting analyses. Our results demonstrate the effectiveness of \n
this family-based approach to support disease gene discovery.\n
Improving Computational Prediction of Clinically Relevant Genome \n
Variation. A. Rychkova, C. Bustamante. Genetics Department, Stanford \n
University, Stanford, CA. \n
Rapid, accurate, and inexpensive genome sequencing promises to trans-

form medical care. A critical hurdle to enabling personalized genomic medi- \n
cine is predicting the functional impact of novel genome variation. This is \n
a particularly pressing problem at “clinically relevant genes” where some \n
mutations are already known to impact phenotype (such as the BRCA1/2 \n
cancer is predicting the functional impact of novel genome variation. This is \n
a particularly pressing problem at “clinically relevant genes” where some \n
mutations are already known to impact phenotype (such as the BRCA1/2 \n
cancer is predicting the functional impact of novel genome variation. This is \n
a particularly pressing problem at “clinically relevant genes” where some \n
mutations are already known to impact phenotype (such as the BRCA1/2
1406T Cancer is a Zero Sum Game between Cells and Cells! A.R. Salehi Chaleshtori1, S. Bamohabat Chafjiri2. 1) Medical Genetic Center of Genome, Isfahan, Iran; 2) Information and Security Lab, Sharif University of Technology, Tehran, Iran.

Cancer is the major problem of our age and massive efforts performed to solve this problem. Cancer is very variable and tumor heterogeneity is documented for many characters, including the production of growth factors, one of the hallmarks of cancer. Combined interactions within the tumor can maintain heterogeneity for the production of growth factors and explain short term effectiveness of some therapies like RNAi. Alternative strategies for evolutionarily stable treatments are discussed, but this study propose that these strategies must be completed by mathematical principles like Game theory. Game theory is a social science whose aim is to understand the behavior of interacting decision-makers. We propose that determination of equilibrium point between growth factors and apoptosis inducers can be very effective in cancer treatment as well as cancer prediction. Cancer is recognized by uncontrolled growth, invasion, metastasis and death in the end. This process is initiate by defect in equilibrium point and defeating the whole system finally. Cells that overexpressing oncopgenic products elevate their growth and divisions to access the impairment success but this process led to whole system defeat permanently. Regarding to the game theorem, we propose that tumorgenesis is a zero sum game (or competitive game) in which cancerous cells evolve the same strategy (tumor growth and division), controlling the game and finally win the game. If other cells select a definite strategy to play against cancerous cells, this cause to control the game by these healthy cells and winning the game eventually. These selfish cancerous cells prefer their success to the system but if these cells possess their balance to reach the equilibrium point through normal expression of oncopgenes and apoptosis inducers and select stable strategy against cancerous cells, system can access the permanent success entirely. It is concluded that game theory is an applied issue in cancer treatment and prediction. Normal cells can control the cancer if these normal cells operate stable and choose the right strategy. These cells require to be trained in this field, win the game and controlling the cancer.

1407S UMD-Predator: A variant annotation masterpiece for NGS pipelines

D. Salgado 1,2,3, J.P. Desvignes 1,2,3, G. Rai 1,2,3, A. Blanchard 1,2,3, M. Milligen 1,2,3, A. Pinard 1,2,3, G. Colloid-Bèrard 1,2,3, C. Béroud 1,2,3, 1) Aix-Marseille Université, GMGF, 13385, Marseille, France; 2) EMBL Australia; 3) Australian Regenerative Medicine Institute (ARMI), Monash University, Building 75, Clayton, Victoria 3800, Australia; 3) Inserm, UMR S 910, 13385, Marseille, France; 4) APHM, Hôpital TIMONE Enfants, Laboratoire de Génétique Moléculaire, 13385, Marseille, France.

Whole Exome Sequencing (WES) technologies are increasingly applied to clinical practice and medical research. Those experiments generate 50 to 90,000 variations per individual from which only one or two are pathogenic mutations responsible for Mendelian diseases. It is thus crucial to differentiate non-pathogenic from pathogenic mutations to limit downstream analysis. To do so, we created the UMD-Preditor system (http://umd-predictor.eu) to annotate all coding variations for their potential pathogenic effect. UMD-Predator contains all nucleotide substitutions from human protein coding transcripts through a combination of data related to splicing signals, evolutionary data, biochemical substitutions matrices, and positioned frequency in the general population. Scoring values range from 0 to 100 and correspond to 4 classes: <50 = polymorphism; ≥50 <95 = probably pathogenic; ≥95 = pathogenic mutations. To evaluate UMD-Preditor’s efficiency we used 4 datasets containing pathogenic (P) and non-pathogenic variations (NP): Varibench with dbSNP (19,335 P; 7,897 NP), UniProt (20,821 P; 36,825 NP), ClinVar (10,669 P; 1,817 NP) and PredictSNP (24,082 P; 19,800 NP). We then compared our system to the commonly used systems SIFT 5.1.1, PolyPhen 2.2.2, Provean 1.1.3, Mutation Assessor 2, CONDEL 1.5, MutationTaster 2 (MT2) and CADD. Various statistical parameters were calculated including ROC curves and Diagnostic Odds Ratio (DOR) which measures the effectiveness of a diagnostic test and the log(DOR) to study the trade-off between sensitivity and specificity. For all parameters, UMD-Preditor gave the best results. For example, with the Varibench/dbSNP dataset it displayed a: ROC AUC of 0.954 vs. 0.834 for CADD; DOR of 86.6 vs. 12.6 for MT2 and a log(DOR) with a 22.7 increase vs. all others. In the NGS context, UMD-Preditor’s webservice analysis time was 45s (± 4s) vs. 420-9740s for other systems. If all systems accurately annotate the pathogenic mutations, UMD-Preditor showed the shortest list of candidate mutations with an average of 739 versus 2027 for other tested predictors (64% reduction of false positives). We believe that the UMD predictor system will be a masterpiece in NGS pipelines both through its very high specificity (0.85), its availability through webservices and speed. It could be combined with other systems such as Human Splicing Finder for WGS.
141OS

Genome and Transcriptome Free Analysis of RNA-Seq Data (GT-FAR) using cloud computing. T. Souaiaia1, K. Vahii2, R. Mayani2, J. Herstein2, O. Evagharoy1, T. Chen1, E. Deelman2, J. Knowles1, 1 Zikha Neurogenetic Institute, USC, Los Angeles, CA, USA; 2) Information Science Institute, USC, Los Angeles, CA, USA.

We present detailed results, best practices and common bottlenecks encountered in the current gold-standard for the evaluation of gene expression, alignment studies involving RNA-seq data frequently suffer from low-mapping rates and incorrect estimation of gene models, especially for model organisms with poorly annotated transcriptomes. To overcome these shortcomings, we have developed a cloud-based pipeline that brings accurate expression analysis to researchers without significant computational resources. GT-FAR consists of three independent modules which (1) Analyze a reference transcriptome and infer possible splice locations, (2) Filter, trim, and align reads to reference transcripts, introns and novel splice junctions and (3) cluster unmapped or raw reads into a compact sequence preserving index to facilitate to detection of significant sequence differences between samples and provide queries to multiple homology databases. The Pegasus WMS for the virtual machine are developing a virtual machine which will allow enabling it to be deployed as a virtual machine in the Amazon EC2 cloud. Art analysis pipeline we have modeled GT-FAR as a Pegasus workflow, offering an efficient combination of CPU and memory availability. Additionally, even when resources are available, the need for on-demand computation may overflow from the necessary computational tasks carried out by many small research groups. To minimize the financial and logistical strain necessary to run a state-of-the-art analysis pipeline we have modeled GT-FAR as a Pegasus workflow, enabling users to compute as a virtual QC, giving the ability to place and run the workflow on their own, or on our server. We are currently developing a custom web interface over Pegasus WMS for the virtual machine are developing a virtual machine which will allow users to upload custom datasets, track computational progress, and be alerted when their output is available for download. The Pegasus WMS represents the workflow in an abstract form that is independent of computational resources or file location resulting in executable workflows that can be deployed on local resources, remote clusters, or computing clouds. The Pegasus WMS also provides the additional advantage of the ability to detect a virus or hardware failure, a rescue workflow is created that performs only the tasks not yet completed.

1411M

Pedigree Reconstruction by PRIMUS using Exome Sequencing Data. J. Staple1, M.H. Cho2, D. Qiao3, E.K. Silverman1, U.W. Center for Mendelian Genomics1, D.A. Nickerson1, J.E. Below1. 1) Genome Sciences, University of Washington, Seattle, WA; 2) Brigham and Women’s Hospital, Boston, MA; 3) Epidemiology, Human Genetics & Environmental Sciences, University of Texas Health Science Center, Houston, TX.

We have developed a software program known as PRIMUS designed to reconstruct pedigrees from variant genotypes generated by genotyping arrays, exome sequencing, or whole genome sequencing. We have verified that the provided samples match the expected pedigrees and help to correct inconsistencies. PRIMUS is an efficient algorithm for accurate pedigree reconstruction, and is particularly useful in the analysis of Mendelian diseases. To illustrate the utility of PRIMUS, we provide the results of simulated pedigrees and demonstrate its utility in reconstructing, validating, and correcting expected pedigrees using simulated and clinical pedigrees obtained from the University of Washington Center for Mendelian Genomics (UW CG). To illustrate the performance of PRIMUS on a broad range of known pedigrees, we simulated pedigrees structures of varying sizes, structures, genotypes, and combinations of missing data for individuals in the pedigrees. Our analysis revealed that PRIMUS reconstructs the true pedigree in 93.1% of the simulations and a partial pedigree for an additional 6.3%. PRIMUS is an essential tool for quality control within the UW CMG, and is used to reconstruct clinically obtained pedigrees submitted to the UW CMG. One dataset consisted of 49 pedigrees containing 351 individuals with Early Onset Chronic Obstructive Pulmonary Disease (EOCOPD). Using PRIMUS, we generated analytical phenotypes for 49 pedigrees per dataset and calculated genome-wide identity-by-decent (IBD) estimates with PLINK and reconstruct each pedigree with PRIMUS. These results confirm that 43 of the 49 pedigrees matched the expected pedigree information collected in the study. PRIMUS provides the capability for the generation of a single report per family that can be used for further manual inspection. PRIMUS has detected inconsistencies in other submitted pedigrees that have altered the disease model used to identify the disease causing gene—ultimately, leading to the identification of the causal gene. Using only genotypic data from arrays or exome sequencing data to verify and correct expected pedigrees, PRIMUS saves time and effort that would otherwise be spent on manual verification. Funding: NSF, NHGRI, and NHLBI.

1412T


NCBI’s dbGaP (database of Genotypes and Phenotypes) has distributed over 5,800 phenotype datasets constituting approximately 4.8 billion phenotypic values to more than 2,000 investigators since 2007. The phenotype data consists of subject consents, subject-sample mapping (SSM), subject phenotypes, and sample attributes. The SSM links IDs found in phenotype data to IDs found in molecular genetic data, such as SNPs, expression, copy number variation, epigenetic data. Phenotypes submitted to dbGaP often contain inconsistencies or missing information, which may skew subsequent analyses of associations between phenotypes and genotypes. We have created a dbGaP submission guide to standardize the formats of all submitted datasets and data dictionaries. All submitted files are initially checked by QC scripts for formatting inconsistencies and data integrity across subjects and samples, subject consents, and the mapping of subjects and samples. The QC process also identifies conflicts in subject gender, affection status, pedigree structure, and unexpected duplicates. Potential HIPAA violations are identified by QC scripts developed to detect the presence of visit dates, birth dates, zip codes, and ages over 89. The potential HIPAA violations reported by the QC scripts are further checked manually to identify true violations. Most of the HIPAA violations detected are dates. Datasets that meet criteria are then compared to their data dictionary to assure that every variable contains a description, that every encoded data value contains a code meaning, and that the values match the units listed and fall within the logical min/max. After the submitted data have passed the QC tests, phenotype summaries are generated and value distributions are displayed on a preview website for review by the submitter. Statistical summaries are further manually checked to detect missing code meanings and extreme data values that could potentially be used to identify individual participants and their families. Additional ID and gender checks are conducted between phenotype and molecular genetic data. This presentation will describe the phenotype quality control process designed to ensure the integrity and accuracy of the phenotype data distributed by dbGaP.
1414M

Background DNA methylation is an important epigenetic modification involved in many biological processes. Reduced representation bisulfite sequencing (RRBS) is a cost-efficient method for studying DNA methylation at base level on a genomic scale, yet analyzing the RRBS data is challenging. Although several programs/pipelines are available for the data analysis, it is not clear which strategy performs the best and there has been no much attention to the contamination issue from artificial cytosines incorporated during the end repair step of library preparation. To address these, we have developed Targeted Alignment and Artificial Cytosine Elimination for RRBS (TRACE-RRBS), which assigns bisulfite sequence reads to MSPI digitally digested reference and specifically removes the end repair cytosines before Cpg methylation summarization. We have tested and compared the algorithms on a simulated and a real dataset in terms of alignment speed, accuracy, and accurate methylation estimate with 7 other RRBS tools and illumina 450k microarray platform. Results By aligning RRBS reads only to the genomic regions where Msp1 digests, TRACE-RRBS demonstrated as the fastest, most sensitive and specific alignment tool among the 7 compared for a simulated dataset. For the real dataset with ~50 million of reads, TRACE-RRBS took about the same time as RRBSmap, a third to a sixth of time needed for Bismark and Novelalign. TRACE-RRBS aligned more reads correctly than most of other tools and achieved the highest correlation with 450k microarray data. The end repair artificial cytosine removal is the most significant improvement in the alignment speed. Conclusions TRACE-RRBS is a fast and accurate alignment and methylation quantification tool for DNA methylation from RRBS. It is implemented using platform independent Java programming language. The package is available for public use (http://bioinformaticstools.mayo.edu).

1414S
Graphical algorithm for integration of genetic and biological data: Proof of principle using psoriasis as a model. L.C. Tsai1*, J.T. Elder2,3, G.R. Abecasis1. 1) Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 2) Department of Dermatology, University of Michigan, Ann Arbor, MI; 3) Ann Arbor Veterans Affairs Hospital, University of Michigan, Ann Arbor, MI.

Pathway analyses to reveal biological mechanisms for genetic association studies have great potential to better understanding of complex traits with major human disease impact. However, current approaches have not been optimized to maximize statistical power to identify enriched functions/pathways, especially when the genetic data derives from studies using platforms (such as Immunochip and MetaboChip) customized to focus on markers near previously identified top-ranking loci. We present here a novel approach, called MEAGA (Minimum distance-based Enrichment Analysis for Genetic Association), with the potential to address both of these important concerns. MEAGA performs enrichment analysis using graphical algorithms to identify subgraphs among genes and measure their closeness in an interaction database. It also incorporates a statistic summarizing the numbers and total length of subgraphs, depicting the overlap between observed genetic signals and defined function/pathway gene-sets. MEAGA uses sampling techniques to approximate empirical and multiple testing-corrected p-values. We show in simulation studies that MEAGA has more power than count-based strategies to identify disease-associated functions/pathways, and the increase in power is influenced by the shortest distances among associated genes in the interactome. We applied MEAGA to the results of a meta-analysis of psoriasis using Immunochip datasets, and showed that associated genes are significantly enriched in immune-related functions and pathways, with increased overlap with protein interaction network. Count-based strategies failed to identify functional enrichment in the same data.

1415T
Unraveling epistatic causal genes of diseases with hyper-sensitive multiple testing procedure. A. Terada*1,2, K. Tsuda*3,4,5, J. Sese*. 1) Department of Computer Science, Ochanomizu University, Tokyo, Japan; 2) Research Fellow of the Japan Society for the Promotion of Science; 3) Department of Computational Biology, The University of Tokyo, Chiba, Japan; 4) Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo, Japan; 5) JST ERATO Minato Discrete Structure Manipulation System Project, Hokkaido, Japan.

Epistatic interactions are key to unraveling the genetic contributions to complex diseases and traits. However, most of the analyses in genome-wide association study (GWAS) perform a statistical test for single SNP or gene. As the reason, the existing multiple testing procedures that are widely used to avoid false discoveries in GWAS, including the Bonferroni correction and the Benjamini-Hochberg procedure, are too conservative to detect such interactions, and finding any statistically significant combinations after applying such corrections is hopeless. We propose a novel multiple testing procedure to identify the epistatic effects of complex diseases, which is called the Limitless Arty Multiple-testing Procedure (LAMP). The LAMP can list any statistically significant combinations without any limit and can be substituted for the Bonferroni correction by rigorously calculating the family-wise error rate (FWER). We demonstrated that the LAMP compute the adjusted P-value by multiplying by a value that is 250 times smaller than the value by the Bonferroni correction or Holm procedure, while maintaining the FWER under the threshold. We also showed that the LAMP could identify significant combinations from existing GWAS data that have the potential to cause diseases, and that were overlooked by the Bonferroni correction. This procedure may contribute to the discovery of new combinatorial effects by the reanalysis of existing data.

1417M
A population- and pedigree-aware alignment strategy for Next Generation Sequencing data. E. Valkanas1*, E. Flynn1, T. Gali2, J. Elson3, A. Brandt4, P. Pemberton1, L. Carmichael4, J. Osman4, S. Leighton4, M. Groner5, D. Adams5,6, W. Gali2, T. Markello2, 1) NIH Undiagnosed Diseases Program Translational Laboratory, Bethesda, MD; 2) Walter Reed National Military Medical Center, Bethesda, MD; 3) Microsoft Research, Redmond, WA; 4) Appistry Inc., St. Louis, MO; 5) Medical Genetics Branch, NHGRI, NIH, Bethesda, MD.

Next Generation Sequencing (NGS) data is conventionally aligned to a standard human reference sequence that provides a common reference for understanding genetic variation. However, since the standard human reference is derived from a limited number of samples, it cannot account for all non-pathogenic variations arising in the human population. This problem confounds alignment and genotyping of NGS data. The NIH Undiagnosed Diseases Program has developed a Diploid Alignment pipeline in order to increase the accuracy of short read alignment and improve variant calling. This strategy applies pedigree information in the form of genome-wide family structure to handle the relatedness and differences in the genetic makeup of the individual. This approach, called MEAGA (Minimum distance-based Enrichment Analysis for Genetic Association), with the potential to address both of these important concerns. MEAGA performs enrichment analysis using graphical algorithms to identify subgraphs among genes and measure their closeness in an interaction database. It also incorporates a statistic summarizing the numbers and total length of subgraphs, depicting the overlap between observed genetic signals and defined function/pathway gene-sets. MEAGA uses sampling techniques to approximate empirical and multiple testing-corrected p-values. We show in simulation studies that MEAGA has more power than count-based strategies to identify disease-associated functions/pathways, and the increase in power is influenced by the shortest distances among associated genes in the interactome. We applied MEAGA to the results of a meta-analysis of psoriasis using Immunochip datasets, and showed that associated genes are significantly enriched in immune-related functions and pathways, with increased overlap with protein interaction network. Count-based strategies failed to identify functional enrichment in the same data.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Forensic DNA profiles of short tandem repeat (STR) loci are currently obtained using PCR followed by capillary electrophoresis (CE). Massively parallel sequencing (MPS) technologies do not rely on size separation and thus relieve the limitations on locus multiplexy. MPS therefore creates extra possibilities within forensics for analyzing degraded samples, mixed samples, etc. Data analysis of raw sequencer reads, resulting in a reliable and usable forensic human identification report is still in early development. Only a couple of bioinformatics methods have been published. Unfortunately, all of these methods are command-line based and thus unsuitable for use by forensic analysts without a bioinformatics training. Recently, we published an open-source bioinformatic framework My-Forensic-Loci-queries (MyFLq) for analysis of MPS forensic data in a generic, platform independent way. The MyFLq framework was successfully applied on an Illumina MiSeq dataset generated from a multi-focus STR PCR on both single contributor samples and multiple person mixture samples.

We now present the latest developments on MyFLq. We created an application with an easy to use graphical user interface in which we have incorporated the MyFLq algorithms. Several new features have been added such as a electropherogram style visual representation of the results, allowing for backward compatibility with CE results. To our knowledge, we are the first to present an open-source GUI application for forensic MPS data analysis. It can run as a standalone web application, or run embedded in the Illumina MyFLq app. In the latter, analysis was performed on the samples on an Illumina machine and starting the MyFLq app to obtain a finished, graphical and interactive human identification report.

We present here a bioinformatics tool to scan raw sequencer reads and stimulate both theoretical and practical researches in systems biology and multivariate data analysis to large scale penalized and functional data analysis. The Swiss pipeline will shift the paradigm of systems biology research from multivariate data analysis to large scale penalized and functional data analysis and stimulate both theoretical and practical researches in systems biology with NGS data.


Bioinformatics, University of Michigan, Ann Arbor, MI.

We present here a bioinformatics tool to scan raw sequencer reads and stimulate both theoretical and practical researches in systems biology research from multivariate data analysis to large scale penalized and functional data analysis and stimulate both theoretical and practical researches in systems biology with NGS data.


Bioinformatics, University of Michigan, Ann Arbor, MI.

We present here a bioinformatics tool to scan raw sequencer reads and stimulate both theoretical and practical researches in systems biology research from multivariate data analysis to large scale penalized and functional data analysis and stimulate both theoretical and practical researches in systems biology with NGS data.

Next generation sequencing (NGS) has revolutionized genetic research, enabling dramatic increases in the discovery of new functional variants in syndromic and common diseases. This technology has been widely adopted by the research community and is rapidly being implemented clinically, driven by recognition of NGS diagnostic utility and enhancements in quality and speed of data generation. Compared to traditional logistic regression, we can only reach 62.25% accuracy of 81.04% in the test sample and the total number of SNP markers is 393,473. Using 65 SNPs in CAD risk prediction study where 1,929 cases and 2,938 controls were sampled and a large-scale ChIP-seq data set, Churchill is able to effectively utilize distributed clusters of computers to allow the processing of large data sets that would have otherwise been too large to process. In our implementation, this we analyzed 1,088 whole genome raw data sets available from the 1000 Genomes Project Consortium (1KG) utilizing AWS resources. The large number of samples, population variant frequencies were produced in less than 7 days for ~$1 per genome. We achieved a high degree of correlation to the 1KG SNP allele frequencies and discovered ~3 million INDELs not reported in the original 1KG analysis. As we look to the future, cloud computing will be indispensable in the analysis of human genome sequences and the 1KG genome will soon be a reality. Churchill solves the sequence analysis computational bottleneck and through use of cloud computing resources enables rapid analysis of population scale sequencing datasets.

Neat-optimal whole genome reconstruction by a small set of genomic variants. M. Xiong, N. Lin, J. Yu, L. Ma, J. Jiang, P. Wang, S. Guo. Dept Biostatistics, Univ Texas Hlth Sci, Houston, TX.

The emerging genomic technologies are producing so large genomic and epigenomic data that traditional technologies and tools are unable to store, transfer, manage and analyze them when the number of sequenced genomes exceeds ten thousands. It is urgent to develop novel concepts, theory and algorithms to handle the mostly irrelevancy among genomic variants and select a set of sufficient genomic variants which can optimally recover whole genome information in the genomic data analysis. To meet these challenges, we propose to formulate genomic and epigenomic data as structured dimers and selection of an essential set of genomic variants for the whole genome reconstruction as a subset selection problem for matrices. We develop matrix approximation theories and error estimation minimized the complete data likelihood. We conducted simulation study and applied the algorithm to real datasets. Comparing with other available tools, our software has been able to identify all the orientations of the bipartite structured dimers as well as monomers from large-scale ChIP-seq data set.

Detecting Nuclear Receptors Using a Finite Mixture Model. M. Xu1, D. Umbach2, Y. Yao1, L. LF1. 1) Intramural Research Program, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health; 2) Intramural Research Program, National Institute of Environmental Health Sciences, Bethesda, MD.

Motivation: Nuclear receptors are a special BiPartite structure which binds to DNA sequences to regulate transcription. The typical orientations of the two BiPartite structure includes direct repeat, reverse direct repeat, inverted repeat, and rever repeat. A large-chip-seq data may contain all or some of these orientations of nuclear receptors as well as half-sites. It is a challenging task to identify those nuclear receptors and their binding sites from ChIP-seq data. We propose a computational hypothesis testing model to estimation of the nuclear receptor binding site; 3) estimate the proportion of each orientation of nuclear receptors. We propose a mixture model framework to model all the orientation of the embedded nuclear receptors and consider Markov model for the background. An EM-based iteration algorithm is used to maximize the complete data likelihood. We conducted simulation study and applied the algorithm to real datasets. Comparing with other available tools, our software has been able to identify all the orientations of the bipartite structured dimers as well as monomers from large-scale ChIP-seq data set.

Bayesian inference for tumour heterogeneity using the Hamming Ball Sampler. C. Yau1,2, P. Kirk1, M. Titisias2. 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxford, United Kingdom; 2) Department of Biostatistics, McGill University, Montreal, Quebec, Canada; 3) Department of Statistics, University of Oxford, Oxford, United Kingdom.

In clinical diagnostics, it is common to observe tumor heterogeneity at many levels including genomic, epigenomic and transcriptomic variation. These differences in tumour states are likely to impact treatment strategies and drug response. To benchmark a range of methods for inferring and modeling the extent of tumour heterogeneity, we conducted a simulation study (1,3) and applied the algorithm to real datasets. Comparing with other available tools, our software has been able to identify all the orientations of the bipartite structured dimers as well as monomers from large-scale ChIP-seq data set.

LVpicking: picking up true, low-frequency variants for studying cancer heterogeneity. J. Zhang1,2, J. Majewski1,2.

Low-frequency, somatic mutations due to cancer subclones are common and important for studying the heterogeneity within a tumour mass that shed light on cancer evolution and can guide therapy. The three challenges are to identify these mutations when they are present at a low frequency and when they are distinguishable from benign mutations.

LVpicker, to pick up true mutations from LFV. This approach uses a binomial test to calculate the probability of observing certain number of mutant reads in total reads in the test sample, given the site-specific error rate that can be generated from pooled control samples. The program also filters variants if all reads supporting the variant (including controls) have mapping quality lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing. Among 117 lower than 30. Using LVpicker, we firstly studied a patient with the blood sample and saliva sample that underwent exome sequencing.
1427T
Efficient and accurate de novo assembly algorithm for paired-end reads and its application in indel calling. L. Zhao1, Y. Guan1,2, L. 1) USDA/ARS Children’s Nutrition Research Center, Baton Rouge, LA, USA; 2) Department of Biochemistry, LSU School of Medicine, Shreveport, LA, USA. De novo assembly of paired-end reads into long contigs has been demonstrated to have markedly greater power in solving genetic related problems. Unfortunately, existing assemblers either suffer from high time complexity or low repeat tolerance. In this study, a novel assembly algorithm is presented to overcome these difficulties, and the in silico results using real data are demonstrated as well. Giving the input of paired-end reads, the model partitions these reads into overlapping groups. Within each group, a Smith-Waterman type of alignment algorithm is designed to align paired-end reads, accounting for the random insert lengths. The alignment forms a temporary assembly object that usually consists of two short contigs. By repeating this procedure, a few temporary assembly objects are assembled. These temporary assembly objects are then aligned together by using a Smith-Waterman type alignment algorithm incorporating seeding to form a larger temporary assembly object. The process continues until a single, gap filled, and long contig is obtained. This bottom-up procedure of assembling contigs from paired-end reads is carried out in parallel to produce many long contigs that span 2000-3000 base pairs. Our method assembles haplotypes, but it retains two/multiple alleles caused by various situations at specific regions, such as the region with different paternal and maternal haplotypes, the repetitive region having multiple alleles, and the region containing paralogs. Based on the assembled contigs, we are able to call indels in the corresponding genome. To validate the power of this method, we have simulated 2000 indels with size ranging from 1 base to 5 bases at the major histocompatibility complex (MHC) region of the Homo sapiens. Preliminary experimental results show that the true discovery rate of our method is 0.88 with the false discovery rate under 0.0025, while at the same false discovery rate level the true discovery rates of Fermi and Velvet (two well-known assemblers) are only 0.37 and 0.23, respectively. Besides the greater indel calling performance, our method is superior to existing ones in the parallelizability per se. These observations indicate that our method has great potential in variants calling and clinical genetic diagnoses.

1428S
PGS: a tool for association study of high-dimensional microRNA expression data with repeated measures. Y. Zheng1, Z. Fei2, W. Zhang2, J. Starren1, L. Liu1, A. Baccarelli3, Y. Li2, L. Hou2,7. 1) Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL; 2) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 3) Institute of Human Genetics, University of Illinois at Chicago, Chicago, IL. Whole-exome sequencing has become a popular approach for disease- and/or phenotype-predisposing variant discovery. Analysis of massive amount of exome sequence data requires efficient bioinformatic approach. Although a number of state-of-art variant calling algorithms enabled accurate variant callings, a reproducible, streamlined pipeline for discovering biologically significant variants in the phenotype of interest is scarce. We have developed a pipeline called FExSeq discovery pipeline that utilizes familial exome sequencing data to 1) perform effective filtration by multiple parallel comparisons to public datasets; 2) perform segregation analysis to identify variants that are family/disease group specific and/or shared among disease group; 3) to prioritize functional variants based on multiple annotation scheme. Here, we present the results from analyzing ~1,500 whole-exome sequenced samples from 710 families of 27 different cancer types. Our comprehensive FExSeq discovery pipeline effectively prioritize list of variants for follow-up functional studies.

1429M
FExSeq: A familial exome sequencing discovery pipeline. C. Chung1,2, Z. Wang1,2, X. Zhang1,2, M. Wang1,2, W. Luo1,2, S. Suman1,2, J. Burdett1,2, J. Boland1,2. 1) Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Frederick, MD; 2) Division of Cancer Epidemiology and Genetics, NCIC, NIH, Rockville, MD. Whole-exome sequencing has become a popular approach for disease-causing and/or disease-predisposing variant discovery. Analysis of massive amount of exome sequence data requires efficient bioinformatic approach. Although a number of state-of-art variant calling algorithms enabled accurate variant callings, a reproducible, streamlined pipeline for discovering biologically significant variants in the phenotype of interest is scarce. We have developed a pipeline called FExSeq discovery pipeline that utilizes familial exome sequencing data to 1) perform effective filtration by multiple parallel comparisons to public datasets; 2) perform segregation analysis to identify variants that are family/disease group specific and/or shared among disease group; 3) to prioritize functional variants based on multiple annotation scheme. Here, we present the results from analyzing ~1,500 whole-exome sequenced samples from 710 families of 27 different cancer types. Our comprehensive FExSeq discovery pipeline effectively prioritize list of variants for follow-up functional studies.

1430T
A novel approach to methylation-Seq data analysis based on functional principle component analysis (FPCA). S. Guo1,2, L. Liu1, J. Jiang1, N. Lin1, M. Chen1, L. Jin1, M. Xiong1. 1) Human Genetics Center, Division of Biostatistics, University of Texas School of Public Health, Houston, TX 77030, USA; 2) State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200433, China. Methyl-CpG binding domain (MBD) protein-enriched genome sequencing (MBD-seq) has been considered as one of the most important approaches for methylation-wide association studies (MWAS). However, the present MBD-seq statistical analysis methods for differential methylated region (DMR), including 1) average reads in certain region or 2) to estimate methylation level by using computational tools, have not possessed generally accepted standards. In our present study, we propose a novel method for the analysis to the MBD-seq data in which the methylation profiles were taken as random functions and functional principle component analysis (FPCA) was introduced to fit the methylation profile for CpG island, CpG shore and CpG shelves. We develop FPCA-based statistics to test for differential methylation. We used sparses hierarchical and K-means clustering to discover biologically meaningful pattern. Genome-wide DNA methylation-seq profile of 24 pairs of colorectal cancer (CRC) and normal tissues were used to validate the performance of the method. Simulation results showed that our proposed method substantially outperforms the traditional methods for testing differential methylation. In real data analysis, our method identified 3,425 hyper-methylation regions (HEMR) and 692 hypo-methylation regions (HOMR) in tumors, in which, 738 HEMR and 224 HOMR cannot be identified by traditional methylation methods, among which the most important CRC associated CpG island methylator phenotypes (CIMP) such as CACNA1G, IGF2, NEUROG1, RUNX3, SOCS1, UCHL1, ADAMTS1 and AOX1 were identified. DMRs identified by FPCA showed significant consistency with the distribution of DMR in tumor and normal tissues. We found more than 39.3% of DMR were distributed in CpG shore and CpG shelves, which suggest the important biological role of CpG shore and CpG shelves in cancer development and progression. In summary, we provided an effective novel method to analyze next-generation MBD-Seq dataset based on Functional PCA (FPCA), which can identify aberrant DMR in colorectal cancer.

1431T
FPCA: A novel statistic to analyze methylation-Seq data. L. Hou1,2, A. Baccarelli1, Y. Zheng1,2. 1) Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL; 2) Department of Biostatistics, University of Michigan, Ann Arbor, MI. Functional principle component analysis (FPCA) is one of the popular approaches to detect the disease associated differential methylation region (DMR). In our present study, we proposed a novel method for the analysis to the MBD-seq data in which the methylation profiles were taken as random functions and functional principle component analysis (FPCA) was introduced to fit the methylation profile for CpG island, CpG shore and CpG shelves. We develop FPCA-based statistics to test for differential methylation. We used sparses hierarchical and K-means clustering to discover biologically meaningful pattern. Genome-wide DNA methylation-seq profile of 24 pairs of colorectal cancer (CRC) and normal tissues were used to validate the performance of the method. Simulation results showed that our proposed method substantially outperforms the traditional methods for testing differential methylation. In real data analysis, our method identified 3,425 hyper-methylation regions (HEMR) and 692 hypo-methylation regions (HOMR) in tumors, in which, 738 HEMR and 224 HOMR cannot be identified by traditional methylation methods, among which the most important CRC associated CpG island methylator phenotypes (CIMP) such as CACNA1G, IGF2, NEUROG1, RUNX3, SOCS1, UCHL1, ADAMTS1 and AOX1 were identified. DMRs identified by FPCA showed significant consistency with the distribution of DMR in tumor and normal tissues. We found more than 39.3% of DMR were distributed in CpG shore and CpG shelves, which suggest the important biological role of CpG shore and CpG shelves in cancer development and progression. In summary, we provided an effective novel method to analyze next-generation MBD-Seq dataset based on Functional PCA (FPCA), which can identify aberrant DMR in colorectal cancer.

Germline copy variant numbers (CNVs) account for a significant fraction of hereditary cancer pathogenic mutations. For example, approximately 5-10% of BRCA1 and BRCA2 pathogenic variants involve large deletions and duplications. Traditional analysis for these rare events was performed by multiplex ligation-dependent probe amplification, quantitative PCR or comparative genomic hybridization. Although germline CNVs can be detected from next-generation sequencing (NGS) data generated using targeted DNA capture technologies (e.g., exomes and other panels), methods for doing so must overcome many technical challenges. Several algorithms have been published to detect CNVs in such data, though they may not yet be adequate for use in diagnostic testing laboratories, particularly for detection of small single-exon CNVs. We have previously described a new method, CNVtiae which is designed to detect single-exon CNV as well as larger regions sequenced using NGS with sensitivity and specificity of >99%. We have tested clinical samples where we performed NGS on 29 genes associated with hereditary cancer syndromes where CNVs were detected. One case is a 38-year-old Christian Arab woman diagnosed with invasive ductal carcinoma at 30 years of age. A duplication involving exons 5-11 of BRCA2 was identified from NGS read count data using CNVtiae. To understand the impact of this duplication on the BRCA2 protein we analyzed sequence alignments from this region for either split-reads or discordant mate pairs and were able to confirm that this duplication occurred in tandem within the gene. The duplication is predicted to cause a Met to Arg change at codon 1594 followed by a frame-shift that ends with a premature truncation at codon 1597. The truncated protein is expected to result in a loss-of-function: a well-documented mechanism for BRCA2 inherited breast cancer susceptibility. In addition to being a novel duplication, this is the first clinically reported duplication in BRCA2 using these new methods of detecting CNVs. We have also identified several other CNVs that are of interest and may warrant genetic counseling for our breast cancer patients.

OncoRep: An n-of-1 reporting tool to support genome-guided treatment for breast cancer patients using RNA-sequencing. T. Meissner, K.M. Fisch, L. Gioia, A.I. Su. Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA, United States.

Introduction Breast cancer is the leading cause of cancer among females. Being a heterogeneous disease, it comprises multiple tumor entities associated with distinctive histological patterns, different biological features and clinical behaviors. Next generation sequencing allows us to study this heterogeneity in multiple dimensions. This provides in depth insight into tumor pathogenesis on the individual patient level paving the path to individualized medicine. The advent of individualized medicine introduces a shift in how cancer patients will be treated in the future, away from the one drug-one disease paradigm towards the idea of bringing the right drug to the right patient. Problem Challenges that arise with this paradigm shift are i) preprocessing and analyzing sequencing data in the n-of-1 setting ii) extracting relevant information from various layers of omics data iii) integrating omics data with drug databases, (iv) presenting the information to the clinician in an understandable manner, and (v) completing these steps in a timely manner to provide clinically relevant and actionable targets to a tumor board or the treating physician. Results To address these challenges we present the extensible and customizable open-source framework OncoRep, a RNA-Seq based n-of-1 reporting tool for patients with cancer. We have applied it to breast cancer, including prospective molecular classification, detection of altered genes, and pathways, identification of gene fusion events, actually actionable mutations and it reports suitable drugs based on identified actionable targets. It presents integration and visualization of these omics data in an approachable html based interactive report as well as a PDF based summary report, providing the clinician and tumor board with a tool to guide the treatment decision making process.


We have identified >70 breast cancer risk loci that have been identified at genome-wide significance, but there may be many more loci with small effect. The challenge is to understand how these loci influence disease risk, and to find read-outs of the combined risks. We are seeking evidence of common mechanisms by mapping the genes associated with these loci onto breast cancer gene regulatory networks. We have examined the regulatory network in breast cancer for enrichment of genes associated with breast GWAS loci. First we used ARACHNe to assign potential target genes, so-called regulators, to transcription factors (TFs) using gene expression data from Metabric. A second step examined whether these regulators are enriched for genes associated with breast cancer GWAS loci, by applying an extended variant set enrichment (EVSE) that uses eQTLs to define the genes associated with the known breast cancer risk loci. We have implemented EVSE on a genome-wide basis and identified 36 TFs significantly associated with risk in both cohorts (995 and 997 samples) of Metabric. Interestingly four master regulators of FGF2 signalising, previously defined as a risk-associated pathway, were found as risk TFs (ESR1, GATA3, FOXA1 and SPDEF). We extensively validated our EVSE results: (1) Negative control GWAS signatures e.g. bone mineral density or random SNPs do not yield a positive association. (2) EVSE with random but size matched regulators does not find a significant association. (3) Lastly, we demonstrate that the eQTL step in the analysis strongly influences the results: eQTLs from ER+ and ER- tumours identify different sets of risk associated TFs. By studying the correlation of gene expression between regulators, we found that risk TFs fall into two distinct, but internally highly correlated groups, each of which may represent a group of TFs relating to a common mechanism. Using the TCGA data set we find that 33% of the risk TFs showed somatic alteration in more than 3% of tumours, while only 0.3% of a set of control TFs were altered to the same extent, providing independent validation of our results and suggesting that similar pathways may be responsible for breast cancer risk as well as disease progression. 1. Fletcher, M et al. (2013) Nat. Commun. 4, 2464-76. 2. Carro, M.S et al. (2012) Nature 463, 318-325. 3. Curtis, C et al. (2013) Nature 486, 346-352. 4. Cowper-Sal Iari, R et al (2012) Nat. Genet. 44, 1191-11985. 5. Michailidou, K et al. (2013) Nat. Genet. 45, 392-398.
A method for the discovery of long-range genomic interactions from 3C-seq experiments. T. Yuan, M. Du, R. Dittmar, S. Xia, Y. Guo, L. Wang, Pathology, Medical College of Wisconsin, Milwaukee, WI.

Abstract: Integration of chromosome conformation capture (3C) with high-throughput sequencing (HTS) (3C-seq) enables us to detect long-range genome-scale chromatin interactions between two genomic regions through counting fragment types, called fragments co-localized in the genome. However, the signals from real chromatin interactions are often overshadowed by a large number of the signals from short-range interactions or accident interactions. It is difficult to determine specific long-range chromatin interactions by setting a threshold of read counts (RC). Here, we present a new algorithm known as Tscore calculation. The algorithm assumes that the most chromatin interactions detected by sequencing are non-specific and randomly distributed, of which density functions can be fitted. For a specific genomic region, the density function of the other genomic regions along this chromosome interacted with this viewpoint (cis-interactions) are calculated through the density function of the normal distribution fitted by multiplying the count numbers of cis-localizations and the distance between the viewpoint and the other regions, and the cumulative probabilities of the regions on the other chromosome interacted with the viewpoint (trans-interactions) are calculated through the density function of the exponential distribution fitted by the counting of trans-localizations between the viewpoint and the other regions. The scores (Tscores), the conversion of the cumulative probabilities, can be used for measuring frequency of cis-trans-interactions. With addition, by comparing various factors, we found that background noise and counting methods have significant effects on the detection of chromatin interactions. We therefore develop a tool known as 3C-analyzer, an integrated analytic approach to 3C-seq analysis. 3C-analyzer integrated all analytic work from raw data in FASTQ format to the detection of significant chromatin interactions from 3C-seq experiments. It provides a user-friendly experience by including graphic user interface, sample management, and primary enzyme setup. The ability of parallel processing is optimized for large data analysis. Our case studies show that 3C-analyzer enables us to detect significant genomic interactions across different 3C technologies (Capture-C and 4C).

Using a reference panel to increase coverage in pooled sequencing experiments. H. Al-Asadi, M. Stephens. University of Chicago, Chicago, IL.

Pooled sequencing is an increasingly popular technique for assessing genome-wide population allele frequencies. Increasing the sequencing depth- or coverage - improves the accuracy of these allele frequency estimates but requires additional cost. Reference panels are often available in pooled sequencing applications and provide additional information about a sample—namely SNP correlations—which allows us to assemble examples along with any others from linked SNPs. In this study, we demonstrate use of an LD-based model which incorporates information from a panel to increase the effective coverage for each SNP. For example, with two perfectly correlated SNPs, our approach allows us to leverage information from one SNP to be used for any one SNP. We test our method on two evolve and re-sequencing experiments.


Genomic DNA (gDNA) is used as starting material in the experimental workflow of many applications in molecular biology. The integrity of the DNA critically affects the success of many downstream experiments like array CGH or sequencing. Initial electrophoretic analysis of the sample is highly recommended as the respective downstream applications can be expensive and time consuming. The Agilent Genomic DNA ScreenTape Assay has been primarily developed for the electrophoretic analysis of genomic DNA samples. A ScreenTape is a pre-packaged microfluidic device designed for performance of electrophoretic migration in a microfluidic chip. It is used in combination with the Agilent 2200 TapeStation instrument. Degradation of gDNA is typically a gradual process in which high-molecular weight DNA is fragmented into smaller species. It can occur either enzymatically, chemically, or mechanically. The integrity of gDNA can be evaluated by visual examination of the electrophoretic trace and is subjective and can be error-prone. In order to standardize this a novel algorithm was developed to score gDNA samples on the 2200 TapeStation. The DNA integrity number (DIN) is calculated from several features obtained from the electrophoretic trace and ranges from 1 to 10. Here we show data demonstrating the reproducibility, scalability and linearity of the DIN. The DIN is independent from instrument, reagent and sample concentration variability and can be used as objective measure for determining the integrity of gDNA.

Genomic susceptibility for cancer prediction by supervised machine-learning methods on SNP-ontology, S. Kim1, M. Kim1, 2 Chemistry, University of California, Davis, CA, USA; 2 Computer Science, University of California, Davis, CA, USA.

It has been widely assumed that human genomic variations are associated with individual’s susceptibility to complex diseases such as cancer. However, extensive association studies have so far had limited success in that the results have low predictive value of practical utility to individuals. We present a prediction process where two supervised learning methods are applied to two different descriptors of each individual’s common genomic variations. Using this approach, we predict an individual’s susceptibility to each of 8 major cancer traits plus healthy trait. The accuracy of the prediction ranges from 33 to 57% depending on cancer type, significantly better than 11% for a random prediction, with probability estimates that may be useful for making practical health-decisions for individuals or for a population.


The majority of observed somatic variants are thought to be passengers and distinguishing them from true functional drivers of cancer remains an ongoing challenge. While great strides have been made in predicting the impact of variants in the context of inherited disease, currently, there are tools specifically designed for annotating somatic variants that have been developed. One cause for this delay has been the lack of reliable datasets for training, but the growing availability of cancer genome data is helping to address this problem. Following the strategy utilized in previous work where recurrent and well-characterized genes lists were used to distinguish putative driver mutations from passengers, we composed a collection of positive and negative examples comprised of variants from the COSMIC, ICGC, TCGA, and UniProt repositories. Using these datasets to evaluate several annotation tools, we observed that while cancer-specific methods performed reasonably well, and large they were outperformed by a composite of predictions from several tools—most of which were intended to assess variants associated with hereditary diseases, not cancer. Second, because most of current methods rely on sequence conservation as a proxy for function, we found them to be less sensitive in regions of intrinsic protein disorder. Finally, after applying these tools to variants predicted from TCGA glioblastoma samples were RNA data was also available, we found that sample-specific expression context provided valuable information, allowing us to de prioritize unexpressed genes while raising the priority of mutations in genes that were overexpressed or exhibited allele-specific expression—trajectories that are known to be associated with cancer. RNA-seq can be useful for inferring functional impact of somatic variants. In light of these observations, we developed a novel composite classifier that incorporates several cancer-focused prioritization algorithms in conjunction with several well-established methods. Additionally, it considers a disorder prediction score, and examines over- and allele-specific expression at a per sample level. Taken together, our novel cancer-focused somatic variant prioritization method performed better than existing algorithms and represents a tool that can facilitate interpretation of somatic variants and understanding of cancer etiology.

Barcode-based template identification of KIR region in human genomes. C. Lo1, S. Zakov1, S. Kim1, B. Halliday2, V. Bathe3. 1 Computer Science and Engineering Dept, University of California, San Diego, San Diego, CA; 2 Biomedical Engineering, School of Science and Engineering, Reykjavik University Reykjavik University. Reykjavik, Iceland.

The immunospecific regions of the human genome, such as the region encoding the killer cell immunoglobulin-like receptors (KIRs) on Chromosome 19 and the region encoding the major histocompatibility complex in humans (HLAs) on Chromosome 6, are among the most important regions functionally, but their hypervariability, in gene copy number and mutations, makes it difficult to assemble and characterize these regions. Previous methods characterizing these regions rely on laboratory methods using traditional and quantitative PCR primers and probes. Here, we propose a computational method to type the KIR region and to determine various allele groups of KIR genes directly from whole genome sequencing data. Our method is based on barcoding (deriving signatures) known KIR templates as well as sequenced fragments, and comparing the two sets of signatures. The method is alignment-free and efficient, and easily scales to large populations of individuals. We demonstrate the method on over simulated data, trios from various populations in the 1000 genomes project as well as 298 Icelandic trios. The results indicate that this method can resolve all KIR types up to an equivalence that can only be broken by long range haplotyping or inheritance patterns.
1441M Network-Augmented Genomic Analysis (NAGA) applied to Cystic Fibrosis studies. S. Loguerco1, D. Marino-Roth2, D. Hutt3, A. Su4, W. Balch5. 1) Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA; 2) Department of Cell Biology, The Scripps Research Institute, La Jolla, CA.

Cystic fibrosis (CF) is an early onset disease characterized by a defect in the apical chloride channel, CF transmembrane conductance regulator (CFTR). The most common disease causing mutation is a 3 base pair deletion resulting in loss of Phe 508 (F508del), which leads to misfolding, endoplasmic reticulum (ER) retention and efficient ER associated degradation of the protein. To elucidate the molecular networks influencing the folding and function of CFTR in CF, we recently screened CFBE410-cells containing F508del-CFTR against a siRNA library of 2500 targets known to be involved in protein homeostasis. In parallel, we generated a high confidence CFTR interactome of F508del-CFTR in the same cell line. Given a list of high-scoring siRNA hits for CFTR rescue of function, and a set of CFTR binding proteins, we sought to connect these datasets through an integrated protein-protein interaction network, and use shortest path analysis to uncover the minimal network structure consistent with both the CFTR interactome and siRNA data. The goal of this approach is to prioritize proteins connecting CFTR with siRNA hits that may act as central “hubs” in cellular processes required for CFTR functional rescue. For each protein in the subgraph, it computes the number of distinct siRNA hits that utilize the protein on its shortest path to CFTR. In order to filter out nonspecific protein hubs, this computation is repeated using a random selection of hits from the original siRNA library. The analysis identified several novel candidates for CFTR rescue of function that could be validated through targeted siRNA screens. In summary, we present here a novel network-based method to integrate functional genomics data (e.g. siRNA screens) with interactomics datasets (e.g. AP-MS, MudPIT), useful for prioritizing novel functional targets and for identifying relevant network modules. It leverages publicly available information on protein-protein interactions but is remarkably generalizable to other scenarios where a connection between functional and biochemical data is sought.


Multiple quantitative genomic data types such as methylation proportion and copy-number are now commonly collected as part of studies seeking to investigate the genetic architecture of human disease. These data types are often extremely high-dimensional with thousands of measurements per subject and can possess a spatial structure with respect to genome position. The presence of these patterns leads naturally to treating the entire profile as the explanatory variable of interest when comparing groups rather than using measurements at single sites. This task can be challenging because profiles must be estimated from the data and are more complicated data types than un-smoothed site-level measurements. We have developed a penalized likelihood approach, Joint Adaptive Differential Estimation (JADE) for simultaneously preforming smoothing, estimation, and cross-group comparison of genomic profiles for quantitative data types with spatial patterns. The method is motivated by a model in which groups have a shared average profile over some of the genome but to differ in a few biologically relevant regions. JADE utilizes the similarity between groups to improve accuracy of group profile estimation while allowing for regions in which the profiles are different. We are also able to create adaptive clusterings which vary along the genome for joint comparison of more than two groups. By taking a unified approach we avoid many of the ad-hoc decisions made in multi-step techniques as well as the challenges of multiple-testing correction and window-size and region pre-selection. JADE is implemented through an efficient dual gradient descent algorithm. We have applied it to the task of identifying differentially methylated regions between different cell types and tissues through ENCODE database of reduced representation bisulfite sequencing experiments. In these experiments methylation level is obtained as a binomial proportion at millions of closely spaced sites. These data have previously been modeled using splines and other smoothing techniques allowing us to compare results using JADE and other methods.


In the last two years, NGS has begun to make significant inroads into the clinic via a set of screening and diagnostic applications such as carrier screening for recessive disease, and testing for inheritable cancer risk. Research tools such as BWA, GATK, FreeBayes, and others make up the current best practices in sequence analysis, but it is rarely appreciated that they are built with various implicit design decisions that are highly specific to the research regime and not appropriate in a clinical setting. We describe in detail KCcall, our whole-exon (155bp-4kb) diploid haplotype caller, which is based on diploid HMM inference over a De Bruijn graph. We describe its performance in various clinical testing applications, using data generated from our own patient stream of thousands of samples per month. KCcall was designed to address the differences between the high-volume clinical and research regimes -- Clinical test designs focus on small genomic regions of ~0.5-1Mb targeted at specific medical questions, rather than whole exome/ genome coverage in typical research studies; Clinically relevant variants are often “hard” sites for NGS and must be called correctly in every patient; research protocols often optimize the average genome-wide accuracy rather than performance in various clinical testing applications, using data generated from our own patient data. KCcall leverages these differences to call whole-exon (155bp-4kb) haplotypes using 100bp reads at near 30x depth sequencing data. Additionally, KCcall replaces heuristic tuning of variant calling parameters with a hands-off calibration procedure using calibration flowcells totaling >30,000x depth in our target area (4 HiSeq flowcells). This eliminates heuristic tuning by modeling error processes site-by-site down to a frequency of 1/1000. The end result of these changes is a tool that behaves with a consistent level of quality across all targets, requires zero subjective tuning steps, and provides a complete and accurate view into the patient’s DNA.
1445T

The current paradigm of clinical sequencing data analysis employs a lengthy discovery approach: aligning reads to the human reference assembly, discovering mutations from base to base and identifying mutations that are likely actionable. Although widely practiced, such a paradigm is not optimal for clinical applications, which demand rapid acquisition of clinically relevant, sensitive, and unambiguous molecular profiles. We developed ClinSeK based on a knowledge-driven inversely-operating paradigm that directly tests well-characterized, clinically-relevant variants from high-throughput sequencing data, without exhaustively aligning and comparing sequencing reads to the human reference genome. We overcome challenges in analyzing repetitive regions and duplicated reads under this new paradigm. Applying ClinSeK to characterize the molecular profile of over 600 deeply sequenced cancer samples indicated that this new approach increases sensitivity in detecting low frequency variants with over 50-fold reduction in processing time than existing approaches that perform sequential alignment and variant calling. ClinSeK can test point mutations, indels and structural variations in single or paired samples. It supports the analysis of both DNA and RNA sequencing data. It can also serve as a quick and independent cross-validation in complement to existing variant discovery pipelines.

1446S

A large portion of the human genome is known to be composed of long tandem repeats that can span several hundred kilobases to multiple megabases. Although the repeat motifs can be sequenced and the amount of repeat material can be approximated by conventional sequencing technologies, it is often difficult or impossible to assemble them into long contigs, so the exact locations and copy numbers of these repeats often remain elusive, especially when the unit length exceeds the read lengths. Without knowing the genomic context of these repeats or the amount of repeat material, it is difficult to attach any biological relevance to them. Thanks to the recent developments in nanochannel technology which allow us to image intact megabase-scale molecules of DNA, repeat regions can be more accurately characterized and put into biological context. Using Irys® technology and novel algorithms designed specifically to investigate long repeat arrays, we were able to find and characterize previously enigmatic repeat regions in the human genome. The genomic positions of these repeats were found by aligning non-repetitive portions of repeat-containing molecules or consensus de novo assemblies to the reference genome. The genomic context then provides insight into the biological significance of the repeats. Here, we have found that molecules containing a prominent 5.6-kilobase tandem repeat are derived from the Lipoprotein(a) (Lp(a)) coding region on chromosome 6, which is linked to atherosclerotic diseases such as heart disease and stroke. Furthermore, since the array is contained on single molecules, we can measure allele array lengths, showing that Irys® technology has potential for aiding in quick, accurate, and cost-efficient prognosis of these and other genetic diseases which are influenced by copy number variations.

1447M
 novoBreak: comprehensively characterizing somatic structural breakpoints in cancer genomes. Z. Chong1, M. Gao2, J. Ruan3, W. Zhou1, T. Chen1, H. Zafar4, J. Chen5, G. Mills5, K. Chen1. 1) Department of Bioinformatics & Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX; 2) Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX; 3) Agricultural Genomes Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China; 4) Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX.

Somatic structural variations (SVs) are a major source of genomic variations in cancer genomes and may play a driving role in cancer genome evolution. However, current algorithms for detecting SVs using Next-Generation Sequencing (NGS) data could not achieve a high sensitivity and specificity. This is mainly due to the challenges in aligning and interpreting short reads spanning SV breakpoints. We have developed a novel algorithm, novoBreak, which minimizes the alignment issues to achieve a high recall rate and a low false discovery rate. novoBreak can characterize almost all types of SV breakpoints including deletions, duplications, inversions, insertions and translocations at base-pair resolution. It first clusters reads around potential breakpoints and then locally assembles the reads associated with each breakpoint into contigs. After aligning the contigs to the reference, it can identify the precise breakpoints and infer the types of SVs. novoBreak performs substantially better than other widely applied solutions, such as BreakDancer, CREST, etc., based both on simulation data and on the whole genome sequencing data in the tumor genome atlas (TCGA) and the 1000 Genomes project. In the recent NCI ICGC-TCGA DREAM 8.5 somatic mutation calling challenge, novoBreak ranks as one of the best-performing methods. The higher sensitivity of novoBreak makes it possible to uncover more novel and rare SVs, which could delineate a more comprehensive landscape of variations in cancer genomes and shed light on cancer genome evolution. The algorithm of novoBreak can be easily extended to SV discovery, which could be used to characterize genome evolution in a long-term view.

1448T
SG-ADVISER: CNV annotation pipeline. G. Erikson, N. Deshpande, A. Torkamani. Scripps Translational Science Institute, La Jolla, CA.

CNVs have been associated with a diverse array of diseases, especially cancer, autism, schizophrenia, and developmental delay. A large number of methods are available for the annotation and interpretation of SNPs and small indels, yet little is available for interpretation of CNVs. The Scripps Genome Annotation and Distributed Variant Interpretation Server (SG-ADVISER) CNV pipeline aims to fill this gap. The SG-ADVISER CNV pipeline is a web server developed at The Scripps Translational Science Institute for the annotation of CNVs. Annotation execution proceeds in highly parallel fashion and includes classes of variant annotations that are entirely independent of one another, serially dependent annotations whose execution are dependent upon the completion and status of prior annotations, and synthetic annotations that generate new information through the combination of multiple annotation outputs. The annotated file includes details regarding location, impact on the coding portion of genes, allele frequency information including allele frequencies from the Scripps Welterdy cohort, and overlap information with other reference datasets including ClinVar, DBS, known syndromes etc. The CNV pipeline accepts variant files in CNVnator, Complete Genomics, or plain tab delimited file formats. A variant classification is produced (ADVISER score) based on the American College of Medical Genetics (ACMG) scoring guidelines with categories 1-5. Variants of category 1 are most clinically relevant. The performance of ADVISER classification scheme was evaluated using data from International Collaboration for Clinical Genomics (ICCG). We annotated both pathogenic and benign variants and the ADVISER classification scheme showed accuracy of 89% specificity and 99% sensitivity. To facilitate the interpretation of the SG-ADVISER CNV pipeline output we added new functionality to the existing SG-ADVISER UI. The SG-ADVISER UI is a visualization tool that allows scientists with little or no programming experience to easily and quickly view, manipulate, sort, and filter the SG-ADVISER CNV output file. For example it is possible to sort or filter CNVs based on Known Disease, Coding Impact or Chromosome Position etc. At any point during the process the tool can provide summary statistics, and output the filtered results to a new file.
1449S
Challenges to CNV Detection in the Clinic using Targeted High Throughput Sequencing Data. S. Sadedin, A. Oshlack. Murdoch Childrens Research Institute, Parkville, Victoria, Australia.

High Throughput Sequencing (HTS) is rapidly gaining clinical acceptance as a cost-effective solution for diagnosis of highly penetrant but genetically heterogeneous diseases. Despite this success, the analysis methods applied are usually limited to detection of point mutations and small insertions, and deletions. Detection of larger structural variations, most notably deletions, is often omitted despite such variants being equally deleterious and common causes of many disorders. While many tools have been published in the literature, in practice, few laboratories are implementing them in their clinical sequencing pipelines. We speculate that this is due to a perceived lack of specificity and robustness. In this work we apply a novel simulation method to explore the reasons for this gap using real data spliced from X chromosomes of male samples into female samples to simulate singlecopy deletions. We find that real world performance of these methods appears to be highly sensitive to a range of factors including: method-specific tuning parameters, exact sequencing technology and methodology, variability in the quality of data, and the number and type of samples sequenced in a batch. We conclude that in order for HTS CNV detection to become clinically accepted, methods must be developed that can work in a highly robust, self-calibrating fashion and that are well tuned to popular sequencing platforms. Along these lines, we work out results from our own method, Angel, which is specialized to achieve high accuracy and robustness on the HaloPlex targeted sequencing platform.

1450M
SAAS-CNV: A joint segmentation approach on aggregated and allele specific signals for the identification of somatic copy number alterations with next-generation sequencing data. Z. Zhang1,2, K. Hao1,2, 1 Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 2 Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY.

Motivation: Cancer genomes exhibit profound somatic copy number alterations (SCNAs). Studying tumor SCNAs using massively parallel sequencing provides unprecedented resolution and meanwhile gives rise to new challenges in data analysis, complicated by tumor aneuploidy and heterogeneity as well as normal cell contamination. While the majority of existing methods utilize total sequencing depth alone for CNV inference, the allele specific signals are undervalued. We proposed a joint segmentation and inference approach using both signal dimensions to address some of the challenges.

Methods: Our method consists of four major steps: 1) extracting read depth supporting reference and alternative alleles at each SNP/Indel loci and comparing the total read depth and alternative allele proportion between tumor and matched normal sample; 2) performing joint segmentation on the two signal dimensions; 3) correcting the copy number baseline from which the SCNA state is determined; 4) calling copy number variation (CNV) and copy-neutral loss of heterozygosity (LOH) for each segment based on both signal dimensions. The method is applicable to whole exome/genome sequencing (WES/WGS) data in a tumor-control study, and is readily extended to SNP array data and tumor-only setting.

Results: We applied the method to a data set containing no SCNAs to test the false positive rate, created by pairing sequencing replicates of a single HapMap sample as normal/tumor pairs, as well as a large-scale WGS data set consisting of 88 hepatocellular carcinoma (HCC) samples along with matched normal samples. SNP array data is also available for the 88 samples to serve as benchmark. Compared with representative methods, such as ExomeCNV and CNAnorm, our method demonstrated high and consistent accuracy, scalability to large cancer studies, capability in handling a variety of platforms, and potential in improving the estimation of tumor ploidy and purity.

Availability: An R package called saasCNV is available at zhang205.u.hpc.cs.msm.edu/saasCNV.

1451T
The new European Variation Archive Resource at EMBL-EBI. I. Medina, D. Spalding, C.Y. Gonzalez, G. Saunders, J. Kandasamy, S. Ur-Rehman, V. Kumanduri, I. Lappalainen, J. Paschall, European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Cambridge, United Kingdom, CB10 1SD.

The European Variation Archive (EVA; www.ebi.ac.uk/eva) is a new ‘one-stop shop’ for all freely available variation data, for analysis of variation from all species. EVA will work with partner databases to guarantee free global access to genetic variation data. The EVA data model focuses on archiving and making available the results of large-scale variant discovery efforts, providing granular access to population-scale allele frequencies paired with study metadata. Data submitted to EVA in VCF format is closely linked with supporting BAM alignments where available in the ENA or EGA databases at the EMBL-EBI. EVA currently contains over 1.7 billion submitted variants, from a range of more than 15 large-scale projects including 1000 Genomes (phase 1 and 3), Exome Variant Server, Genome of the Netherlands and UK10K. We also mirror the data stored at both dbSNP and ClinVar permitting our users easy access to this data. As an example of our commitment to non-human species, EVA includes data from the NextGen project that has compiled more than 250 million livestock variants including rich meta-data on sub-strain and environment. With regards structural variation, EVA builds upon the existing Database of Genomic Variations (DGV; www.ebi.ac.uk/dgva) that, in collaboration with dbVar at NCBI, currently contains in excess of 10 million Supporting Structural Variants (SSVs) in 120 studies, covering a range of 28 species. The web portal of EVA aims to provide a dynamic and visually interactive set of queries and filters based on modern web technologies such as HTML5. Using a browsing and explore our study or variation catalogue, visualize variations or search for a gene. This front-end is made modular and scalable through use of a RESTful web service interface to the backend data store, allowing EVA data to be accessible programmatically for a variety of applications such as annotation pipelines. To provide a high-performance and scalable solution EVA is based on NoSQL databases such as MongoDB and is designed to accommodate data derived from tens of thousands of whole genomes. Data mining and visualisation tools allow direct access to data, as well as the granularity of a given submitted study. Complex datasets can be built up based upon entry points including gene, study, genomic location, variation type and consequence type. All software is released as open source.

Please direct questions and submissions to eva-helpdesk@ebi.ac.uk.

1452S

Reliable prediction methods are needed to analyze NGS datasets. Due to the large volume of identified variations only computational approaches can handle the datasets. Genomes contain millions of variants and typically over 10% of them lead to amino acid substitution. The applied methods have to be both reliable and fast. We have developed a completely new machine learning-based method PON-P2, which fulfills these criteria. It has excellent performance (accuracy 0.87, MCC 0.77) and it can predict very large datasets in reasonable time. It is more reliable and faster than our previous predictor, PON-P1 and competing methods. PON-P2 is for amino acid substitutions. The method is based on extensive feature selection and training with a large benchmark dataset from VariBench [2]. The method is implemented with random forest. PON-Diso is a method for predicting effects of amino acid substitutions on order/disorder status of proteins [3]. Several proteins contain disordered regions or are completely disordered, i.e. without regular ordered structure. Changes to the structure in these regions can be related to diseases. We tested the performance of a number of existing disorder prediction methods, but they were found unsuitable for this task. We developed a novel tool that has a success rate on 70% in cross validation. A related tool was trained for mismatch repair (MMR) gene variations [4]. PON-MMR has accuracy of 0.87. When applied to 758 unclassified variants in InSiGHT database, it could classify 248 cases as pathogenic or benign. InSiGHT variation interpretation committee [5] classified 1370 variants out of 2380 investigated. 46 of those we had predicted to be either benign or pathogenic. 44/46 of them were correct (96%). These tools have been tested with independent data sets and show they have the highest performance currently available. Available at http://structure.bmc.lu.se/.

1453M
HLA-Genotyper Prediction of HLA Genotypes from Next Generation Sequencing Data. J. Farrell1, G. Jun1, L.A. Farrell1, A. DeStefano2, P. Sebastiani3, 1 Biomedical Genetic- Evans 218, Boston Univ Med, Boston, MA; 2) School of Public Health, Boston University, Boston, MA.

Background: HLA genotyping of next generation sequencing data would be useful for testing the association of HLA alleles with adverse drug reactions, auto-immune diseases, infectious diseases and genetic diseases.

Methods: We have developed HLA-Genotyper for the prediction of HLA genotypes from next generation sequencing data. HLA-Genotyper is a software tool which performs 4-digit HLA allele prediction using a novel Naive Bayes algorithm. The evaluation of the novel algorithm for MHC 1 and II Classical HLA loci (A, B, C, DRB1, DOA1, and DQB1) was conducted using whole genome, whole exome and RNA-Seq data from 51 European and 50 Yoruba samples from the 1000 Genomes Project. To validate the predictions, the precision and recall of the predicted HLA genotypes were compared to a “gold standard” based on SSO and SBT HLA typing.

Results: For 560 HLA alleles predicted in 51 low coverage (3-7x) European samples, the precision was 0.96 with a recall of 0.59. For the 597 HLA alleles predicted in 50 low coverage Yoruba samples, the precision was 0.90 with a recall of 0.49. For the 51 European samples with whole exome sequencing (50-100x coverage), the precision of the HLA genotype predictions increased to 0.97 with a recall of 0.96. For the 50 Yoruba samples with whole exome sequencing, precision decreased to 0.90 with a recall of 0.90. In 45 European samples, the predictions of 518 HLA alleles from RNA-Seq data were near complete concordance with the “gold standard” with a precision of 1.0 and recall of 0.99. The RNA-Seq validation results of 453 assayed alleles from 38 Yoruba samples also had very high precision and recall rates of 0.98. Conclusion: With modest coverage and read lengths between 75 and 101 bp, the HLA-Genotyper software accurately predicted HLA genotypes from next generation sequencing data commonly used for research. The predictions were best from the RNA-Seq data with near concordance with the gold standard and better than PCR-SSO accuracy. Using HLA-Genotyper, researchers may readily unravel the association of HLA alleles with many diseases from next generation sequencing experiments without the expensive and laborious HLA typing of thousands of subjects in diverse ethnic populations. This will improve researchers ability to understand biological processes such as T-cell recognition and biological illness and adverse drug reactions.

1454T

Next generation sequencing (NGS) technologies enable the detection of novel rare variants in genome wide scale. Haplotype phasing of these variants is important for rare variant association studies because variants are often grouped into exon or gene level. Effects of multiple variants and phase-dependent interactions such as compound heterozygosity and co-inheritance are considered for the analysis. In addition, phased haplotype information is required for estimating population genetic parameters by considering a coalescent tree or an ancestral recombination graph based inference. In the process of haplotype phasing, various approaches have been proposed by considering the linkage disequilibrium between variant sites. These types of approaches can provide accurate phasing results for common variants, but their accuracies for low-frequency variants or variants around recombination hotspots tend to be low. The other type of approach uses NGS reads that have phase-informative reads, spanning multiple heterozygous variant positions (called phase-informative reads). Although this approach is promising for phasing low-frequency variants, the applicable length of the estimated haplotypes is limited due to the current length of NGS reads (around hundred bases). However, since the length of sequence reads is growing rapidly, the rate of heterozygous sites phased by phase-informative reads is expected to increase. The quality of estimated variants is crucial in haplotype phasing based on NGS reads, while accurate variant calling is still challenging due to errors on sequencing and read mapping. In particular, methods for calling phased variants can be corrected by considering haplotype phasing, simultaneous estimation of variants and haplotypes is important. Here, we introduce a statistically unified approach for variant calling and haplotype phasing named HapMonster, in which haplotype phasing information is used for reducing the number of false positive variants, improving the recall of variant calling and the improved variant calls are used for more accurate haplotype phasing. From the comparison with other existing methods on simulation and real sequencing data, we confirm the effectiveness of HapMonster in both variant calling and haplotype phasing.

1455S

Background: Structural variation in the genome can have a profound effect on gene function. In particular, short tandem repeat (STR) expansions are implicated in a wide array of genetic disorders, such as Huntington Disease. We developed software able to call genotypes of complex variations from next generation sequencing (NGS) with short paired-end reads that current software packages fail to assemble and call. These failures have been attributed to 1) most de novo sequencing algorithms utilize De Brujin graphs and are not well-suited to assembling low-complexity sequences due to k-mer limitation, and 2) de novo sequencing algorithms often utilize sequence error correction techniques that remove information from low-complexity sequences and interfere with STR sequence assembly.

Methods: We developed Kragle, software that combines and adapts three known algorithms: paired-end aware read recruitment, consensus-overlap de novo sequence assembly and a statistical framework for genotypes. These components were combined to take full advantage of the reads and assemble low-complexity sequences with repetitive content as long as the read length. The statistical framework then allows us to call diploid genotypes from the assembled haplotype sequences and assign confidence based on alternative genotype hypotheses. Results: We processed and analyzed 31 positive control samples representing 11 distinct STR-related loci and inherited conditions in technical duplicate. We used Kragle to genotype the samples at 9 of these 11 loci (2 loci failed PCR) and at 10 additional STR loci. The resulting matrix of 51 samples x 19 genotypes showed complete agreement within samples and at least one correct genotype in 19 of 21 samples (90.5%). This translates to 100% sensitivity and 99.2% specificity. The reproducibility of genotype calls between technical replicates was 97% when alleles were allowed to differ by +/-1 repeat. These genotyped loci include STR’s in the following genes: AR, ATN1, ATXN1, ATXN2, ATXN3, ATXN7, DMPK, FXN, and HTT. In another experiment we were able to call a heterozygous large deletion in BRCA1 gene, and confirm it with PCR.

Conclusions: The above experiments confirmed that Kragle is capable of assembling and genotyping STR sequences up to the length of an NGS read, and assembling junctions resulting from deletions, insertions and chromosomal rearrangements. This will enable clinical diagnosis of many genetic conditions using short-read NGS data.

1456M
Pipeline and Varant annotation tool for identifying causal variants in inherited rare disorders. K. Kundu1, S. Rana1, A. Chellappan1, U. Sunden2, J.M. Puck3, S.E. Brenner3, R. Sinkivasan1. 1) Innovation Labs, Tata Consultancy Services, California, CA 94720, USA; 2) Department of Pediatrics, University of California, San Francisco, Box 0519, CA, USA; 3) University of California, Berkeley, CA 94720, USA.

We have developed a pipeline for the analysis of genomic variant data, which identifies features enabling solving numerous clinical cases related to SCID (Severe Combined Immunodeficiency) and related diseases. The first several steps of the pipeline employ standard tools for mapping and processing of the mapped data in preparation for variant calling, but integrate innovative and efficient features to yield high quality sets of variants. Quality metrics for mapping, gene coverage, and called variants are generated throughout the run to ensure confidence in downstream analyses. Our Varant tool (available under an open source license) provides extensive variant annotations, such as genes and transcripts affected by the variant, the type of effect, possible functional consequences including pathways that may be affected by the variant, known disease associations from a variety of data sources such as ClinVar, OMIM, GAD, the mouse phenotype database and the inherited rare disorders. We benchmarked Varant to ensure that it not only has all the features present in other widespread tools, such as Annovar and snpEff, but also overcomes the errors made by other tools and is more liberally licensed. For cases where pedigree or phasing information is available the Varant tool classifies variants as following a compound heterozygous model, de novo, uniparental disomy, or X-linked recessive with potential clinical importance. The annotations are combined with predictions from programs such as ToppGene and Endeavour to further prioritize short-listed variants.

Our pipeline was able to identify likely causal variants in several cases where routine protocols would have been expected to fail. In one example of a patient with abnormal T-cell receptor excision circles (TREC) we were able to identify and later confirm mutations in the NBN (nibrin) gene that led to the identification of Nijmegen syndrome. In another case, we were able to identify the large homozygous mutation in the DBC1 gene in a patient with SCID; other tools missed a key variant because the reference genome incorporates a frameshift mutation.
SVSI: A Fast and Powerful Set-Valued System Identification Approach to Identifying Rare Variants in Sequencing Studies for Ordered Categorical Traits. W. Bi\textsuperscript{1}, G. Kang\textsuperscript{1}, Y. Zhao\textsuperscript{1}, Y. Cui\textsuperscript{2}, S. Yan\textsuperscript{1}, Y. Li\textsuperscript{2}, C. Hartford\textsuperscript{2}, W. Leung\textsuperscript{2,7}, J. Zhang\textsuperscript{1.1} 1) Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences; 2) Department of Biostatistics, St. Jude Children’s Research Hospital, Memphis, Tennessee; 3) Department of Statistics and Probability, Michigan State University, East Lansing, Michigan; 4) Department of Genetics, Department of Biostatistics, University of North Carolina, Chapel Hill, NC; 5) Department of Computer Science, University of North Carolina, Chapel Hill, NC; 6) Department of Bone Marrow Transplantation and Cellular Therapy, St. Jude Children’s Research Hospital, Memphis, Tennessee; 7) Department of Pediatrics, University of Tennessee Health Science Center, Memphis, Tennessee.

For phenotype-genotype association studies that involve a phenotype with ordered multiple response categories, we usually either regroup multiple categories of the phenotype into two categories (“cases” and “controls”) and then apply the standard logistic regression (LG) model, or apply non-parametric method of spearman rank correlation or parametric method of ordered logistic (oLG) regression model which accounts for the ordinal nature of the phenotype. However, these approaches may lose statistical power or may not control type I error rate if the underlying genetic variants are rare or sample size is limited. Here we propose a set-valued (SV) system model, which assumes that the underlying continuous phenotype follows a normal distribution, to identify genetic variants associated with an ordinal categorical phenotype. We couple this model with a set-valued system identification method to identify all the key underlying system parameters. Simulated studies show that SV well controlled the type I error rate. In the comparison among LG, SV and oLG methods, LG had significantly lower power than SV and oLG due to disregarding of the ordinal nature of the phenotype, and SV had similar or higher power than oLG. Additionally, the association parameter estimate was 2.72–2.87 fold less variable compared to the oLG association parameter estimate. Less variability in the association parameter estimate translates to greater power and robustness across the spectrum of minor allele frequencies. These advantages are most pronounced for rare and low penetrance variants common in genetic association studies for ordered categorical phenotype.

A comprehensive empirical evaluation of linear mixed models for GWAS. D. Heckerman\textsuperscript{1}, C. Widmer\textsuperscript{1}, O. Weissbrod\textsuperscript{2}, N. Fusi\textsuperscript{3}, C. Kadie\textsuperscript{3}, R. Davidson\textsuperscript{2}, J. Listgarten\textsuperscript{1}, C. Lippton\textsuperscript{1}. 1) Microsoft Research, Los Angeles, CA; 2) Computer Science Department, Technion, Israel; 3) Microsoft Research, Redmond, WA.

The linear mixed model (LMM) has recently emerged as the model of choice to correct for confounding structure, including population structure and family relatedness, in GWAS. At their core, mixed models rely on the estimation of a genetic similarity matrix (GSM), which encodes the pairwise similarity between every two individuals in a cohort. These similarities are estimated from SNPs or other genetic variants. Traditionally, the GSM for an LMM has been estimated from all available variants. Recently, however, it has been proposed that only a subset of variants should be used. The motivation behind this suggestion is based on a mathematical equivalence between the LMM and linear regression. In particular, an LMM is equivalent to a form of linear regression in which the variants that determine the GSM in the LMM view are covariates in the linear-regression view. Thus, taking the linear-regression view, one should only estimate the GSM using variants that are relevant to the phenotype. When all variants have an influence on the phenotype, this understanding leads to the traditional approach of including all of variants in the GSM. When some variants have absolutely no influence on the phenotype, then they should be omitted. But what happens when some variants are only marginally relevant, either having an extremely weak causal effect on the phenotype or an extremely weak association with the phenotype through confounding factors? We investigated variant selection with an empirical study of synthetic data sets spanning a wide range of population structure, family structure, and polygenicity, measuring performance in terms of control of type I error and power. Generally, we find that, when population and/or family structure is present, variant selection alone does not perform well. In particular, when population structure is present, the inclusion of principal components (PCs) as covariates to the model improves performance. When family structure is present, variant selection improves performance only when it is used to create a second GSM in addition to a GSM based on all variants, yielding a dual GSM. When population structure and family structure are both present, a combination of family PCs and family structure is required to recover the association between SNPs and the phenotype. When population and/or family structure is present, use of variant selection is only beneficial when there are a relatively small number of causal SNPs with relatively large effect sizes.

Accelerating curation of the catalog of GWAS by automatic text mining. C. Hsu. UC San Diego, La Jolla, CA.

The Catalog of GWAS is an important resource containing published association between SNPs and phenotypes identified by Genome-Wide Association Studies (GWAS), a well-defined study approach. GWAS studies have been successfully producing discovery and replication of many new disease loci. The number of GWAS is growing rapidly. There is a need for a database that allows researchers to easily query and search for previous results. Such database has been created and maintained by the National Human Genome Research Institute (NHGRI), called “A Catalog of Published Genome-Wide Association Studies” (Catalog of GWAS). The catalog has led to interesting characterization of previous results in GWAS and NHGRI has continued to update and curate the catalog regularly. However, curation of the catalog is current performed by expert curators. Though this will ensure the quality but new publications in GWAS really outpace any human curation team can possibly handle. The goal of the project GWASTool is to develop a new tool to automatically extract the information from research articles for the curation of the catalog of GWAS. Our proposal is to use the curated data currently available from NHGRI as the training examples and apply novel machine-learning algorithms to train information extractors to allow accurate automatic extraction. Machine learning is particularly suitable for this project because a sufficiently large set of training examples is available from the Catalog of GWAS. The idea is to match the records to the text of the article to mark where they were mentioned and should be extracted from. Moreover, data fields for GWAS studies are relatively well defined, unlike other biological text mining tasks, such as extracting protein-protein interaction. The information extractors to be developed will consist of a tagging module to tag the information in full-text of GWAS research articles and a link module to link the tagged values into entries. A special tagger, based on template matching, will be developed to tag and link the information given in tables. We will demonstrate how the automatic information extraction tool works through a Web-based online tool and present remarkable statistics of the GWAS studies from the extracted data.
1460T

GACT: A Tool for Predicting and Converting Genome Build and Allele Definition during Imputation and Meta-analysis of SNP Genotype Data. A. Sulovari1,2, D. Li1,3,4. 1) Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, VT; 2) Cell, Molecular and Biomedical Sciences Graduate Program, University of Vermont, Burlington, VT; 3) Department of Computer Science, University of Vermont, Burlington, VT; 4) Neuroscience, Behavior and Health Initiative, University of Vermont, Burlington, VT.

Meta-analysis is one of the commonly used methods in genetic studies of complex diseases as it can increase the statistical power to identify new disease-associated variants by combining genotypes from multiple GWAS. This approach requires same allele definition and genome build among individual studies. Similarly, genotype imputation, commonly used prior to meta-analysis, requires the same consistency. However, the genotypes from various GWAS are generated using different genotyping platforms, arrays or SNP-calling approaches, resulting in use of different genome builds and allele definitions. Incorrect assumptions regarding allele definition combined with genome-wide meta-analysis and imputation of genotypes show no significant decrease of imputation quality (even significantly higher when compared to the imputation with singletons in the reference), especially for rare SNPs. Based on our GWAS data, an imputation of a set of SNPs with a number of 600,000 markers to be sufficient for high quality genome-wide imputation of rare SNPs (high quality assayed SNPs may compensate for low true-genotypic density). In conclusion, we have developed a new, powerful, and user-friendly tool with both command-line and interactive online version that can accurately predict, and convert between any of the common allele definitions and between genome builds for genome-wide meta-analysis and imputation of genotypes from SNP-arrays or deep-sequencing. Availability: GACT (http://www.umm.edu/genomics/software/gact).

1461S

Applying compressed sensing to genome-wide association studies. S. Vattikuti1,2, J.J. Lee3, C.C. Chang4,5, S.H. Hsu6, C.C. Chow1, 1) Mathematical Biology Section, Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD; 2) Department of Psychology, University of Minnesota Twin Cities, 75 East River Parkway, Minneapolis, MN 55455, USA; 3) BGI Hong Kong, 16 Dai Fu Street, Tai Po Industrial Estate, Tai Po, Hong Kong; 4) Cognitive Genomics Lab, BGI Shenzhen, Yantian District, Shenzhen, China; 5) Office of the Vice President for Research and Graduate Studies, Michigan State University, 426 Auditorium Road, East Lansing, MI 48824, USA.

The aim of a genome-wide association study (GWAS) is to isolate DNA markers for variants affecting phenotypes of interest. This is constrained by the fact that the number of markers often far exceeds the number of samples. Compressed sensing (CS) is a body of theory regarding signal recovery when the number of predictor variables (i.e., genotyped markers) exceeds the sample size [1]. Its applicability to GWAS has not been investigated. Using CS theory, we show that all markers with nonzero coefficients can be identified with an error that is small relative to the number of markers sufficiently few in number (sparse) relative to sample size. For heritability equal to one ($h^2 = 1$), there is a sharp phase transition from poor performance to complete selection as the sample size is increased. For heritability below one, complete selection still occurs, but the transition is smoothed. We find for $h^2 \approx 0.5$ that a sample size of approximately thirty times the number of markers with nonzero coefficients is sufficient for full selection. This boundary is only weakly dependent on the number of genotyped markers. Practical measures of signal recovery are robust to linkage disequilibrium between a true causal variant and markers residing in the same genomic region. We have a predicted that the number of genotyped markers are strongly correlated with height-associated markers identified by the GIANT Consortium [2].


1462M

Cross-Phenotype Analysis of GWAS (CPAG): A powerful tool for detecting shared genetic architecture among human traits and underlying shared pathways. L.Y. Wang1,2, D.C. Ko3,4. 1) Duke University, Department of Molecular Genetics & Microbiology, Durham, NC; 2) Department of Medicine, Duke University, Durham, NC; 3) Center for Human Genome Variation, Duke University, Durham, NC.

Recent large-scale meta-analyses of GWAS have demonstrated that cross-phenotypic associations and pleiotropy are abundant among GWAS hits for various diseases and traits. However, the degree of disease similarity as well as the underlying genetic basis driving such connections remains poorly understood. By integrating GWAS results, we developed a novel approach to estimate the degree of disease similarity, identify informative disease clusters, and highlight the potential underpinning pathways. A probability-based method was used to estimate similarity of traits by correcting unobserved shared genetic components and is therefore less sensitive to small and/or even unique shared similarities. Comparing nearly a dozen similarity methods, we determined which was most optimal and robust. Unlike previous methods such as polygenic score, linear mixed-effect model and network model, our approach is based on analyses of summary statistics without requiring more genotype/phenotype data. Using data from the NHGRI GWAS Catalog, we constructed human disease trees with more than 600 traits by using similarity matrices based on either SNPs, SNPs corrected for linkage disequilibrium, and implicated genes among pairwise diseases. For pairwise comparisons that showed significant similarity, we implemented gene set enrichment analysis to identify possible associated pathways. Our method categorized related diseases into disease clusters well consistent with pre-defined groups but also revealed novel and unexpected disease connections. We validated the cross-phenotype analysis of GWAS (CPAG) method computationally using two different approaches by comparing the CPAG generated categorization to pre-defined categories. All of this functionality for disease similarity, visualizing clustering, and pathway analysis has been implemented into a flexible stand-alone software program, named CPAG. CPAG also accepts user-defined datasets to predict potential disease connections and shared pathways. Our approach will become increasingly powerful with the ever-expanding identification of novel causal genetic variants leading to different clinical findings and expanding our understanding of the genetic architecture of complex traits and additional insights into the pathophysiology of diseases. Indeed, we are currently experimentally testing novel human trait connections discovered from our analysis using zebralift.
1464S
A fast and accurate method for detection of IBD shared haplotypes in genomewide SNP data. D. Bjelland1, U. Lingala1, N. Lapinski1, M. Jones2, M. Keller1,2,1 Institute for Behavioral Genetics University of Colorado, Boulder, CO; 2) Department of Psychology University of Colorado, Boulder, CO.

Several programs have been developed to detect shared identical-by-descent haplotypes between individuals using SNP data. For detecting short shared haplotypes (e.g., 2-3 cM) in large datasets (e.g., >10,000 individuals), the performance of these programs suffers either in terms of accuracy or in computational time. With this in mind, we developed a program, FISHR, that utilizes initial shared haplotype output from GERMLINE, quickly post-processes this data, and more accurately discovers the true shared haplotypes. FISHR was compared to GERMLINE and BEAGLE Refined IBD by simulating 15 megabases of genotypic data for 8,000 individuals with realistic levels of SNP and phasing errors. GERMLINE was the fastest program analyzed, with times of 9.0, 34.7, 152.3, and 601.0 seconds for samples of 1,000, 2,000, 4,000, and 8,000 individuals, respectively. Run times on equivalent sample sizes were 13.0, 48.3, 189.7, and 753.0 seconds for FISHR (including the initial GERMLINE runs) and 400.4, 2064.0, 29,840.0, and 921,605.7 seconds (or 10.6 days) for BEAGLE Refined IBD. When discovering shared haplotypes with lengths between 2 and 3 cM, FISHR consistently had ~8% higher positive predictive value for matched sensitivity when compared to GERMLINE and ~10% higher positive predictive value for matched sensitivity when compared to BEAGLE Refined IBD. Similar results occurred when analyzing shared haplotypes between 3 and 4 cM. Thus, FISHR is more accurate than both GERMLINE and Refined IBD, and is fast enough to be usable for detection of identical-by-descent shared haplotypes in very large genome-wide datasets.

1465M

The Expectation-Maximization (EM) algorithm for haplotype inference has been superseded for accuracy by newer programs, e.g. PHASE and SHAPEIT, but remains useful for rapid analysis. We have developed an EM implementation, “HapFerret”, that is characterized by flexibility and ease of use, notable its use of a natural format for input genotypes and output haplotypes. Genotypes may be input as lines of comma separated alleles, where the alleles may be any alphanumeric; i.e. for HLA data, for the loci MogCA—DQA1—rs9273349—DQB1—rs1894407—TAP2—TAP1, a hypothetical genotype input line: 9,12 101,301 C,T 302,503 A,A 3.2,4.2 3.1,3.1. HapFerret contains a bootstrapping procedure, comparing inferred haplotypes between successive bootstrap replicates, a feature that addresses some of spurious precision. We show with data with known haplotypes that bootstrapped inference from HapFerret has an accuracy intermediate between standard EM and PHASE. HapFerret may be downloaded from: https://ccrod.cancer.gov/confluence/display/BCGC/BCGC-Software.

1466T
A novel approach to craniofacial gene discovery: SysCLFT (Systems tool for Cleft lip/palate gene discovery). I. Saadi1, D. Anand2, D. Bjelland3, T.W. Kho1, S.A. Lachke1, 1) Anatomy & Cell Biol, Univ Kansas Med Ctr, Kansas City, KS; 2) Biological Sciences, Center for Bioinformatics and Computational Biology, University of Delaware, Newark, DE; 3) Victor Chang Cardiac Research Institute, Sydney, NSW, Australia; 4) Faculty of Medicine, University of New South Wales, Sydney, Australia.

Although study of mammalian craniofacial (CF) development has resulted in a detailed functional understanding of several genes associated with orofacial clefts (OCFs), many additional loci remain unsolved. We recently created a novel innovative systems approach that is highly effective in identifying genes critical to the development and/or homeostasis of specific tissues. Successful application of this approach identified several novel genes with functional significance in eye tissue and associated defects (Lachke Science 292:1701-1706, 2011). Here, we present an approach to develop SysCLFT (Systems tool for Cleft lip/palate gene discovery), a new interactive web resource for efficient prioritization of candidate genes associated with OCFs and construction of CF developmental GRN. We analyzed 67 microarray gene expression datasets of various CF tissues, available publicly in FaceBase and NCBI-GEOM, using in silico subtraction with embryonic whole body (WB) reference datasets. This approach generates systematic gene ranks based on CF-tissue enrichment - instead of simply high expression - and offers a new method of prioritization of candidate genes associated with OCFs. Indeed, with WB subtraction, the top 500 highly ranked genes from the palate, frontalonasal, mandibular or maxillary process datasets are highly enriched in gene ontology (GO) categories for palate development, cell adhesion, cell proliferation, epithelium development and skeletal system development. In concordance with WB subtraction, i.e. as data currently exist, SysCLFT compared to only 4 of 45 (9%) genes without WB-subtraction. Major genes with established role in CF development, including Irf6, Pax7, Pax9, Pbx2, Pbx3, Mnx2, Satb2, Tbx22, are highly enriched in WB-subtracted datasets from concordant tissue sets. Most significantly, we show that in silico subtraction method successfully identifies a majority (85%, n=45) of known OCF genes within the top 5 minRank genes in mean chromosomal intervals of 13Mb, each containing 106 genes on average. Thus, SysCLFT is a new, highly effective approach to facilitate prediction of genes and prioritization of targets in a large body of GWAS and exome-based CF gene discovery efforts.

1467S
Phenolyzer: prioritizing candidate genes from disease/phenotype descriptors. H. Yang, K. Wang, Z. Wu. Institute for Behavioral Genetics University of Colorado, Boulder, CO; 2) Department of Psychology University of Colorado, Boulder, CO.

Whole-genome/exome sequencing and copy number variation studies can generate many candidate genes for specific diseases or phenotypes. However, to prioritize these candidate genes, traditional approaches typically do not consider the prior knowledge on associations between genes and specific phenotypes. To address this problem, we developed an integrative tool named Phenolyzer, which takes as input a discrete list of phenotype terms and generate a list of candidate genes weighted by their association with the phenotype, even in the absence of any genotype data. Phenolyzer works by following a biologist’s logic - 1) search databases in CTD’s (Comparative Toxicogenomics Database) disease vocabulary and Disease Ontology for a certain disease/phenotype term, interpret the term into multiple specific disease names, find all the associated genes and related information in OMIM (Online Mendelian Inheritance in Man), Orphanet (a journal for rare disease), NCBI’s ClinVar, GeneReviews (an expert-authored, peer-reviewed disease descriptions) and GWAS (Genome-Wide Association Studies) databases, then generate the seed gene set with conditional probability as scores, 2) grow the seed gene set in the HPRD (Human Protein Reference Database) protein interaction, NCBI’s Biosystem, HGNC (HUGO Gene Nomenclature Committee) gene family, and HTRI (Human Transcriptional Regulation Interactions) databases and retrieve an augmented gene set, 3) integrating all the information to score genes. The input can be one or several disease/phenotype terms, while users can optionally supply a gene list or genomic region to further limit the search. Additionally, Phenolyzer provides an intuitive visualization of gene-phenotype network and gene-network on the fly. Furthermore, combined with ANNовар, users can prioritize disease variants from whole-genome/exome sequencing data, significantly expediting discovery of diseases. The great performance of Phenolyzer was also compared with other similar tools like Phenofinder and PolyPhen. In addition to a command line and software tool, we implemented Phenolyzer as a user-friendly web server to facilitate easy access for users without informatics skills. The server is available at http://phenolyzer.ucsd.edu.
1468M

Determining the Number of Contributors using Forensically Relevant STRs: Effects of Template Mass and Complexity on the Ability to Correctly Identify the Number of Contributors. L. Alfonse1, H. Swaminathan2, D.S. Lun2, M. Medard3, C.M. Grgicak1. 1) Biomedical Forensic Sciences, Boston University School of Medicine, Boston, MA; 2) Department of Computer Science, Rutgers University-Camden, Camden, NJ; 3) Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA.

The objective of this project was to evaluate NOCIt - a software that can assess the number of contributors (NOC) to DNA samples. The likelihood that n unrelated individuals contributed to a sample is computed via a Monte Carlo process during which genotypes for the n contributors are randomly chosen based on the allele frequencies, and the ratio of the n contributors to the mixture is randomly selected from a uniform distribution. Modeling of baseline noise, reverse and forward stutter proportions, dropout rates, and allele heights is performed. The likelihood of observing the peak heights at each locus given the genotypes of the contributors, the mixture ratio, and the template DNA mass is computed using single source calibration samples of known genotypes; likelihood values are 'averaged' between iterations and multiplied across loci to give the overall likelihood given n contributors. The output is presented as the probability distribution over 0 to 5 contributors. The performance of NOCIt was tested on 1-, 2-, 3-, 4- and 5-person mixtures of various contributor ratios, amplified with the AmpFlstr® Identifiler® Plus Amplification kit (29 cycles). Samples were amplified at target masses ranging from 0.25 to 0.008 ng and analyzed at 1 RFU. Artifacts such as -A, dissociated dye, bleed through, and spikes were manually removed. Allele frequencies from the Caucasian population provided in the AmpFlstr® Identifiler® Plus manual were utilized. The accuracy of NOCIt, calculated by comparing the software's output to the known NOC, was dependent upon 1) the true NOC in the sample and 2) the total mass of DNA in the sample. For example, the accuracy decreased from 100% to 20% for the 1- to 5-person mixtures, respectively. However, when sufficient levels of DNA were amplified, i.e. 0.25 ng, the accuracy was 100% for 1- to 4-person mixtures and 67% for the mixtures containing DNA from 5 persons. The minimum from any one contributor required to correctly estimate the NOC was ~10 cells. These findings are important for forensic applications, as they provide data regarding the number of samples needed to calibrate NOCIt.

1469T

Detection of autozygous segments in exomes of inbred individuals. M. Vigeland1, K.S. Gjøtterud1, M. Medard2, C.M. Grgicak1. 1) Dept of Medical Genetics, Oslo University Hospital, Oslo, Norway; 2) Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, As, Norway.

We present AutEx, a new algorithm for detecting autozygous segments directly from whole-exome variant files. Our approach is based on the hidden Markov model developed by Leutenegger et al. (2003), taking as input the level of inbreeding, allele frequencies, variant distances and sequencing error probabilities. Using both simulated data and real inbred exomes we compare our method to existing software detecting runs of homozygosity, including PLINK. Overall AutEx performs better, in particular when the level of inbreeding is accurately known. The AutEx algorithm is implemented as part of FULTUS, a lightweight GUI program aimed at non-bioinformaticians, providing a framework for filtering, exploring and statistical analysis of exome data. Combined with these features, AutEx becomes an effective and easy-to-use tool for narrowing down the search for causal variants in exomes of inbred individuals with recessive disorders.

1470S

A systems biology approach for enriching genetic association studies of metabolite profiles with pathway knowledge. K. Willems van Dijk1, H.K. Dharuni1, P. Henneman2, A. Demirkiran1, A. Isaacs3, C.M. van Duijn1, P.A.C. 't Hoen1, J.B. van Klinken1. 1) Human Genetics, Leiden University Medical Center, Leiden, Zuid Holland, Netherlands; 2) Clinical Genetics, Academic Medical Center, Amsterdam, Netherlands; 3) Genetic Epidemiology Unit, Epidemiology, Erasmus Medical Center, Rotterdam, Netherlands.

Plasma metabolite levels have been proposed to reflect the interaction of the genome with the environment and may thus provide insight into the etiology of complex metabolic disorders. Genome-Wide Association Studies (GWAS) on plasma metabolite levels have led to the discovery of many novel loci, but they have been relatively unsuccessful in providing insight into the etiology of complex metabolic disorders. Here, we propose Elementary Flux Mode (EFM) analysis as an approach to assist in the functional interpretation of GWAS results on metabolomics data and to explain the mechanistic link between genotype and metabolotype. EFM analysis consists of enumerating the complete set of steady state fluxes that are possible in a metabolic network, thus providing a comprehensive picture of the network's functional capacity. We assessed the potential of EFM analysis to explain a set of genetically determined metabolotypes (GDMs) concerning amino acid metabolism in humans that were reported in the OMIM and GWAS catalogue database. We performed EFM analysis on a condensed version of the genome-scale stoichiometric model (GSSM) of the hepatocyte developed by Gille et al. (2010). The resulting model yielded 8.5•10^9 EFMs that were involved in the degradation, synthesis or conversion of amino acids. Subsequently, for each reaction and amino acid, we determined the essentiality of that reaction in the metabolism of the amino acid by counting the number of EFMs that both contained the reaction and were involved in the conversion of the amino acid. Comparing these results with the selected GDMs, we found that the known genotype-metabolotype relationships could be predicted with a high degree of accuracy (ROC curve AUC = 0.93). In contrast, predictions that were made based on co-occurrence of the gene and metabolite in KEGG pathway gene sets or on the shortest distance between the gene and the metabolite in the metabolic network had a low degree of specificity and were less accurate (AUC = 0.79 and 0.74 resp.). Thus, we propose the use of EFM analysis for metabolite data in a genome-wide setting using a Systems Biology approach. A workflow has been developed to 1) compute the EFMs in a GSSM, 2) map GWAS results to the EFMs, and 3) visualise enriched EFMs. Our approach integrates downstream analysis with mathematical models of human metabolism and has the potential to provide new mechanistic insights into the causes of metabolic disease.
1472T Enhanced statistical methods to detect cross-population heterogeneity at GWAS risk loci. M. Roytman1, B. Pasanen2,3,2,2, 1) Bioinformatics Interdepartmental Program, UCLA; 2) Dept of Human Genetics, David Geffen School of Medicine, UCLA; 3) Dept of Pathology and Laboratory Medicine, David Geffen School of Medicine, UCLA.

GWAS have identified thousands of disease-associated variants to date, many of which have successfully been used to identify biological pathways for drug targets. However, a large majority of these risk alleles were identified in individuals of European ancestry making their transferability to other ethnic groups uncertain. Here we present methods for quantifying how well the known GWAS risk variants extend to individuals of non-European ancestry. We introduce methods that quantify the difference in effect size of a particular risk variant across two populations and propose a log likelihood ratio test for assessing the presence of heterogeneity at a given risk locus. We investigate the power of our approach at different levels of heterogeneity and show that our methods has high power of detecting realistic differences in effect size, while maintaining well-calibrated statistics, closely following a chi-squared distribution with one degree of freedom under null data simulations (where effect size does not differ between the populations). We extend our approach to admixed populations (e.g. African American) and show how locus-specific can be integrated to detect heterogeneity in the allelic effect across different ancestries.

1473S Regulatory network constructed from the epigenome of normal cells reveals functional connections between disease genes. R.F. Lowdon1, G. Elliott1, B. Zhang1, J.B. Cheng2, S.J. Fisher3, J.F. Costello4, T. Wang5. 1) Genetics, Washington University School of Medicine, Saint Louis, MO, 63108; 2) Department of Dermatology, University of California San Francisco, CA, 94143; 3) Department of Obstetrics, Gynecology & Reproductive Sciences, School of Medicine, University of California San Francisco, San Francisco, CA, USA; 4) Dept. of Neurological Surgery, Helen Diller Family Comprehensive Cancer Center, UC-San Francisco, CA 94143. It is a daunting but critical challenge to understand the functional connection between human disease genes and the function of disease-associated non-coding variants. Uncovering the connections between these genes and variants is crucial for understanding normal and aberrant regulation of human disease genes. We hypothesized that cell type-specific gene regulatory instructions for a transcription factor and its target genes are encoded in the epigenomic landscape around the genes. To test this hypothesis, we developed a regulatory analysis pipeline to construct gene regulatory networks based on a cell’s epigenome. Our regulatory analysis approach has the following components: first, by profiling and contrasting epigenomes of different cell types, we defined cell type-specific regulatory elements. Then, we identified enrichment of binding site motifs of specific transcription factors in these regulatory elements. Next we used custom region-gene association tools to identify candidate genes for each element. We also used co-variation of transcription factor expression and target gene expression to increase the confidence of our prediction. Finally, we queried publicly available TF-target gene interaction data to quantitatively connect our results with known TFs and target genes. This approach allowed us to construct a gene regulatory network of a cell type of interest. We successfully captured many known connections in our reconstructed networks, as well as novel predicted connections. Surprisingly, this approach revealed functional connections across disease genes, predicted potential key genes that are not yet implicated in disease, and explained functional consequences of some non-coding, disease associated variants. In our first test, we constructed a gene regulatory network of the embryonic surface ectoderm using epigenomes of cell types of shared developmental origin. We confirmed co-regulation of six genes that encode hemidesmosomes, a structural complex in the epidermis in disease, and explained functional consequences of some non-coding, disease-associated variants. In our second test, we confirmed co-regulation of six genes that encode hemidesmosomes, a structural complex in the epidermis, and explained functional consequences of some non-coding, disease-associated variants. We confirmed co-regulation of six genes that encode hemidesmosomes, a structural complex in the epidermis, and explained functional consequences of some non-coding, disease-associated variants. This underscores the power of our approach at different levels of heterogeneity and shows that our methods has high power of detecting realistic differences in effect size, while maintaining well-calibrated statistics, closely following a chi-squared distribution with one degree of freedom under null data simulations (where effect size does not differ between the populations). We extend our approach to admixed populations (e.g. African American) and show how locus-specific can be integrated to detect heterogeneity in the allelic effect across different ancestries.

1474M Dynamic changes in RNA modifications (epitranscriptome) localization drives new regulation of cancer cells. Y. Salitore1,2,2, T. Fernando4, A. Melnick1, L. Cerchietti1, C. Mason1,2, 1) Department of Physiology and Biophysics, Weill Medical College, Cornell University, New York, NY 10065, USA; 2) The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY 10065, USA; 3) Tri-Institutional TRiD, The Rockefeller University, The Sloan-Kettering Institute, Weill Cornell Medical College, New York, NY 10065, USA; 4) Department of Hematology and Oncology, Weill Cornell Medical College, New York, NY, USA.

The epitranscriptome is a very nascent, exciting field that has revealed a complex middle layer of cellular regulation that is mediated by modified RNA bases. Methylated RNA ImmunoPrecipitation-Seq (MeRIP-seq) utilizes an antibody specific for the RNA modification methyl-6-adenosine (m^6A) to enrich for RNA fragments that contain the epitope. Combined with next-generation sequencing (NGS), these immunoprecipitated fragments (IP) are sequenced with RNA-Seq controls to identify peaks: genomic regions of interest where putative m^6A sites exist. However, the original MeRIP peak-finding methods (like MeRIPPeR) utilized the non-parametric Fisher’s exact test to find peak regions that were statistically significant in enrichment for the IP relative to the control. While this was successful in identifying thousands of peaks in both human and mouse samples, m^6A is known to be a dynamic modification and m^6A sites and methylation levels change in response to cell stimuli. Here we introduce MeRIPPeR 2.0, which uses the negative binomial distribution to better model the high dispersion in the RNA-Seq data and applies a probabilistic graphical model to elucidate peak methylation from the IP efficiency. This not only enables better identification of epitranscriptomic peaks, but also more accurate characterization of how those peaks change between conditions. We then applied MeRIP-seq and MeRIPPeR 2.0 to Ly1 diffuse large B cell lymphoma samples to further examine the dynamic nature of m^6A in the context of cancer, using normal cells, proliferating cells, and those cells after heat shock, as well as investigating nuclear and cytosolic fractions separately. Heat shock has previously been demonstrated to have an impact on m^6A sites and also simulates B-cell activation. Previous epitranscriptomic studies have demonstrated an enrichment of peaks near the stop codon, which was expected since CpG sites are highly enriched in the 3’ UTR. In concordance with these results, we identified thousands of m^6A peaks throughout the transcriptome and an enrichment near the stop codon. However, we also discovered a small set of m^6A peaks in the 5’ UTR that are highly enriched in the transcription start site and enriched at the stop codon. This indicates that not only the degree of RNA methylation, but also the coordinates of foci within genes can shift, allowing a completely different regulatory avenue for m^6A that has not been seen before.

1475T An integrated method to predict functional impact of non-synonymous SNVs in human genome. M. Wang1, L. Wei1,2, 1) Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, P.R. China; 2) National Institute of Biological Sciences, Beijing, P.R. China.

Correct interpretation of large scale genome variants functional impact and implications in human disease is essential for modern biomedical research and clinical genetics. In silico methods predicting whether a non-synonymous SNV is deleterious or neutral facilitate the discovery of disease causal mutations and classification for SNVs of unknown clinical significance. Most of these tools are built upon machine learning approaches, which classify the two types of SNVs mainly based on conservation information, sequence features and protein structures. However, correct assessment of disease relevance for a SNV requires informative and comprehensive evaluation, which should take various sources of evidence into consideration. We propose a novel method to predict functional impact of non-synonymous SNVs in human genome by integrating multiple sources of evidences. Genetic evidence including co-segregation of a SNP with disease phenotype in a family and population case-control associations provide direct evidence linking SNPs to a phenotype. Such heterogeneous factors were integrated in our method by a unified Bayes model, which gives a combined posterior probability that the SNV is pathogenic given all the evidences. Performance evaluation on benchmark data sets showed the integrated method outperforms those widely employed tools. We also demonstrated its use in disease causal gene discovery from next-generation whole exome sequencing. This method would facilitate rapid annotation and reliable classification for SNPs discovered by NGS in a large scale.
1476S
Identifying the master regulators of complex autoimmune disease susceptibility in Alopecia Areata with reverse-engineered regulatory networks. J.C. Chen1,2, A.M. Christiansen3,4, N. Basavaraj1, H. Liu5, Y.T. Fang6, L.C. McB. H. Tilly4, N. Bapat1, N. Asadpour1,3, M.B. Gerstein5, M. Snyder5, E. Schadt4, W.H. Wong3,4, K.F. Au2, H.Y.K Lam3, 1) Department of Electrical Engineering, Stanford University, Stanford, CA; 2) Department of Bina Engineering, Bina Technologies, Redwood City, CA; 3) Department of Genomics and Bioinformatics, Bina Technologies, Redwood City, CA; 4) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 5) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 6) Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT; 7) Department of Statistics, Stanford University, Stanford, CA; 8) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA; 9) Department of Internal Medicine, University of Iowa, Iowa City, IA.

Emergence of next generation sequencing has brought a new era in transcriptome analysis with RNA-Seq. Although Second Generation Sequencing (SGS) is a high-throughput and high-fidelity technology, full-length isoform analysis is still a challenging problem due to the relatively short read length of SGS, the imperfect performance of existing tools, and the complexity of alternative splicing. Recent studies have suggested to use Third Generation Sequencing (TGS) to improve isoform detection, but it is missing the essence of the analyses enabled by SGS.

Here we present a comprehensive transcriptome profiling pipeline, the RNA-Seq Cocktail pipeline, integrating both SGS and TGS RNA-seq data along with a number of best-of-breed algorithms. TopHat2, SpliceMap, and STAR are used to align RNA-Seq reads from SGS and detect splice junctions. IsoTree approach has given us a novel means to study the mechanistic regulation behind the complex gene expression networks associated with aberrant autoimmune response in the context of AA.

1477M
Global profiling of condition specific transcription factor binding with ATAC-seq. R. Pique-Regi, D. Water, M. Estill, F. Luca. Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI.

Deciphering the regulatory sequences which control gene transcription is a critical step in understanding both cellular and condition-specific regulatory programs encoded in the human genome. Transcriptional response is typically regulated by transcription factors (TFs) which are known to bind specific regulatory sequence motifs. Profiling the environmental binding activity of these factors can be quickly accomplished at a genome-wide scale with the recently developed technique ATAC-seq, which utilizes the Tn5 transposase to fragment and tag accessible DNA. When coupled with an advanced computational method such as CENTIPEDE binding models for TFs with known motifs can be generated across the genome. To date, there are no methods that efficiently incorporate the information provided by paired-end sequencing which allows both the identification of the library fragment length as well as the two cleavage locations that generated the fragment.

First, we have extended CENTIPEDE to utilize fragment length information, and we are working on extending the model to include the joint statistics of cleavage pairs. Our results indicate that paired-end sequencing provides a more informative footprint model for ATAC-seq libraries which leads to greater accuracy in predicting TF binding. These results were validated with ChIP-seq data (ENCOD Project) for multiple factors including CTCF, NRFSF, NRF-1, and NFKB.

We then assayed TF activity in lymphoblastoid cell-lines (LCLs) across multiple treatments (selenium, copper, retinoic acid and iron) for which we previously determined significant differences in gene expression levels. From our initial sequencing results we were able to resolve 383 actively bound motifs (Z-score > 5) across all conditions. We were also able to characterize 5236 regions that have significantly changed accessibility (FDR < 10%) in response to both copper and selenium. No major change in accessibility was detected for retinoic acid but condition-specific response was detected for iron. It uses CENTIPEDE binding models for TFs with known motifs to identify accessible DNA, which is then used to generate a set of TFBSs. The CENTIPEDE algorithm has up to 33% more detected junctions compared to the junctions reported by Cufflinks2 (51%) compared to the aforementioned databases (<10%), suggesting limitations in the use of short read SGS data for detecting long isoforms. Our previous study has already shown that at the same specificity (5% FPR), IDP had a much higher sensitivity (62% TP) than Cufflinks (20% TP). We have also applied our pipeline to other RNA-seq data, including the MCF-7 breast cancer cell line. Our RNA-Seq Cocktail pipeline is by far the most comprehensive RNA-Seq pipeline with very high accuracy.

1478T

Alopecia areata (AA) is an autoimmune disease affecting the hair follicle in which T-cell infiltrates aberrantly localize to and destroy the follicle, resulting in hair loss. While most research has focused on the immunologic etiology of the disease, little has been done to investigate the contributing genetics of the affected end organ e.g. the skin and hair in AA, yet this is an indispensable driver to the etiology of this autoimmune disease. Patients with autoimmune disease do not uniformly present with symptoms across multiple tissues, indicating that genetic components in the end organ specific to a disease are required for autoimmune response. We have taken a systems biology approach to the study of acquired autoimmunity in AA, and modeled it as a genetic, regulatory disease phenotype using reverse-engineered cellular networks. The goal of this study was to take a global, unbiased approach to understanding the regulation of key molecular programs mediating the AA autoimmune response by microarray profiling of scalp biopsies from 60 individuals across three categories of AA clinical presentation and unaffected controls. Through integrating our network-based analysis, with Gene Set Enrichment analysis and unsupervised clustering, we have identified that varying presentations of AA are mediated by the unique expression of six key regulators, with increased severity in presentation coinciding with the involvement of a higher fraction of these regulators. The manipulation of these master regulators is sufficient to induce cytotoxicity in cultured human dermal papillae cells and keratinocytes. This approach has given us a novel means to study the mechanistic regulation behind the complex gene expression networks associated with aberrant autoimmune response in the context of AA.

1478S
Typing of PRDM9 in childhood cancers from Next Generation Sequencing. R. Afshar1, J. Hong2, E. Awada1. 1) Bioinformatics Dept, University of Montreal, Montreal, Quebec, Canada; 2) Sainte-Justine Research Center, Montreal, Quebec, Canada.

Childhood cancers differ remarkably from those seen in the adult population, and their unique biology suggests that the genetic factors underlying the malignant transformation in the developing tissues differ from those generally identified in adults. The identification of these genetic factors may promote the development of biomarkers contributing to early cancer detection.

Our research group suggests that patients affected with childhood acute lymphoblastic leukemia show a higher frequency of the rare C allele of PRDM9, which is found in only ~1% of the European population. PRDM9 protein are composed of a repetitive zinc finger array at their C terminal tail, which binds to specific motifs of the DNA during recombination. This type of repetitive structure confers instability to the genetic sequence and has been shown to evolve rapidly, explaining partially the extreme polymorphism of the PRDM9 gene in the human population. Moreover, this structure poses important challenges in typing PRDM9 alleles from Next Generation Sequencing (NGS) data, and current computational typing methods are unsuccessful at identifying novel alleles. Here, we present a computational method to accurately type PRDM9 alleles from NGS data, and therefore make possible the identification of alleles in childhood cancer NGS patients. This study provides a method for the challenge of typing PRDM9 alleles from NGS data. Furthermore, the association of rare PRDM9 alleles with particular childhood cancer types may shed light on the unique molecular mechanisms underlying these diseases.
1480M
Integrated Variant Comparison Using Three Different DNA-seq Analy-

Technology advancement and increased adoption of next generation sequencing has introduced a large amount of raw data that requires extensive analysis to obtain relevant biological information. However, the costs and duration of the subsequent bioinformatic steps including read mapping, variant calling and contextualization can rarely be pre-determined and often cannot be accurately assessed until a substantial portion of the analysis is complete. Here we demonstrate an approach with pre-defined costs and extensive scalability that uses the most highly considered, peer-reviewed, open source techniques and generates complete documentation of the process. The new DNA sequencing analysis pipeline on the Maverix Analytic Platform incorporates three best-in-class algorithms for variant analysis: GATK, FreeBayes, and SAMTools. We illustrate the few steps required to upload FASTQ files and to configure and launch an analysis. All data are extensively protected with security standards that are HIPAA compliant. The key results are summarized in graphs, while individual variants are annotated and presented in an interactive tabular format that is linked to a private version of the UCSC Genome Browser which can display statistics or supporting reads for a selected variant call. The common and distinct variants called by different analysis methods are visually displayed, and the resulting variants can be extensively filtered on many characteristics. Results can also be downloaded in the tabular format or exported in VCF format for use in other downstream applications. The intermediate results, processing logs, and command lines as executed are available for those who wish to review each step of the pipeline in detail. Use of this approach accelerates the availability of the findings contained in the sequencing results without a loss of quality or control.

1481T
Predicting splicing mutations by information theory-based analysis in rare and common diseases: performance and best practices. N.G. Caminsky1, E.J. Mučaki1, P.K. Rogan1,2. 1) Biochemistry, Western Univer-

Disease-causing mutations that affect pre-mRNA splicing are common. We review the literature citing information theory-based mRNA splicing mutation analyses (n=376). This type of analysis has been applied by many groups to study both common Mendelian conditions, and rare syndromes. While the Shannon entropy framework remains the foundation of these analyses, the software for performing these studies has rapidly progressed. Currently, the Automated Splice Site and Exon Definition Analysis server (splice.cytog-
nomix.com) is used to comprehensively analyze single splicing and regul-
atory mutations. The Shannon pipeline (shannonpipeline.cytognomix.com) analyzes mutations in exomes, genomes, and targeted gene sets. These resources compute individual or total exon information contents (RT or R(Total) in bits) of normal and mutant sequences, which correspond to thermody-
namic differences that account for splicing effects. Primary applications include interpretation of novel variants, detection of benign versus deleterious muta-
tions, deleterious alleles (including leaky, inactivating, and cryptic sites), and quantitative evaluation of changes in RT or R(Total) that account for phenotypic variability and disease severity. Our results, based on a sample of these studies (n=106) report 157 natural splice sites, 19 active cryptic sites, and 9 are regulatory splicing mutations. Of those affecting natural splice sites, 20 cause leaky and 50 inactivate splicing. Canonical dinucleotide mutations in splice sites comprised 28% but adjacent positions occurred at 41% and 31% within introns and exons, respectively. 82.7% of the studies (n=29) that validated predicted mutations with in vitro or in vivo expression assays confirmed the information theory-based analyses. Genome-scale information analyses of complete cancer genomes (3 cell lines) detected 17 to 31 splicing mutations, of which 6 to 17 abolished natural splicing, 1-5 were leaky, and 9-13 activated cryptic sites. In 447 tumors of breast cancer patients, on average, 11.8 splicing mutations were present per exome, where 3.1 were inactivating, 6.2 were leaky and 2.6 activated cryptic splicing. Splicing mutations in these tumors have been validated with RNA-Seq data by automating this process (veridical.org). In summary, we propose guidelines for detection of splicing mutations, distinction of polymorphisms, interpre-
tation of constitutive and regulatory binding site mutations, and exon definition of splice isoforms.

1482S
PedigreeAnnotator: a GATK walker to annotate variants based on pedi-

Family-based exome/genome analyses have been effectively used to iden-
tify genetic defects in rare Mendelian disorders. Several available tools, with their own advantages/disadvantages, have the ability to either integrate pedigree information or add genomic annotations but no single tool combines all the steps required to make biologically/clinically relevant variant prioritiza-
tion. Previously we have described a set of R-scripts and a workflow to perform pedigree-aware analyses of next-generation sequencing data. To improve the accuracy and performance of our existing method we reimple-
mented our R code and workflow with the open-source version of GATK. PedigreeAnnotator accepts, as input, a PLINK format pedigree file and a VCF file containing multiple samples from a single or several families. The pedigree information is used to predict the mode of inheritance, recalibrate genotypes, compute Mendelian error probabilities & inbreeding coefficients and classify variants accordingly (e.g. denovo). Compound heterozygosity and loss of heterozygosity (LOH) annotations are also computed. All avail-
able families can be analyzed simultaneously in a computationally efficient way. Additionally, detailed QC is performed including confirmation of family relationships, gender check and cross-sample contamination. Family data is then integrated with publicly available annotation resources (e.g. OMIM, ENCODE). Furthermore, to aid variant prioritization, a scoring scheme is implemented, taking into account different attributes of the annotated variant. A summary report is generated to include QC, variant statistics and a short list of variants based on the prioritization score. PedigreeAnnotator is designed to perform extended variant analysis accurately and efficiently while being user-friendly.

1483M
GenAP workbench: aiding variant classification in clinical diagnostic settings. M.C. Eike1, E. Skørve2, T. Håndstad2, H. Fontenelle3, K.S. Gjønnes1, J. Borsting2, M. Aanesland2, A.L. Culén1, T. Grünfeld1, D.E. Undlien1. 1) Department of Medical Genetics, Oslo University Hospital, Oslo, Norway; 2) Department of Informatics, University of Oslo, Oslo, Norway.

Classifying genetic variants in terms of clinical significance is a complex process where multiple sources of information of variable quality are reviewed, weighted and combined. With the introduction of high-throughput sequencing technologies into the clinic, this process can quickly become a daunting task without the help of supporting tools. We have therefore develop-
oped a workbench for rules-based, semi-automated analysis of sequence variation data for diagnostic use, as part of the project Norwegian clinical genetic Analysis Platform (genAP).

The underlying system allows rich annotation of sequence variants from external and internal resources, and the workbench displays this information in a series of structured tabs, highlighting key information and guiding the analyst through an analysis workflow based on standard operating proce-
dures at our department. This includes frequency data, mutation databases and prediction tools for missense variants, as well as tools for splice sites. In addition, we have developed a literature evaluation module, where relevant references are classified by the analyst according to quality and conclusions. For the final report, the system uses the annotation information and analyst input to generate suggested classifications for each variant. Approved results may be exported and added to the in-house database.

We are currently testing a prototype of the workbench, with the aim to introduce it for full scale use by the end of 2014. This presentation will demonstrate the logic of the system, with specific examples and some of our experiences from user testing.
Benchmarking of Strand NGS variant caller using a whole genome sample NA12878 and data from Genome in a bottle consortium. R. Harirhan, R. Gupta, P. Gupta, A. Narayanan, S. Aditya, S. Katragadda, V. Veeramachaneni. Strand Life Sciences, Bangalore, Karnataka, India.

**Background and Objectives:** To realize the potential and promise of Next Generation Sequencing (NGS) technology towards research and clinical applications, computational approaches are essential to call variants and translate them into actionable knowledge. Many algorithms are developed for variant calling, however often times they rely on user defined predictions. In this work, we’ll present Strand NGS (formerly Avadis NGS) variant calling approach and benchmarking results on a 1000 genomes CEU female sample, NA12878. Strand NGS is capable of calling different types of variants including SNPs, InDels, structural variations (SVs) and copy number variations (CNVs). We compare our variant calls (SNPs and InDels) with those from GATK UnifiedGenotyper and highly confident variant call set from NIST - Genome in a Bottle Consortium. **Results:** Strand NGS and GATK identify a total of 6,393,054 and 6,105,466 variants respectively with very similar Het/Hom and Ti/Tv ratios. We observed a high overlap (93%) in these variant calls. When variant calls from Strand NGS are filtered using quality metrics like % supporting reads, variant score, read coverage, strand bias and other PV4 biases, the overlap between the variant calls increases to ~98%, making Strand NGS and GATK very similar. We also compared the Strand NGS variant call set to highly confident variants from Genome in a bottle consortium and found a significant overlap between them. We observed that application of the filters in our previous paper decreased the overlap due to selection bias in the overall quality metrics helps to selectively reduce false positive variant calls, thereby increasing both sensitivity and specificity. **Conclusions:** Due to several issues with the raw sequencing data, variant calling is still a challenging problem. Although numerous open-source and proprietary algorithms are available, there are well known pitfalls associated with all the tools. The data sets remain a challenge. Assessment of different quality metrics like supporting reads %, variant score, read coverage, strand bias and other PV4 biases, provides a useful way to filter the likely false positive variant calls. The authors proposed in this paper that the variant call quality metrics are a powerful and flexible approach to call variants and provides a visually appealing way to assess their quality using a variety of quality metrics.
In silico prediction of splice-altering single nucleotide variants in human genome. X. Jian, X. Liu. 1) Division of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX; 2) Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX.

Purpose In silico tools have been developed to predict mutations that may have an impact on pre-mRNA splicing. The major problem that prohibits the use of these tools in research and clinical practice is the difficulty in interpreting the output, partially due to lack of large-scale evaluation analysis of these tools. We aimed to compare some of these tools on a much larger scale and to provide a more accurate and directly interpretable prediction score for splicing mutations. Methods Two groups of single nucleotide variants (SNVs) within splicing consensus regions (−3 − 8 at the donor site and −12 − −2 at the acceptor site) that have known impact on splicing from public databases were used to evaluate eight in silico tools by receiver operating characteristic (ROC) analysis with ten-fold cross-validation. Two ensemble learning methods, adaptive boosting and random forests, were used to construct new models that take advantage of individual tools and to compute an ensemble prediction score for a given mutation. An additional test set was used to validate the new models. Results We collected 2,959 SNVs, among which 1,164 were positive and 1,795 were negative. Four tools were excluded from ROC analysis due to high missing rate. Among the remaining four tools, the Position Weight Matrix model and MaxEntScan outperformed other tools. We combined scores of these four tools as well as four conservation scores using two ensemble methods to construct two new models. Both models significantly improved the predictive performance (area under the curve > 0.97, accuracy > 93%) and were further validated by an additional test set with 45 labeled splicing mutations. We pre-calculated our ensemble scores for all possible SNVs within splicing consensus regions across the human genome and applied them to the mutations from the Catalogue of Somatic Mutations in Cancer database. Analysis showed that predicted splice-altering variants are enriched in both recurrent mutations and known cancer genes. Conclusions Our results demonstrated that some in silico methods are powerful tools in predicting the impact of splicing mutations. Furthermore, the development of computational tool that will enable visualization and interpretation of results. In this case study, we will enable hands-on exploration of RNA-seq data tracks. The results include examples of potential biomarkers identified by using this integrated analysis approach that illustrates the untapped opportunities for discovery in large-scale RNA-seq analysis.
1492M PARADIGM-SHIFT predicts the functional impact of 'driver modules' in multiple cancers using pathway impact analysis. S. Ng1, C. Benz2, D. Hauessler1, J. Stuart1. 1) Biomolecular Engineering, UC Santa Cruz, Santa Cruz, CA; 2) Buck Institute for Research on Aging, Novato, CA.

The major mechanism by which cancer arises is through genomic alterations. These alterations can lead to changes in gene regulation, protein structure, and function. Individual tumors can contain hundreds to thousands of alterations. It is critical to distinguish alterations that have an important role defining the cancer - drivers - from alterations that are unimportant to the tumor - passengers. Driver genes can lead to significant changes in their pathways; however, alterations in a single gene may not be sufficient to explain all the pathway perturbations across patients. Additional alterations could range from DNA copy number changes, gene-fusions, or even lesser understood non-coding mutations. Identifying these 'driver modules' is essential for understanding cancer disease mechanisms, which can help guide treatment decisions as well as identify novel targets for treatment.

We have developed a functional impact prediction method called PARADIGM-SHIFT based on integrated pathway analysis to discriminate loss-of-function, neutral, and gain-of-function alterations. Utilizing the set of regulatory interactions annotated for a given gene, we can detect a shift in the downstream effects of an altered gene compared to what is expected from its upstream influences. Additionally, since these shifts in pathway signal can be detected for all samples, PARADIGM-SHIFT can be used to identify additional genomic alterations that lead to similar changes to the altered pathway to form 'driver modules.' Application of our method to the TCGA Pan-Cancer cohort identifies many genes with significant alterations that lead to loss- and gain-of-function. PARADIGM-SHIFT then identifies several additional genomic alterations, including link-fusions and non-coding mutations, which are significantly implicated in these pathway changes. This analysis offers insight into the mechanism of gene-fusions and non-coding mutations that cannot be assessed by most conventional methods.


Biomatters' Geneious R7 is a bioinformatics software platform that allows researchers the command of industry-leading algorithms and tools for their genomic and protein sequence analyses. Using a glass-box approach for software design, Geneious R7 offers a comprehensive suite of peer-reviewed tools that enable researchers to be more efficient with their bioinformatic workflows. Researchers at all levels can easily manage, analyze, and share their sequence data via a single intuitive software application. R7 provides tools for next-generation sequence analysis, sequence alignment, molecular cloning, chromatogram assembly, and phylogenetics. New features for this major version release include tools for Gibson & TALEN assembly, TOPO cloning along with algorithms for RAXML, FastTree, Garli, LastZ, Bowtie2, as well as a number of new plug-ins. R7 affords real-time dynamic bly, TOPO cloning along with algorithms for RAxML, FastTree, Garli, LastZ, molecular cloning, chromatogram assembly, and phylogenetics. New features for this major version release include tools for Gibson & TALEN assembly, TOPO cloning along with algorithms for RAXML, FastTree, Garli, LastZ, Bowtie2, as well as a number of new plug-ins. R7 affords real-time dynamic interaction with sequence data and empowers biologists to produce stunning publication quality images to increase the impact of their research. By utilizing Geneious R7, biologists can easily improve their sequence analysis workflow efficiencies to free up more time for their research. This poster aims to demonstrate the new features and benefits of the highly integrated Geneious R7 tool-suite.


The UCSC Genome Browser is one of the most popular resources in genomics, allowing users to view a large collection of annotations on human, mouse and many other vertebrate assemblies released by NCBI. The Genome Browser also allows users to upload their own annotations as custom tracks, to create custom Track Hubs which can contain thousands of tracks of custom annotations.

Two new features of the UCSC Genome Browser facilitate its use in situations where privacy is a priority, and where the sequence being visualized is not part of UCSC’s standard set of assembly browsers.

The first technology is called Genome Browser in a Box (GBiB) which uses virtual machine technology to allow users to easily mirror the genome browser on a protected network, or in situations where private annotation cannot be uploaded to UCSC for privacy reasons. To install GBiB one need only download VirtualBox, which is freely available for all major operating systems, and the GBiB virtual image which is a pre-installed copy of the UCSC Genome Browser. The Genome Browser is open source and free for non-commercial users.

The second new feature of the UCSC Genome Browser is Assembly Hubs, an extension of the Track Hub mechanism introduced in 2012. Track Hubs allow users to create their own annotation on assemblies supported at UCSC. Assembly Track Hubs allow users to provide their own DNA sequences (which the Track Hub annotation can be viewed and non-coding assemblies) that require only a network accessible location where the sequence and annotations are stored in a flat-file directory hierarchy. Currently Track Hubs support the following annotation formats: BigBed, BigWig, BAM, HAL, and VCF/tabs.

These two technologies combined make it possible for an end-user to create a UCSC browser instance with custom sequence and annotations without passing sensitive data over the Internet. Using GBiB and Track Hubs, users can view their own annotations aligned with other annotations curated at UCSC.

1495M Comparative transcriptome analysis reveals a pro-angiogenic compensatory mechanism for increased placental vascularization in women with reduced vasodilation. L.M. RODRIGUEZ1, L. RISHISHWAR1, W. WANG2, A.C. AGUILAR1, K. JORDAN3, A. CASTILLO1. 1) UNIVERSIDAD DEL VALLE, CALI, VALLE DEL CAUCA, COLOMBIA; 2) GEORGIA INSTI-
TUTE OF TECHNOLOGY, ATL, GA, USA.

Purpose: The activation of angiogenic pathways is essential for placental development in pregnant women. Flow mediated dilatation (FMD) is a metric commonly used to measure endothelial response to the increased blood flow characteristic of placental vascularization during pregnancy. For this study, we used a genome-wide expression (i.e. transcriptome) analysis approach to identify genes that may help to mediate the link between endothelial response and the construction of the placental vascular network during pregnancy. Methodology: RNA-seq was used to characterize genome-wide expression profiles in placental tissue for cohorts of pregnant women that had either high versus low-FMD pregnant women in an effort to understand how endothelial response to vasodilation relates to the process of placental angiogenesis and vascularization during pregnancy. Results: We applied a novel transcriptome analysis approach (GFOLD) that allows for characterization of differentially expressed genes using low sample sizes; in this case, we had sample sizes of n=2 each for the high and low-FMD cohorts. Using this approach, placenta samples from women with low-FMD showed significant down-regulated expression for the igfbp1 (GFOLD score = -1.164) and lep (GFOLD score = -0.079) genes, which encode for the Insulin-like growth factor (IGF)-binding proteins (IGFBP1) and Leptin (LEP) respectively. These genes are both commonly used in placenta and particularly important to vasodilation in low-FMD women. This would account for the fact that the low-FMD women in our cohort do not show any apparent pregnancy complications related to placental vascularization. However, the down-regulation of these genes, and the effect on their pathways, may lead to subsequent problems with predisposition to diabetes and macrovascular disease in low-FMD women.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
bioinformatics and genomic technology

1496T Pilot from the virtual genomics clinic (VGC). J. V. Thakur1, 2, 3, 1 Massachusetts General Hospital, Boston, MA; 2 Theopleis, Cambridge, MA; 3 Veritas Genetics, Cambridge, MA. While the number of published reports demonstrating medical utility of exome sequencing in patients suspected of monogenic disease continues to grow, data are still limited on the medical utility of exome sequencing in healthy, asymptomatic individuals. Over the past year, ACMG recommendations to return results on a panel of 56 genes in both adult and pediatric cases whenever clinical exome or whole genome sequencing is performed, highlights the importance of gathering this data. Initial reports on the number needed to screen (NNS) in the general population before a suspected pathogenic mutation is identified varies widely from < 1% to > 20%, but is usually reported in the 1.5%-5% range (excluding potential reproductive utility from knowledge of carrier status for recessive disorders) - and is highly dependent on variant filtering stringencies. Preliminary pilot phase data from over 80 exomes with ~60% average on target coverage from the Virtual Genomics Clinic (VGC) study parsed against the 56 ACMG IF (incidental findings) genes and undergoing secondary individualized analyses based on voluntarily provided personal and family medical histories, is reported. The pilot cohort was primarily recruited from the Young Presidents’ Organization (YPO). Study consent, pre- and post-test genomic counseling, as well as return of results, are generally performed remotely via tele- or videoconferencing. All study participants are given the option to receive an easily portable, compressed copy of their exome data on thumbdrive drive and may opt in for indefinite biobanking of their samples for ongoing research and clinically focused reanalysis of data on a yearly basis on average and guided by evolving medical genetics knowledge. Alternatively, participants may opt to have their samples discarded and backup data deleted after receiving their compressed data and initial medical interpretation of results. Data privacy is enhanced using several “off the grid” systems such as hybrid systems utilizing onsite bioinformatics infrastructure for storage and processing, and systems utilizing cloud-based systems for data analysis and storage. This system provides a quick and easy platform to facilitate gathering of data. Future efforts, including the use of physical decryption keys such as Giambattista della Porta modified Alberti cipher disks for off-site collaborators (eg, for access to corresponding phenotype data), will add additional layers of data security in the VGC.

1497S Mega2: data reformatting for facilitating genetic linkage and association analyses. D.E. Weeks1, 2, R.V. Baron1, C. Kollar1, N. Mukhopadhyay1, 1 Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 2 Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 3 Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA. A comprehensive genetic study of a complex disease invariably requires reformatting the data into the precise input format required by each of many analysis programs used. Our Mega2 software facilitates this process by automating common data reformatting tasks, thereby markedly reducing the chance of data reformatting errors, making data setup much less error-prone, and saving the user the time of writing, debugging, and maintaining data conversion scripts. We have recently extended the capabilities of Mega2 in a number of ways. In addition to accepting LINKAGE-format input files, it now accepts VCF, BCF, and compressed VCF files and can apply common ‘VCFtools’ filters. While continuing to support conversion to commonly-used formats such as Merlin, Mendel, SimWalk2, and SOLAR, PLINK, Cranefoot, IQLS, FBAT, MORGAN, BEAGLE, Eigenstrat, and Structure, Mega2 has now been extended to support data conversion to PLINK/SEQ format. Mega2 can now pass through non-numeric alleles to analysis programs that will accept them, instead of recoding them to numeric alleles as was done previously. For some output options, Mega2 can generate high-quality plots of the results using our nplplot R package, as well as generate custom track files for use within the UCSC genome browser. When controlled by a batch file, Mega2 can be used in an automated manner within data analysis pipelines. Mega2 also supports organisms other than humans. Mega2 is open source and freely available and has extensive documentation and a tutorial available online at http://schema.bioinformatics.pitt.edu. This work was supported by NIH grant R01 GM076667 (PI: Weeks). Earlier contributions to our initial code base were made by Lee Almasy, Mark Schroeder, and William P. Mulvihill.

1498T Single Cell RNA-Seq analysis of Tumor Composition. I. Ragoussis1, 2, Y.-C. Wang1, 2, E. Iaccuci1, 2, L. Letourneau1, 2, F. Savage1, A. Monast2, N. Bertos3, A. Omeropolou3, M. Park4, 5, 1 Department of Human Genetics & McGill University and Genome Quebec Innovation Centre, McGill University, Montreal, Quebec, Canada; 2 The Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, Quebec, Canada; 3 Department of Pathology, McGill University, Montreal, Quebec, Canada; 4 McGill University and Genome Quebec Innovation Centre, McGill University, Montreal, Quebec, Canada. Single-cell RNA-Seq experimentation and representation is an advanced and powerful tool for cutting-edge human genetics research as we can now monitor the whole transcriptome at the single-cell level. The ability to interrogate control and disease states of single cells allows for specific and unadulterated characterization of these conditions. Here we provide the issues surrounding the analysis of individual and bulk breast cancer cells derived from a tumor biopsy. In brief, live single-cell mRNA to cDNA libraries were prepared using the Fluidigm C1 Single-Cell microfluidic system. cDNA libraries were then converted to sequencing ready libraries using the Illumina Nextera XT kit and sequenced on the HiSeq2500 using Rapid Paired End 150bp mode. In parallel, control cDNA libraries were prepared from extracted total RNA and 200 cells serving as benchmarks to systematically evaluate the sensitivity and accuracy of our single-cell RNA-seq approach. We sequenced > 100 cells from a Her2+ breast tumor and detected between 4501 and 7168 genes in the single cell samples and between 8782 and 13512 genes in the bulk samples. The analysis was carried out via a process which included filtering, clustering, as well as various comparative and pathway analysis. Clustering (based on read counts and read sequence similarity) revealed a few primary clusters and subgroup profiles (clusters). Statistical analysis between these groups of cells revealed several significant genes, of which five were known breast cancer stem cell markers (CD44, TGFB2, MUC1, KRT5, and IGTA6). Pathway analysis revealed enrichment of some kinase related genes, as well as groups (clusters) believed to be responsible for the cellular states of the groups as mediated by known cancer associated genes (BCAS2, BRIX1, BRE, and TRIPT1). These approaches demonstrate our ability to classify single cell samples, identify potential biomarkers, and elucidate which pathways were involved in the aberrant cell growth. Complementary to this, single-cell samples showed cell state specific clustering leading to the conclusion that this approach allows for the characterization of cellular heterogeneity, both important in understanding the biology of tumors. With this setting, we have been able to move towards distinguishing between different cell populations in the tumor sample, characterizing the tumor purity, and identifying distinct expression patterns.

1499T Standardized phenotyping enables rapid and accurate prioritization of disease-associated and previously unreported sequence variants. W.P. Bone1, D.R. Adams2, M.J. Davis3, D. Draper1, E.F. Flynn1, R.A. Goodfellow1,1, H.M. Haendel1,5, M. Haendel5, M. Nehrebecky1, 2, P.N. Robinson3, M. Sicari4, D. Smedley1, C.J. Triff1, C. Toro1, E. Valkana1, C. Wahl1, N.L. Washington4, L. Wolfe1, C.F. Boekeloe1, W.A. Gahl1, 1) Undiagnosed Diseases Program, National Human Genome Research Institute, Bethesda, MD; 2) National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA; 3) Mouse Informatics group, Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, United Kingdom; 4) Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA; 5) Institute for Medical and Human Genetics, Charité-Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany; 6) University Library and Department of Medical Informatics and Epidemiology, Oregon Health and Sciences University, Portland, Oregon 97239, USA. The NIH Undiagnosed Diseases Program (UDP) uses exome analysis for generating hypotheses of potential genetic causes of disease. This approach, which incorporates Mendelian segregation and population frequency filters and prioritizes on rarity and predicted deleteriousness, has enabled several diagnoses. However, this method is laborious, even when the various UDP websites are consulted. Here we report on a cohort of 23 undiagnosed UDP patients, Exomizer 2.0 ranked all eleven diagnostically relevant variants within the dataset within 40 minutes. Further, we have developed a web site http://urinfo3.org which provides functional annotations of disease genes and provides an interactive and effective method of screening for variants in known disease causing genes and possibly for identification of new disorders.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1500S


Genome-wide association studies (GWAS) have revealed numerous risk loci associated with type 2 diabetes (T2D) and obesity. However, GWAS signals are rarely traced to the disease-causing variants, given the still incomplete annotation of the human genome, especially for non-coding variants that account for the majority of disease associations. This is particularly complex given the uncertainty on the specific causal variant (given numerous variants in linkage disequilibrium, LD), the uncertainty on the cell type of action, the target genes and upstream regulators of non-coding variants. Here, we overcome these challenges using a combined computational and experimental approach. We use regulatory annotations across 127 tissue/cell types to prioritize functional non-coding variants in regulatory regions, and a comparative genomics across multiple species to recognize likely driver regulatory motifs. We next use diverse experimental assays to validate the cell-type specific regulatory activity for 16 predicted variants and their binding regulators. In particular, for the FTO obesity/T2D risk locus, we report multiple lines of evidence supporting that the intronic variant rs1421085 is causal for the FTO association, acts in adipocytes, affects ARID5B binding, and alters IRX5 gene regulation. Briefly: 1) rs1421085 harbors a functional regulatory ARID5B motif based on conservation of a ARIDSB-CART-GATA motif cluster across species; 2) rs1421085 is located in a candidate adipocyte enhancer based on chromatin state analysis across 127 reference epigenomes; 3) rs1421085 shows allele-specific binding of ARID5B using electrophoretic mobility shift assays; 4) rs1421085 shows cell-type specific enhancer effects for SGBS adipocytes using luciferase assays; 5) the rs1421085 risk allele increases IRX5 expression in human adipocytes, based on qRT-PCR; 6) Lastly, increase in IRX5 mRNA levels depend on the risk allele and regulation by ARID5B, based on ARIDSB orRNA knockdowns. This suggests that a non-coding variant acting in adipocytes underlies the FTO association via its IRX5 knockdown effects, which we previously thought were driven by the brain. More broadly, our results suggest a general method for the computational discovery and experimental dissection of disease variants and have important implications on the study complex traits, which can help bridge the genotype-to-phenotype gap between genetic variants, molecular mechanisms, and cellular and organismal phenotypes.

1501M

Fast and Accurate Site Frequency Spectrum Estimation from Low Coverage Sequence Data. E. Han, J. Shmulevich, J. November. Biostatistics, UCLA, Los Angeles, CA; 2) Human Genetics and Biomatics, UCLA, Los Angeles, CA; 3) Ecology and Evolutionary Biology, UCLA, Los Angeles, CA; 4) Human Genetics, University of Chicago, Chicago, IL.

The distribution of allele frequencies across polymorphic sites, also known as the site frequency spectrum (SFS), is of primary interest in population genetics. It is a complete summary of sequence variation at unlinked sites and more generally, its shape reflects underlying population genetic processes. One practical challenge is that inferring the SFS from low coverage sequencing data in a straightforward manner by using genotype calls can lead to significant bias. To reduce bias, previous studies have used a statistical method that directly estimates the SFS from sequencing data by first computing the site likelihood vector (i.e. the likelihood a site has a each possible allele frequency conditional on observed sequence reads) using a dynamic programming (DP) algorithm. Although this method produces an accurate SFS, computing the site likelihood vector is quadratic in the number of samples sequenced. To overcome this computational challenge, we propose an algorithm we call the "adaptive K-restricted" algorithm, which is linear in the number of genomes to compute the site likelihood vector. This algorithm works because in a lower triangular matrix that arises in the DP algorithm, all non-negligible values of the site likelihood vector are concentrated on a few cells around the best-guess allele counts. We show that our adaptive K-restricted algorithm has comparable accuracy but is faster than the original DP algorithm. This speed improvement makes SFS estimation practical when using low coverage NGS data from a large number of individuals.

1502T

Parallelization of genome-wide local ancestry inference. R. Johnson, G. Nelson, C. Winkler. 1) BSP CCR Genetics Core, Frederick National Laboratory, Frederick, MD; 2) Basic Research Laboratory, Frederick National Laboratory, Frederick, MD.

Admixture linkage analysis has proven to be a powerful method to map disease genes and other phenotypic traits. The ALDsuite package in R uses a hidden Markov model (HMM) to infer local ancestry in both sparse and dense marker data, includes statistical tools to map disease genes in admixed populations and generates graphical output for visualizing results. One common drawback of HMM frameworks is their computational inefficiency, relative to other algorithms. While ALDsuite has been designed to maintain computational efficiency with marker sets of increasing density, the use of multiple Markov chains in parallel processes can offer significant time savings. During burn-in iterations, each independent chain’s parameter state is combined and a global parameter state is maintained. At the completion of each iteration, each local parameter state is updated to be more like the global parameter state, using a weighted sum of the two. Early in the burn-in phase the global parameter state is more heavily weighted to push all chains to a more likely region of the parameter space. Later in the burn-in phase the local parameter state is more heavily weighted to allow more independent starting points for follow-on iterations. Each chain is then allowed to independently sample the parameter space until the desired number of samples has been reached. ALDsuite was developed to provide geneticists with easy to use software to augment admixture mapping studies with powerful insights gained from an analysis of admixture linkage disequilibrium in two-way and poly admixed populations. This parallelization scheme allows for a near-linear increase in speed per additional processor.

1503S

CliniCall - Bridging the Gap From High-Throughput DNA Sequencing to Actionable Variants. S. McGee, G. Jimenez, T. Kolar, M.O. Dorschner, J.D. Smith, D.A. Nickerson. Genome Sciences, University of Washington, Seattle, WA.

CliniCall is a stand-alone application that bridges the information gap between patient DNA sequencing data and clinical relevance. With the explosion of clinical DNA sequencing tests available there is a critical need for analysis tools to help filter and prioritize clinically actionable variants. This has created a significant challenge for clinical laboratories leveraging these new technologies, ranging from weighing the importance of rare variation to cross-referencing multiple annotation databases. To facilitate analysis, interpretation and reporting, we have developed a user-friendly tool to view sequence alignments, quality metrics and variant annotation. A combination of BAM and VCF files - or simply a GVCF file - can be easily imported for variant identification within user-defined gene sets. With CliniCall's new multi-sample display, variants can be easily compared across samples and filtered on a user-defined set of metrics such as allele frequency, conservation score, presence/absence in specific databases and annotation functionality or clinical-significance. The CliniCall platform clearly summarizes data quality for the sequenced samples, and simultaneously links to existing databases. In addition, CliniCall now not only graphically summarizes the data, but allows for filtering on a constantly-updated set of ClniVar variants. In addition it provides an up-to-date set of Targeted Gene Panels - such as the ACMG 56 Incidental Genes - that can be automatically pre-loaded into the analysis scheme for full reporting. This all is designed to make quantitative decisions regarding the validity and categorization of the identified variants easy for the clinical geneticist. This is done by utilizing an extensive set of exome sequences to report a statistical measure of confidence for each newly discovered rare variant based on the quality of all previous calls. The ClniCall tool provides an easy and efficient way to integrate a myriad of data into a single viewable format for clinical genetics.
A major problem in the analysis of human genome sequences is the miss-mapping of short read nextGen sequences to nearly equivalent gene loci. i.e., Loci homologous to the sequence miss-mapped. This is a function of the number and density of the difference between any one human sequence and the human reference genome used as the alignment template. In order to correct the alignment template to be a closer match to the sequence that the short reads were derived from, we customized the canonical reference sequence with high confidence differences from the canonical sequence that were derived from SNP chip hybridization results. BEAGLE imputed 1000 genome haplotypes, and results from one iteration cycle of the short read data results of genotyped calls by Haplotype Caller according to Broad Institutes best practice pipeline. Diploid alignment exome data has already been performed in the UDP and is being run on a cloud-based service (Appistry Inc.). We now describe this technique as applied to the whole genome sequencing, using low amplification 105bp Illumina HiSeq2000 short read data. This data is approximately 20x larger than exome data, and presents challenges to the current NIH Biowulf super cluster computing environment, and to the current generation of analytic and data storage resources of the UDP. We present our current results on the first 12 genomes and compare these to the exome data from the same family. With this further progress we plan to extend our analytic search for rare unknown disease causal variants beyond the exomic part of the human genome.
1508T

The abundance of mitochondrial DNA (mtDNA) differs between cell types and changes in copy number (mtDNA CN) have been associated with complex diseases such as cancer. Mitochondria are organelles responsible for the vast majority of energy production within the cell and are centrally involved with processes including signaling, calcium regulation, and programmed cell death. Mitochondria retain a small circular genome which codes for 13 genes, 22 tRNAs, and 2 rRNAs which are essential for respiration. A single cell contains hundreds to thousands of mtDNAs. Because of its central role in respiration, changes in mtDNA CN can have severe effects on metabolism and cellular function. In particular, drastic changes in mtDNA CN have been repeatedly observed in a variety of cancers. Measurement of mtDNA CN has traditionally been performed using real-time quantitative PCR; however, data generated from next-gen sequencing provides an opportunity to quickly and easily quantify mtDNA CN without the need for additional experimental techniques. Several methods have been described to measure mtDNA CN using sequencing data, but the accuracy of these methods has not been well verified. We use a simulation study to compare the accuracy of various methods for measuring mtDNA CN and determine an optimal approach for measuring mtDNA CN using next-gen sequencing data. Accurate estimation of mtDNA CN from sequencing data will provide a tool to incorporate mtDNA CN into association studies and measure changes in mtDNA CN during disease progression.

1509S
An accurate and integrative computational approach for cancer genome studies. L.T. Fang1, M. Mohiyuddin1, J.C. Mur2, P.T. Afsah3, A. Kiani4, N. Bani Asadi5, L. Bullinger1, A. Dolnik1, C. Yau1, W.H. Wong2, H.Y.K. Lam3; 1) Department of Bioinformatics, Bina Technologies, Redwood City, CA; 2) Department of Electrical Engineering, Stanford University, Stanford, CA; 3) Department of Engineering, Bina Technologies, Redwood City, CA; 4) Department of Internal Medicine III, University Hospital of Ulm, Ulm, Germany; 5) Wellcome Trust Centre for Human Genetics, Oxford University, Oxford, UK; 6) Department of Statistics, Stanford University, Stanford, CA; 7) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

Identifying somatic mutations is a key analysis and challenge in cancer research. Its complexity lies in the impure and heterogeneous nature of the samples, rendering each data set a unique problem. An algorithm in one mutation detector may work well for one data set but poorly for another. In this regard, we take an integrative approach to identify and rank the most clinically important mutations based on a combination of different algorithms, sequencing features, and prior knowledge.

We incorporate MuTect, SomaticSniper, VarScan2, and JointSNVMix2 in our cancer analysis pipeline for somatic SNP/Indel detection. For copy number aberration (CNA), we incorporate Control-FREEC, BIC-seq, and OncoSNP-SEQ. We have also developed a brand-new ranking method, OncoRank, which ranks the mutations in each tumor-normal study based on a number of features such as consensus calls, sequencing features (e.g., coverage, strand bias, and allele frequency) and knowledge base (e.g., dbSNP, COSMIC, and Cancer Gene Census). Here, we present our results for a breast cancer cell line (HCC1143) where the samples were pure, as well as two challenging leukemia (AML) data sets where heavy cross-contamination was expected.

For the HCC1143 study, a total of 189,142 somatic single-point mutation candidates were called, of which 90 candidates were COSMIC v54 entries and 38% (71,879) were dbSNP138 entries. Among those, 14,418 candidates were reported by all four detectors with 81 in COSMIC and 8.8% in dbSNP. Using our ranking method, we reported a similar size of call set with a score of at least 6/10. The call set had 13,146 candidates with 87 in COSMIC and 2% in dbSNP, indicating our method was more sensitive (with more COSMIC mutations) and more specific (with fewer known common calls) than the consensus of all four detectors. With our copy number analysis, we correctly detected the ploidy of the HCC1143 cell line as 3.8. Copy number gains were also detected by all the three detectors for breast cancer associated genes (TP53, CDKN1, and AKT1). For the AML study, our pipeline successfully reported top-scored causative variants such as the NPM1 insertion, DNMT3A mutation, and KRAS mutation in two leukemia patients. These results showed that our approach can increase both the sensitivity and specificity of mutation detection, accurately impute causal mutations to handle vastly different data types from pure tumor cell line to challenging liquid cancers.

1510M
Accurate Randomized Dimension Reduction with Applications to Linear Mixed Model Correlations of eQTL Data. G. Darrell1, S. Georgiev2, S. Mukherjee3, B. Engelhardt4; 1) Computational Biology and Bioinformatics, Duke University, Durham, NC; 2) Genetics Department, Stanford University, Palo Alto, CA; 3) Department of Statistical Science, Duke University, Durham, NC; 4) Department of Biostatistics and Bioinformatics, Duke University, Durham, NC.

Expression quantitative trait loci studies have power to detect many causal variants associated with gene expression. In many studies, particularly those of model organisms such as mice, complex population structure can overwhelm true signal and cause spurious associations. Linear mixed models have been shown to partition data in genome-wide association studies into fixed and random effects, accurately accounting for population structure and sample substructure (Zhou et al. 2012). Using sophisticated dimension reduction based on randomized singular value decomposition (SVD), we show it is possible to correct for confounding in many eQTL studies where computational burden made it previously infeasible. We show up to a 50% increase in speed when using our method for randomized SVD compared to optimized out of the box libraries for SVD, while maintaining accuracy above 99%. Our method applied to many eQTL datasets from the GTEx project yields results significantly faster than current methods such as GEMMA (Zhou et al. 2012).

1511T
Evaluation of a genotyping array design for tagging common variation across Africa. T. Carstensen1,2, GDAP Investigators. 1) Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 2) Cambridge University, Cambridge, United Kingdom.

Background: Most current genotype arrays have been ascertainment largely on European populations, and are not designed to capture genetic variation across several populations in Africa. Designing genotype arrays to efficiently capture common genetic variation in Africa could have important implications for large-scale GWAS, including power to identify novel susceptibility loci in the region. Here, we explore the first such design to capture common genetic variation across Africa. Methods: For tagging, we used whole genome sequence (WGS) data generated through the African Genome Variation Project (320 samples sequenced at 4x coverage from Ugandans, Zulu in South Africa and Ethiopians) and phase 1 of 1000 Genomes (97 Luhya from Kenya and 88 Yoruba from Nigeria). To design the African SNP array we used a greedy pairwise tagging algorithm, written to be computationally efficient and scalable across the whole genome. Tagging was carried out across populations using a window size of 500kb flanking each SNP, at an linkage disequilibrium (LD) threshold of 0.80 to capture variants with a MAF≥0.05. To examine efficiency of the genotype array, we carried out imputation using the tagged markers, and calculated the correlation between imputed genotypes and WGS genotypes. We calculated coverage across the genome as the proportion of common variants imputed with r2≥0.80 across all populations. Results: Using our tagging algorithm, we show that approximately 2.7 million SNPs are required to tag all common variation across the five African populations. With the most informative 1 million tagging SNPs identified using the tagging algorithm, we were able to tag more than 80% of common variation across 5 populations in Africa. Conclusions: We present the first exploration of a genotype array design to capture common variation across Africa. We show that while tagging common variation across Africa requires a reasonably large number of variants, even 1 million tagging variants can capture a substantial proportion of common genetic variation across African populations. Our findings suggest that designing a genotype array for large-scale GWAS in Africa is feasible, and can be made cost-effective so as to scale to large sample sizes. We envisage that using hybrid approaches for design that include cyclical tagging and imputation will further improve the efficiency of such an array.
1512S

Increasing demand for NGS tests in our clinical testing laboratory has driven the need for a clinical interpretation process with three key features: performance, efficiency and accuracy. The first is comprehensiveness: handling of different types of panels, disease indications, variants, phenotypes, and family situations, a plethora of bioinformatics predictions and knowledge of genes and variants reported in literature. The second is ergonomics: the ability of an interpretation team to collaboratively assess literature and bioinformatics evidence, review it and then create accurate reports systematically on a large number of cases. The third is efficiency: minimising the time needed to go from reads to report. Thus StrandOmics, a clinical interpretation and reporting platform was created. We illustrate the clinical utility of this platform using some interesting cases comprising:

a) analysis of variants from whole-genome, whole-exome and targeted panel data
b) analysis of single individual as well as more complex multi-individual cases
c) molecular diagnosis for affected individuals and risk prediction for unaffected individuals
d) troubleshooting false negatives by Sanger sequencing

StrandOmics combines knowledge from our internal curated literature content with various publicly available data sources and bioinformatics prediction tools to integrate genomic, phenotypic, structural and functional information. This integrated knowledge is then used to automatically prioritize a list of variants based on ACMG guidelines, the inheritance model and disease phenotypes. All the information needed for interpreting the clinical significance of a variant is displayed in a user-friendly manner, which greatly reduces the time for interpretation. Our team of clinical interpreters has thus far used StrandOmics to analyze over 500 cases with a variety of clinical indications and phenotypes. After its introduction, the average end-to-end time for interpretation, review and reporting was reduced by 67%. In addition, StrandOmics also reports the biologically curated genes and to the identification of likely causative genetic variants in affected individuals. A significant challenge in this approach is identification of the clinically relevant mutation(s) among the millions of variants uncovered by whole genome sequencing. As part of a research study of rare congenital anomalies, we are evaluating the effectiveness of publicly available bioinformatics methods in prioritizing causative mutations. Initially we are testing a genotype to phenotype approach, in which variants are filtered based on inheritance patterns, population frequency, and predicted deleteriousness. The highest ranked variants are then evaluated for congruency with the clinical phenotype. Significant reductions in the number of variants requiring manual assessment come from adapting the tools for combined quality filtering and inheritance-based analysis of family trios, and from comparison to our internal database of whole genome sequences from an ethnically diverse cohort of >1,500 healthy family trios representing >80 countries of birth. Implementation of this approach is in progress; preliminary analyses revealed causative variants for 5 probands enrolled in our congenital anomalies study, all of which have been validated.

1515S
Comparisons on whole-exome capture homogeneity among different versions of capturing kits and populations. M.G. Borges1,2, C. Rocha1,2, B.S. Carvalho1,2, I. Lopes-Cendes1,2. 1) State University of Campinas, Campinas, Sao Paulo, Brazil; 2) Brazilian Institute of Neuroscience and Neuro-technology - BRAINN, Brazil.

coding regions of the genome, which correspond to less than 2% of its entirety, are known as exome: the portion of the human genome believed to concentrate most of the disease-causing mutations. It has been shown that when performing whole genome sequencing of a single individual we expect to find approximately three million variants; however, if one focuses only on coding regions this number drops to less than twenty thousand. In this context, exome sequencing is thought to be a cost-effective strategy for high performance molecular diagnosis applied to genomic medicine and large scale population studies. However, to best apply this approach to different populations it is important to determine whether ethnic differences can affect sensitivity and specificity of the method. These effects could have a negative impact in exome capturing efficiency and homogeneity among different populations according to the necessity of protocol adjustments when analysing patients from a mixed ethnic background. In order to investigate the impact of ethnicity in exome homogeneity, we selected 120 individuals from the 1000 Genomes Consortium. These exomes were sequenced and aligned at three different time-points. For each of those temporal subsets of data there are 40 individuals: ten of each of the four considered populations (ACB - African Caribbean from Barbados; GBR - British in England and Scotland; YRI - Yoruba in Nigeria; JPT - Japanese from Tokyo). We obtained the mean homogeneity for each exon within the human exome as well as the homogeneity at specific genes considered to be clinically relevant by the American College of Medical Genetics and Genomics, as presented at “ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing”. After using multidimensional scaling to reduce dimensionality and capture variability, we identified patterns of exome homogeneity that can be related to time progression and, therefore, protocol advancement. We did not observe evidence for association between ethnicity and mean homogeneity at both levels of data analysis (whole exome and clinically relevant genes). Additionally we perceived that for the clinically significant genes the homogeneity density tends to assume a more normalized distribution as a newer version of capture kits is used. This heterogeneity over time can represent an important issue in large sequencing projects, which are to be developed over the next years.

1513M
Performance survey of protein mutational prediction methods. D.A. Baird1, A. Bierzynska1, I.N.M. Day1, M.A. Saleem2. 1) Bristol Genetic Epidemiology Laboratories (BGE), School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol, B58 2BN; 2) Academic Renal Unit, School of Clinical Sciences, University of Bristol, Dorothy Hodgkin Building, Whitson Street, Bristol, BS1 3NY.

Point mutations are the most common form of variation within the human genome. Next Generation Sequencing (NGS) has accelerated research into the effects of these mutations in human disease. The major and fundamental challenge facing researchers in this area is determining the pathogenicity of mutations amongst a background of many neutral variants. Protein prediction tools are often used as a guide for this. Due to the diverse ways in which an amino acid residue change may impact protein structure, mutational prediction is challenging and a wide range of tools using a variety of computational approaches exist. There is a lack of literature jointly reviewing the performance of tools, with most comparisons done on an individual ad hoc basis during the development stage of the tool. This study extends and updates the surveys of Thusberg et al [1] and Shihab et al [2], by comparing the performance of 9 recent and commonly used web based prediction tools. Mutations annotated as neutral and pathogenic in SwissProt was selected as the benchmark for the performance comparison. The benchmark dataset was submitted to each prediction tool and performance statistics (accuracy, specificity, sensitivity and Matthews Correlation Coefficient (MCC)) and ROC curves calculated. The overall best performing tools were Condel 2 and FATHMM. 1.Thusberg, J., Olatubosun, A. & Vihinen, M. Performance of mutation pathogenicity prediction methods on missense variants. Hum. Mutat. 32, 358-368 (2011). 2.Shihab, H. A. et al. Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. Hum. Mutat. 34, 57-65 (2013).
1516M Prediction consequences of amino-acid substitutions in the IDS gene using in silico tools. A.C. Brusius-Facchin, R. Giugliani, S. Leistner-Segall, S. Sagat. 1 Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; 2 Programa de Pós-Graduação em Medicina: Ciências Médicas, UFRGS; 3 Departamento de Genética, UFRGS.

Computational methods are used to predict the molecular consequences of amino-acid substitutions in the IDS gene and other genetic mutations. The prediction of amino-acid positions, sequence homology, protein folding, crystal structure and information from a database of hotspot mutations in a specific protein, in order to analyze novel missense mutations found in the IDS gene. PANTHER scores of -3 infer a 50% probability of being deleterious, while -12 is associated with a 100% probability of being deleterious. The POLYPHEN scoring predicts three outcomes for mutations: ‘benign’ (most likely lacking any phenotypic effect), ‘possibly damaging’ (may affect protein structure or function) and ‘probably damaging’ (high degree of confidence that protein structure function will be affected). The SIFT scores range from zero to one, with zero predicted to be the most deleterious mutation and one the least deleterious. PMUT classifies as neutral or pathogenic. The results obtained after in silico analysis was: PANTHER, mutations p.D45V; p.D45G; p.S61Y; p.Q80R; p.C84Y; p.L160H; p.L314H; p.D308H; p.D334Y; p.H138Y; p.P160H, p.R95S, p.H342P showed scores between -3 to -10, and the mutations p.Q81Y; p.R95S; p.H138Y; p.N265K; p.V503D showed scores below -3. SIFT: the mutations p.D45V, p.D45G, p.S61Y; p.Q80R; p.C84Y; p.L160H; p.L314H; p.D308H; p.D334Y showed scores between zero to one, with zero predicted to be the most deleterious mutation and one the least deleterious. Our results showed agreement to predict the molecular consequences were neutral except for p.C84Y that was predicted as pathological. Three mutations p.L314H, p.D308H, p.D334Y, p.D334V and p.H342P were predicted as damaging and the mutations p.Q81Y, p.V503D, p.E344K were predicted as tolerated, Polyphen: the mutations p.D45V, p.D45G, p.S61Y; p.Q80R, p.C84Y, p.L160H, p.L314H, p.D308H, p.D334Y and p.H342P were predicted as probably damaging and the mutations p.N265K, p.V503D, p.E344K were predicted as possibly damaging, PMUT: analysis predicted that all mutations were neutral except for p.C84Y that was predicted as pathological. Three software tools showed agreement to predict the molecular consequences and can be considered good tools to characterize the novel alterations found in the IDS gene. The mutations p.N265K, p.E344K and p.V503D in all softwares did not show severe consequences in the protein which should be correlated with a specific (mild) phenotype.

1517T Investigating the relationship between allele frequency of benign variants used in training mutation impact predictors and their stringency at calling deleteriousness. A. Carroll, E. Rozenfeld, J. Price, T. Behrens, T. Behrens, C. Benoist, A. Regev, M.N. Leeb, T. Rajapakshe, T. De Jager,3,5, C. Benoist,3,5, T. Bhangale,4, W. Ortmann5, T. Behrens6, H. Nacohen7,8, A. Regev1,7,8, 1) Broad Institute, Cambridge, MA; 2) Division of Health Sciences and Technology, MIT, Cambridge, MA; 3) Harvard Medical School, Boston, MA; 4) Massachusetts General Hospital, Boston, MA; 5) Genentech Inc., South San Francisco, CA; 6) Department of Biology, MIT, Cambridge, MA; 7) Howard Hughes Medical Institute.

RNA-sequencing (RNA-seq) has emerged as an important tool for profiling gene expression variation because it allows a comprehensive characterization of the transcriptome, including transcript levels, splicing, and allele specific events. Several large scale studies have now used RNA-seq to identify the genetic basis of gene expression variation in humans across different cell types and cell stimuli. Although a number of computational approaches have been developed to (i) align reads, (ii) assemble or annotate transcripts, and (iii) quantify transcript abundances, different computational pipelines aimed at defining quantitative traits from RNA-seq have not been systematically compared.

Here, we evaluate three pipelines for identifying the genetic basis of gene expression in humans. We: (1) align to the transcriptome and estimate fraction of transcripts using RSEM; (2) align to the genome, assemble and estimate fraction of transcripts using TopHat and Cufflinks; and (3) align to the genome, and directly count reads using HISEQ. We apply each pipeline to generate expression or isoform usage quantitative traits for three datasets: (i) the GEUVADIS lymphoblastoid cell line data across 462 individuals, (ii) whole blood across 922 individuals using single end RNA-seq, and (iii) immortalized human neural stem cell data from six cell lines. We evaluate the performance of each pipeline based on its ability to produce a reproducible set of genetic associations to expression (eQTL) or isoform ratio (iQTL) within each dataset through cross validation or across datasets. We compare the results obtained from the three pipelines to those obtained using RSEM to estimate relative abundances from the annotated transcriptome, produced highly reproducible expression levels as compared to microarrays and Nanostation measurements on the same samples. Furthermore, within each transcript, the quantitative traits obtained from the three tools of transcripts, such as transcripts per million and isoform ratio, produced highly reproducible eQTLs and iQTLs across all three datasets. Overall, our analysis highlights that a computational pipeline based on estimating fraction of transcripts from multiply mapped reads is a robust method for discovering variants that control expression and splicing from RNA-seq data.
that facilitates the detection of point mutations, indels and duplications. Custom designs can be used as a targeted NGS resequencing approach. Were successfully identified with a high coverage (>1000 fold average exon validation using known positive control samples comprising samples with analyzed 45101 bp covered by 2999 amplicons. The Haloplex assay was endocrine neoplasia for the validation of the Haloplex assay. Altogether, we comparing the coverage of regions in the sample of patients with control samples. Mustab, Panther and Provean. The amino acid sequence of human ARSB was taken from Uniprot entry P15949 and the crystallographic coordinates were obtained from the PDB database (ID: 1FSU) To prediction of the effect of point mutations on protein stability was performed using structural models generated by FoldX. The analysis of residue conservation was conducted with the ConSurf server and for illustration of conserved amino acids in the sequence was used the Jmol visualization tool. Bead completion of the protein structure of the effect of the change c.1618A-C (p.H111P) in the function on the protein, finding a highly deleterious and a potentially pathogenic variant. Histidine is a e-amino acid with an imidazole functional group and is considered a proline reactive amino acid. The proline change interrupts the conformal reactions of enzymes. The proline change interrupts the conformal reactions of the side chain and makes the chain direction changes abruptly. The presence of proline disrupts the formation of any regular repeating structure Useful bioinformatics programs was observed that this amino acid is conserved in all species; this variant causes a conformational change in the protein and affects the catalytic site seriously, responsible for the enzyme deficit.

Simultaneous detection of copy number variations (CNV) and point mutations with next generation sequencing (NGS) using Agilent HaloPlex custom designs. C. Haag, K. Hauschulz, J. Strüb, E. Schütze. 1) Endocrine Practice and Metabolics Section, Heinrich-Heine University, Germany; 2) Agilent Technologies Sales & Services GmbH & Co. KG, Germany; 3) JSI medical systems GmbH, Kippenheim, Germany.

Copy number variation (CNV) is a form of structural variation in the genome. Usually, CNV refers to the duplication or deletion of genomic segments larger than 1 kbp. Copy number variations have been recognized as pathogenic mutations for many years, for example in mental retardation disorders or Duchenne muscular dystrophy (DMD), where single or multiple exons are affected. The detection of CNVs is important for clinical diagnostics also a common feature in different cancers and the detection of these changes shows promise for the diagnosis of a disease and also for therapeutic or prognostic purposes. Two different methods are traditionally used for mutation detection: one for CNVs and one for point mutations. For the detection of copy number changes in selected regions array Comparative Genomic Hybridization (aCGH) or PCR-based methods such as Multiplex Ligation-dependent Probe Amplification (MLPA) are well established, whereas Sanger-based sequencing or next generation sequencing are the standard methods for the detection of point mutations and small deletions or insertions. We used the Agilent HaloPlex Target Enrichment system (on the Illumina MiSeq platform) for the combined detection of point mutations and copy number variations. Copy number variations are detected by comparing the coverage of regions in the sample of patients with control samples. The copy number variation analysis was performed with the CNV function of the Sequence Pilot software (JSI Medical Systems GmbH, Kippenheim, Germany). We used three different HaloPlex custom designs including 21 genes related to both mental retardation disorders and cancer as well as 2 genes related to endocrine neoplasia for the validation of the HaloPlex assay. Altogether, we analyzed 45101 bp covered by 2999 amplicons. The HaloPlex assay was validated using known positive control samples comprising samples with positive or negative controls. In addition, the disease likelihood analysis was successfully identified with a high coverage (>1000 fold average exon coverage) and 99.2% of the targeted bases were covered with at least 50 reads. Regions with no coverage account for 0.5%. In summary, HaloPlex custom designs can be used as a targeted NGS resequencing approach that facilitates the detection of point mutations, indels and duplications/ deletions in parallel.


Genome wide association studies (GWAS) have been very successful in identifying loci associated with complex disease, however, the genotyping arrays were not designed to capture low frequency and rare variation. As a result, the role of variants in that category of allele frequency remains largely unexplored for a large number of diseases. One possibility would be to perform whole genome sequencing data on a large number of samples, however, the cost of large-scale studies remain prohibitively expensive. Genotype imputation presents a viable alternative for sample sets with existing genotypes from dense microarray platforms. Currently a large number of public projects such as the 1000 Genomes project and the NHLBI exome sequencing project (ESP) can be used as reference datasets to impute missing variation into the sample genotype data. Recent improvements in imputation algorithms and the expansion of reference datasets have improved the accuracy of imputation for even low minor allele frequency variants. Imputed variants can then be assessed against binary phenotypes or quantitative laboratory values derived from patients’ electronic medical records. The Center for Applied Genomics (CAG) maintains a biorepository of over 150,000 genotyped samples, 45,000 of which are pediatric samples randomly recruited from the Children’s Hospital of Philadelphia (CHOP) with complete electronic medical records. As a proof of principle, we imputed missing variants into a subset of just under 100,000 samples that had been previously genotyped in CHOP and validated using a gold standard of genotyping using the Illumina BeadChip platform. Results: Genotype imputation of 170,000 untyped markers (~39M) was carried out using IMPUTE2 against the 1KG Phase I integrated variant set after prephasing with Shapeit. Imputed genotypes were converted to a binary encoding and uploaded to QIAGEN’s Ingenuity Variant Analysis platform for annotation and gene-based analysis using the SKAT algorithm. As a proof of principle we carried out an association against EMR-derived LDL values from 2,500 African American children. The results of the gene-based analysis showed highly significant association of the LDL (P=3.4x10^-10) and PCP (P=3.34x10^-7) genes both of which have recently been associated with LDL in African Americans through genotyping on exome array chips. Imputation of low frequency and rare variation into genotyped data thus represents a viable alternative to sequencing and re-imputing on exome array chips for the study of this class of variant in complex disease.
1524S Impact of statin on gene expression in human lung tissues. J. Lane, B.Y. Aminou, S. vanEeden, D. Sin, M. Obeidat, S. Tébultti, W. Timens, D.S. Postma, M. Laviollette, P. Paré, Y. Bosé (1,2,4,1) Institut universitaire de cardiologie et de pneumologie de Québec, Québec, Quebec, Canada; 2) University of British Columbia Center for Heart Lung Innovation and Institute for Heart and Lung Health, St. Paul’s Hospital, Vancouver, Canada; 3) University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; 4) Department of Molecular Medicine, Laval University, Quebec, Canada.

Objectives: Statins inhibit 3-hydroxy-3-methylglutaryl-coenzyme A reductase, which is involved in the synthesis of cholesterol and about 1/3 of COPD patients are actively using statins. Statins may improve the respiratory health of patients with asthma, COPD and lung cancer by reducing the adverse impact of chronic inflammation and cardiovascular co-morbidities. We studied the effect of using statins on gene expression in human lung in order to identify new molecular pathways underpinning the effects of statins. Methods: Human non-tumor lung tissues were obtained from patients who underwent lung resection. Gene expression was measured on a custom Affymetrix array in a discovery cohort (n=408) and two replication sets (n=282 and 341). Genes differentially expressed between patients taking or not taking statins were tested by linear regression, adjusting for age, gender and smoking status. A propensity score substitution or addition of the effect of the individuals included in the analysis. The results of each cohort were combined in a meta-analysis and biological pathways were studied using Ingenuity Pathway Analysis, DAVID and Gene Set Enrichment Analysis. Results: The discovery set includes 267 statin users. Fifteen genes were found significantly up-regulated in the lungs of statin users (FDR < 0.05). Twelve of these genes were mapped exclusively to the rare network in the second (p-value < 0.05). The meta-analysis improved the significance of the 15 up-regulated genes. Biological pathways analyses provided 21 significant pathways (FDR < 0.05). The first pathway suggested that statins up-regulate genes involved in cholesterol synthesis and receptor substrate/protein kinase B pathway were also affected by statins.

Conclusion: To our knowledge this is the first study to report the effect of statins on the transcriptome of the human lung. The results suggest that statins may improve respiratory health by modulating pulmonary expression of genes involved in biological pathways associated with asthma, COPD and lung cancer.


Schizophrenia, bipolar disorder, and major depressive disorder are psychiatric disorders that affect millions of people every year. The etiology of these disorders in unknown, though each have strong hereditary components. Genome wide association studies associated over 400 single nucleotide polymorphisms (SNPs) with these disorders with a majority of these variants lying in non-coding regions of the genome. In order to characterize the function of these non-coding SNPs, we used epigenomic data from the Encyclopedia of DNA Elements (ENCODE) Consortium and the Roadmap Epigenomics Project to define regulatory regions overlapping each of the variants. In agreement with previously reported results, we found a majority of variants and their linkage disequilibrium (LD) partners overlapped active regulatory elements, as Spliceosome and Long-term potentiation. We remapped reads from ENCODE and Roadmap datasets to observe whether the change in allele affected binding of transcription factors, DNAse I hypersensitivity, and gene expression. We then predicted gene targets for each of these SNPs using the Gene Network tool. These target genes were enriched in similar pathways as the major pathways including growth factor beta and interleukin 2, both of which were previously reported as dysregulated in psychiatric disorders.
1528M
A network-based approach to dissect the cilia/centrosome complex interactome. M. Morleo1, R. Amato1, L. Giaquinto1, D. di Bernardo1, B. Franco1,2, 1) TIGEM, Telethon Institute of Genetics and Medicine, Naples, Italy; 2) Department of Computer and Systems Engineering, University of Naples "Federico II", Naples, Italy; 3) Department of Medical Translational Sciences, Federico II University, Naples, Italy.

Background Cilia are microtubule-based organelles protruding from almost all mammalian cells that, when dysfunctional, result in genetic disorders called "cilipathies". High-throughput studies have revealed that cilia are composed of thousands of proteins. However, despite many efforts, much remains to be determined regarding the biological functions of this increasingly important complex organelle. Results We derived an online tool from a systematic network-based approach to dissect the cilia/centrosome complex interactome (CCCI). The tool is able to integrate all current data available into a model which provides an "interaction" perspective on ciliary function. We generated a network of interactions between human proteins organized into functionally relevant "communities", i.e. a group containing genes that are both highly interconnected and strongly co-expressed. We then combined sequence and expression data in order to identify the transcription factors responsible for regulating genes within their respective communities. Our analyses revealed communities significantly specialized for delegating specific biological functions such as mRNA processing, protein translation, folding and degradation processes that had never been associated with ciliary proteins until now. Conclusions CCCI will allow us to clarify the roles of previously unknown ciliary functions, elucidate the molecular mechanisms underlying ciliary-associated phenotypes, and apply our knowledge of the functional roles of relatively uncharacterized molecular entities and disease phenotypes to new clinical applications.

1529T
Identification and Clinical Assessment of Deletion Structural Variants in Whole Genome Sequences of Acutely Ill Neonates. A.C. Noll1,2, L.D. Smith1, N.A. Miller1, C.J. Saunders2, I. Thiffault1, A. Newton1, K. Detherage1, J. Hoang2, L.K. Willig2, L.D. Cooley2, S.D. Fiedler2, E.G. Farrow2, S.F. Kingsmore1,2, 1) University of Kansas Medical Center, Kansas City, KS; 2) Center for Pediatric Genomic Medicine, Children's Mercy Hospital, Kansas City, MO.

Effective management of acutely ill newborns with likely genetic diseases requires rapid and comprehensive identification of causative haplotypes. While whole genome sequencing (WGS) can identify pathogenic nucleotide variants in newborns in less than 50 hours, deletion structural variants (DSVs) >50 nucleotides are also an important component of the mutation burden for genetic diseases and are missed by many variant detection pipelines. Concomitant identification of DSVs in rapid WGS data for diagnosis of acutely ill newborns would be highly valuable. Here we provide the development of DSV detection methods that combine consensus calls from two WGS DSV detection tools (Breakdancer and GenomeStrip) with a novel filtering and clinical strategy. WGS simulation data demonstrated Breakdancer and GenomeStrip consensus calls had 93% sensitivity, 99% positive predictive value, and high precision. Upon inspection of read depth, and alignment distance of paired reads in the integrated genome viewer (IGV), DSV consensus calls overlapping with SNP array calls were found to be 95% true positive. Consensus DSV calling with parameterized filtering was implemented in an 8-hour computational pipeline called SKALD (Screening Konsensus and Annotation of Large Deletions). IGV evaluation of SKALD results in a tetrad demonstrated more than 80% were true positives, however the sensitivity was low. To assess the clinical utility of SKALD, WGS from 12 familial trios with an acutely ill newborn proband in which causative nucleotide variants had not been identified were analyzed. A heterozygous deletion of exons 1-3 of MMP21 (NC_000010.11 g.127,460,714_127,461,028del) was found in trios with a heterozygous frameshift deletion (p.Met122SerX55) in two siblings with transposition of the great arteries and arterial ductal heterotaxy (TAH). These variants are expected to be deleterious to MMP21, resulting in a truncated or degraded protein. Mmp21 -/- mice exhibit TAH, supporting MMP21 as a novel disease gene. In the newborn female with dysmorphic features, ventricular septal defect, and persistent pulmonary hypertension, SKALD identified the breakpoints of a heterozygous, de novo 1p36.32p36.13 deletion (NC_000001.11 g.4,848,728_18,503,068del). In sum, SKALD has the potential to increase the diagnostic yield of WGS in acutely ill infants and to discover novel disease genes.

1530S
Disease functional domain annotation in single gene association in Parkinson Disease. K. Nuytemans1,2, V. Inchausti1,2, L. Maldonado1,2, A. Mehta2,3, N.A. Miller1,2, 1) Human Genomics, University of Miami, Miami, FL; 2) Miami Uddall Center, University of Miami, Miami, FL.

Although whole exome sequencing (WES) allows identification of all variants in the coding regions of the genome, it was needed to assess the potential impact of any of those variants on the protein and thus potentially the disease mechanism. So far, many studies have used amino acid change prediction programs based on conservation or 3D structure, but these programs lack the protein’s functional domains which will augment the potential effect of variants to be included in the analysis. Most protein databases currently available, however, do not allow for annotation on the genomic level. As a proof-of-principle we used the Pfam database available as a UCSC track to filter variants for analysis. We performed WES in 410 PD patients and 229 controls. Functional annotation of identified variants for effect in protein and presence in protein functional domains (UCSC Pfam track) was performed through ANNOVAR. Rare variants (MAF<5%) were identified using the 1000Genomes and the NHLBI Exome Sequencing Project datasets. Association with risk for PD was assessed for sets of variants in genes using the optimal Sequence Kernel Association test (SKAT-O). Genes with p-value <0.01 were included in the enrichment analyses in KEGG pathways using WEB-based Gene Set Analyses Toolkit (webgestalt). In total, 37,965 protein sequence altering variants (nonsense or missense) were annotated to a protein functional domain in 9736 genes. The top 10 genes (p<0.002) using rare or all variants were similar and included genes previously reported in PD. Variants in or involved in "PD pathway" genes may be involved in the PD pathway. Additional analyses showed that patients reported in PD pathway analyses. These proof-of-principle analyses confirm the validity of using presence in a protein functional domain as a variant inclusion requirement. The analyses presented here are limited by the data available on gene sets and PD (Pfam) (predicted) protein domain database. Therefore, we will set out to evaluate this approach of protein domain annotation for genomic variants including different algorithms (e.g. TIGR, PANTHER,..) to widen the search for potentially damaging variants.
**1532T**

Novel bioinformatics driven imaging-genetics approach exploring the aetiology of Alzheimer's disease. S. Patel1,2,3, M. Park1, J. Pipitone2, M.M. Chang2, S. Chakravarty1,3,4,5, R. Madhani1,2,6,7,2. Author for correspondence. Alzheimer's Disease Neuroimaging Initiative. 1) Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada; 2) Institute of Medical Science, University of Toronto, Toronto, Canada; 3) Kimel Family Translational Imaging-Genetics Research Laboratory, the Centre for Addiction and Mental Health, Toronto, Canada; 4) Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Canada; 5) Department of Psychiatry, University of Toronto, Toronto, Canada; 6) Douglas Mental Health University Institute, Verdun, QC, Canada.

**Introduction:** Alzheimer’s disease (AD) is a devastating illness, affecting over 35 million people worldwide and expected to increase to 115 million by 2050. Our goal is to identify biomarkers for AD progression based on structural brain magnetic resonance imaging (MRI) and transcription expression data, systematically and bioinformatically-driven approach to shed light on the genetic burden associated with AD. **Methods:** T1-weighted MR images were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database. Hippocampal segmentation was carried out using the MASS1 brain algorithm on 162 AD subjects, 317 with mild cognitive impairment and 183 healthy controls [Pipitone et al., 2014]. Hippocampal volumes were used for association testing with both genotypes from the Human 610-Quad BeadChip and imputed genotypes (n=5,706,974) and the following covariates were controlled for: gender, age, first principle component from multidimensional scaling, baseline diagnoses, APOE status and intracranial volume. Stratified false discovery rate (sFDR) [Sun et al., 2006] was used to prioritize SNPs in genes selected using Gene Ontology (GO) networks. A list of shared genes with AD was extracted from a previously published study [Lambert et al., 2013]. Common GO terms within the gene list were identified and Cytoscape was used to visualize the relationships of the GO terms in a network format. The GO network was then pruned to enrich for GO terms containing specific enrichment in age, sex and cognitive measures. We identified all the SNPs that are associated to all the GO terms in the nervous system development GO network which formed our stratum for sFDR. **Results:** A total of 249,001 out of 5,706,558 SNPs were selected from 1146 genes from the nervous system development GO network to be up-weighted in sFDR. After sFDR correction, 84 significant SNPs were found to be associated with hippocampal volume. **Conclusion:** SNPs in the nervous system development stratum list were not significantly associated with hippocampal volume. SNPs in this stratum may not play a specific role in hippocampal volume, for example SNPs in genes related to growth arrest-specific protein 7 is involved in neuronal development and is mainly expressed in mature cerebellar Purkinje cells [Ju et al., 1998]. Further modifications to prune the nervous system development GO network to be up-weighted in sFDR. The functional and biological characterization of these significant SNPs were found to be associated with hippocampal volume. **Objective:** Identify, classify non-coding RNAs and detect transcriptional expression for the two different types of neuronal tissues (dorsal root ganglia (DRG) and proximal sciatic nerve (SN)) following sciatic nerve injury. **Method:** Microarray probe sequences for the differentially expressed genes (DRG) and proximal sciatic nerve following sciatic nerve injury at 4 od, 1d, 4d, 7d and 14d post-sciatic nerve injury in rats were obtained for DRG and SN from the gene expression omnibus (GEO) generated by the Agilent-014879 Whole Rat Genome Microarray 4x44K G4131F. Using these datasets and the expression change within different time points, probe sequences were grouped by tissue, gene and category of differential ncRNAs (DRG and SN) at 4 different time points in nerve injury. In the primary analysis, the extracted probe sequences were aligned against mouse reference to identify the homologous sequences between the two rodent species. In the secondary analysis, the raw microarray probe sequences were analysed using GeneSpring 12.6 GX to identify differentially expressed genes across the different time points and potentially, understand the gene regulatory effect related to nerve injury. **Results:** From the primary analysis, we detected increased expression of more than 1000 ncRNAs and over 200 pseudogenes in DRG tissues from 1d-14d post-injury. In SN tissues, we detected about 190 lincRNAs at 1d which increased to 310 lincRNAs at 7d post-injury. Comparatively in SN, more ncRNA categories (N=15) were identified than DRG. For instance, the expression of lincRNA is the first report of the detection of different categories of ncRNAs and transcription expression underlying NP following sciatic nerve injury. Interestingly, the increase in the expression levels of long ncRNAs at extended day post-injury poses many questions related to tissue repair or other mechanisms related to NP.

**1534M**

Unrevealing the genomic architecture of chromosomal breakpoint region using multiparametric computational approach. R.M. Rawal1,2,3, A. Kleinfelder1,2,3,4, R.K. Chelliah1,2,3,4, T. Mercier1,2,3,4, J. Pate1,2,3,4, J. Bernier1,2,7,2. Author for correspondence. 1) Ludwig Institute for Cancer Research, The Gujarat Cancer & Research Institute, Ahmedabad, Gujarat, India.

Recent non-random chromosomal translocations are hallmark characteristics of leukemogenesis however molecular mechanisms underlying these rearrangements are less explored. The fundamental question is, why and how chromosomes break and reunite so precisely in the genome. Meticulous understanding of mechanism leading to chromosomal rearrangement can be achieved by characterizing breakpoints. To address this hypothesis, a novel multiparametric computational approach for characterization of six most frequent leukemic translocations within and around breakpoint region was performed. To best of our knowledge, this study is unique in finding the presence of Segmental Duplications (SDs) flanking breakpoints of all major leukemic translocation. Breakpoint islands were also analyzed for other complex genomic architecture and physical properties e.g. SIDD, repetitive elements, recombination signal sequence, base composition, Topo-II binding site & restriction endonuclease cleavage site etc. Our study distinctly emphasizes on the probable role of SDs and various genomic features in the occurrence of breakpoints. Further, it also highlights what those potential features may be which play a crucial role in causing double-strand breaks leading to translocation.

**1535T**

Assessing the hidden genome architecture of structural variants with Globus Genomics Galaxy pipelines. A. Rodriguez1, S.G. Potkin2, R. Madhani1,2,6,7,2. F. Macciardi2,3,5,6,7, S. Gaud1,2,3,5,6,7,4, R. Madhani1,2,6,7,2. 1) Computation Institute, University of Chicago, Chicago, IL; 2) MCS, Argonne National Laboratory, Argonne, IL; 3) Department of Psychiatry & Human Behavior, University of California, Irvine, School of Medicine, Irvine, CA; 4) Division of Child and Adolescent Psychiatry, Department of Psychiatry, Columbia University/ NYSPI, New York, NY; 5) Center for Autism Research and Treatment (CART), University of California, Irvine, California; 6) Center for Epigenetics and Metabolism, University of California, Irvine, California; 7) Department of Pharmacological and Biomolecular Sciences, University of Milan, Milan, Italy; 8) Department of Infectious, Parasitic and Immune-Mediated Diseases Istituto Superiore di Sanità Rome, Italy.

Next-generation sequencing (NGS) opened a new phase of discovery of structural variants. The list of programs and algorithms to assess insertions, deletions, duplications, inversions and mobile elements is in continuous expansion. To get a comprehensive picture of the structural variants genomic architecture, researchers need to use different programs with advanced programming-based query interface, which is often demanding due to the large number of required tools and considerable computational infrastructure and needs. We used cloud services and Globus Genomics platform for development of comprehensive pipeline for assessing the “hidden” genomic architecture of structural variants that includes all steps from processing and quality control of data generated by NGS to functional annotation and data visualization. Globus Genomics enabled us to share our analysis pipelines, collaboratively build the pipeline, share the results of the analysis with our collaborators and ultimately make our published results accessible to the research community at large thus promoting transparency and reproducibility.

Our research platform integrated Globus for data transfer and management, Galaxy for tool and workflow management and Amazon Web Services (AWS) for compute resources. In addition, we leveraged the Swift parallel language into the execution of tools for seamless parallelization of datasets at a chromosomal level. We optimized each tool to use best available Amazon Web Services resource based on memory and computation needs. Optimizations include identification of tools that can be parallelized to minimize execution time and costs which allowed us to reduce the execution time factor by 5. At the time of this writing, we implemented in the pipeline the most up to date, publicly available tools: ForestSV used for detection of deletions and duplications; INDELS and mobile elements insertions are detected using Retrospec and Tango; reference-based mobile elements (LINE, ALUs and HERVs) are detected with RepeatMasker.

Finally, we leveraged the Meticulous understanding of mechanism leading to chromosomal rearrangement for other complex genomic architecture and physical properties e.g. SIDD, repetitive elements, recombination signal sequence, base composition, Topo-II binding site & restriction endonuclease cleavage site etc. Our study distinctly emphasizes on the probable role of SDs and various genomic features in the occurrence of breakpoints. Further, it also highlights what those potential features may be which play a crucial role in causing double-strand breaks leading to translocation.

Our objective is to perform integrative analysis of genetic variations such as mutations, gene fusions, copy number, and transcript abundance in tumor samples from patients and cultured tumor cell lines. We envision that these analyses will help identify cancer cell lines best represent a given primary tumor type. We will also discover candidates for biomarkers and therapeutic targets. We process saw RNA sequencing data for identification of gene fusion events. We obtain other processed genetic data from TCGA Firehose and from ICGC Data Portal. We devise mapping strategies specific to each sequencing data type; differential mutation and differential gene and transcript expression (from RNA sequencing data); identification of genetic variations shared by patient samples and cell lines representing the corresponding primary tumor; compound screening panels on cancer cell lines for selected tumor types. Where applicable, we utilize in silico analysis pipelines to handle mapping and variant calling (RNASeq and WXS), gene fusion detection, splice junction detection, transcript-level expression, and isoform usage (RNAseq). Our effort leads to several candidate biomarkers and targets. Experimental validations will be needed to confirm the utility of these candidates. We can also use our integrated data for selecting cell lines to use in compound screenings.

Identification of differentially methylated genes potentially associated with neurological diseases. W. Souza, B. Carvalho, D. Dogini, I. Lopes-Cendes. Department of Medical Genetics, UNICAMP, Campinas, Brazil.

Epigenetic marks, including DNA methylation and histone modifications could play a role in neurological disorders with complex inheritance, such as epilepsy and stroke. However, this issue has not been sufficiently investigated. In this context, our aim was to perform an epigenome-wide association study (EWAS) using a public dataset (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE37579), to assess methylation profile across distinct tissues of patients with neurological disorders and compared this to that obtained from a control group. We used principal component analysis (PCA) for data quality control and empirical Bayes methods to adjust for systematic differences allowing better comparisons between tissues and disease-related conditions. Our results show a latent variable that stratified control samples into two subgroups and could represent a potential source of bias for downstream analyses. After further investigation we observed that the two clusters were tightly related to the different ethnicities of the individuals in the control group. After correcting for this variable we applied recent methods for methylation profiling using bumphunter Bioc conductor package, which allows one to search for differences in methylation at the regional level rather than at the single-CpG level. We identified differentially methylated regions (DMR) when comparing DNA methylation profiles from different tissue groups. Finally, we compared the list of differentially methylated regions to a curated list of genes known to be associated with neurological disorders. We found that regions of a set of genes related with epilepsy and stroke (including RELN, ARHGEF9 and BDNF) were differentially methylated. In conclusion, we identified and corrected a source of bias in the dataset, performed an analysis for differential methylation and identified DMRs that included genes potentially associated with epilepsy.

Genome-Wide Association Study Of Cerebrospinal Fluid Prostatic Acid Phosphatase Levels. L.A. Staley1, P.G. Ridge1, M.H. Bailey1, C. Cruceanu2, C. Groop1, J.S.K. Kauwe1. 1) Biology, Brigham Young University, Provo, UT, USA; 2) Biology and Biomedical Sciences, Washington University in St. Louis, St. Louis, MO, USA; 3) Psychiatry, Washington University School of Medicine in St. Louis, St. Louis, MO, USA.

Prostatic Acid Phosphatase (PAP) is an enzyme that is produced in the prostate in males, and its physiological function is thought to deal with the liquefaction process of semen. Understanding the genetics of this enzyme is very important because this enzyme may be found in highly increased levels in men who have prostate cancer and moderately increased levels for bone diseases, such as Paget’s disease or hyperparathyroidism and blood cell diseases, such as sickle disease, or multiple myeloma or lympho-somal storage diseases, such as Gaucher’s disease. For the initial association analysis in each series we used PLINK to perform linear regression and evaluated the association between the additive model for 5.8M SNPs and each phenotype. Age, gender, and the principle components from Eigensoft were included as covariates. Meta analysis was performed using default settings in METAL. Genomic inflation factor scores (GIF) were estimated using the R package GenABEL. The initial threshold for significance was P<5×10^-8. SNPs that met this threshold were further filtered using the following criteria. First, we rejected markers where the direction of the effect was different in the Knight ADRC and ADNI datasets. Second, we removed all markers where the minor allele frequency was less than 5% (unless they were directly genotyped or had a clear functional annotation). Finally, we also rejected all associations with phenotypes where the genomic inflation factor was greater than 1.03 (GIF was calculated without SNPs where MAF was different in the Knight ADRC and ADNI datasets. We found 233 genome-wide significant markers (P<5×10^-8). We will analyze each of these markers individually and study the biological functions of each, to discover which is casual of the increased levels of this enzyme for men who have cancer or other diseases.

European Psoriasis Differences are Defined by Variation in the Epidermal Differentiation Complex. C.E. Tanes, L. Jackson, A. Tozeren. Biomedical Sciences, Engineering and Health System, Drexel University, Philadelphia, PA, USA.

Psoriasis is an autoimmune skin condition affecting 2-3% of Americans. Interestingly, European Americans are almost twice as likely to have psoriatic plaques as African Americans (3.6% and 1.9% respectively). Previous studies have shown that psoriasis has a strong genetic signature in human populations, leading us to examine the role that genetic variation might be playing in these prevalence differences. We harness the numerous curated genes identified through single gene and genome wide association studies to broaden the investigation of candidate polymorphisms in ethnic populations. Psoriasis genes curated through the National Center for Biotechnology Information (NCBI) Gene were mapped onto the human genome and regions with an overabundance of psoriasis genes were identified. These genomic hotspots were projected onto gene ontology categories and cellular pathways to draw a bioinformatics portrait of psoriasis. HapMap (11 populations) and Human Genome Diversity Panel (51 populations) polymorphism data were analyzed for the psoriasis hotspot regions. Our analyses identify six psoriasis hotspots located on chromosomes 1, 2, 5, 6, 17 and 20. Functional annotation of these hotspots showed significant (p<0.05) biological processes consistent with keratinization, B cell proliferation and antigen presentation, core disruptions consistent with previous psoriasis studies. Population based polymorphic analyses identified an European specific pattern in genes making up the epidermal differentiation complex (1q21), which bear further investigation as the potential cause of population differences in psoriasis prevalence.
1540M
Exploiting whole exome-seq data for variant discovery from highly divergent regions in the human genome. S.L. Tian, H.H. Yan, S. Slager, Maxo Clinical, Boston, MA and Amelieff Corporation, Tokyo, Japan.

Whole Exome Sequencing (WES) has been widely used to characterize genetic variations within human genomes. However, mutation rate varies substantially across the human genome, which is a key determinant for variant calling. For example, human leukocyte antigen (HLA) region contains the most polymorphic genes in the human genome and is associated with over 100 diseases. Accurate identification of variants from such highly divergent regions (HDRs) often relies on Sanger sequencing. While framework algorithms have been established for genome-wide variant calling for WES data, their feasibilities in HDRs have yet to be assessed. Using simulated and real data generated from our CLL (chronic lymphocytic leukemia) project, we have conducted a comprehensive assessment of numerous pipelines for variant calling in HDRs. The results suggest a straightforward model with a high level of accuracy and specificity over a wide range of divergence. Five short-read aligners were selected, including bwa, novoalign, gsnap, NextGenMap, and stampy. From simulated data we found that four of the mappers were highly comparable in mapping accuracy (>99%) for reads with <1% divergence; novoalign was 3-4% lower instead, a result largely caused by its extensive use of ‘soft clipping’ to mask unaligned portion in a read. At an increased divergence of 5%, gsnap, NextGenMap and stampy still achieved ~99% mapping accuracy, versus up to 20% unmapped rate by bwa. 10% and 15% divergence in particular, stampy represents the most effective mapper, followed by NextGenMap and gsnap. We next investigate the overall performance of five selected variant callers: GATK UnifiedGenotyper, GATK HaplotypeCaller, freebayes, platypus and SAMtools/mpileup. GATK UnifiedGenotyper showed the highest sensitivity for SNP calling (mismatch error) in the HDRs, while from simulated reads mapped by gsnap (at >1% divergence rates) or by bwa and stampy (at <1% divergence rates). More importantly, we observed similar trends in performance for these mappers and callers when applied to the CLL data. We provide comparison of the calls made from our in-house Illumina iSelect genotype array data and to the known variants in dbSNP. Our findings highlight several key factors and combinations recommended to use in order to have a successful WES-based variant discovery in HDRs.

1541T
An in silico Post-GWAS Analysis of C-Reactive Protein Loci: a Pipeline of Sequential Bioinformatics-Based Approaches. A. Vaez1, H. Janzen2,3, E. Bonné2, L. Dahlborg2, M. Blennow2, E. de Geus4,5, D. Boomsma2,3,5, B. Penninx2,3,5, I. Nolte1, H. Snieder1, B. Alizadeh1, 1) Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; 2) Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands; 3) Neuroscience Campus Amsterdam, The Netherlands; 4) Department of Psychiatry, VU University and VU University Medical Center, Amsterdam, The Netherlands; 5) Department of Biological Psychology, VU University, Amsterdam. The Netherlands; 5) EMGO+ institute, VU University and VU University Medical Center, Amsterdam, The Netherlands.

Genome-wide association studies (GWASs) have successfully identified a number of Single Nucleotide Polymorphisms (SNPs) associated with serum levels of C-reactive protein (CRP). An important limitation of GWASs is that the number of Single Nucleotide Polymorphisms (SNPs) associated with serum CRP levels of Caucasian individuals, with a focus on non-synonymous SNPs as potentially functional variants. In the second phase of eQTL analysis, it attempts to identify all nearby genes whose expression levels are associated with the component GWAS SNPs. These two phases generate a number of relevant genes that serve as input to the next phase of functional network analysis. We applied this pipeline to the 18 SNPs that had previously been associated with CRP at a genome wide level. Our in silico sequencing analysis using the 1000 Genomes Project data identified 3,801 linked variants, including 25 non-synonymous SNPs. Our eQTL analysis, based on one of the largest single datasets of genome-wide expression probes (n=5000) assessed in participants from the Netherlands twin register (NTR) and the Netherlands study of Gene-Environment interaction (GENIE) who had provided 23 annotated expression probes belonging to 15 genes (FDR<0.01). The initial phase of functional network analysis, which was based on 40 relevant genes identified in the previous phase, revealed significant biological function biomarkers (FDR<0.01). Our post GWAS analysis of CRP GWAS SNPs confirmed the previously known overlap between CRP and lipid biology. Additionally, it suggested an important role for interferons in the metabolism of CRP.

References:

1542S
The Variant Characterization of 211 Whole Genome Sequences: The Cache County Study on Memory Health and Aging. M.E. Wadsworth1, B.D. Pickrell2, J.B. Miller3, J.D. Duece4, R.G. Mungur5, C.D. Concannon6, J.T. Tschanz7, M.C. Norton7, J.S.K. Kauwe7, K.V. Voelkerding8, P.G. Ridge9, 1) Department of Biology, Brigham Young University, Provo, UT, USA; 2) The Cache County Study on Memory and Aging, Utah State University, Logan, UT, USA; 3) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, USA; 4) Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, USA.

The 1000 Genomes Project is the current standard for population based variation data, which has been used for genome-wide association studies (GWAS) for a variety of complex traits. With the development of new bioinformatics methods, these datasets are now available for further analysis. We investigated the whole genome sequences (WGS) from 211 Caucasian individuals from the Cache County Study on Memory Health and Aging. Each of the genomes is from elderly (>65 years old), healthy individuals of European descent. Many of the CEU genomes in the International HapMap study were derived from individuals in the Cache study and are comparable to the European genomes from the 1000 Genomes Project. Due to the high coverage of these genomes it is possible to accurately characterize rare variants. We created a variant profile for each of the 211 individuals and the Cache dataset as a whole. In total, we identified 23,073 non-synonymous mutations (a SNP changing a nucleotide in the gene) in the 211 genomes, as well as the calls made from our in-house Illumina iSelect genotype array data and to the known variants in dbSNP. Our findings highlight several key factors and combinations recommended to use in order to have a successful WES-based variant discovery in HDRs.

1543M
Genome-wide haplotype-based association study in Chinese Han population identified novel susceptibility locus for systemic lupus erythematosus. Y. Wang1, Y. Zhang2, J. Yang1, Y. Lai3, W. Yang2, 1) Department of Paediatrics and Adolescent Medicine, LKS Faculty of Medicine and Queen Mary Hospital, The University of Hong Kong, Hong Kong, China; 2) Centre for Genomic Sciences, LKS Faculty of Medicine and Queen Mary Hospital, The University of Hong Kong, Hong Kong, China; 3) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, USA.

Systemic lupus erythematosus (SLE) is a heterogeneous and chronic autoimmune disease, resulting in recurrent inflammatory response and multi-organ damage. Family and twin studies suggested that genetic factors contribute to the disease. The first wave of Genome-wide association studies have identified more than 40 susceptibility loci robustly associated with SLE. However, all of them together only explained a small fraction of disease heritability. In this study, we carried out a genome-wide haplotype-based association study, using the previously published WGS data from two Chinese Han cohorts, Hong Kong (Southern China) and Anhui (Northern China), respectively. Our analyses identified a novel region (meta haplotype-based p-value 6E-07) in chromosome 2q22.1 (106,072,254-110,767,502) with a maximum p-value of 1.2E-17. After detection, known variations, variations of low quality, and variations with low impacts are filtered out. We analyzed publicly available NGS data of tumor-normal pair samples to verify the accuracy of the associations identified in the novel locus.

1544T

Next-generation sequencing technique has dramatically improved the efficiency and the speed of large scale sample analysis. As a result, picking out variations that have a significant impact on the disease of interest from a massive number of detected variants becomes a difficult challenge. To overcome this problem, we established an analysis workflow to detect realistically verifiable number of somatic SNVs. In this workflow somatic SNVs are detected by SomaticSniper, the existing software designed to detect somatic SNVs. After detection, known variations, variations of low quality, and variations with low impacts are filtered out. We analyzed publicly available NGS data of tumor-normal pair samples to verify the accuracy of our workflow.
Identification of recurrent drive gene fusions in melanoma using RNA-Seq data. T. Zhang, M. Xu, P. Johansson, A. Pritchard, N.K. Hayward, K.M. Brown. 1) National Cancer Institute, Bethesda, MD; 2) QIMR Berghofer Medical Research Institute, Brisbane, Australia.

Global gene expression profiling by microarray or RNA-seq has been used successfully to discover prognostic and predictive gene signature biomarkers in a number of cancer types and therapies. Translation from pre-clinical studies to the clinic involves sharing common characteristics of known ‘driver’ events. In total, we observed 295 potential fusion events, 112 of which involved neighboring genes located within 1 Mb on the same chromosome and 101 involving more distant same-chromosome partners. Only 82 of potential fusions were inter-chromosomal. In summary, we detected 19 fusion events involving PRKAR1A, RBMS3, CBX3, PTPRO, RERG, PTEN and BRAF recently occurring in at least 3 samples with the same partner genes. Notably, we identified multiple fusions involving known melanoma oncogenes or tumor suppressors, including BRAF, PTEN, GRIN2A and NF1. The Oncofuse pipeline identified the most significant fusions in the TCGA melanomas. We found that the fusion gene gain of chromosome 22p (MECP2 and IGF1R, FDR<0.01). In addition, we also identified and summarized fusion events in the TCGA Skin Cutaneous Melanoma dataset and found the recurrent/driver fusions in melanoma cell lines were also identified in TCGA samples. Particularly, PTEN, MAPKAPK2, RERG, PTPRO and BRAF fusions in melanoma cell lines were also observed in multiply TCGA samples with the same breakpoint. There were a total of 9 BRAF fusions in our cell lines and TCGA samples, and 7 of them were BRAF wild type with protein kinase domain retained. We were able to functionally characterizing several of these gene fusions, including BRAF, PTEN, ATP2, and PRKAR1A to establish their role in driving melanoma and as potential novel therapeutic targets.


Molecular targets for effective therapeutic intervention in melanoma have recently been identified within the RAS-RAF-RAF-MEK-ERK and, to a less extent, PI3k-AKT pathways. Despite the promise of approaches targeting these pathways, therapeutic resistance remains a significant problem, and the need to identify new molecular targets remains a critical need. Gene fusion events resulting from inversions, interstitial deletions, or translocations occur commonly in human tumors and are thought to be a result of the fusion of two genes with unique oncogenic properties. In order to identify fusions driving melanoma development and/or progression, we performed paired-end RNA-Seq of 72 melanoma cell lines using the Illumina HiSeq 2000. After filtering out low quality data, TopHat was used to align RNA-Seq reads to the human genome (hg19), and TopHat-Fusion was subsequently applied to identify potential fusions. In order to assess the potential of these fusions to act as driver events, we used a naive Bayesian classifier Oncofuse to highlight fusions with common characteristics of known ‘driver’ events. In total, we observed 295 potential fusion events, 112 of which involved neighboring genes located within 1 Mb on the same chromosome and 101 involving more distant same-chromosome partners. Only 82 of potential fusions were inter-chromosomal. In summary, we detected 19 fusion events involving PRKAR1A, RBMS3, CBX3, PTPRO, RERG, PTEN and BRAF recently occurring in at least 3 samples with the same partner genes. Notably, we identified multiple fusions involving known melanoma oncogenes or tumor suppressors, including BRAF, PTEN, GRIN2A and NF1. The Oncofuse pipeline identified the most significant fusions in the TCGA melanomas. We found that the fusion gene gain of chromosome 22p (MECP2 and IGF1R, FDR<0.01). In addition, we also identified and summarized fusion events in the TCGA Skin Cutaneous Melanoma dataset and found the recurrent/driver fusions in melanoma cell lines were also identified in TCGA samples. Particularly, PTEN, MAPKAPK2, RERG, PTPRO and BRAF fusions in melanoma cell lines were also observed in multiply TCGA samples with the same breakpoint. There were a total of 9 BRAF fusions in our cell lines and TCGA samples, and 7 of them were BRAF wild type with protein kinase domain retained. We were able to functionally characterizing several of these gene fusions, including BRAF, PTEN, ATP2, and PRKAR1A to establish their role in driving melanoma and as potential novel therapeutic targets.

Genome-wide association study of serum metabolites using non-targeted metabolomics to identify new metabolic loci. L. Lind, S. Gustafsson, S. Stahlholm, A. Ganna, R. Cohn, C.D. Broeckling, J. Pinner, E. Ingelsson. 1) Dept of Medical Sciences, Uppsala University, Uppsala, Uppsala, Sweden; 2) Department of Medical Sciences, Molecular Epidemiology and Science for Life, Karolinska Institutet, Stockholm, Sweden; 3) Uppsala University, Uppsala, Sweden; 4) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA; 5) Genetic and Genomic Medicine Centre, Umeå University, Umeå, Sweden; 6) Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, Colorado, USA.

Background: One way to understand human metabolism in more detail is to identify genes involved in metabolic pathways. The most recent efforts have linked around 150 genetic regions to different metabolites via genome-wide association studies (GWAS) of the metabolome. The present study aimed to extend this knowledge using three large independent cohorts in which metabolites were measured using metabolome-wide association profiling methods. Methods: In the ULSAM (n=1,028), PIVUS (n=970) and TwinGene studies (n=1,670), up to 10,000 metabolic features were measured in serum/plasma using the same technical platform across studies (UPLC and MS/MS). Genotyping was performed using the Illumina Human OmniExpress or Omni2.5M array and subsequently imputed using 1000 Genomes CEU data. A GWAS was performed for each metabolic feature using linear regression adjusting for age, sex, and population-based principal components. We used TwinGene as the discovery dataset and attempted replication in PIVUS for the genome-wide association signals identified in the ULSAM and PIVUS cohorts. Results: We could confirm 42 of the previously reported associations between genetic regions and metabolites. Top findings included associations with the PYROXDH locus (p-value=4.0×10-323) and the CYP2D6/3A7 locus (p-value=2×10-310). In total, we identified 8 novel loci with strong associations between genetic variants and circulating metabolites (p-value range 1.1×10-9 to 6.1×10-93). Likely genes in those 8 regions were CYP3A43, LFIN, LOXL4, FDX1, CYP2D6, UGT2B17, SLC22A1, and CYP3A43. Conclusions: While a GWAS of serum metabolite data, we could identify eight new genetic loci being related to metabolic traits, and replicate a large number of previously reported loci. Our findings will contribute towards an improved understanding of human metabolism.
1549M
A β-cell specific protein subnetwork significantly enriched for association with GLP-1 stimulated insulin secretion: A DIRECT study. V. Gudmundsdottir1, H.K. Pedersen1, L.M. Hart2, K. Banasik3, A. Deforce3, E. de Geus5, M. Eekhoff6, M. Diamant7, M. McCarthy7, E. Pearson4, C. Workman1, R. Gupta1, S. Brunak1. 1) Technical University of Denmark, Lyngby, Denmark; 2) Leiden University Medical Center, Leiden, The Netherlands; 3) University of Oxford, Oxford, UK; 4) University of Copenhagen, Copenhagen, Denmark; 5) VU University Amsterdam, The Netherlands; 6) VU University Medical Center, Amsterdam, The Netherlands; 7) Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford, UK; 8) University of Dundee, Dundee, UK.

The heritability of insulin secretion response (ISR) during a modified hyperglycemic clamp with glucose, glucagon-like peptide 1 (GLP-1) and arginine stimulation has been shown to be considerable (h²=0.52 to 0.80) and it is of great interest to identify genetic variants influencing these responses. Evaluating ISR with a modified hyperglycemic clamp is not feasible in larger cohorts of sizes preferred for genome-wide association studies (GWAS). When sample sizes are small, GWAS may be complemented with systems biology approaches to aid the prioritization of genetic variants. Within the DIRECT consortium a GWAS was performed on GLP-1 stimulated ISR and data integration used to add biological context to the results and facilitate variant prioritization. GLP-1 stimulated ISR was measured with a modified hyperglycemic clamp in 130 twins and sibs from the Netherlands twin registry. The cohort was genotyped using the Illumina HumanCore Exome BeadChip and association analysis was performed using the QTassoc software and adjusted for age, sex, familial relationships and insulin sensitivity index. Gene-based P-values were mapped onto a β-cell specific protein-protein interaction (PPI) network, which was created by pruning high confidence PPIs from InWeb 3.0 using published β-cell RNAseq data. Connected components in the network enriched for high scoring genes were identified with the Cytoscape plugin [ActiveModules] and their significance evaluated by comparing to 10,000 degree-preserved randomly sampled subnetworks from the β-cell PPI. None of the variants tested in the GWAS reached a genome-wide significance of P ≤ 5.0E-8. However, the top scoring subnetwork (25 genes) had a significantly higher combined z-score than expected by random chance (z = 4.60, P = 3.78E-5). The top-scoring subnetwork was most strongly enriched for the Gene Ontology terms "cell junction" (P = 3.6E-4), "plasma membrane part" (P = 2.2E-4) and "cell projection" (P = 4.9E-3) and contained a number of genes known to affect β-cell mass and function (FOXO1), insulin secretion (WFS1, RYR2) and "cell projection" (P = 4.9E-3) and contained a number of genes known to affect β-cell mass and function (FOXO1), insulin secretion (WFS1, RYR2) and possibly damaging (missense, nonsense, frameshifting indels, canonical splice site). These findings were then experimentally validated using high-throughput methods.

1550T
Analysis of Whole Exome Datasets to Test the Hypothesis of Digenic Inheritance in Stargardt Disease. K. Lee1, D.S. Marchuk1, M.J. Prieur2, C. Bizon1, D. Young3, D. Gilis3, P. Owen4, K. Wilhelmson5, K.E. Weck1,2, S. Garg3, J.P. Evans1, J.S. Berg1. 1) Dept Genetics, Univ North Carolina, Chapel Hill, NC; 2) Greenwood Genetic Center, Greenwood, SC; 3) Renaissance Computing Institute, Chapel Hill, NC; 4) Dept Pathology and Laboratory Medicine, Univ North Carolina, Chapel Hill, NC; 5) Dept Ophthalmology, Univ North Carolina, Chapel Hill, NC.

Whole exome sequencing (WES) is a valuable tool in determining the molecular etiology for retinal disorders given their genetic heterogeneity. We performed WES in 45 patients with various retinal disorders. Two patients with simplex Stargardt disease (SD; [MIM 248200]) harbored one pathogenic variant in ABCA4 and one pathogenic variant in a second retinal disease gene (CNGB3 or GUCA1A) suggesting the possibility of digenic inheritance. Digenic inheritance has been reported in other retinal disorders, but not in SD. Approximately one-third of simplex cases have only one identifiable mutation in ABCA4; digenic inheritance may explain a subset of these cases and would be supported by finding an excess of mutations in other retinal disease genes in such patients. We evaluated the frequency of double carrier status for ABCA4 and another retinal disease gene in the general population by searching for carriers of ABCA4 pathogenic variants in WES datasets not enriched for retinal disorders (1,189 individuals). ABCA4 variants previously reported as pathogenic, or rare (<1% minor allele frequency) and possibly damaging (missense, nonsense, frameshifting indels, canonical splice site), were selected for further analysis. Pathogenicity was assessed by allele frequency, conservation data, in silico modeling and co-segregation data. For the 40 (3%) ABCA4 carriers identified, we analyzed 448 variants within 213 genes associated with retinal disorders. No individual had two pathogenic ABCA4 variants. One of 8 variants of possible pathogenicity was identified in 9 ABCA4 carriers within PDE6A, EYS, ALMS1, USH2A, BBS9, BBS10, ABC6 or NR2E3, two individuals harbored the same variant previously reported as a pathogenic mutation in the NR2E3 gene. However, no carrier had a known pathogenic variant in other genes associated with SD, including PROM1, PRPH2, ELOVL4 and CNGB3, or GUCA1A. The finding of 2 heterozygous variants associated with SD among ABCA4 carriers is compatible with a digenic inheritance for SD. WES datasets are an effective method to evaluate heterogeneous disorders for the presence of double carrier status as a means to test the hypothesis of digenic inheritance.

1551S
Efficiency of exome sequencing for the molecular diagnosis of Pseudo-xanthoma Elasticum. M.J. Hosen1, F. Van Nieuwerburgh2, D. Deforce2, L. Martin3, G. Letteris4, A. De Paepe1, P.J. Coucke1, O.M. Vanacker1. 1) Dept Genetics, Ghent University Hospital, Ghent, Belgium; 2) Department of Pharmaceutics, Laboratory of Pharmaceutical Biotechnology, Ghent University, Ghent, Belgium; 3) Department of Dermatology, Angers University Hospital, Angers, France; 4) Department of Vascular Biology and Sports Medicine, Angers University, Angers, France.

The molecular etiology of pseudo-xanthoma elasticum (PXE), an autosomal recessive connective tissue disorder, has become increasingly complex as not only mutations in ABCC6 but also ENPP1 and GCDC can cause resembling phenotypes. Identification of modifier genes, such as VEGFA, has further contributed to the molecular heterogeneity of PXE. In such heterogeneous diseases, Next Generation Sequencing allows to perform mutation screening of several genes in a single reaction. We explored whole exome sequencing (WES) as an efficient diagnostic tool to identify the causal mutations in ABCC6, GCDC, ENPP1 and VKORC1 in 16 PXE patients. WES identified the causal mutation in 27 of 32 alleles (ABC6 or GCDC, with no causal mutations in ENPP1 or VKORC1). Exomes with insufficient reads (<20 depth) and patients with no or single mutations were further evaluated by Sanger sequencing (SS) and MLPA, but no additional mutations were found. The potential of WES is to explore multiple genes at a time, the ease and reliability of the approach, but the opportunity to search for novel genes when targeted analysis is negative. Together with low cost, rapid and less laborious workflow, we conclude that WES complemented with SS can provide a tiered approach to molecular diagnostics of PXE.
1552M

Unraveling Genetic Architectures Spanning Mendelian and Complex Phenotypes with Data Driven Electronic Medical Record Validation. B.S. Glicksberg1, L. Li1, R.Z. Castellanos1, J. Hakenberg1, W. Cheng2, S. Khader1, M. Ma1, L. Shi1, H. Shah1, J.T. Dudley1,2, R. Chen1. 1) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY; 3) Department of Policy and Evaluation, Icahn School of Medicine at Mount Sinai, New York, NY.

While Mendelian and complex phenotypes are traditionally organized into distinct categories, emerging studies suggest there may be pervasive shared genetic connections within and across these categories. To further explore this intriguing relationship, we have compiled and analyzed a comprehensive, integrated knowledgebase comprised of known published mutation-gene-phenotype relationship representing over 2,500 Mendelian phenotypes with around 4,000 unique variant-associated genes collected from the Online Mendelian Inheritance in Man (OMIM) and Human Gene Mutation Database (HGMD) online sources. Additionally, we compiled genetic disease associations for >700 complex phenotypes and >3,500 variant-associated genes from the open Genome- Wide Association Studies (GWAS) Catalog as well as a proprietary database ActiVar which provided additional GWAS findings from manual curation of ~13,000 publications. We performed a statistical analysis based on the Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine similarity method to test for enrichment between Mendelian and complex diseases based on shared genetic associations. This analysis identified 1,116 significantly enriched Mendelian-complex disease pairs at q-value < 0.05 and 609 enriched pairs at q-value < 0.01. We further evaluate these novel Mendelian-complex disease pair enrichments by analyzing disease-morbidizing SNPs across Electronic Medical Records (EMR) for 3.5 million unique patients comprising more than 2 million clinical encounters in the Mount Sinai Electronic Medical Record Warehouse (MSDW). Of the genetically enriched pairs (q<0.01), we found 23 that were also significantly enriched in the MSDW using a Bonferroni corrected Fisher’s Exact test (p<0.0001). There were 25 Mendelian phenotypes hubs and 44 complex disease hubs that had 10 or more significantly enriched connections. These results from our analyses have been incorporated into a comprehensive phenotype browser in the linkage analysis software, aiming to accelerate the discovery of Mendelian and complex disease genes.

1553T

Integration of GWAS signals, measures of polymorphic structure and linkage disequilibrium to discover clinically relevant biomarkers and improve identification of causal variants. M.W. Lutz1, R. Sauth2, O. Chiba-Falek1, D.K. Burns1, A.M. Saunders1, A.D. Roses1,2,1. 1) Department of Neurology, Duke University School of Medicine, Durham, NC; 2) Polymeric DNA Technologies, Alameda, CA; 3) Zinfandel Pharmaceuticals, Inc., Durham, NC.

Variant in non-coding RNA, epigenetic regulators, introns, promoters, and distal regulatory elements are associated with a variety of complex phenotypes including longevity and neurodegenerative diseases. We have developed a computational pipeline that integrates published GWAS findings from manual curation of ~13,000 publications. We performed a statistical analysis based on the Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine similarity method to test for enrichment between Mendelian and complex diseases based on shared genetic associations. This analysis identified 23 that were also significantly enriched in the MSDW using a Bonferroni corrected Fisher’s Exact test (p<0.0001). There were 25 Mendelian phenotypes hubs and 44 complex disease hubs that had 10 or more significantly enriched connections. These results from our analyses have been incorporated into a comprehensive phenotype browser in the linkage analysis software, aiming to accelerate the discovery of Mendelian and complex disease genes.

1554S

The Utah Genome Project: Successfully discovering and diagnosing genetic disease using VAAST, pVAAST and Phevor. M.V. Singleton1, B. Kennedy1, Z. Kronenberg1, M.G. Reese2, M. Yandell2,1. 1) Department of Human Genetics, Eccles Institute of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT 84112, USA; 2) Utah Science, Technology, and Research Center for Genetic Discovery, University of Utah, Salt Lake City, UT 84112, USA; 3) Omicia Inc., 1625 Clay Street, Oakland, CA 94612, USA.

The Utah Genome Project (UGP) is a large-scale, intramural genome-sequencing endeavor, the aim of which is to discover new disease-causing genes and diagnose inherited diseases. What sets the UGP apart from similar projects is its variant interpretation and diagnostic pipeline that includes innovative bioinformatics tools, including VAAST, pVAAST and Phevor. VAAST and pVAAST are probabilistic disease-gene finders capable of prioritizing every variant and gene in the human genome. Single affected and case control cohorts are analyzed using VAAST while pVAAST adds linkage information for additional power in large pedigrees. Successful molecular diagnosis using an exome or genome sequence hinges on accurate association of damaging variants to the patient’s phenotype. Unfortunately, many clinical scenarios (e.g. single affected or small nuclear families) have little power to confidently identify damaging alleles using sequence data alone. Today’s variant interpretation tools are simply underpowered for accurate diagnosis in these situations, limiting successful diagnoses. Phevor alleviates this lack of power by dynamically incorporating the patient’s phenotype into the disease-gene search. Phevor works by expanding known phenotype associations using knowledge resident in biomedical ontologies, like the Human Phenotype and Gene Ontologies, and combining these results with the VAAST/pVAAST outputs. Phevor accurately re prioritizes candidates identified by VAAST/pVAAST in light of knowledge found in the ontologies. Phevor is not limited to a fixed set of genes or phenotypes. This enables Phevor to discover disease-gene associations for established and previously un-described or atypical phenotypes. The UGP pipeline, including Phevor, has been successfully applied to several clinical cases. The UGP provides successful diagnosis of genetic disease for those with known and novel disease-alleles, novel disease-phenotype associations and family specific mutations discovered in large pedigrees. Presented are several clinical cases where the UGP provided a molecular diagnosis where the common technique (filtering and querying disease databases) failed. Free from the limitations imposed by disease databases, the UGP analysis pipeline promises to improve diagnostic rates for inherited disease.
1555M
GWAS analysis of epigenetic age acceleration. A. Lu1, R. Ophoff1,2,3,4, K. Eijk1, J. Bell1, I. Erte1, P. Tsai1, T. Spector1, K. Hao2, S. Dracheva1, S. Horvath1,2, 1 Human Genetics, UCLA, Los Angeles, CA; 2 Center for Neurobehavioral Genetics, Department of Psychiatry, UCLA, Los Angeles, CA; 4 Rudolf Magnus Institute of Neuroscience, Department of Genetics, University Medical Center Utrecht, Utrecht, The Netherlands; 5 Department of Twin Research and Genetic Epidemiology, King’s College London, United Kingdom; 6 Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine; 7 Department of Psychiatry, Mount Sinai School of Medicine.

Recent articles describe methods for estimating the age of a tissue sample based on DNA methylation levels. By comparing estimated age, referred to as DNA age, with chronological age, one can arrive at a measure of age acceleration. Age acceleration is highly heritable (estimated to be around 40%). Age acceleration is a quantitative measure that can be used in QTL analysis to identify SNPs. Here we report the findings of a large meta-analysis involving over a dozen data sets comprised of both DNA methylation and SNP data. Imputation was performed for each study using 1000 Genome reference panels newly released in December 2013. This yields ~7 M reliable markers with common variants, minor allele frequency (MAF) >0.05 in at least one study, available for association analysis. A genome-wide QTL analysis was performed in each dataset. We used a meta-analysis to aggregate the results from the different data sets identifying 34 SNPs (1 called and 33 imputed) in 8 genes. The imputed markers passed stringent quality controls (genotype information measure > 0.74 and mean MAF = 0.16). When we restricted the analysis to participants who are younger than 70, we found 5 more SNPs in two additional genes. Overall, our results demonstrate that epigenetic age acceleration is under strong genetic control.

1556T
Differentially Expressed Genes in Asthma Differ by Tissue-type. T.B. Mersha1, E. Geirr1, D. Ghosri2, J.A. Bernstein3, G.K.K. Hershey1. 1) Cincinnati Children’s Hospital Medical Center, Cincinnati, OH; 2) Department of Internal Medicine, University of Cincinnati, OH.

Several studies have identified differentially expressed genes (DEGs) between healthy versus asthmatic patients using microarray technology. This approach, however, has led to considerable variation in the lists of genes generated by various groups due to variation from random noise to biological samples. There are two objectives of this study a) conduct a comprehensive cell-type specific expression profiling analysis in order to identify the target genes, pathways and networks that are enriched in each cell or tissue-type, b) develop cell and tissue-type specific predictions models. Based on cell/tissue origin, a total of 400 samples and 22,590 DEGs were divided into bronchial epithelia (BEC), nasal epithelia (NEC), bronchial fibroblasts (FB), and lung fibroblasts (FIB). Multivariate analysis of heatmaps of DEGs from similar cells shows a clear separation between asthmatics and controls. Moreover, DEGs from various tissue-specific datasets share more commonality at the pathway level than at the gene level as reflected by the binary similarity index of 0.65 between BEC and NEC. These results also highlight the importance of a cell-type based approach when analyzing multiple gene expression data. Our study identified novel asthma-related genes and pathways as future therapeutic targets and prediction models.

1557S
Smoking-related miRNAs and mRNAs in human peripheral blood mononuclear cells. M.W. Su1,2, K.Y. Tung1, S.L. Yu1, C.H. Tsai1, Y.F. Wu1, Y.J. Lai2, C.H. Liu3, Y.L. Lee3,1, 1) Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan; 2) Department of Clinical Laboratory Sciences and Medical Biotechnology, National Taiwan University College of Medicine, Taipei 100, Taiwan; 3) Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

Teenager smoking is of great importance in public health which may increase the incidence of asthma and allergy. Tobacco smoke would be sensed by macrophages as an injurious stimulus and induce the expression change of genes associated with immune, and inflammatory. However, the regulatory mechanisms and the benefits of smoking reduction in alteration of immunotoxicological system remained poorly understood. In 2011, we initiated a community-based trial aiming to evaluate the effectiveness of smoking reduction in teenagers, to examine the causal relationship between smoking reduction and alterations in immunotoxicological system. A total of 100 teenagers with active smoking habits were recruited and 12 teenagers who reported a substantial reduction in smoke quantity and an increase in cotinine/creatinine ratio were included in genomic analyses. We examined the whole genome mRNA expression in peripheral blood mononuclear cells (PBMC). A total of 292 genes with false discovery rate adjusted p-value < 0.01 were deemed to be differentially expressed in TSRT. The Gene Ontology biological process annotations based on Gene Set Enrichment Analysis was conducted on the 292 differentially expressed genes. The smoking reduction induced expression changes mainly involved in the signal transduction (31 genes), RNA binding, and NLRP3 inflammasome. Smoking induced the expression of immune genes in PBMC of TSRT. The Gene Ontology biological process annotations based on Gene Set Enrichment Analysis was conducted on the 292 differentially expressed genes. The smoking reduction induced expression changes mainly involved in the signal transduction (31 genes), RNA binding, and NLRP3 inflammasome. Smoking induced the expression of immune genes in PBMC of TSRT. A genome-wide QTL analysis was performed in each dataset. We used a meta-analysis to aggregate the results from the different data sets identifying 34 SNPs (1 called and 33 imputed) in 8 genes. The imputed markers passed stringent quality controls (genotype information measure > 0.74 and mean MAF = 0.16). When we restricted the analysis to participants who are younger than 70, we found 5 more SNPs in two additional genes. Overall, our results demonstrate that epigenetic age acceleration is under strong genetic control.

1558M
The mitochondrial mutational landscape of human cancer and its impact on tissue- and tumor-specific gene expression. S. Grandhi1, N. Dhami1, C. Bosworth1, T. Latramboise. Department of Genetics and Genome Science, The University of Cincinnati, Cincinnati, OH.

Unlike that of the nuclear genome, the regulatory landscape of the mitochondrial genome remains largely uncharacterized, particularly in the context of cancer. In contrast to normal cells, which rely primarily on mitochondrial oxidative phosphorylation for energy production, cancer cells instead rely heavily on aerobic glycolysis, even in the presence of oxygen. Although there is an accumulation of somatic mitochondrial DNA (mtDNA) mutations in cancer, the impact of these mutations on the role of mitochondria in reprogramming cancer cell bioenergetics is unclear. Even in the context of normal cells, mitochondria are highly sensitive to their cellular environment: heart, muscle, kidney, and liver cells have higher mitochondrial gene expression levels than lung, adrenal, and thyroid cells, likely in response to tissue-specific energy demands. Here we investigate the role of the mitochondrial genome in tumor cell bioenergetics while addressing tissue specificity. To achieve this, we mine publicly-available next-generation sequencing (NGS) data across 15 cancer types from The Cancer Genome Atlas (TCGA) and the University of California, Los Angeles, Center for Compartmentalized Lysosomes (CCLG), and the Encyclopedia of DNA Elements (ENCODE). Using these datasets, we survey mitochondrial DNA mutational patterns, mRNA expression levels, and mtDNA transcription factor occupancy. Our analysis of TCGA RNA-seq data from approximately 1,800 cancer patients suggests that the mitochondrial-encoded gene expression is significantly dysregulated in cancer cells and also clusters distinctly based on the tumor’s tissue of origin. Additionally, we observe that patterns of differential gene expression in tumor as compared to matched normal samples vary among tumor types. To interrogate potential mechanisms underlying these gene expression patterns, we integrate them with mRNA copy number, mtDNA mutations, and mtDNA-binding transcription factors, while addressing the overall somatic mutation frequency in mtDNA across cancer types.

This pan-cancer query of the mitochondrial regulatory landscape distinguishes tumor- and tissue-specific mitochondrial aberrations on the DNA and RNA levels by integrating multiple NGS platforms from disparate sources and analyzing them in tandem.

Large intergenic non-coding RNAs (lincRNAs) are emerging as key regulators of diverse cellular processes. Recent advances in RNA sequencing (RNA-Seq) and computational methods allow for a comprehensive analysis of lincRNAs. There have been no comprehensive studies involving the presence and role of lincRNA in the eye tissues and eye related pathologies till date. To build a comprehensive catalogue of novel lincRNAs expressed in eye, we performed RNA-Seq on eight human post-mortem eyes and generated a comprehensive set of data for both region-based (macular versus peripheral retina) and tissue-based (retina versus RPE/choroid) studies. We have processed RNA-seq data and identified ~2000 potential novel lincRNAs per sample which are not present in the catalog of the Human Body Map lincRNAs. Briefly, we first assembled the transcripts using Cufflinks and Scripture. To differentiate true lincRNA transcripts from assembly artifacts, we used the software SebNif, which implements a series of filtering steps such as filters of transcript length, expression level and coding potential. Each lincRNA has been annotated with genomic features around its promoter and gene body, such as 36 of histone H3 (H3K4me3, and H3K36me3), Expressed Sequence Tag (EST) and Cap Analysis of Gene Expression (CAGE) tags. We also looked for colocalization and coexpression of the novel lincRNAs with known AMD GWAS hits to identify possible links between lincRNA and AMD genes.

15562 Gene variant modification in keratinocyte cell samples irradiated by UV using RNA-seq, V. Mijatovic1, G. Patuzzo1, C. Bettencourt1,2, I. Reyneke, M. Ombrello, R. Goldbach-Mansky.

Intramural Research Program, NIAMS, National Institute of Health, Bethesda, MD.


Adult-onset Still’s disease (AOSD) is a severe, inflammatory disease that is characterized by high, spiking fever, arthritis, a characteristic skin rash and increased inflammatory markers. The molecular mechanism underlying AOSD is unknown, but given its apparent lack of heritability, we hypothesize that somatic mutations or mosaicism may contribute to its pathogenesis. To investigate whether somatic mutations play a role in AOSD, we have sequenced matched tissues from whole body (including bone marrow) from 5 patients. Using fibroblasts as the control tissue to determine germline variation, we test the hypothesis that monocytes or another cell population in whole blood might harbor somatic mutations that contribute to the disease. Whole exome sequencing was performed on germline and somatic derived DNA from fibroblasts (Otogenetics) using Agilent SureSelect V5 (51Mb) capture and Illumina HiSeq instruments. The average coverage of each sample is about 60X with greater than 95% of the exome having at least 10X coverage. Four different comparison methods have been utilized for detection of somatic mutations: WAI-PICARD-GATK, MuTect, VarScan2 and Diamand. While the first three tools require BAM files generated by aligning sequence reads to the reference genome, Diamand directly compares the reads and thus doesn’t require the reference genome for variant detection. The number of potential somatic mutations called by each method ranges from fewer than 100 to over 1,000 per matched pair. Among the methods used, Diamand reports the most candidate somatic mutations, while MuTect identifies the least. The results generally don’t overlap between the different methods and are thus unlikely to be due to low quality sequencing, suggesting that the majority of the somatic calls are likely to be false positives. This observation is consistent with the 0.5% to 5% discordant rates observed between technical replicates and the assumed low mutation rates in the affected tissues. The few mutations called by the different methods are being evaluated and could provide important clues on the molecular etiology of AOSD.

1560S Weighted gene co-expression network analysis suggests white matter might play a role in epilepsy and episodic motor disorders. L. Silveira-Morinaya1,2,3,4, M. Ryten1,2,4, P. Farabaugh2, C. Bettencourt1,2,4, A. J. Lees4,5, H. Houlden6,7, J. Hardy1,8,4, T. Warner1,2,4,1, UK Brain Expression Consortium. 1) Reta Liwa Weston Institute, UCL Institute of Neurology, London, London, United Kingdom; 2) UCL Institute of Neurology, UCL For1 Brain for Neurological Studies, London, London, UK; 3) Neurology Department, University of Campinas, Campinas, Brazil; 4) Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK; 5) Department of Medical and Molecular Genetics, King’s College London, London, UK; 6) Istituto di Ricerca Genetica e Biomedica CNR, Cagliari, Italy.

PURPOSE: There is a significant phenotypic overlap of genes causing epilepsy and episodic motor phenomena (such as episodic dyskinesia, episodic ataxia and episodic encephalaplegia). We used gene-wise expression (GWE) data to study co-expression of these genes. METHODS: GWE data generated from 788 samples of 101 adult control human brains (10 distinct brain regions each) was collected as part of the UK Human Brain Expression (HBBE) project. Whole-exome sequencing analysis was used to group all expressed genes into modules in an unsupervised manner. Target genes were: “epileptic” genes (ATP1A2, SCN1A, SLC1A3, ATP1A3, KCNMA1, CACNB4, CACNA1A, KCNA1, GCH1, SLC2A1, PNKD and PRRT2) and “epilepsy” genes (PRRT2, ARX, CDK5R2, CHRNA2, CHRNB2, GABRG2, GABRA1, KCNQ2, KCNQ3, LG1, PCDH19, SCN1A, SCN1B, SCN2A, STXBPI, NDE1, TBC1D24, GOSR2, AFG3L2, GRN, KCNT1, CHRNA4, PLCB1, and SLC2A1). The overrepresentation of each group in modules was assessed using chi-squared tests with Yates correction. Gene ontology (GO) and KEGG pathway enrichment analysis were performed on variants called by at least 100 to over 1,000 per matched pair. Among the methods used, Diamand reports the most candidate somatic mutations, while MuTect identifies the least. The results generally don’t overlap between the different methods and are thus unlikely to be due to low quality sequencing, suggesting that the majority of the somatic calls are likely to be false positives. This observation is consistent with the 0.5% to 5% discordant rates observed between technical replicates and the assumed low mutation rates in the affected tissues. The few mutations called by the different methods are being evaluated and could provide important clues on the molecular etiology of AOSD.


Large intergenic non-coding RNAs (lincRNAs) are emerging as key regulators of diverse cellular processes. Recent advances in RNA sequencing (RNA-Seq) and computational methods allow for a comprehensive analysis of lincRNAs. There have been no comprehensive studies involving the presence and role of lincRNA in the eye tissues and eye related pathologies till date. To build a comprehensive catalogue of novel lincRNAs expressed in eye, we performed RNA-Seq on eight human post-mortem eyes and generated a comprehensive set of data for both region-based (macular versus peripheral retina) and tissue-based (retina versus RPE/choroid) studies. We have processed RNA-seq data and identified ~2000 potential novel lincRNAs per sample which are not present in the catalog of the Human Body Map lincRNAs. Briefly, we first assembled the transcripts using Cufflinks and Scripture. To differentiate true lincRNA transcripts from assembly artifacts, we used the software SebNif, which implements a series of filtering steps such as filters of transcript length, expression level and coding potential. Each lincRNA has been annotated with genomic features around its promoter and gene body, such as 36 of histone H3 (H3K4me3, and H3K36me3), Expressed Sequence Tag (EST) and Cap Analysis of Gene Expression (CAGE) tags. We also looked for colocalization and coexpression of the novel lincRNAs with known AMD GWAS hits to identify possible links between lincRNA and AMD genes.
Transcriptional regulatory networks (TRNs) that represent regulatory relationships from transcriptional factors (TFs) to target genes are important resources to study genetic regulatory programs for complex phenotypes in higher organisms including human. While many human TRNs inferred from various large-scale experiments are publicly available, only few reference databases for human TRNs have been published. Here we present an extensive database of human reference transcriptional regulatory network, TRUSTR (Transcriptional Regulatory Relationships Unraveled by Sentence-based Text-mining), which contains 5,460 regulatory relationships between 646 TFs and 1,823 target genes identified by manual curation of 18,446 relevant sentences extracted from more than 15,000 abstracts. To collect the relevant sentences, which might contain information about human transcriptional regulation, we scan more than 20 million PubMed abstracts for a set of key words implying human transcriptional regulation. TRUSTR is the largest reference database for TF-target links based on literature curation to date to our best knowledge. TRUSTR provide disease & biological process ontology terms that are enriched with target gene sets of TFs. Also we constructed a network of TF cooperativity, which might be useful to study co-regulatory complex. TRUSTR will provide a reference database of TF-target relationships for validation of large-scale mapping of TF-target links and learning new TF-target links from biological data mining. TRUSTR data is freely available at http://www.inetbio.org/trtrust.

MokaSeq: Initial validation of the sequence analysis module of an NGS software platform for clinical diagnostics. J.W. Ahn1, N. Chandler2, N. Parkin3, M.A. Simpson2, S.C. Yau2, C. Mackie Ogilvie2, K.J.P. Ryan2, 1) Genetics, Guy’s and St Thomas’ NHS Foundation Trust, London, United Kingdom; 2) Genetics, Viapath, London SE1 9RT, United Kingdom; 3) Division of Genetics and Molecular Medicine, Kings College London, London SE1 9RT, United Kingdom.

Guy’s and St Thomas’ NHS Foundation Trust and Viapath represent a centre of excellence in clinical genetics, servicing the South East Thames region population of ~6 million. We currently offer NGS-based testing for a number of heterogeneous genetic disorders. For each of these tests, DNA libraries are prepared targeting specific genes using custom Agilent SureSelect kits and sequenced on the Illumina MiSeq and HiSeq systems. As part of the newly formed bioinformatics core within Viapath, we are developing a bespoke NGS software platform (Moka) to perform high quality analysis within a robust and scalable framework that can be adapted in line with diagnostic demands. The Moka platform will provide an end-to-end system from patient test request to diagnostic report, and will include components for LIMS, sequence analysis, a variant database, variant interpretation, and cross-referencing/auditing our resulting data. Moka’s sequence analysis module (MokaSeq) uses the Novoalign aligner alongside Picard Tools to carry out BAM file processing, and Samtools for variant calling of SNVs and small indels. Variants not relating to the regions of interest specified by our library panels are filtered out with Vcftools and the remaining variants are then annotated using Annovar, and a compendium of variant attribute repositories. CNV detection is performed using the R package ExomeDepth. Here we report findings from our initial validation of the sequence analysis module; this validation investigated the efficacy of detecting SNVs and CNVs previously identified by NGS and confirmed by Sanger sequencing/MLPA. To date, Moka has shown robust detection of all Sanger confirmed SNVs (n=114) at a coverage of 30x or greater. Similarly all previously confirmed CNVs have also been detected, albeit from high coverage data (~6000x). Down-sampling samples by a factor of 5, 10, 50 and 100 to simulate coverage depths of 1200x, 600x, 120x and 60x respectively did not affect the capacity to detect any of the CNVs. However, reduction in coverage to 120x and 60x did result in the detection of a number of spurious false positives.
Reducing INDEL errors in whole-genome and exome sequencing. H. Fang1,2, G. Narzisi1, J.A. O’Rawe1,2, Y. Wu1,2, M.C. Schatz1,2, J.G. Lyon1,2.
1) Stanley Institute for Cognitive Genomics, One Bungtown Road, Cold Spring Harbor Laboratory, NY, USA; 2) Stony Brook University, 100 Nicolls Rd, Stony Brook, NY, USA; 3) Simons Center for Quantitative Biology, One Bungtown Road, Cold Spring Harbor Laboratory, NY, USA, 11724.

INDELs, especially those disrupting protein-coding regions of the genome, have been associated with human diseases. However, there are still many errors with INDEL variant calling, with these errors being driven by library preparation, sequencing biases, and algorithm artifacts. We have recently developed and reported a new INDEL-calling algorithm, Scalpel, with substantially improved accuracy. Results: We have analyzed simulation, whole genome sequencing (WGS), whole exome sequencing (WES), and PCR-free sequencing data from the same samples to investigate the properties of false-positive and false-negative INDEL errors. Simulation data show that Scalpel can achieve sensitivity of 90% at 30X mean coverage, while maintaining a reasonable false discovery rate (0.06%) for large INDELs. We developed a classification scheme for INDELs utilizing extensive high-depth sequencing validation data, and we show that low-quality INDELs have ~2.7-fold higher error rates than high-quality INDELs (32% vs. 12%). The mean concordance of INDEL detection between our WGS and WES data is ~52%, while WGS data uniquely identifies ~10.8-fold more high-quality INDELs, and WES data uniquely identifies ~1.9-fold more low-quality INDELs. INDELs called within both datasets are designated as “high confidence INDELs”. Low-quality homopolymer A/T (Poly-A/T) INDELs are enriched in the WES-specific call set (82.6%), relative to the high confidence (44.2%) and WGS-specific call sets (45.1%). Concordance of INDEL detection between with-PCR and PCR-free data is ~71%, while PCR-free data uniquely yields ~4.4-fold more high-quality INDELs and with-PCR data uniquely yields ~6.3-fold more low-quality INDELs. We demonstrate that these types of INDEL errors are significantly reduced with a PCR-free library protocol, implying that these errors are introduced with PCR amplification. We have also calculated that ~95% of INDELs can be detected with 60X mean coverage WGS data using 2’100 bp reads from the Illumina HiSeq 2500 platform. Accurate detection of heterozygous INDELs naturally follows from this result.

INDELs: Conclusions: Homopolymer A/T INDELs are a major source of low quality INDELs and multiple signature events, and these are highly enriched in the WES data. We recommend WGS for human genomes at 60X mean coverage with PCR-free protocols, which can substantially improve the quality of personal genomes.

1568T
Single-nucleotide mosaicsisms in whole-genome sequences of clinically unremarkable individuals. Y. Huang1, X. Xu2, Y. Ye3, Q. Wu1, L. Yan1, B. Zhao1,4, X. Yang2, H. He3, S. Wang2, B. Gu2, H. Zhao1, M. Wang1, H. Gao1, G. Gao1, Z. Zhang2, X. Yang2, X. Wu2, Y. Zhang2, L. Wei1,3. 1) Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, People’s Republic of China; 2) Peking University First Hospital, Peking University, Beijing 100034, People’s Republic of China; 3) National Institute of Biological Sciences, Beijing 102206, People’s Republic of China; 4) Graduate School of Peking University Medical College, Beijing 100730, People’s Republic of China.

Postzygotic single-nucleotide mosaicsisms (SNMs) have been studied in cancer and a few other overgrowth human disorders at whole-genome scale and found to play critical roles. However, in clinically unremarkable individuals, SNMs have never been identified at whole-genome scale largely due to technical difficulties and lack of matched control tissue samples, and thus the genome-wide characteristics of SNMs remain unknown. We developed a new Bayesian mosaic genotyper and a series of error filters, using which we were able to identify 17 SNM sites from whole-genome sequencing of peripheral-blood DNAs from three clinically unremarkable adults. The SNMs were thoroughly validated using pyrosequencing, Sanger sequencing of individual cloned fragments, and multiplex ligation-dependent probe amplification. In addition to peripheral-blood samples, the identified SNMs were also detected in the majority of other samples non-invasively obtained from the same donors, implying the whole-body presence of SNMs.
1569S
A Next-Gen Sequencing Software Workflow for Gene Panel Validation Control. M. Keyser1, J. Carville1, T. Schwei1, T. Durfee PhD1, A. Pollack-Berti PhD1, D. Nash1, J. Steren1, S. Baldwin1, R. Nelson PhD1, K. Dullea1, J. Schroeder1, P. Pinkas PhD1, G. Plunkett III PhD1,2, F. Blatter PhD1,2,3,4
1) DNASTAR. Inc., Madison, WI; 2) University of Wisconsin, Department of Genetics, Madison, Wisconsin, USA; 3) Scarb Genomics LLC, Madison, Wisconsin, USA.
DNASTAR offers an integrated suite of software for assembling and analyzing sequence data from all major next-generation sequencing platforms. DNASTAR software supports a variety of key workflows on a desktop computer. A new Gene Panel Validation Control workflow supports several types of data sets, including Ion Torrent AmpliSeq™ Comprehensive Cancer Panel, Illumina TruSight Cancer Panel, as well as custom gene panels and evaluates the efficacy of gene panel targeting and the accuracy of variant calling. The accuracy of gene panel targeting is determined by multiple factors, including specificity of primers and probes used for gene panel design, efficiency of the sequencing technology, assembly of accuracy, SNP calling, and the SNP filters applied. DNASTAR’s SeqMan NGen and ArrayStar programs provide an accurate alignment algorithm and variant caller. They then utilize a validated SNP set in the form of a VCF file and a BED or Manifest file that specify targeted regions, to calculate SNP-calling sensitivity, specificity, and accuracy. By utilizing this workflow, users can ultimately validate their entire process to verify that their known variants are being identified.

1570M
Whole Genome Sequencing of 30 Admixed Brazilians. M. Machado1, R. Moreira1, A. Pereira2, M. Barreto3, B. Horta4, M.P. Lima-Costa4, A. Hori-moto5, N. Esteban5, M. Sclar5, G. Soares-Souza5, R. Zamudio5, M. San-tollalía1, C. Zöllni1, F. Kehdyl1, M.R. Rodrigues1, W. Magalhães1, E. Tapa-zona-Santos1, The Brazilian EPIGEN Consortium. 1) Universidade Federal de Minas Gerais, Belo Horizonte, Brazil; 2) Universidade Federal da Bahia, Brazil; 3) Universidade Federal de Pelotas, Brazil; 4) Fundação Oswaldo Cruz, Centro de Pesquisa René Rachou, MG, Brazil; 5) Instituto do Coração, Universidade de São Paulo, Brazil.
As part of the Epigen-Brazil Initiative on population genomics and genetic epidemiology, we resequenced the complete genome of 30 Brazilian admixed individuals randomly selected from each of three Brazilian population-based cohorts with different histories of admixture: Salvador from North East (n=10), Bambuí from the South East (n=10) and Pelotas from Southern Brazil (n=10). Salvador has predominant African ancestry (51%), 43% of European and 6% of Native American ancestry. Bambuí and Pelotas are predominantly European (>76%), with 16% of African and 7-8% of Native American ancestry. Sequences were obtained by paired-end reads strategy using Illumina technology. On average, each genome was sequenced 42.7X, with 128 Gbases that successfully passed filter and aligned to the reference HumanNCBI37_UCSC genome, 82% of bases that showed a quality QScore=30. 96% of Non-N reference bases with a coverage ≥10X. Agreement with illumina HumanOmni2.5 genotyping was 99.26% and with HumanOmni1 was 99.53%. We identified 15,033,510 SNVs and 1,479,746 of these are new (not present in dbSNP nor in 1000Genomes databases on May 2014). Importantly, 989 of these have a non-reference allele frequency >0.50 in our sample of 60 haploid genomes. Most of SNVs were classified as intergenic (58.03%) or intronic (34.88%) variants, the remaining variants were classified in other functional categories. Considering only exonic SNPs, similar proportions of non-synonymous and synonymous SNPs were identified: 49.91% and 47.88% respectively, a result that matches other studies. By comparing the allele frequency spectra obtained by WGS with those from the Illumina HumanOmni2.5 on the same cohorts, we observed that for this array, rare SNPs are under-represented in the population with higher African ancestry (Salvador). Thus even if this array was inspired in a multiethnic database (1000Genomes) and enriched for rare alleles, a population-based bias persists. Finally, we observed 9,143 deleterious variants (CONDEL score > 0.51) and we are now analyzing how admixture and the demographic history of these populations affected the additive and recessive genetic load in the Brazilian sample. Funding: Brazilian Ministry of Health/FINEP.

1571T
IntSplice: A tool to predict aberrant splicing of an SNV at intronic positions -50 to -3. K. Ohno, A. Shibata, T. Okuno, M.A. Rahman, Y. Azuma, A. Masuda. Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan.
Precise spatiotemporal regulation of splicing is mediated by splicing cis-elements on pre-mRNA. Single nucleotide variations (SNVs) affecting intronic cis-elements potentially compromise splicing, but no efficient tool has been available. First, we analyzed the effect size of each intronic nucleotide on alternative splicing events annotated in the ENSEMBL database. We found that nucleotides at intron -13 to -5 (Int-13:Int-5), Int-3, exon +1 (Ex+1), and Ex+2 were critical splicing determinants. We further extracted a total of 111 parameters that possibly dictated the strength of splicing signals. Second, we calculated percent-spike-in (PSI) at all 3’ splice sites of 14 RNA-seq data of normal human tissues and generated support vector regression (SVR) models with the 111 parameters. However, the correlation efficiencies between the calculated and predicted PSI’s were less than 0.3. Third, we generated support vector machine (SVM) models to directly differentiate pathogenic SNVs in the human gene mutation database (HGM) and normal SNVs in dbSNP using the 111 parameters as well as PSI’s predicted by the SVR models above. HGMD had 1,162 SNVs (mutations) at Int-50:Int-3. We randomly chose 1,162 SNVs among 63,605 SNVs in dbSNP. A total of 2,324 SNVs were randomly divided into five groups, and four were employed to generate a SVM model and the remaining one was used to validate the efficiency. We repeated the modeling and validation five times. The generated SVM models yielded sensitivities of 0.800 ± 0.041 (mean and SD) and specificities of 0.849 ± 0.021. We compared efficiencies of our models with PSSM and MaxEnt, and found that sensitivity as well as specificity of our models were better than those of PSSM and MaxEnt. We serially introduced “A” nucleotide at Int-11:Int-3 in RAPSN minigene, and introduced them in cultured cells. RT-PCR analysis revealed that three out of the nine mutants caused aberrant splicing. We found that these mutants lost binding to U2AF65. We tested these nine mutants with our IntSplice model and found that splicing consequences were properly predicted in eight mutants. We created a web service program, IntSplice (http://www.med.nagoya-u.ac.jp/neurogenetics/IntSplice) to predict sum of sensitivities and specificities of our models better than those of PSSM and MaxEnt. First, we analyzed the effect size of each intronic nucleotide on alternative splicing events annotated in the ENSEMBL database. We found that nucleotides at intron -13 to -5 (Int-13:Int-5), Int-3, exon +1 (Ex+1), and Ex+2 were critical splicing determinants. We further extracted a total of 111 parameters that possibly dictated the strength of splicing signals. Second, we calculated percent-spike-in (PSI) at all 3’ splice sites of 14 RNA-seq data of normal human tissues and generated support vector regression (SVR) models with the 111 parameters. However, the correlation efficiencies between the calculated and predicted PSI’s were less than 0.3. Third, we generated support vector machine (SVM) models to directly differentiate pathogenic SNVs in the human gene mutation database (HGM) and normal SNVs in dbSNP using the 111 parameters as well as PSI’s predicted by the SVR models above. HGMD had 1,162 SNVs (mutations) at Int-50:Int-3. We randomly chose 1,162 SNVs among 63,605 SNVs in dbSNP. A total of 2,324 SNVs were randomly divided into five groups, and four were employed to generate a SVM model and the remaining one was used to validate the efficiency. We repeated the modeling and validation five times. The generated SVM models yielded sensitivities of 0.800 ± 0.041 (mean and SD) and specificities of 0.849 ± 0.021. We compared efficiencies of our models with PSSM and MaxEnt, and found that sensitivity as well as specificity of our models were better than those of PSSM and MaxEnt. We serially introduced “A” nucleotide at Int-11:Int-3 in RAPSN minigene, and introduced them in cultured cells. RT-PCR analysis revealed that three out of the nine mutants caused aberrant splicing. We found that these mutants lost binding to U2AF65. We tested these nine mutants with our IntSplice model and found that splicing consequences were properly predicted in eight mutants. We created a web service program, IntSplice (http://www.med.nagoya-u.ac.jp/neurogenetics/IntSplice) to predict sum of sensitivities and specificities of our models better than those of PSSM and MaxEnt.
1572S  
Next generation sequencing approach to molecular diagnosis of autoimmune diseases: from gene panel design to variant call. M. Rusmini, F. Caroli, A. Grossi, I. Geretti, R. Ravazzolo, A. Martin, M. Gattorno, I. Ceccherini. 1 U.O.C. Genetica Medica; Istituto G Gaslini, Genova, Italy; 2 Dipartimento di Neuroscienze, Psicologia, Area del Farmaco e Salute del Bambino, Ospedale Pediatrico Meyer, Firenze, Italy; 3 Dipartimento di medicina genetica e ambientale e clinica, Università degli Studi di Firenze, Firenze, Italy; 4 Dipartimento di Neuroscienze, Oftalmologia, Genetica e Materno Infantile (DINOGMI), Università di Genova, Genova, Italy; 5 Pediatrica II, Reumatologia, Istituto G Gaslini, Genova, Italy.  
Auto-Inflammatory Disorders (AIDs) are a heterogeneous group of monogenic diseases caused by primary dysfunction of the innate immune system. Previous diagnosis by Sanger sequencing performed on a restricted number of genes has resulted time and cost consuming in addition to fail in detecting mutations in around 86% of patients referred to our unit. Clinical misdiagnosis, mutations in untested gene regions, genetic heterogeneity are possible explanations. We thought the Next Generation Sequencing approach could be the solution. By using the Ion Amplicon™ Designer (LifeTech) online tool, 203 amplicons were included in the screening panel, for a total of 121 exons and 22 Kb of target DNA. The design has been performed on 11 genes mainly involved on AIDs, as reported in the database for these disorders, the Nfesvers database. Ion PGM™ has been used for runs and the mean coverage has turned out to be 347X, with 92.5% of amplicons at > 20X and 79.5% at > 35X. The analysis from FastQ to VCFs was carried out using 3 different workflows: 1) Ion Torrent Alignment and Ion Reporter™ 4.0, specific for data generated by PGM II in-house pipeline based on open source tools like BWA and GATK, iii) CLC Bio software. We first focused on a set of 50 DNA samples already genotyped for the respective causative genes with the intention of comparing the three pipelines of analysis and assessing their sensitivity and specificity. Variants detected in genes which were not analyzed before have been validated by Sanger sequencing and differences in the variant calls from the three pipelines investigated. Through our analysis we found that the in-house pipeline has returned the most reliable results, missing only two of the 65 expected variants. However, this approach detected 34% of false positives, compared to the in-house pipeline. We then called variants using a new pipeline that detected 92% of the false positive variants. Moreover, only 2 of the 77 unexpected allelic variants detected with all the three pipelines and sequenced by Sanger have not been validated. Genetic diagnosis of patients affected by one of the Auto-Inflammatory Disorders (AIDs) can now be made by these NGS approaches. Moreover, the complete analysis of all the genes present in the panel will allow us to correlate genotypes with the vast range of phenotypes, in addition to the assessment of allele frequencies of the variants detected and possibly involved in the pathologies.

1574T  
A practical method to detect SNVs and indels from whole genome and exome sequencing data and an importance of in-house data for variant filtering. D. Shimemizu, A. Fujimoto, S. Akiyama, T. Abe, K. Nakano, K.A. Boroievich, Y. Yamamoto, M. Furuta, M. Kubo, H. Nakagawa, T. Tsunoda.  
In the recent development of massively parallel sequencing technology has allowed the creation of comprehensive catalogs of genetic variation. However, due to the relatively high sequencing error rate for short read sequence data, sophisticated analysis methods are required to obtain high-quality variant calls. Here, we developed a probabilistic multinomial method for the detection of single nucleotide variants (SNVs) as well as short insertions and deletions (indels) in whole genome sequencing (WGS) and whole exome sequencing (WES) data for single sample calling. Evaluation with DNA genotyping arrays revealed a concordance rate of 99.98% for WGS calls and 99.99% for WES calls. Sanger sequencing of the discordant calls determined the false positive and false negative rates for the WGS (0.0068% and 0.17%) and WES (0.0036% and 0.0084%) datasets. Short indels were also identified with high accuracy (WGS: 94.7%, WES: 97.3%). Currently we have applied this variant calling method to approximately 1,400 Japanese individuals and developed our own in-house database. Our own in-house database efficiently played a key role in variant filtering for exome sequence analysis. We believe our variant method and own in-house database can contribute to the greater understanding of human diseases.

1575S  
Group-based Variant Calling for a Large Cohort of Human Whole Genomes Leveraging Next-Generation Supercomputing. K. A. Standish, T. M. Carland, G. K. Lockwood, W. Pfeiffer, M. Tattner, C. C. Huang, S. Lamberth, S. Rajagopal, G. K. Lockwood, E. Jaeger, L. Smith, S. Szalma, G. Rajagopal, M. Curran, N. J. Schork. 1) Biomedical Sciences Graduate Program, University of California, San Diego; 2) Department of Human Biology, J. Craig Venter Institute, La Jolla, CA; 3) The Scripps Research Institute, La Jolla, CA; 4) San Diego Supercomputer Center, University of California, San Diego; 5) Janssen R&D, LLC.  
Recent advances in DNA sequencing technologies have increased the efficiency of sequencing, leading to an expanding deluge of high-quality data. However, DNA sequence reads have limited biological utility without relevant downstream processing and analysis, including read-quality assessment, alignment to a reference genome, variant identification, and individual genotyping. While the tools for performing these steps have improved, processing a whole genome from reads to variants remains an expensive and time-consuming aspect of sequencing studies. Furthermore, the accuracy of genotyping depends on accurate assessments of sequencing errors, which can be enhanced by processing large numbers of sequencing reads simultaneously. Here, we describe an efficient approach for obtaining high-quality variant calls from 438 whole genomes sequenced on an Illumina Hi-Seq 2500 platform. We exploited a group calling approach to minimize specific genotype assignment errors that arise from ignoring sequencing error rates and inconsistent coverage. To accommodate the enormous computational and storage requirements associated with the scale and approach of this study, we used SDSC’s Gordon supercomputer. It has 1,024 compute nodes, each with 16 cores, 64GB of DRAM, and 300GB of solid-state flash memory. To minimize cost, we optimized each step involved in genotyping to run as efficiently as possible on Gordon, parallelizing steps as appropriate during alignment and subsequent processing steps. To circumvent an I/O issue that arose while sorting the aligned reads, we used two “BigFlash” nodes with 4TB of SSD flash memory for temporary storage. Variants were then called with GATK’s HaplotypeCaller in groups of 20-24 genomes. We found evidence that the ancestral makeup of a group influenced the accuracy of variant calls so groups were comprised of individuals with similar ancestral background. This approach provided increased power for variant detection while avoiding the prohibitively long computing time associated with large group variant calling. Our results suggest that this approach yields high-quality variant calls and genotype assignments in an efficient manner and highlights the need for sophisticated computational strategies in analyzing large numbers of human genomes.

Bioinformatics and Genomic Technology

351
1576M
Repeat-Aware Hidden Markov Models for the Comprehensive Joint Calling of SNPs, Indels, and Short Tandem Repeats. A. Tan, H.M. Kang. University of Michigan, Ann Arbor, MI.

Accurate detection of genetic variants in Next Generation Sequencing data has always been complicated by the presence of low complexity regions of the genome. While many indels located in regions with high sequence diversity are easy to assay, the challenge is in the large fraction (~50%) of indels that are located in low complexity regions of the genome, mostly in the form of short tandem repeats (STRs), often with inexact repeat units. In the vicinity of indels in low complexity regions, it is also difficult to classify nearby base mismatches as a SNP or as part of an inexact repeat within an STR.

To be an effective short variant caller, the calling algorithm must account for the inherent heterogeneous nature of indels. We thus developed a variant calling algorithm based on a set of novel repeat-aware Hidden Markov Models (raHMMs) that locally aligns a pair of sequences allowing for mismatches or indels, with an arbitrary number of repeat units.

For variant discovery, our approach first identifies the repeat unit in a candidate variant allele; using this candidate repeat unit, we then perform pairwise alignments using a pair of raHMMs to determine the most appropriate flanking sequences, even in the presence of inexact STRs. This procedure allows us to systematically define variants by their flanks and earmarks regions that a SNP should not be called indiscriminately.

For variant genotyping, we apply another raHMM that explicitly models both flanking sequences and the repeat unit. This model allows us to determine if a read contains the defining flanks of an indel, counts the number of repeats observed and keeps track of alternative alleles that are not explicitly modeled.

This calling algorithm not only allows us to better classify the variant types in repeat rich regions but also genotypes STRs with inexact repeat units. We demonstrate that our method comprehensively detects and genotypes SNPs, indels and STRs using the deeply sequenced trio in the 1000 Genomes Project. This is implemented in vt and is available from http://genome.sph.umich.edu/wiki/Vt.

1577T

Breakpoints involved in translocation and chromothripsis are traditionally described using ISCN nomenclature based on chromosomal banding patterns (1). The sequence variation nomenclature guidelines of Human Genome Variation Society (HGVS, http://www.hgvs.org/mutnomen) traditionally focused on simple variants not requiring specific rules for detailed description of genetic rearrangements. This changed with the introduction of new technologies allowing rapid discovery of breakpoint sequences from complex structural rearrangements including translocations. The description of such complex variants challenges the existing guidelines. Previously, we suggested extensions for simple translocations (2). Here, we suggest extending the HGVS nomenclature guidelines to facilitate unambiguous description of more complex structural rearrangements including chromothripsis. A main requirement for the description is that precise chromosomal breakpoint sequences can be derived easily. The suggested format should provide sufficient flexibility and consistency limiting alternative interpretations and ambiguous descriptions. The new rules can be combined with those proposed previously for complex changes, which included: i) nesting to support description of changes within inversions and duplications, ii) composite changes to support concatenation of inserted sequences (3). We have applied the rules in practice by describing complex cases involving many breakpoints. The specifications should allow easy implementation in sequence variant nomenclature checkers (e.g. Mutalyzer, https://Mutalyzer. er.nl). We are planning to extend the functionality of Mutalyzer to incorporate the latest version of the HGVS sequence variation nomenclature guidelines as part of the development of curational tools for gene variant databases (Locus-Specific DataBases, LSDBs).


2) http://www.hgvs.org/mutnomen/STrans_HGVS2013_PT.pdf

1578S

Targeted gene panels, whole exome sequencing (WES), whole genome sequencing (WGS), and transcriptome sequencing (RNA-Seq), are revolutionizing cancer research. Many analyses, such as the comparison of tumor and normal cells, include the comparison of variant calls in the resulting VCF files. Here, we illustrate the benefits of leveraging read mappings in the comparison of WES and RNA-seq data using the CLC Cancer Research Workbench. WES and RNA-seq reads (paired-end, Illumina GA II, 100bp) from two patients (MM065 and MM089) are analyzed with built-in workflows in CLC Cancer Research Workbench. The sequences are available from SRA (bioproject 1823451). When filtering the results for tumor-specific variants inducing an amino acid change 42 and 43 variants from WES and 827 and 1,497 variants from RNA-seq are identified in MM065 and MM089, respectively. A function in the Cancer Research Workbench, “Compare Shared Variants within a Group of Samples” identifies variants from both WES and RNA-seq. Only two and five shared WES/RNA-seq variants, respectively, are identified for the two patients. “Identify Known Mutations from Sample Mappings”, searches RNA-seq read mappings for variants previously called from the exome data. This approach identifies 3 exome variants in the RNA-seq data from patient MM065 and 11 of the exome variants in the RNA-seq data of patient MM089. In contrast to comparing VCF files, the “Identify Known Mutations from Sample Mappings” tool provides a more reliable approach to recover previously detected variants in the WES data directly from the RNA-seq read mappings. Using this approach, 3 out of 42 variants for MM065 and 11 out of 43 variants for MM089 are confirmed by RNA-seq. 1Harbour J, et al., “Frequent mutation of BAP1 in metastasizing uveal melanomas”, Science, 2010 Nov 4;330(6009):1410-3.

1579M
Combining sets of indels with improved specificity and sensitivity using BAYSIC. D. Weaver1, B. Cantarel1,2, G. Benstead-Hume1, A. Mackey1, J. Reese1. 1) Genformatic, Austin, TX; 2) Baylor, Scott & White Baylor Institute of Immunology, Dallas, TX; 3) University of Virginia. Charlottesville, VA.

Insertions and deletions (“Indels”) comprise a source of human genome variation as significant as SNVs, and frequently cause clinically relevant phenotypes (Mulaney, et al., 2010). Moreover, the additional complexities of indel variation make their detection and accurate representation a more technically challenging problem than SNV detection. Many methods exist of the detection of indels, and as with SNVs, the concordance among sets of indels detected by different methods is poor. We previously described BAYSIC, a method of integrating variant calls using Bayesian statistics and latent class analysis to combine sets of SNVs with improved specificity and sensitivity. We extend this strategy here to combine sets of indels with increased sensitivity and specificity when compared to other currently popular methods for indel detection, e.g., GATK. BAYSIC provides the user with a posterior probability threshold that can be specified by the user to tune BAYSIC’s performance according to the user’s tolerance for false negative and false positive indel calls. Using genome data from the 1000 Genomes project, we evaluated the performance of BAYSIC and demonstrate BAYSIC’s enhanced accuracy when compared to other popular indel detection methods.
1580T
Clinical assessment of patients at Mass. Eye and Ear includes diagnostic testing for pathogenic genetic variants. We assess both single nucleotide changes that result in missense or frameshift alterations, and structural/copy number changes that result from deletions or duplications. Our diagnostic rate based on SNP analysis alone is ~ 50%. In order to increase this rate we have used microrna chip based genotype analysis on the Illumina iScan platform. However, we wish to reduce costs by evaluating our exome sequence data to detect CNVs. To do so we compared CNV detection applications based on the following criteria: a) paired-end versus read-depth analysis, b) ease of installation and use, c) inclusion of baseline or matched-control samples, d) use of exome versus targeted exome sequence data, and e) detection of positive control samples. We have collected a number of data sets that have been found to include CNVs which have been confirmed to be associated with specific traits. We used these data sets as positive controls for detection of CNVs. We have so far compared XHMM, BreakDancer, and WaveCNV, with many more comparisons planned. We were unable to use WaveCNV due to its difficult installation and lack of documentation. We have evaluated whole exome sequence data in two ways: 1) comparison of family members against a group of "baseline" samples, and 2) comparison of family members against each other without baseline samples. •XHMM identified a deletion in samples known to have that particular deletion. However, XHMM was unable to identify the duplication in samples that had a known small duplications. •We evaluated the same family samples with BreakDancer. Neither the known deletion, nor the known duplication were identified by BreakDancer. BreakDancer generates a large amount of SV results that we filtered to search for known CNVs. •Inclusion of baseline samples did not improve CNV detection.

1581S
Cloud-based variation analysis using SRA sequencing data directly. C. Xiao, E. Yaschenko, S. Sherry. National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 45 Center Drive, Bethesda, MD 20892.
Variation analysis plays an important role in elucidating the causes of various human diseases. The drastically reduced costs of genome sequencing driven by next generation sequence technologies now make it possible to analyze genetic variations with hundreds or thousands of samples simultaneously, but currently with the cost of ever increasing local storage requirements. We have evaluated whole exome sequence data in two ways: 1) comparison of family members against a group of "baseline" samples, and 2) comparison of family members against each other without baseline samples. •XHMM identified a deletion in samples known to have that particular deletion. However, XHMM was unable to identify the duplication in samples that had a known small duplications. •We evaluated the same family samples with BreakDancer. Neither the known deletion, nor the known duplication were identified by BreakDancer. BreakDancer generates a large amount of SV results that we filtered to search for known CNVs. •Inclusion of baseline samples did not improve CNV detection.

1582M
Low false-positive rate chromosomal structural variation detection procedure with statistical comparisons between case and control using paired-end reads. K. Yamashita1, A. Yamashita2, Y. Yokubu2, J. Takeda2, J. Sese1. 1) Ochanomizu University, Bunkyo-ku, Tokyo, Japan; 2) Osaka University, Suita-shi, Osaka, Japan.
Recent high-throughput sequencing technology has uncovered the high frequency of both somatic and germline chromosomal structural changes. To detect the changes, various software tools such as BreakDancer and GASV have been developed. However, most of those tools tend to yield a high rate of false-positives, which hampers the full understanding of the chromosomal changes in the cells. We here introduce a structural variation (SV) detection method with an extremely low false-positive rate by using the statistical comparison between case and control sequences. To detect SVs between two individuals, most existing tools need to perform the detection on each individual, and subsequently compare them with each other. However, this strategy may suffer from the problem of how to assess the differences in sequencing depth, quality and base biases between the samples. Our method first generates lists of candidate breakpoints using multiple sources of read information such as paired-end distances and the number of reads on each direction, and then makes statistical comparison of the candidate regions between samples. To evaluate our method's performance, we carried out SV analysis of mouse embryonic stem cell (ESC) lines by comparing gamma-ray irradiated (case) and unirradiated (control) cell populations. The frequency of irradiation-induced SVs was enhanced by increased genomic instability of the engineered ESCs in which the expression of the Blm helicase gene was transiently suppressed by doxycycline administration (Yamanishi et al, Genome Research 2013). We performed whole-exome sequencing on each of both lines and 55 paired-end reads of which turned out to be false-positives by referring to the mapping results. In contrast, our method detected 21 deletions, 22 inversions, and 125 translocations; GASV detected 12.5 deletions, 86 inversions, and 138 translocations, most of which are false-positives by referring to the mapping results. This result indicates that our method largely contributes to the reduction of false-positive detection of chromosomal SV in WGS data. In our presentation, we also show the result of SV analysis when our method is applied to human sequences in 1000 genomes database.

1583T
Fully phased allele-level sequencing of highly polymorphic HLA genes is greatly facilitated by SMRT sequencing technology due to its long read length and ability to sequence an amplicon representing a gene allele in isolation. In the present work, we evaluate DNA barcoding strategies for efficient pooling of multiple samples and loci for simultaneous sequencing of these genes. Multiplex sequencing using symmetric and asymmetric barcode-tailed HLA amplification primers were tested. Full-length HLA class I genes HLA-A, -B, -C were evaluated, using 8 different 16-bp barcode sequences in symmetric and asymmetric pairing. Eight barcodes generated 28 unique asymmetric pairings and allowed for simultaneous sequencing of 28 genome reference DNA samples; symmetric barcodes could be used for unique tagging of 8 genome reference DNA samples. Following amplification, the symmetric and asymmetric barcode-tagged libraries were pooled into two separate libraries for sequencing. The data was analyzed both separately and together to evaluate the ability of the barcode tags to uniquely de-convolute the samples. Consensus sequences were generated using LAA protocol in SMRT analysis 2.2. Allele sequences were typed using GenDx NGSengine HLA-typing software. We demonstrate the use of a DNA barcode strategy for multiplex full-length HLA gene sequencing to provide allele-level genotyping along with SNP phasing information.
1584S Detection of Copy Number Variations in Cancer Genomes from High-Throughput Sequencing Data. G. Klambauer, S. Hochreiter. Institute of Bioinformatics, Johannes Kepler University Linz, Linz, Upper Austria, Austria.

"Copy Number estimation by a Mixture Of PoissonS" (cn.MOPS), is a well established and widely used method for detection of germline copy number variations (CNVs) in high-throughput sequencing data. cn.MOPS showed excellent performance at the detection of CNVs in HapMap samples, as well as in genomes of bacteria, fungi and plants. Since cn.MOPS constructs a model across samples for each genomic position, it is not affected by read count variations along chromosomes, and, therefore, geared to targeted sequencing. In a comparative study, cn.MOPS was the best performing method at the detection of CNVs in targeted sequencing data. However, the detection of somatic CNVs in cancer genomes is still challenging due to admixture of normal and tumor tissue, nondiploidy and very large copy number variations that affect normalization. Therefore, preprocessing, normalization, and the core algorithm of cn.MOPS have been optimized for CNV detection in cancer genomes. We demonstrate the improved performance of the enhanced cn.MOPS algorithm for cancer genomes on whole genome sequencing data from the International Cancer Genome Consortium (ICGC). cn.MOPS has been optimized for computation time and parallelized, which makes the method perfectly suited to analyze data sets of hundreds of cancer samples within a few hours.

1585M Efficient variant pipeline for diagnosis of inherited cardiomyopathies associated genes using Ion Torrent PGM™ platform. L. Cerdeira1, T.G M. Oliveira1, A. Pereira1, M. Mitne-Neto1. 1) Research and Development, Fleury Group, São Paulo, SP, Brazil; 2) University of São Paulo - Heart Institute, São Paulo, SP, Brazil.

Hypertrophic cardiomyopathy (HC) is a primary cardiac disease characterized by hypertrophy of the left ventricle (LV) without dilation, usually asymmetrical and predominantly septal, in the absence of any other cardiac or systemic disease that can cause myocardial hypertrophy. Typically, CH is caused by mutations in genes encoding sarcomeric elements. Currently 19 genes have been discovered and linked to the CH spectrum, besides the filaments of the sarcomere, additional subgroups can be classified as related CH, as Z disc genes and calcium transport. Diagnosis is mainly clinical and usually only identified after the symptoms beginning. For that reason molecular genetic tests came up as a differential tool for the discovery of the mutations causing the phenotype. This study developed a bioinformatics pipeline for accurate molecular diagnosis of CH using Ion PGM data. The pipeline was developed using CLCBio Genomic Workbench 6.5 workflow and had as a first step a mapping assessment, with the 5 nucleotides at the 3'end trimmed and a Phred ≥ 20 used for quality control. Alignment against the human genome HG19 version was done using standard thresholds, followed by identification of variants by coverage and quality positioning. The identification of known variants was validated against the databases: dbSNP and clinicalvar and for further evaluation a prediction of splice site differences between tumor and normal samples.

1586T Assessing novel centromeric repeat sequence variation within individuals by long read sequencing. K.H. Miga1, J. Chin2, A. Bashir3,4. 1) Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA 95064, USA; 2) Pacific Biosciences, Inc, 1380 Willow Rd, Menlo Park, CA 94025, USA; 3) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA; 4) Icahn Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA.

Centromeres and other heterochromatic regions are commonly enriched with long arrays of near-identical tandem repeats, known as satellite DNA, that offer a limited number of variant sites to differentiate individual repeat copies across millions of bases. This substantial sequence homogeneity challenges available assembly strategies, and as a result, centromeric regions are vastly underrepresented in genomic studies. Further, as these sites are known to be variable among individuals in the population, it is necessary to not only characterize the sequence organization of these regions in a single genome, but to develop high-throughput methods to study this new source of human sequence variation among individual genomes. To advance characterization in these regions we have designed alpha-CENTAUROI (centromeric automated repeat identification for alpha satellite DNA) that takes advantage of Pacific Biosciences’ long reads from whole-genome sequencing. Long reads allow direct determination of satellite higher-order repeat structure as opposed to using indirect inference methods, like assembly, with reads shorter than the underlying lengths of the high order repeat unit. Here we demonstrate a comprehensive assessment of higher-order repeat patterns for two human cell lines: NA12878 (diploid) and NA19243 (tetraploid). First, we show the reliability of the method by validating consistency with existing centromere repeat references. Additionally, we are able to identify changes in repeat unit directionality that exist within arrays and between individuals. Based on this analysis, resolution of higher-order repeat patterns could be readily performed at low depth and reasonable cost across a population, or in genomes without high-quality references. This study demonstrates the methods to generate a sequence survey for regions enriched in satellite DNA that are typically omitted from genomic studies. We believe it establishes a foundation to extend and improve genomic characterization of any higher-order repeat structure using long reads.

1587S Anchored Assembly: An algorithm for large structural variant detection using NGS data. J. Brustle, B. Drees. Spiral Genetics, Seattle, WA.

Statement of purpose Characterizing large indels, inversions, and multinucleotide variants is important for understanding cancer, bacterial pathogens, and neurological disorders. Standard pipelines often miss these variants because of inexact alignment and lower sensitivity. Spiral Genetics has developed Anchored Assembly, a novel method for the analysis of long reads using direct, de novo read overlap assembly to accurately detect variants from next-generation sequence reads. Anchored Assembly’s range of detection and low false discovery rates may be useful for characterizing structural differences between tumor and normal samples.

Methods used Anchored Assembly was evaluated against Pindel and BWA + GATK using simulated read data. Datasets were generated by populating chromosome 22 of the human genome reference sequence with a set of SNPs, insertions, deletions, inversions, and tandem repeats.

Summary of results On human chromosome 22 data, Anchored Assembly detected over 90% of indels and structural variants up to 50 kbp and SNPs with false discovery rates well below 1%. In comparison, Pindel and BWA + GATK had overall false discovery rates of 10% and 9%, respectively. We detect, on average, over 90% of indels and structural variants up to 30 kbp in non-repetitive regions. The ability to detect deletions and structural variants is diminished by variant size, and the ability to accurately detect and assemble insertions continues well into the 30 kbp range.
1588M
Short inversion detection by splitting and re-aligning poorly mapped next-generation sequencing reads. R. Chen, W. Yang, Y. Lau. Paediatrics & Adolescent medicine, The university of Hong Kong, Hong Kong, Hong Kong.

Rapid development of sequencing technology has enabled routinely discovery of deletions and insertions. Unlike these two kinds of structural variation, characterization of inversions is left behind. Summary of public databases and researches shows that little short inversions with length smaller than 500bp have been detected for now. Because in contrast to small insertions and deletions, which are considered by gap alignment and recorded in primary mapping files, inversions short enough to interrupt alignment probably cause poorly mapped reads, which are mostly left out of consideration for structural variation detection by existing methods. And as a result, signals of short inversions are lost. Here, we introduce SRInversion, which tries to reuse those poorly mapped reads by splitting and re-aligning them for inversion detection and in this way, improve the resolution of inversion detection to less than 10bp. Comparison with previous structural variation detection methods, i.e. Pindel, BreakDancer, and DELLY, using simulated data indicates that SRInversion performs much better when inversion size is smaller than 100bp. Testing on chromosome 21 of a high-coverage parent-child trio (NA12878, NA12891, and NA12892) from 1000 genomes project shows that SRInversion is the only method that is able to detect inversions smaller than 100bp. For inversions with length from 100bp to 1000bp, besides SRInversion, Pindel is the only method that have output, which only include 7 regions, indicating relatively low power in detecting small inversions. PCR will be performed on randomly selected inversions detected by SRInversion on each individual of the trio for further validation. This method was also tested on a single African individual (NA18507) with both high-coverage and low-coverage whole genome data to check the effect of coverage, which will helps broader the spectrum of inversion at the same time.

1589T
Reproducible and repurposable toolkit of structural variant callers applied to 3,751 whole genomes and 10,940 whole exomes. S. Ma, S. Ambrette, A. Carroll, A. Sabo, P. Mishra, W. Salerno, A. English, N. Veeraraghavan, E. Boerwinkle, 1) DNAnexus, Mountain View, CA; 2) Baylor College of Medicine, Human Genome Sequencing Center, Houston, TX; 3) University of Texas Health Science Center, Human Genetics Center, Houston, TX.

Despite the significant influence human structural variation and copy number variation has on mendelian disease and cancer genomics, accurate calling of these variants remains a difficult problem. Great efforts have been made in understanding the concordance of SNPs and small indel calls, allowing for groups to harden these pipelines for large scale research and clinical applications. A similar systematic approach to structural variant calls has until now remained lacking.

Baylor College of Medicine’s Human Genome Sequencing Center (HGSC) has collaborated with DNAnexus to package together SVPipe which runs BreakDancer, CNVnator, CREST, DELLY, Genome STRIP, Pindel, and Tiresias for whole genome analysis, converting each output to a common format. A pair of CNV callers, CoNIFER and XHMM were also set up for exome analysis. The SVPipe multi-algorithmic approach was run on 3,751 low coverage whole genomes and the CNV callers were run on 10,940 exomes from the CHARGE project. These call sets provide a deeper understanding into the overlap between the tools and the amount of structural variants present in a large human population.

Thanks to HGSC’s work, all of these callers are publicly available on the DNAnexus platform. Users can run any caller without the need to install or configure software and each has been optimized to improve the speed and cost of its execution. These tools are provided with a number of benchmarks to validate the results of pipelines which use them.

1590S
Orthogonal Resequencing Support of Structural Variation in a Personal Genome. W.J. Salerno, A.C. English, C. Gonzalez-Jaureguiz, D.A. Hampson, S. Ambrette, D. Ritter, S. White, C. Davis, P. Mishra, Y. Liu, C.R. Beck, M. Dahdouli, N. Veeraraghavan, A. Hawes, D.A. Wheeler, J.G. Reid, D.M. Muzny, J. Rogers, K. Wohley, A. Sabo, W.J. Salerno, J.R. Lupski, E. Boerwinkle, R.A. Gibbs. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA.

Personal genome sequencing is becoming standard in research and clinical settings, and discovery of genetic disease etiologies increasingly requires to genomic characterizations beyond single nucleotide variants and small insertions and deletions. The human genome is subject to large (>100 bp) deletions, insertions, inversions, translocations, and complex combinations of these events. While the number of such structural variant (SV) events is far fewer than the number of single nucleotide variants, SVs collectively affect a comparable or even larger fraction of the genome, directly impacting gene function and regulation. Any effort to robustly characterize structural variation must address the broad variation in SV size, type, and complexity.

We have developed and implemented a novel approach for integrating whole-genome array comparative genomic hybridization (aCGH) and short-read next-generation sequencing (NGS) data with long-read (Pacific BioSciences RSII), long-insert (illuminata Nextera), and whole-genome architecture (BioNano Irys) data. These methods are illustrated via analysis of the genome of a phenotypically well-characterized individual with autosomal recessive Charcot-Marie-Tooth neuropathy. The integration of these orthogonal methods was performed with Parliament, a consensus SV-calling infrastructure that merges and evaluates the results from existing and novel SV-detection software and multiple input data sources. In this individual, Parliament identified 33,882 loci (composing 2.65% of the genome) that are inconsistent with the hg19 reference assembly. Of these, 7,460 are supported as putative SVs by local hybrid assembly and 3,440 are supported by long-read force calling or multi-source heuristics. These 10,900 putative SVs have an aggregate length of approximately 20 Mbp (0.6% of the genome), including 4,427 deletions, 4,442 insertions, and multiple complex SVs. The full genome was divided into individual SV-calling settings, and discovery of genetic disease etiologies increasingly requires to genomic characterizations beyond single nucleotide variants and small insertions and deletions. The human genome is subject to large (>100 bp) deletions, insertions, inversions, translocations, and complex combinations of these events. While the number of such structural variant (SV) events is far fewer than the number of single nucleotide variants, SVs collectively affect a comparable or even larger fraction of the genome, directly impacting gene function and regulation. Any effort to robustly characterize structural variation must address the broad variation in SV size, type, and complexity.

We have developed and implemented a novel approach for integrating whole-genome array comparative genomic hybridization (aCGH) and short-read next-generation sequencing (NGS) data with long-read (Pacific BioSciences RSII), long-insert (illuminata Nextera), and whole-genome architecture (BioNano Irys) data. These methods are illustrated via analysis of the genome of a phenotypically well-characterized individual with autosomal recessive Charcot-Marie-Tooth neuropathy. The integration of these orthogonal methods was performed with Parliament, a consensus SV-calling infrastructure that merges and evaluates the results from existing and novel SV-detection software and multiple input data sources. In this individual, Parliament identified 33,882 loci (composing 2.65% of the genome) that are inconsistent with the hg19 reference assembly. Of these, 7,460 are supported as putative SVs by local hybrid assembly and 3,440 are supported by long-read force calling or multi-source heuristics. These 10,900 putative SVs have an aggregate length of approximately 20 Mbp (0.6% of the genome), including 4,427 deletions, 4,442 insertions, and multiple complex SVs. The full genome was divided into individual SV-calling settings, and discovery of genetic disease etiologies increasingly requires to genomic characterizations beyond single nucleotide variants and small insertions and deletions. The human genome is subject to large (>100 bp) deletions, insertions, inversions, translocations, and complex combinations of these events. While the number of such structural variant (SV) events is far fewer than the number of single nucleotide variants, SVs collectively affect a comparable or even larger fraction of the genome, directly impacting gene function and regulation. Any effort to robustly characterize structural variation must address the broad variation in SV size, type, and complexity.

We have developed and implemented a novel approach for integrating whole-genome array comparative genomic hybridization (aCGH) and short-read next-generation sequencing (NGS) data with long-read (Pacific BioSciences RSII), long-insert (illuminata Nextera), and whole-genome architecture (BioNano Irys) data. These methods are illustrated via analysis of the genome of a phenotypically well-characterized individual with autosomal recessive Charcot-Marie-Tooth neuropathy. The integration of these orthogonal methods was performed with Parliament, a consensus SV-calling infrastructure that merges and evaluates the results from existing and novel SV-detection software and multiple input data sources. In this individual, Parliament identified 33,882 loci (composing 2.65% of the genome) that are inconsistent with the hg19 reference assembly. Of these, 7,460 are supported as putative SVs by local hybrid assembly and 3,440 are supported by long-read force calling or multi-source heuristics. These 10,900 putative SVs have an aggregate length of approximately 20 Mbp (0.6% of the genome), including 4,427 deletions, 4,442 insertions, and multiple complex SVs. The full genome was divided into individual SV-calling settings, and discovery of genetic disease etiologies increasingly requires to genomic characterizations beyond single nucleotide variants and small insertions and deletions. The human genome is subject to large (>100 bp) deletions, insertions, inversions, translocations, and complex combinations of these events. While the number of such structural variant (SV) events is far fewer than the number of single nucleotide variants, SVs collectively affect a comparable or even larger fraction of the genome, directly impacting gene function and regulation. Any effort to robustly characterize structural variation must address the broad variation in SV size, type, and complexity.

We have developed and implemented a novel approach for integrating whole-genome array comparative genomic hybridization (aCGH) and short-read next-generation sequencing (NGS) data with long-read (Pacific BioSciences RSII), long-insert (illuminata Nextera), and whole-genome architecture (BioNano Irys) data. These methods are illustrated via analysis of the genome of a phenotypically well-characterized individual with autosomal recessive Charcot-Marie-Tooth neuropathy. The integration of these orthogonal methods was performed with Parliament, a consensus SV-calling infrastructure that merges and evaluates the results from existing and novel SV-detection software and multiple input data sources. In this individual, Parliament identified 33,882 loci (composing 2.65% of the genome) that are inconsistent with the hg19 reference assembly. Of these, 7,460 are supported as putative SVs by local hybrid assembly and 3,440 are supported by long-read force calling or multi-source heuristics. These 10,900 putative SVs have an aggregate length of approximately 20 Mbp (0.6% of the genome), including 4,427 deletions, 4,442 insertions, and multiple complex SVs. The full genome was divided into individual SV-calling settings, and discovery of genetic disease etiologies increasingly requires to genomic characterizations beyond single nucleotide variants and small insertions and deletions. The human genome is subject to large (>100 bp) deletions, insertions, inversions, translocations, and complex combinations of these events. While the number of such structural variant (SV) events is far fewer than the number of single nucleotide variants, SVs collectively affect a comparable or even larger fraction of the genome, directly impacting gene function and regulation. Any effort to robustly characterize structural variation must address the broad variation in SV size, type, and complexity.
1592T
A Convergent Clinical Exome Pipeline Specialized for Targeted Gene Analysis. J. Plazzer1, A. Oshlack2, A. Gaff2, N. Thorne2, G. Taylor1, H. Dashnow4, A. Lone5, M. Bahlo5, T. Bakker6, D. Bauer1, K. Siemering7, P. James1, S. Sadedin5, Melbourne Genomics Health Alliance. 1) The Royal Melbourne Hospital, Parkville, Victoria, Australia; 2) Murdoch Childrens Research Institute, Parkville, Australia; 3) Melbourne Genomics Health Alliance, The University of Melbourne; 4) Centre for Translational Pathology, The University of Melbourne, Parkville, Australia; 5) Victorian Life Sciences Computation Initiative, Carlton, Australia; 6) Walter and Eliza Hall Institute of Medical Research, Parkville, Australia; 7) Division of Computational Informatics, CSIRO, Sydney; 8) Australian Genome Research Facility, Parkville, Australia.

Efforts to move high throughput sequencing into the clinic must confront many challenges including meeting clinical standards for cost, reproducibility, quality, ethics and privacy concerns. The Melbourne Genomics Health Alliance was formed from a diverse group of institutions with the aim of sharing the burden of these challenges through a common sequencing and bioinformatics platform. Critical to this concept is that regardless of disease, all patients undergo full exome sequencing, and targeting to a specific disease is performed in silico. While this has many benefits, it also adds significant complexity to the analysis pipeline as it must meet the differing diagnostic needs of diverse disease cohorts. In this work we present the Melbourne Genomics Health Alliance bioinformatics analysis pipeline, which has been specifically designed as a clinical grade exome sequencing pipeline with built in support for multiple sub-target regions. The pipeline is based on the Bpipe platform and includes support for customized sets of targeted regions, prioritized genes, regions blacklisted against incidental findings, annotations from a variety of other sources, and population variants, PDF provenance and quality reports. It allows genes to be prioritized at both a disease cohort and individual level and combines this with output from annotation tools to produce a clinically interpretable report that is available in a standard spreadsheet format. This allows for additional data to be added, such as LOVD (Leiden Open Variant Database) instance for curation. Being based on Bpipe, it offers powerful features for running many samples in parallel either on a cluster or dedicated computing resources as well as full traceability through auditable provenance of every file.

1593S
An ensemble variant calling approach to 10,000 low coverage whole genomes. Z. Huang1, N. Rustagi1, L.A. Cupples2, D. Muzny1, R. Gibbons2, E. Boerwinkle3, T. F. Yu1, CHARGE Consortium. 1) Human Genome Sequencing Center and Molecular and Human Genetics Department, Baylor College of Medicine, Houston, TX 77030, USA; 2) Boston University School of Public Health, Boston, MA 02118, USA and the National Heart, Lung and Blood Institute (NHLBI) Framingham Heart Study, Framingham, MA, 01702, USA; 3) Human Genetic Center, University of Texas Health Science Center, Houston, TX 77030, USA.

Accurate calling of common and rare variants is critical for population genetic studies, and novel gene discovery, as is consideration of single nucleotide variants, indels and short tandem repeats (STRs). There is growing consensus of the practical utility of large sample sizes having both deep whole exome sequencing (WES) and low coverage whole genome sequence (WGS), that are preferably obtained from the same library preparation. The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium is aiming to identify novel genes (and other noncoding genomic motifs) influencing health and disease, and is accruing deep WES and low coverage WGS (6) in ~10,000 deeply phenotyped individuals. In order to call and integrate variants across a large sample of low coverage genomes, multiple variant calling pipelines were used to avoid caller specific bias, and consensus filtering was adopted to select high quality variants. We utilized GATK UnifiedGenotyper (UG) and HaploTypeCaller (HC), GotCloud and SNPTools for SNP calling; GATK-UG, GATK-HC, Pindel for INDEL calling; and LobSTR for STR calling. We compared two consensus strategies: the 3-out-of-4 and a machine learning based consensus filtering. We built an automated pipeline including individual and consensus calling and quality assessment. This pipeline is computationally intensive, requiring a fully parallelizable design in a cloud environment (e.g. DNAnexus) to achieve high efficiency speed. We benchmarked the method using chromosome 20 data extracted from 3,751 whole genomes. With the 3-out-of-4 consensus strategy, we identify 240 million SNPs with a transition-transversion ratio of 2.83. Among them, 40.2% are novel compared to 1000Genomes phase3, and 91.2% of novel SNPs have MAF<0.1%. Compared to CHARGE WES from the same individuals, the WGS call set achieved a sensitivity of 61.1%, specificity of 99.9%, and 95.4% of the missed SNPs are singletons and doubletons. The SNP rediscovery rate in exome region is 38.3% for singletons and 62.2% for doubletons, which is higher than in 1000G phase1. We also identified 8,141 unique STR loci on chromosome 20, with a rediscovery rate of 97.8% compared to 1000G phase1 results. The entire process, including genotype/haplotype inference and WES integration, requires 3000 CPU-month for the Chromosome20 pilot. Deploying the pipeline in distributed computing environments (i.e. the Cloud) holds promise to make the ensemble method approachable to the community.

1594M
Likelihood-based filtering of indels and structural variants by leveraging Mendelian inheritance and transfer learning. H. Kang1, D. Absher1, A. Tan2, C. Quick1, A. Locke1, Z. Chen1, G. Jan1, J. Vickers1, C. Pato1, M. McInnis1, S. Zoieller2, G. Brenn3, S. Levy4, W. Iacono5, M. McCue6, L. Scott7, S. Vrieze1, M. Boehnke1, G. Abecasis1, The BRIDGES Consortium. 1) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Ann Arbor, MI; 2) HudsonAlpha Institute for Biotechnology, Huntsville, AL; 3) Center for Addiction and Mental Health, Department of Psychiatry, University of Toronto, ON, Canada; 4) Department of Psychiatry and Behavioral Sciences, University of California, Los Angeles, CA; 5) Department of Psychiatry, University of Michigan, Ann Arbor, MI; 6) University of Michigan Depression Center, Ann Arbor, MI; 7) Medical Research Council Social, Genetics and Developmental Psychiatry Center, Institute for Psychiatry, King's College London, UK; 8) Psychology Dept, University of Minnesota, Minneapolis, MN.

Accurate detection and filtering of short insertions and deletions (indels) and structural variants (SVs) is very challenging. Unlike SNPs, which can be validated using array-based genotyping, indels and SVs have fewer reliable resources to evaluate quality. Assessment of the pattern of genotype calls between members of the same family can provide an alternative way to evaluate variant quality, as the vast majority of variants should follow Mendelian inheritance (MI) patterns. However, even at truly variant sites, genotype calls contain errors and uncertainties due to variable read depth or quality, limiting the ability to distinguish between true and false variants. We develop a variant filtering method (MiFT) by quantifying the statistical support for a MI pattern of each variant in family sequence data. First, we compare the likelihood of sequence reads under models with and without family structure to obtain a Bayes’ Factor that quantifies the degree of support for MI. Second, we train a support vector machine (SVM) classifier to differentiate between variants with very high or low Bayes’ Factors, using the variant features annotated by each caller. Third, we can transfer the SVM classifier to predict the MI pattern of any independent dataset without related individuals. We used MiFT to filter SNP, indel, and SV calls produced by GotCloud, UnifiedGenotyper (UG), HaploTypeCaller (HC), GenomeSTRIP, Pindel, and CBS on whole genome sequencing data from 811 families from the Melbourne Genomics Health Alliance. We then applied the prediction method to 2,833 unrelated subjects sequenced in BRIDGES. We observed that existing SNP filtering methods are highly concordant to results from MiFT (e.g. 97% for UG). For indels, concordance to MiFT was less prediction with standard filtering was lower (e.g. 78% for UG). When the filters disagree, variants exclusive to MiFT were significantly more concordant to independent validation results from Affymetrix exome arrays (e.g. OR=8.3, p=7.6x10-8 for UG). For structural variants, we observed that GenomeSTRIP effectively filters out Mendelian inconsistent variants, but also filters out 30% of the Mendelian consistent variants preserved by MiFT. MiFT provides a general framework to robustly filter SNPs, indels, and SVs. MiFT allows us to evaluate variant call sets against a single standard and will facilitate systematic integration of multiple call sets while achieving high sensitivity and specificity.
1595T Strength in Numbers: Efficiency and Quality Improvements in Clinical Whole Genome Interpretation. E. Ramos, C. Mead, J. Sardina, A. Khoury, D. Mote, S. Pond, A. Crawford, S. Ayaj, J. Sihavy, S. Chadhury, T. Hambuch. Illumina Clinical Services Laboratory, San Diego, CA. As the cost of whole genome sequencing rapidly decreases and technical limitations are surmounted, one of the most significant remaining challenges of clinical whole human genome sequencing (WGS) is the interpretation of newly discovered variants. Since October 2012, the Illumina Clinical Services Laboratory has sequenced and interpreted the genomes from approximately 500 primarily healthy adults; with interpretation focusing on 1600 genes associated with 1221 of the most commonly tested monogenic conditions. All single nucleotide variants were assessed to determine clinical significance by a team of trained geneticists and genetic counselors, in accordance with the American College of Medical Genetics & Genomics guidelines. This process includes a manual review of the literature for all variants detected within that subset of 1600 genes. As the Illumina Variant Database has grown to greater than 60,000 variants, including more than 650 variants with potential clinical significance, the average number of variants per genome that are novel to our laboratory has decayed significantly. Currently, of the approximately 5000 single nucleotide variants detected within the subset of 1600 interpreted genes, an average of 65 novel variants is found per genome. This is a six-fold decrease from the 360 novel variants identified per genome in May 2013, resulting in a substantial concordance reduction in manual effort. This reduction has been enabled through custom software developed to facilitate the clinical process of classifying, interpreting and reporting on variant disease relationships. These two factors have contributed to an overall decrease in time required for a single WGS sample, from 35 hours in May 2013 to less than ten hours today. Continued growth in the total number of variants within the database, as well as improvements to the software that aids in the annotation and interpretation process, is anticipated to result in a continued diminution of interpretation effort, which will improve the quality of ongoing interpretation and mitigate the cost of offering WGS in the clinic. Although challenges involving the clinical interpretation of whole genome sequence data will persist, these advances are anticipated to facilitate the implementation of WGS in the clinical laboratory testing arena.

1595M Computational validation of NGS variant calls using genotyping data. M.A. Taub1, S. Shringarpure2, R.A. Mathias3, R.D. Hernandez4, I.D. O’Connor1, Z.A. Szpiech2, R. Torres2, F.M. De La Vega2, C.D. Bustamante2, K.C. Barnes1. 1) Biostatistics, Johns Hopkins, Baltimore, MD; 2) Genetics, Stanford University, Palo Alto, CA; 3) Medicine, Johns Hopkins, Baltimore, MD; 4) Bioengineering and Therapeutic Sciences, UCSF, San Francisco, CA; 5) Medicine, University of Maryland, College Park, MD. Variant calling from next-generation sequencing (NGS) data is susceptible to false positive calls due to sequencing, mapping and other errors. We present a method that uses machine learning techniques, specifically Random Forests, for computationally validating variant calls obtained from a sample of individuals. While existing methods use site quality information from known samples such as HapMap and dbSNP for training classifiers to distinguish between true and false variant calls, our method uses genotype data from the same samples to learn a more accurate classifier. We demonstrate our method on a set of variant calls obtained from 643 high-coverage African-American genomes from the The Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA), sequenced to high coverage (~30X). On variant calls obtained using Illumina’s single-sample caller CASAVA, our method has a True Positive Rate of 97.5% (at a False Positive Rate of 5%). On variant calls obtained from Real Time Genomics’ multisample variant caller, our method obtains a True Positive Rate of 95% (at a False Positive Rate of 5%). Since most NGS sequencing data is accompanied by genotype data for the same samples, our method can be trained on each dataset to provide a more accurate computational validation of site calls than using generic methods. In addition, our method can allow for adjustment based on allele frequency (e.g., allow a different set of criteria to determine quality for rare vs. common variants) and thereby provide insight into sequencing characteristics that most clearly indicate data quality for variants of different frequencies. We have also applied our classifier to compare call sets generated with different calling methods, including both single-sample and multi-sample callers, and we have found that allele frequency is an important determinant of which calling method makes the most accurate calls.

1596S Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. E. Smith1, A.R. Carson1, T. Solomon1, H. Matsui1, S.K. Braekkan2, K. Jepsen2, J.B. Hansen2, K.A. Frazer1,2,3. 1) Department of Pediatrics and Rady Children’s Hospital, University of CA, San Diego, CA; 2) Department of Clinical Medicine, Hematological Research Group, University of Tromso, Tromso, Norway; 3) Division of Internal Medicine, University of North Norway, Tromsø, Norway; 4) Clinical and Translational Research Institute, University of California, San Diego, USA; 5) Department of Clinical Medicine, University of Tromso, Tromso, Norway. Genotype data from next generation sequencing studies contain errors which can significantly impact the power to detect signals in common and rare variant association tests. These genotyping errors are not explicitly filtered by the standard GATK Variant Quality Score Recalibration (VQSR) tool and thus remain a source of errors in whole exome sequencing (WES) projects that follow GATK’s recommended best practices. Therefore, additional data filtering methods are required to effectively remove these errors before performing association analyses with complex phenotypes. Here we empirically derive thresholds for genotype and variant filters that, when used in conjunction with the VQSR tool, achieve higher data quality than when using VQSR alone. The detailed filtering strategies improve the concordance of sequenced genotypes with array genotypes from 99.33% to 99.77%; improve the percent of discordant genotypes removed from 10.5% to 69.5%; and improve the TiTv ratio from 2.63 to 2.75. We also demonstrate that managing batch effects by separating samples based on different target capture and sequencing chemistry protocols results in a final data set containing 40.8% more true- than-quality variants. In addition, imputation is an important component of WES studies and we use estimate common variant genotypes to generate additional markers for association analyses. As such, we demonstrate filtering methods for imputed data that improve genotype concordance from 79.3% to 99.8% while removing 99.5% of discordant genotype calls. In addition, an unbiased analysis of our data has revealed that the error rate associated with the clinical filtering algorithms is very small (~3% for every 5% bacterial DNA in the sample) and does not result in a difference in the number of variants called. Concordance differences between blood and saliva are minor and can be eliminated by ensuring adequate sequencing coverage.

1598T Blood vs Saliva: Analysis of the Effect of Sample Type on Variant Calling Confidence for Human Whole Genome Sequencing. M. Tayeb1, A. Mijaikovic2, M. Kovacevic2, M. Popovic2, S. Wernicke2, C. Dillane3, A. Deluca3, R.M. Ivasiow2. 1) DNA Genotek Inc., Ottawa, Canada; 2) Seven Bridges Genomics Inc, Cambridge, Massachusetts; 3) Medline, San Diego, CA. Saliva, obtained using the Oragene self-collection kit, as a source of genomic DNA has grown in popularity due to participating compliance and the high quality and utility of the collected DNA. Although Orangene/saliva is supported by over 1000 peer-reviewed publications, to date few papers have been published on Whole Genome Sequencing (WGS) from saliva. Here we present a systematic, multi-sample analysis of the effect of sample type (blood vs. saliva) on variant calling confidence and the effect of bacterial DNA in saliva on sequence alignment. Blood and saliva were collected from seven individuals (two families). The bacterial content in each saliva sample was determined by 16S qPCR. DNA was extracted and sequenced on the Illumina HiSeq2000 using standard protocols from Illumina. Variant calling and filtering was performed on the Seven Bridges Genomics platform for bioinformatics analysis using a standard BWA and GATK pipeline in accordance with the Broad Institute’s best-practices guidelines. The percentage of bacterial DNA was found to closely correlate with the number of reads that could not be mapped to the human genome. Those reads were aligned to bacterial and viral genomes obtained from the Human Microbiome Project (HMP). A significant portion of unaligned reads mapped to the HMP database reference (72% in blood, 32% in saliva). No significant difference in the total number of variants called in blood and saliva were observed. Only a slight difference in concordance between sample types was observed (<0.15% for SNPs and <1% for INDEL). Downsampling blood samples to equal coverage virtually eliminated these differences as filtering and its effect on concordance in saliva was primarily due to coverage effects. Additional analysis was performed to check for existence of ambiguous regions that act as baits for bacterial reads causing saliva samples to accumulate false positive mutations in those regions. Further analysis using a higher sequencing coverage (70X for saliva instead of 40X) in both blood and saliva samples, and we have found that allele frequency is an important determinant of which calling method makes the most accurate calls.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1599S
High resolution HLA genotyping software for exome and whole genome sequencing data. K. Kryukov, S. Nakagawa, T. Imanishi. Department of Molecular Life Science, Tokai University School of Medicine, Isehara, Kanagawa, Japan.

The MHC region on chromosome 6 contains human leukocyte antigen (HLA) genes, which are some of the most polymorphic genes in the human genome. Many HLA alleles are known to be associated with susceptibility to various infectious and autoimmune diseases. Genotyping HLA loci correctly is challenging due to abundant repeats, high sequence similarity among HLA genes, as well as the presence of highly similar pseudogenes. Normally, sequence-based genotyping is done by highly targeted sequencing of specific loci. Here we present a genotyping software that allows predicting HLA genotypes based on the whole exome or the whole genome sequencing data. Our method is based on sequence similarity searches of the raw reads to the database that combines HLA alleles from the IMGT/HLA database, the known MHC haplotype sequences from Sanger Center, and the entire reference human genome. This allows to accurately map reads to specific loci. We then focus on each locus and compare the mapped reads with the known allele sequences to determine the genotype. Our method has accuracy of 90% in 2-digit typing and 74% in 4-digit typing for Class-I and Class-II genes, on Ion Torrent sequence data targeted to specific HLA genes.

1600M
Using haploid human DNA to design and evaluate the HiSeq X data processing strategy. M.O. Pollard, T.M. Keane, S.A. McCarthy, J.C. Randall, R.M. Durbin. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, UK.

The Illumina HiSeq X is a new sequencing platform that promises a reduction in sequencing cost to below $1000 per whole human genome. The dramatic increase in throughput that this platform will deliver has required changes to chemistry, software, and output. Of interest to dry lab scientists are several key changes in the data produced including an increase in read length to 150 base pairs and the binning of base quality scores into eight distinct quality bins in order to reduce the quantity of data that is stored.

To take full advantage of this new platform, we have designed a series of systematic tests to evaluate the sequence data produced by the HiSeq X and determine the most appropriate subsequent processing protocols. We will test and compare different combinations of mapper, variant caller, and reference; using well characterised samples in order to examine the effects that this technology has on each and evaluating its utility for variant detection.

In this experiment, we plan to sequence the 1000 Genomes CEU Trio and two human haploid cell lines: CHM1hTERT and HAP1. We will evaluate the effects of various read alignment tools (BWA backtrack, BWA MEM, and SNAP), variant callers (samtools, GATK Unified Genotyper, and GATK Haplotype Caller), and choice of reference (GRCh37 vs GRCh38) on the sensitivity and specificity of SNP/InDel identification. The CEU Trio is a well-studied family, which has been used in many previous studies such as HapMap and 1000 Genomes providing a corpus of knowledge to compare against. Both the CHM1hTERT and HAP1 cell lines are effectively haploid, meaning that any heterozygous variants found are likely to be either a sequencing error or data processing artifact. In addition, when HAP1 is mixed with CHM1hTERT in equal proportions it should allow us to evaluate heterozygous calling and phasing performance. The findings of these experiments will be valuable for determining the optimal data processing strategy for the HiSeq X.
1603M
Single molecule reconstruction and variant detection less than 1 genome in 1000. K.R. Covington, M. Wang, D.M. Muzny, H. Doddapaneni, R.A. Gibbs, D.A. Wheeler. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX., USA.

Genetic mosaicism contributes to phenotypic diversity both in physiology and disease including phenotypic diversity in monozygotic twins, brain develop- opment and disease, and cancer. Conventional sequencing approaches to variant discovery rely on repeated sampling of variants, from the population of genomes in a given sample. Standard sequencing methods fail to address rare, subclonal mutations due to confounding influence of sequencing errors, even at high sequence depth. We address this problem by adapting pub- lished proof-of-concept methodologies for deep sequencing using single molecule reconstruction. Single molecule reconstruction relies on repeatedly sequencing the PCR products of single molecules of DNA and using the sequence data to reconstruct the sequence of the initial molecule. With this strategy we turn PCR duplicates into an advantage to repeatedly sample fragments which can be traced to a single DNA molecule. We used a combination of 6 to 24 Illumina barcode adapters pooled to individually tag sheared DNA libraries. DNA libraries were then captured using custom cancer capture designs. Libraries were sequenced using either Illumina MiSeq (for initial testing) or HiSeq platforms. Sequencing data was aligned to the human genome hg19 build. We developed custom software to reas- semble sequencing reads into their original single molecule DNA fragments and for further analysis. Over 90% of non-reference bases sequenced were observed only in one read of a duplex set, indicating that these non-reference bases are merely sequencing errors. Our ability to recreate the initial DNA molecule allowed us to detect variant genomes at less than 1-1000 variant allele fraction in a series of spike-in tests. Interestingly, over 50% of consen- sus variants were called in only one molecule, even in normal tissues. These data suggest the presence of an extensive reservoir of somatic variation in normal tissues, which may contribute to human disease.

1604T

Identification of variants from cancer samples using NGS target enrichment is becoming increasingly popular in clinical research. Cancer samples are challenging for variant detection due to tumor heterogeneity, low abundance of variant bases, mixed genotypes, multi-allelic loci, copy-number variations, variable background mutation rates and contamination from normal cells.

After evaluating available methods it was imperative to develop an algorithm in-house. SNPPET was developed in Agilent SureCall software to detect cancer variants with high confidence. SNPPET has two basic steps - first an efficient search for variants and then a careful examination of the variant's neighborhood to confirm the call. In the first step, the reads and bases are filtered for quality and coverage criteria such as mapping qualities, base qualities and read depths. SNPPEET's initial model evaluates each locus and all non-reference alleles are assigned to sequencing error. Then a second model considers each non-reference allele to be a true variant. The models allow presence of more than one allele at each locus. A log-odds score is calculated between the two models and if the score passes a specified threshold is considered to be called. If the allele is more likely to be a variant than a sequencing error. If multi-allelic option is chosen, all potential variant alleles are retained. Otherwise only the one with max- nimo allele fraction is retained. In the second step, a local re-assembly is done to evaluate all potential variants combinations as haplotypes. The haplotypes are scored by a Bayesian model and all variants passing the threshold are retained. SNPPET allows one to examine loci of interest irrespective of whether a “true” variant is present or not. This allows one to ascertain whether there is low significance event proportionality at a target which did not pass the statistical tests, but still may be indicative of potential variation in a site of interest. Dilution series experiments were performed to benchmark the sensitivity and specificity of SNPPET and to compare the variant calls with other state of the art methods. A cancer cell line (HCT116) was mixed with normal HapMap (NA10831) at varying concentrations. At concentration of <=5%; SNPPET detected expected variants that are missed by GATK unified genotyper and SAMTools. SNPPET demonstrated greater sensitivity than any other methods compared, keeping the specificity at par with these methods.

1605S
Identification of common non-synonymous SNPs in proteomic data-sets and their use to obtain measures of individualization and biogeo- graphic background. G. Parker1,2, D. Anex, T. Leppert2, L. Baird3, N. Matsunami2, M. Leppert2, B. Hart1. 1) Utah Valley University, Dept of Biol- ogy, Orem, UT; 2) Forensic Science Center, Lawrence Livermore National Laboratory, Livermore, CA; 3) Department of Human Genetics, University of Utah, Salt Lake City; UT. We have developed methodology to extract identifying genetic information from proteomic datasets. DNA-typing has revolutionized forensic practice and jurisprudence, however DNA often is degraded due to biological, chemi- cal or environmental factors. Protein is considerably more stable, more abundant than DNA, and persists in the environment for a longer period.

Protein also contains genetic information in its primary structure, the result of non-synonymous SNPs (nsSNPs). These single amino-acid polymorphisms are accessible to shotgun tandem mass spectrometry. We have identified nsSNP-containing peptides from 35 alleles in 26 genes expressed in the forensically informative hair shaft proteome. We obtained complex proteomic datasets from trypsin digests of the hair shafts of 54 validated European American individuals. Peptides corresponding to nsSNPs expressed in this protein population were identified and collated for each individual. The com- bined probability of each individual nsSNP profile was calculated using genotypic frequencies of each allelic combination in the European population (1000 Genomes Project) and the “product-rule”. The power of genetic dis- crimination ranged from 1 in 1,002 to 1 in 9,000. The average power of discrimination was 1 in 280. The power of discrimination increased as a function of proteomic dataset quality (r² = 0.82, n = 58, p < 0.0001). When the power of discrimination is calculated using genotypic frequencies from the African population, increased powers of discrimination are achieved. This is consistent with an increased likelihood that the samples originate from a European relative to African origin. Likelihood measurements range from 1 to 780 with an average of 50, a median of 18, and a standard deviation of 116. (n = 64). Direct validation of the imputed status of each nsSNP allele was achieved with Sanger sequencing. A total of 430 genotype determina- tions were made from the proteomic data and 426 assignments were con- firmed (specificity = 99.1%, FPR = 0.93%). The average sensitivity was 96.5%. We have established an approach to using proteomic datasets as a source of identifying genetic information, allowing measures of identity and biogeographic background to be made from forensic or anthropological protein. This study also demonstrates that protections should be enacted to ensure the privacy and confidentiality of human subjects when providing tissue for proteomic studies.

1606M
SNP and CNV Detection in Trisomy 21 Individuals Using a First-Princi- ples Approach. Y.A. Jakubek, D.J. Cutler. Human Genetics, Emory Univer- sity, Atlanta, GA.

Genotyping arrays are used to simultaneously detect single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) across the genome. Current array-based methods use empirical data to call SNPs; therefore, they are not well suited for studies with a small number of samples and for SNPs with a low minor allele frequency. These methods drop 20% to 33% of the targeted SNPs and are susceptible to batch effects. We have developed a SNP and CNV detection algorithm for Affy 6.0 arrays that is based on a low-level model of hybridization on the array. In it we model binding between all probe-DNA duplexes that form on the array. Further- more, we explicitly model batch effects through a set of parameters that describe binding for each array experiment: DNA concentration, probe errors, salt concentration, wash stringency, and scanner settings. This bio- chemical approach allows for the individual analysis of each chip. Since we directly model DNA concentration, our algorithm can type SNPs and CNVs directly. This approach uses a low-level model of hybridization on the array. We analyzed data from Down syndrome and normal samples. 13% of tar- geted SNPs show significant cross-hybridization. 84% of SNPs on diploid chromosomes and 57.5% of SNPs on aneuploid chromosome 21 had QS > 0.99. Our algorithm directly estimates DNA concentration for each probe and then uses these values to detect CNVs. This approach is fundamentally different from current approaches that rely on probe intensity. We called an average of 50 CNVs per individual and 68% of the CNVs called were in the database of genomic variants (DGV). This data was previously analyzed; validation was attempted for 64 CNVs of which 59 were validated. We method only called the 59 validated CNVs. This method can be easily adapted to call SNPs within CNVs and for aneuploid chromosomal in gen- ically diverse samples using only first-principles our method detects genetic variants with high accuracy.
error differed among programs. At this point, there is no substitute for the
take phenotypic predictions. For splice outcomes incorrectly predicted by
outcome prediction. Because for COL1A1 mutations, the relation between
is the first step. If splice site mutations in the same gene result in a range
HSF correctly predicted splice outcome in 42% of cases,
1G>A changes always resulted in frameshifts when the first nucleotide of
of the resulting mRNA(s). Stable mRNA transcripts produce more severe
(I) chain of type I collagen, alter mRNA splicing. The Osteogenesis
Pathology, University of Washington, Seattle, WA; 2) Medical Genetics,
Data from Patients with Osteogenesis Imperfecta.

Approximately 20% of pathogenic mutations in COL1A1, which encodes
the pro(1) chain of type I collagen, alter mRNA splicing. The Osteogenesis
model and measure sensitivity and specificity as a function of expression
levels, including the impact of allele-specific expression on sensitivity. We
assess performance across a range of RNA-seq data-types, including differ-
ent sample preparations techniques and sequencing depth, showing false
positive rates ranging from 1% to 5%. Integrating matched DNA and RNA
data, we demonstrate how RNA-seq can be used to improve accuracy of
DNA variant call sets by rescuing up to 5% of marginal calls. We derive
allele-specific expression and between-exon phasing information to guide
variant interpretation.

An Evaluation of Splice Prediction Software Accuracy Using in vivo
Data from Patients with Osteogenesis Imperfecta. J. Schleit1, S.S. Bailey1,
T.T. Tran1, D. Chen1, S. Slowers1, U. Schwarze1, P.H. Byers1,2, 1) Pathology,
University of Washington, Seattle, WA; 2) Medical Genetics, University of
Washington, Seattle, WA.

Whole exome sequencing (WES) is now a standard method to detect
rare and common genetic variants. Whole genome sequencing (WGS) is
becoming an attractive alternative approach with an affordable price. Here
we compared WES and WGS for the detection of rare and common variants
in the coding regions of the human genome using the most recent
technologies. We performed WES (using the SureSelect Human All Exon
Kit 71 Mb) and WGS (using the TruSeq DNA PCR-Free Sample Preparation
Kit 71 Mb) on 51 unrelated French individuals. The average raw read pairs
and reads of ~100 base pairs. A total of 132,108 SNVs were detected by either WES
or WGS within the regions targeted by the WES kit, 86.7% by both methods,
4.5% by WES, only and 8.8% by WGS only. While these variants were present
in a total number of reads which were similar in WES (~5.8 millions)
and in WGS (~4.7 millions), the distribution of the coverage depth was
strikingly different between the two methods. In WES data, the distribution
of depth was lognormal-like with a mode at 5x and a median at 26x, while
the distribution was uniform in the WGS data with a modal distribution at 38x,
showing a much more homogeneous coverage by WGS. Among the 114,586
SNVs coded by both platforms, we noted discordant genotypes in 3,676
variants (3.2%) that were in all cases called as homozygous by WES and
as heterozygous by WGS. We then filtered the variants on depth (> 8x),
generation family (≥2) or variant frequency at the same position on
≥20%). Only 69.4% of WES variants passed the filter, while 97.1% of WGS
variants did. After filtering, a total of 82,331 SNVs were called by both
platforms (99.97% of those variants had the same genotype), 3,857 variants
by WGS only and 1,41,468 by WES only. Our sample set was selected exclu-
sively to WES by the Integral Genomic Viewer tool showed that these variants are
also present in WGS data but not called by Unified Genotyper. The utilization of
an alternative/additional variant caller on WGS data should call these
variation. In contrast, variants were correctly called by WGS only in WES
data. These variants mapped to coding regions, UTRs, splice sites and non-
coding RNAs. We conclude that WGS is more powerful and more efficient
than WES to detect SNVs even in exonic regions, as it calls about ~50% more high quality SNVs than WES. This work will have major impact to
define the most cost-effective strategy to identify the genetic etiology of rare
and common diseases.

Evaluation of the Illumina NextSeq500 for Rapid Whole Genome
Sequencing. S. Dames1, J. Durtschi1, R. Mao1,2, K.V. Voelkerding1,2, 1) ARUP Labs, Salt Lake City, UT; 2) University of Utah Department of
Pathology.

Introduction: Rapid sequencing of the whole genomes offers a new diag-
nostic tool in clinical settings, including newborns. A new next generation
sequencing platform, the Illumina NextSeq500, generates 120 Gb of
sequence in 29 hours using 2 X 150 base length reads. This represents
sequence in 29 hours using 2 X 150 base length reads. This represents
30-fold average coverage of the human genome. The current study evalu-
ates the quality of whole genome sequencing data generated by the Next-
Seq500. Methods: HapMap sample NA12878 whole genome Illumina librar-
ies were processed using a Beckman Coulter SPRi-TE. The ARUP validated
clinical pipeline comprised of BWA, SAMtools, GATK, and Annovar was
employed. NextSeq500 NA12878 whole genome data was compared to
NA12878 NIST and NA12878 exome data. For the gene structure variation
and loss/gain analysis, alternate percentage and reference allele percentage
results were plotted against chromosomal location to visualize
CNVs that alter the expected allele read frequency. Results: Two independ-
ent runs of NA12878 yielded 68 Gb at 87.9 percent >Q30 and 134.5 Gb at
70.4 percent >Q30. Whole genome data filtered by exome bed file coordi-
nates was intersected with NIST NA12878 data. 66,975 variants were
detected in whole genome data, with 66,492 >4-fold (99.3 percent) and
65,232 >8-fold coverage (98.1 percent), average 21-fold coverage. The
NIST high confidence data contains 55,176 variants in these regions, where
53,958 (97.8 percent) were concordant between whole genome data and
NIST. For rare variants (≥20%) were present between the exome data and
NIST data. The copy number gain/loss results were concordant with high
density SNP array. Conclusions: Preliminary analysis demonstrates an
approximate 98 percent concordance between variant calls derived from
whole genome data. For rare and common genetic variants of NA12878.
Whole genome data filtered with NIST NA12878 variant data More sophisti-
cated analysis comparing different alignment and variant calling algorithms to improve alignment and variant
call accuracy and speeds will be presented. Further analysis is underway
to determine the cause of the high concordance of genome wide
whole genome data. Improvements in the speed of analysis and quality metrics
will improve turn around time, and may preclude the requirement of Sanger
sequence for a subset of variants.
**1611S**


Strand Life Sciences, Bangalore, Karnataka, India.

**Background and Objectives:** Next generation sequencing technology has led to the generation of millions of short reads at an affordable cost. Aligning these short reads to a reference genome is a crucial task for most of the downstream analysis. However aligning the short reads to a reference genome is a non-trivial task because of the large size of the data and many complex regions in the genome. These challenges necessitate the need for sophisticated algorithms that are both accurate and computationally efficient. In this work, we will briefly discuss Strand NGS (formerly Avadis NGS) alignment algorithm and present the benchmarking results on several simulated data sets and a real whole-genome data to compare it with other state-of-the-art algorithms.

**Results:** Multiple aligners like Strand NGS, BWA, BWA-Mem, Bowtie2 and Novaalign3 are compared for accuracy and computational efficiency using 4 simulated data sets from the GCAT website and a real Illumina HiSeq 2500 whole-genome paired-end data of 1000 genomes CEU female sample, NA12878. Strand NGS and Novaalign3 showed comparable accuracy in terms of % correctly mapped reads and receiver operating curves (ROC). They also seem to outperform other algorithms especially on data sets with longer indels. Our aligner, Strand NGS has been designed to detect longer indels by excluding the longest gap in a read match while computing the percent identity (alignment score) for the match. This makes it possible to give a higher cut-off for the percentage of gaps allowed, without reducing the percent match cut-off. For reads potentially originating from complex genomic locations like repeat regions, Strand NGS aligner produces a higher true positive rate compared to Novaalign3, when aligned reads are filtered based on the associated assigned mapping qualities. As for the performance comparison based on computational efficiency, other than minor differences, practically all the included algorithms showed comparable performance.

**Conclusions:** Alignment of millions of short reads to a large reference genome with many complex regions still is a hard and almost all current algorithms adopt some form of strategy to trade-off accuracy and computational efficiency. The benchmarking results presented in this study suggest that Strand NGS aligner is a powerful approach for DNA short reads. It either compares well or even outperforms other state-of-the-art algorithms.

---

**1613T**

**A comprehensive comparison of RNA-seq and microarray in transcriptome profiling of rat livers exposed to a broad range of agents.** C. Wang, B. Gong, P.B. Bushel, J. Thierry-Mieg, D. Thierry-Mieg, J. Xu, F. Fang, H. Hong, L.J. Lancashire, C. Furlanello, L. Shi, R.S. Paulus, S. Auerbach, W. Tong. 1 Center for Genomics, Loma Linda University, Loma Linda, CA; 2 National Center for Toxicological Research, Jefferson, AR; 3 National Institute of Environmental Health Sciences, RTP, NC; 4 NCBI, NLM, National Institutes of Health, Bethesda, MD; 5 Thomson Reuters, London, UK; 6 Fondazione Bruno Kessler, Trento, Italy; 7 Fudan University, Shanghai, China.

Next-generation sequencing technologies have revolutionized the genomic research and allow the genome and transcriptome of any organism to be explored without a priori assumption. Compared to microarrays, RNA-seq is able to provide single-nucleotide resolution, strand specificity, and short-range connectivity through paired-end sequencing. However, emerging transcriptomics technologies should be evaluated by ways in which they create opportunities to advance the understanding of complex biological systems. An extensive and systemic investigation based on a comprehensive study design was performed to investigate the strengths/weaknesses and comparability of the two platforms in biologic elucidation. RNA extracted from livers of rats exposed to 27 agents eliciting a broad range of transcriptional response, comprised of seven modes of action (MOAs) with varying degree of biological complexity. was used to profile gene expression using Illumina RNA-seq and Affymetrix microarrays and to assess concordance. We found that RNA-seq detected more differentially expressed genes (DEGs) than microarray and hence enriched more biological pathways. While RNA-seq was superior in detecting DEGs with low abundance and could also detect differential expression of transcripts, non-coding RNAs, and exon-exon junctions in a treatment-dependent manner, prediction performances of classifiers derived from both platforms were comparable. Overall, the cross-platform concordance in terms of DEGs, enriched pathways, or modes of action is highly correlated with treatment effect size, gene-expression abundance, and the biological complexity of the mode of action.

---

**1614S**

**Implementing an NGS Bioinformatics Pipeline: Making the Transition from Research to Clinical.** L. Watkins, K. Hetrick, D. Snyder, H. Ling, S. Griffith, J. Goldstein, M. Mawhinney, J. Romm, E. Hsu, G. Lowe, K. Roberts, K. Doheny, B. Craig. Center for Inherited Disease Research (CIDR), Johns Hopkins University, Baltimore, MD.

The Center for Inherited Disease Research (CIDR) provides high quality sequencing and genotyping services and statistical genetics consultation to investigators working to discover genes that contribute to disease. The CIDR NGS bioinformatics analysis pipeline, continually developed since 2009 to keep pace with the rapid changes in the predominant sequencing analysis tools, is designed and tuned for large-scale research projects with huge numbers of analyses that must be done in parallel and fault-tolerant fashion. In winter 2013-14 CIDR partnered with a molecular diagnostic lab at our institution to migrate a traditional genetic test panel to NGS, wherein CIDR would run the NGS bioinformatics analysis and they would handle everything else. Despite a limited budget and tight timeline the effort was successful, with the first clinical results reported out in March 2014. In some regards the implementation of a bioinformatics pipeline for a clinical bioinformatics pipeline is similar to that of a product development (e.g., the need for reliability and strict quality control) while others can be quite different (e.g., the small numbers of samples, absolute need for quick turn-around time, strict adherence to well-defined regions of interest, an imperative to never report a false positive, the specific and limited relevant annotations), in particular the requirement to lock down all aspects of the pipeline after validation and to document it in detail and track any associated changes in accordance with new CAP guidelines for next-generation sequencing and bioinformatics analysis. This mitigates against sharing bioinformatics pipelines and infrastructure between research and clinical, so the decision was made to establish completely separate systems. This poster describes the practical aspects and related considerations of the subsequent effort, including getting separate equipment set up and tested/validated, the logistics of running real samples in their dose and in their sequence, the execution of run on 15 Illumina sequenced whole genomes. Each genome has an average read depth of 37. For our research we test the following algorithms: Bowtie2, BWA, GSNAP, Mosaik, Novaalign, Soap2, and Stampy. Each of these is free for academic use and is currently maintained. Ultimately, this will lead to more accurate experimental and analytical results.
1615M
AmpliSeq based 16S ribosomal RNA Sequencing and Species Identification. J. Dickman, J. Risinger, M. Toluoe. Bio Scientific Corporation, Austin, TX, USA.

Next generation sequencing analysis of 16S ribosomal RNA is commonly used to identify bacterial species and perform taxonomic studies. Bacterial 16S rRNA genes contain nine hyper-variable regions with considerable sequence diversity among different bacterial species and can be used for species identification. Rapid determination of highly complex bacterial populations through targeted amplification can provide an accurate gauge of diversity at taxonomic hierarchies as low as the genus level. A single 16S rRNA hyper-variable domain does not have enough sequence diversity to distinguish all bacterial species. With increased read lengths of MiSeq sequencing chemistry, Bio Scientific has expanded the common analysis of the fourth hypervariable domain (V4) of prokaryotic 16S rRNA to V1, V2 and V3 regions simultaneously. Optimized preparation through a streamlined standardized procedure allows for high-quality, reproducible libraries. This optimization can be applied to different windows of 16S rRNA as well as other relevant prokaryotic taxonomic markers.

1616T
Super-resolution imaging technique mbPAINT for DNA optical mapping. J. Chen1, L. Kisley2, A. Bremauntz2, C.F. Landes2. 1) Department of Chemistry & Biochemistry, Ohio University, Athens, OH; 2) Department of Chemistry, Rice University, Houston, TX.

Optical mapping has been used to assist the de novo assembly of whole genomes. This fluorescent imaging technique measures the restriction maps of very long DNA molecules that are stretched on a solid support. The resolutions of the traditional techniques are limited by the diffraction limit of light. Here I am reporting a new optical mapping strategy that allows sub-diffraction limit imaging of genome optical maps using a super-resolution optical nanoscopy based on single-molecule localization method, which has been named motion blur Point Accumulation for Imaging in Nanoscale Topography (mbPAINT). The strategy is demonstrated with model target and probe DNA sequences, and an optical map of a randomly chosen repeated 7-basepair sequence in the phage λ-DNA.

1617S

High-throughput short read sequencing has revolutionized re-sequencing applications. However, they have had a more limited impact on generating high-quality de novo assemblies of eukaryotic genomes, principally because the short reads are insufficient for disambiguating complex repeats in the genomes. However, in re-sequencing of human genomes with short reads, structural variations may be missed or be ambiguous because of the need to compare to a reference without the structural variants. A combination of rapid advances in Pacific Biosciences long-read sequencing chemistry and improved assembly methods are converging to make very high quality whole genome assemblies possible. We recently sequenced multiple strains of S. cerevisiae and S. pombe. By using the BluePippin (Sage Sciences) platform to produce long fragment libraries and the P5 enzyme / C3 chemistry combination, we obtained over 80x coverage of reads exceeding 10,000 bp, and maximum read lengths extending beyond 35,000 bp. We are able to produce near-perfect assemblies of the genomes such that every chromosome was assembled into a single contig or a very small number of contigs. We also made an assembly of the IR64 strain of rice (430Mbases genome) using only Pacific Biosciences long read data. The N50 contig size of this assembly is 600kb, doubling our previous assembly N50 obtained with a mix of Illumina short read and PacBio long reads.

1618M

Transposases are enzymes that facilitate movement (transposition) of mobile genetic elements (transposons) within and between genomes. High frequency transposition is believed to be detrimental to survival as transposases encoded in prokaryotic and eukaryotic genomes are notably inactive. For example, the best characterized transposase, hyperactive Tn5, required extensive engineering before it could be applied to in vitro DNA manipulations, such as cloning, mutagenesis and next generation sequencing library preparation. Based on our knowledge of Tn5, we were surprised to find that a transposase encoded in the genome of bioluminescent marine bacterium Vibrio harveyi is at least as active in vitro as hyperactive Tn5 transposase. Unlike Tn5, Vibrio harveyi transposase was expressed in E.coli to a high level of about 25% of soluble protein of total cell lysate. We purified the enzyme to homogeneity, proven it to be stable at -20°C for over 3 years and determined that it could be immobilized to a solid support. Vibrio harveyi transposase is included in the SureSelectX Oxford library prep kit. The method of this kit generates consistent fragmentation and uniform sequencing coverage from 50ng of genomic DNA compared to traditional time-consuming adapter-ligation methods that typically require 1-3ug input. In less than 1.5 hours (with only 30 minutes hands-on time) and using an automation-friendly workflow, one can prepare dual-indexed libraries for sequencing on the Illumina HiSeq or MiSeq instruments. Advantages of this transposase as a model for transposition studies and in applications for single-cell analysis, microfluidics and microarrays are discussed.

1619T

Assays of high sensitivity and minimal turnaround time that generate unambiguous results from minimal sample input are becoming essential. Coupled with the single-molecule resolution afforded by next-generation sequencing, target enrichment provides a cost-effective solution to enable confident analysis of various genomic targets by increasing assay sensitivity. In this study, we enable an accelerated workflow to generate exome-enriched libraries while providing uncompromised variant calling performance. Briefly, fragmented and adapter-tagged DNA libraries are produced in a single step with only 50ng genomic DNA using Agilent’s transposase-based library prep technology. DNA libraries are then taken through a novel 1.5-hour fast hybrid process, that provides accelerated capture of targets using Agilent’s long 120-mer cRNA baits. Combining transposase-based library preparation with this novel in-solution capture technology enables the preparation of up to 16 dual-indexed (post-capture) libraries for sequencing on the Illumina HiSeq or MiSeq platforms in less than 8 hours. Optimized for transposase-generated libraries, SureSelect fast hybridization technology is fully automatable and can be used with bait libraries as small as 0.2Mb up to full exome + UTR libraries, achieving true “exome in a day” results. With this streamlined protocol, we routinely obtain high performance libraries with >80% on-target, >98% coverage at 1x depth, >75% at 20x depth, and <15% duplicates with the Human All Exon V5 and V5-UTR bait sets. Further performance metrics such as library complexity, SNP sensitivity and concordance will be presented for additional bait sets and competitor kits.
1620S

Whole exome sequencing (WES) has become a commonly used method for both research and clinical applications. The most widely used enrichment types are in-solution hybridization and amplification based methods for exome sequencing. WES allows the detection of causative genetic changes and candidate genes in Mendelian and complex disorders as well as germ-line genetic alterations in cancer. We report a comparison of solution based capture methods (Agilent SureSelect V4 and v5, NimbleGen SeqCap EZ 3.0, Illumina Nextera Rapid and Expanded Capture) as well as an amplification based method (Life Technologies Ion AmpliSeq Exome) and a restriction digest mediated capture method (Agilent Haloplex Exome). The kits included in this study are commonly used and up to date capture methods. All captured samples were sequenced on an Illumina HiSeq2000 except for the Ion AmpliSeq exome samples which were sequenced on a Life Technologies Ion Proton system. We used a well characterized family trio from the Center for Applied Genomics Biorepository and the Genome in a bottle (NIST) human HapMap sample NA12878 available from Coriell to compare enrichment and genetic variant metrics of the different methods. We found significant differences in target coverage, on target bases, uniformity of coverage, detected SNVs and indels, and trio discordance. While generally all methods included in this study performed well, our results have important implications and should be considered especially for applications of clinical nature.

1621M

The Center for Genome Science (CGS) in Korea National Institute of Health (KNIH) drew up the Korean Reference Genome project (KRG) in 2012. Total 622 DNA samples of study subjects for the KRG were obtained from three different resources. The whole genome sequencing was conducted by illumina HiSeq2000 sequencer. The sequencing reads trimmed by the Sickie trimming and mapped to human reference genome (hg19). The mapped sequences were analyzed by SAMtools to calling the variants. VCFtools were used to merged the population variant sites, and compute the variants stats. Annovar were used to understand the proportion of the Known variants in the dbSNP138 and 1000G and the variant positions of gene regions. Over 97% of the SNVs identified in each genome were found in dbSNP database or 1000 Genome Project variants. Over 55% of the indels in each genome were located in the 1000 Genome project. The insertion/deletion ratio was 0.75 and approximately 0.1% of the indels were located in the coding region. In this presentation, we will describe the sequencing subjects, raw sequencing stats, and variant calling results. The all variants presented in this paper can be found from our web-browser.

1622T
Resolving the ‘Dark Matter’ in Human Genomes through Long-Read Sequencing. J. Korchak. Pacific Biosciences, Menlo Park, CA.

Second-generation sequencing has brought about tremendous insights into the genetic underpinnings of biology. However, there are many functionally and medically important regions in the genome that are too complex to represent by a single path or at which there are variants that have sequence not represented on the chromosome. We report here the reference assembly GRCh37 (hg19) contained just 9 alternate loci scaffold at 3 genomic regions. Released in December 2013, GRCh38 (hg38), the current reference assembly, now contains 261 alternate loci scaffolds at 178 distinct genomic regions. These alternate loci provide the only reference representation of more than 3 Mb of sequence, including more than 100 genes. We will present analyses showing how use of alternate loci in alignment target sets can reduce off-target alignments, which may improve variant calling and reveal a need for the development of alignment and variant calling tool chains that make use of the full reference assembly. We will highlight other key improvements in GRCh38, including base updates, sequence additions and the use of single haplotype resources to refile complex regions, which corrected errors and improved annotation. We will describe modeled centromere sequences that have been added to the chromosomes and are expected to serve as mapping targets for α-satellite DNA and benefit studies of human variation, epigenomic regulation and centromere biology. We will present tools and resources developed at NCBI for migrating data to and utilizing GRCh38 and discuss the evolution of the new reference genome assembly as we enter the era of personal genomics.
Third generation sequencing and analysis of complete mitochondrial genomes. E.P. Hoffman1, K.B. Getting2, K. Kiesler2, P.M. Vallone2, L.B. Davenport1, S. Dodgari, K. Panchapakesan1, S. Knoblach1, J.M. Devaney1.

1) Research Ctr Genetic Medicine, Children’s Natl Medical Ctr, Washington, DC; 2) National Institute of Standards and Technology, Gaithersburg, MD.

Next-generation sequencing (NGS) has enhanced investigators’ ability to conduct biomedical research examining genetics. Mitochondrial genetics has benefitted from NGS because technology now allows for screening of all 16,569 base pairs of the mitochondrial genome simultaneously for single nucleotide polymorphisms and low-level heteroplasmy. We utilized Single-Molecule Real Time (SMRT) sequencing on a PacBio RS II (Pacific Biosciences) to sequence NIST Standard Reference Material (SRM) 2392 component A (lymphoblastoid cell culture line CHF) and component B (lymphoblastoid cell culture line 9947A), and SRM 2392-I (promyelocytic cell line HL-60). These SRMs are intended to provide quality control when sequencing of human mitochondrial DNA (mtDNA) for forensic identification, medical diagnosis, phylogenetics, and anthropology. We amplified with three overlapping PCR primers sets (6kb PCR products) the entire mtDNA sequence. Then we processed the three mtDNA SRMs using a DNA Template Prep Kit 2.0 (3kb to 10kb), Polymerase Binding Kit P4, and DNA Sequencing Kit 2.0. SMRTbell libraries were constructed from pooled, full-length 6kb PCR amplicons tiling the entire 16.6kb mtDNA genome. Each pooled sample was placed on SMRT cell V3 and a 180-minute movie was performed. The data was processed using SMRT Portal with the revised Cambridge Reference Sequence (rCRS). Average read lengths for each SRM were around 3,300 bp with an average coverage of 6,673X. For 2392 component A, we discovered 36 variations from the rCRS including three insertions of C’s at positions 309.1, 315.1, and 16193.1. For 2392 component B and 2392-I, we were able to detect 20 variations and 33 variations, respectively, in each SRM from the rCRS including insertions of C’s at mtDNA positions 309.1 and 315.1. The number of variants discovered in the three SRMs using the PacBio RS II matched previously published data. We are now investigating methods of analysis to detect heteroplasmy that is occurring in these mtDNA SRMs. This study shows that a third-generation sequencing system can be used to generate entire mtDNA sequences using a three-primer system for amplification. The current NGS approach used in this abstract has the potential to be beneficial in a forensic setting or for medical diagnostics.

The International Fanconi Anemia Registry (IFAR) was established at The Rockefeller University in 1982 to serve as a longitudinal phenotyping system, source of DNA for genotyping studies, and a biorepository. In 1998, a separate public database was created to curate variants in all known FA genes, now in LOVD v.2.0 (www.rockefeller.edu/fanconi/mutate/). Here we describe our newest web-based application whereby clinical and molecular data for all 16 currently identified FA genes will be able to be queried via the internet. The ultimate goal of this new IFAR application is to help move FA research forward by collecting as much genotypic and phenotypic data as possible, displaying these data in a comprehensible, user-friendly manner, and making the data set available to other FA clinicians and researchers. The enhanced IFAR application uses Java for the Application Server programming, chosen because it is fast, secure and reliable. We are using Oracle as the database to store and retrieve related information. The web interface for the client was developed using JSP, HTML and JavaScript, allowing for multiple users to access the system concurrently. The special FA ontology was developed using Protégé for data modeling with the files in OWL format. The “Fanconi Anemia Ontology” is publically available in the NCBO BioPortal (http://bioportal.bioontology.org/ontologies/IFAR?p=summary). This Ontology was created using an OWL file provided by Dr. Ada Hamish and Francois Schiettecatte at the Centers for Mendelian Genetics, with their permission. Novel classes pertaining to FA were added and modifications were made using HPO, OMIM, NCI, and SNOMED. The ontology, while intended for the IFAR specifically, can be easily modified for other disease systems. We are currently migrating all of the clinical and molecular data from the original IFAR database created in FileMaker Pro, to this new ontology-based application. After this is accomplished, access to the de-identified data will be available to others, following rules set by the Rockefeller University Hospital regarding patient privacy. There are currently more than 1300 FA patients with this rare disease enrolled in the IFAR. The new display screens and navigational tools make the application easy to use, and should eventually enable researchers and physicians worldwide to learn more about genotype/phenotype correlations in FA and related diseases, and the role of the FA proteins in DNA damage repair and cancer.

RD-Connect platform and standardized exome-phenome analysis pipeline: application to 20 use cases. S. Beltran1, D. Salgado2,14, V. de la Torre1,3,7,13, D. Paschall1, S. Laurin1, J.P. Desveignes2,3, A. Top2, M. Calissano12, L. Zaharieva6, F. Muntoni10, M. Roos7, P. Lopes10, M. Girdea11, C. Kingswood1, M. Vázquez5, J.M. Fernández González2, R. Thompson7, M. Brundo11, V. Straub7, H. Lochmüller7, M. Bellgard12, J.L. Oliveira12, P.A.C. Hobden9, A. Valencia5, C. Béroud6, I. Gut1. 1) Centro Nacional de Análisis Genómico (CNAG), Barcelona, Catalonia, Spain; 2) Aix-Marseille Université, Marseille, France; 3) Inserm, UMR S 910, Marseille, France; 4) EMBL Australia, Australian Regenerative Medicine Institute (ARMI), Monash University, Clayton, Victoria, Australia; 5) Centro Nacional de Investigaciones Oncológicas, Madrid, Spain; 6) European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge, United Kingdom; 7) Institute of Genetic Medicine, MRC Centre for Neuromuscular Diseases, Newcastle University, UK; 8) Dubowitz Neuromuscular Centre, UCL Institute of Child Health and Great Ormond Street Hospital for Children, London, United Kingdom; 9) Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; 10) CETIB/IEETA, University of Aveiro, Portugal; 11) Centre for Computational Medicine, Hospital for Sick Children and University of Toronto, Toronto, ON, Canada; 12) Centre for Comparative Genomics, Murdoch University, Perth, Western Australia; 13) APHM, Hôpital TIMONE Enfants , Laboratoire de Génétique Moléculaire, Marseille, France.

Around 300 million people worldwide are estimated to suffer from one of the 6000+ known rare diseases. Rare disease research faces particular challenges because patient populations, clinical expertise, and research communities are small in number and highly fragmented both geographically and in terms of medical specialty. Therefore, discovery is slowed down due to limited access to sufficient patients, high quality information and results and would benefit from greater coordination of patient registries, biobanks, data repositories, and best practices in deployment of new -omics technologies and analysis methods. The EU FP7-funded RD-Connect project is building a platform to harmonize and securely integrate databases, registries, biobanks, clinical bioinformatics and -omics data generated with standardized pipelines. Platform design has consisted of a formal use-case analysis process to define clinical and research use cases that could benefit from RD-Connect infrastructure. The use cases were then converted into data models, integrated in the platform, and would benefit from greater coordination of patient registries, biobanks, data repositories, and best practices in deployment of new -omics technologies and analysis methods. The EU FP7-funded RD-Connect project is building a platform to harmonize and securely integrate databases, registries, biobanks, clinical bioinformatics and -omics data generated with standardized pipelines. Platform design has consisted of a formal use-case analysis process to define clinical and research use cases that could benefit from RD-Connect infrastructure. The use cases were then converted into data models, integrated in the platform, and would benefit from greater coordination of patient registries, biobanks, data repositories, and best practices in deployment of new -omics technologies and analysis methods.

The International Fanconi Anemia Registry (IFAR) was established at The Rockefeller University in 1982 to serve as a longitudinal phenotyping system, source of DNA for genotyping studies, and a biorepository. In 1998, a separate public database was created to curate variants in all known FA genes, now in LOVD v.2.0 (www.rockefeller.edu/fanconi/mutate/). Here we describe our newest web-based application whereby clinical and molecular data for all 16 currently identified FA genes will be able to be queried via the internet. The ultimate goal of this new IFAR application is to help move FA research forward by collecting as much genotypic and phenotypic data as possible, displaying these data in a comprehensible, user-friendly manner, and making the data set available to other FA clinicians and researchers. The enhanced IFAR application uses Java for the Application Server programming, chosen because it is fast, secure and reliable. We are using Oracle as the database to store and retrieve related information. The web interface for the client was developed using JSP, HTML and JavaScript, allowing for multiple users to access the system concurrently. The special FA ontology was developed using Protégé for data modeling with the files in OWL format. The “Fanconi Anemia Ontology” is publically available in the NCBO BioPortal (http://bioportal.bioontology.org/ontologies/IFAR?p=summary). This Ontology was created using an OWL file provided by Dr. Ada Hamish and Francois Schiettecatte at the Centers for Mendelian Genetics, with their permission. Novel classes pertaining to FA were added and modifications were made using HPO, OMIM, NCI, and SNOMED. The ontology, while intended for the IFAR specifically, can be easily modified for other disease systems. We are currently migrating all of the clinical and molecular data from the original IFAR database created in FileMaker Pro, to this new ontology-based application. After this is accomplished, access to the de-identified data will be available to others, following rules set by the Rockefeller University Hospital regarding patient privacy. There are currently more than 1300 FA patients with this rare disease enrolled in the IFAR. The new display screens and navigational tools make the application easy to use, and should eventually enable researchers and physicians worldwide to learn more about genotype/phenotype correlations in FA and related diseases, and the role of the FA proteins in DNA damage repair and cancer.

The International Fanconi Anemia Registry (IFAR) was established at The Rockefeller University in 1982 to serve as a longitudinal phenotyping system, source of DNA for genotyping studies, and a biorepository. In 1998, a separate public database was created to curate variants in all known FA genes, now in LOVD v.2.0 (www.rockefeller.edu/fanconi/mutate/). Here we describe our newest web-based application whereby clinical and molecular data for all 16 currently identified FA genes will be able to be queried via the internet. The ultimate goal of this new IFAR application is to help move FA research forward by collecting as much genotypic and phenotypic data as possible, displaying these data in a comprehensible, user-friendly manner, and making the data set available to other FA clinicians and researchers. The enhanced IFAR application uses Java for the Application Server programming, chosen because it is fast, secure and reliable. We are using Oracle as the database to store and retrieve related information. The web interface for the client was developed using JSP, HTML and JavaScript, allowing for multiple users to access the system concurrently. The special FA ontology was developed using Protégé for data modeling with the files in OWL format. The “Fanconi Anemia Ontology” is publically available in the NCBO BioPortal (http://bioportal.bioontology.org/ontologies/IFAR?p=summary). This Ontology was created using an OWL file provided by Dr. Ada Hamish and Francois Schiettecatte at the Centers for Mendelian Genetics, with their permission. Novel classes pertaining to FA were added and modifications were made using HPO, OMIM, NCI, and SNOMED. The ontology, while intended for the IFAR specifically, can be easily modified for other disease systems. We are currently migrating all of the clinical and molecular data from the original IFAR database created in FileMaker Pro, to this new ontology-based application. After this is accomplished, access to the de-identified data will be available to others, following rules set by the Rockefeller University Hospital regarding patient privacy. There are currently more than 1300 FA patients with this rare disease enrolled in the IFAR. The new display screens and navigational tools make the application easy to use, and should eventually enable researchers and physicians worldwide to learn more about genotype/phenotype correlations in FA and related diseases, and the role of the FA proteins in DNA damage repair and cancer.
PATHWAY APPROACHES TO STRENGTHEN GENETIC VARIATION ANALYSIS.

Introduction. Pathway and network analysis is used to understand genomics data from transcriptomics, proteomics and metabolomics experiments. It uses prior knowledge of biological relationships and can evaluate related effects that strengthen each other. Genetic variants analysis (such as SNPs, indels...) would also benefit from this approach because of the evaluation of related variations in different genes, and integration of variant data with other omics results. Currently, genetic variations cannot easily be combined in pathway representations. It is also not clear how to better visualize and interpret variation data once we connected it to pathway content programmatically. In this paper, we describe a way to integrate genetic information in the open source pathway analysis program PathVisio (pathvisio.org). Design and analysis ideas. We connected variants with related genes, using bioinformatic mapping approaches (see bridgeDb.org). With this we can visualize variants on a pathway. We now aim for meaningful presentations showing some design studies: - Variants information for selected genes is shown in a separate panel of PathVisio page. - Genetic variations with statistical values from large genetic datasets (e.g. GWAS) can be highlighted for the annotated variants (colors in the panel and clickable values). - Specific information about variants like HapMap environment or function predictions can be shown as popups or via linkouts to genome browsers. We will also map variants and genes for statistical approaches using larger sets of biological pathways (e.g. from WikiPathways), to find those where many variants occur in specific phenotypes. This is comparable to the over representation and gene set enrichment analyses. Finally, we will investigate the availability of existing advanced tools like sift and polyphen2 to integrate predictions of biological functional changes. Pathways as a bridge to networks. Pathway and network biology can be easily combined: our tools convert one into the other and we can use the same mappings between database identifiers in both. Variations are evaluated in pathway as networks, increasing network extensions (e.g. with miRNA or drugs) and topology studies. This will combine network connectivity, betweenness and shortest path analysis with the occurrence of deleterious variants in disease phenotypes.

PATHWAY APPROACHES TO STRENGTHEN GENETIC VARIATION ANALYSIS.

- C. Bizon
- S.E. Plon
- L.J. Hayek
- H. Chen
- A. Jackson
- S. Hayek
- L. Babb
- C. Bizon
- D. Maglott
- E.M. Ramos
- H.L. Rehm
- J.M. Cherry
- C. Bustamante
- S.E. Pion
- A. Miloavijevic
- J. Marchini
- P. Sneddon


Random effect models have been used in genetics to control for aggregate genetic structure in association studies, estimate heritability and predict the latent genetic contribution to phenotypic variation. Unfortunately, the ordinary random effect model does not generalize well to multivariate phenotypes; the resulting high-dimensional parameter estimates are both statistically unstable and computationally intractable. Moreover, missing observations plague multivariate data, as typical methods must throw away all samples with even one missing phenotype. Motivated by these concerns, we propose a new random effect model to decompose genetic and environmental contributions to multivariate phenotypic variance. The method naturally incorporates missing data and, in particular, can be used to impute unobserved phenotypes for subsequent analyses. In terms of imputation, this new method outperforms all tested competitors, including standard univariate methods from genetics and state-of-the-art imputation tools from mainstream statistics, on both simulated data and real human glucocorticoid data. A second application is to estimate the latent genetic contribution to phenotypes, or breeding value of study individuals. On simulated data and real data from chicken breeding, the proposed method produces the best estimates of the future breeding values. The model is fit with variational Bayes, enabling this method to be run on thousands of individuals and tens of phenotypes.

ClinGen database for curation of clinically relevant genomic variants.

- X. Feng
- T.P. Sneddon
- H. Chen
- A.R. Jackson
- S. Plankenhorn
- L.J. Babb
- C. Bizon
- D. Maglott
- E.M. Ramos
- H.L. Rehm
- J.M. Cherry
- C. Bustamante
- S.E. Pion
- A. Miloavijevic
- J. Marchini
- P. Sneddon


Random effect models have been used in genetics to control for aggregate genetic structure in association studies, estimate heritability and predict the latent genetic contribution to phenotypic variation. Unfortunately, the ordinary random effect model does not generalize well to multivariate phenotypes; the resulting high-dimensional parameter estimates are both statistically unstable and computationally intractable. Moreover, missing observations plague multivariate data, as typical methods must throw away all samples with even one missing phenotype. Motivated by these concerns, we propose a new random effect model to decompose genetic and environmental contributions to multivariate phenotypic variance. The method naturally incorporates missing data and, in particular, can be used to impute unobserved phenotypes for subsequent analyses. In terms of imputation, this new method outperforms all tested competitors, including standard univariate methods from genetics and state-of-the-art imputation tools from mainstream statistics, on both simulated data and real human glucocorticoid data. A second application is to estimate the latent genetic contribution to phenotypes, or breeding value of study individuals. On simulated data and real data from chicken breeding, the proposed method produces the best estimates of the future breeding values. The model is fit with variational Bayes, enabling this method to be run on thousands of individuals and tens of phenotypes.

ClinGen database for curation of clinically relevant genomic variants.


Random effect models have been used in genetics to control for aggregate genetic structure in association studies, estimate heritability and predict the latent genetic contribution to phenotypic variation. Unfortunately, the ordinary random effect model does not generalize well to multivariate phenotypes; the resulting high-dimensional parameter estimates are both statistically unstable and computationally intractable. Moreover, missing observations plague multivariate data, as typical methods must throw away all samples with even one missing phenotype. Motivated by these concerns, we propose a new random effect model to decompose genetic and environmental contributions to multivariate phenotypic variance. The method naturally incorporates missing data and, in particular, can be used to impute unobserved phenotypes for subsequent analyses. In terms of imputation, this new method outperforms all tested competitors, including standard univariate methods from genetics and state-of-the-art imputation tools from mainstream statistics, on both simulated data and real human glucocorticoid data. A second application is to estimate the latent genetic contribution to phenotypes, or breeding value of study individuals. On simulated data and real data from chicken breeding, the proposed method produces the best estimates of the future breeding values. The model is fit with variational Bayes, enabling this method to be run on thousands of individuals and tens of phenotypes.

Gene variant databases (LSDBs) using the LOVD platform (Leiden Open-source Variation Database, http://www.LOVD.nl) collect and share information about genes, variants and phenotypes (diseases) from research and diagnostic labs. LOVD version 3 also facilitates exome and genome sequence data analysis. LOVD can be queried and adapted for many different purposes. The Dutch and Belgian working group for Breast Cancer DNA Diagnostics (LOD) has decided to share >7500 variants detected in the BRCA1/2 genes in breast cancer families since 1997. For this, the data, almost evenly split among both genes, have been submitted to the LOVD and shared gene variant database installation (1,2). Others include: world-wide BRCA variant data sharing (Human Variome Project and Global Alliance for Genomics and Health); Insight Consortium colon cancer variant database; country-specific data views (Finnish Disease (FinDis) portal). LOVD system advantages: simple standardized submission of new data, instant updates after curation, easy maintenance and automatic backups. Although most data are publicly accessible online, some data (detailed phenotype information) are shared by consortium members only using the new LOVD access level, designated “collaborator.” Others can see whether such information is available (password protected file links), giving them the option to contact the submitter for further details. Members can contribute their opinions about variant classification, increasing its consistency, but being aware of potential misinterpretation they have reservations sharing this information. Data are stored variant-by-variant and connected to each individual patient and submitting diagnostic lab. Users can perform queries per gene or individual, use other linked resources of interest, view data tracks in genome browsers and use web services to access variants stored in other gene variant databases. For databases curated by us both phenotype descriptions and/or gene variants can be submitted with the request to assign the so-called VIP-status for both variants and phenotypes demanding specific attention. Within 3 weeks, this match-making feature successfully brought researchers into contact, cracking rare disease cases and resulting in a high-impact publication in a high impact journal. LOVD is used at many major cancer hospitals in Europe and the US, including MD Anderson and Dana Farber in Boston, and many more in the Netherlands. LOVD has granted LOVD the recommended system status for variant collection.


This talk will discuss challenges and recent developments in implementing informatics infrastructure to support clinical use of genomics, specifically in the area of oncology. While much focus has been placed on the ability to properly call variants, the “last mile” in taking processed variant calls and producing a useful, annotated clinical report with therapeutic recommendations is critical for enabling the clinical use of genomics data. Generating, delivering, and updating a clinical genomic sequencing report that a physician can use in routine care has proven to be a challenge, driven by reliance on outdated software tools. To bring genomics data to point of care, labs and clinics must be able to handle three types of data with very different properties and requirements: evolving biomedical knowledge, complex medical data, and high-volume genomics data. This talk will present a software solution, based on semantic computing principles, that enables the integration of all three data types for the purposes of clinical reporting. We will present the results of clinical pilots focused on maintenance of an evolving variant interpretation knowledge base, sharing of variant interpretation across institutions, and the use of this knowledge base to automate clinical genome report generation and updating, as knowledge changes. We will describe the interpretation of identified variants through a shared knowledge base, automated assembly of annotated clinical reports, and delivery of reports to clinicians at point of care. Results of a pilot program to integrate these technologies into practice within an integrated community health care system will be presented. We will also present the results of a physician adoption of a web-based, interactive clinical genome report format that incorporates clinical care guidelines with genomic report data.

Functional interpretation of noncoding somatic variants from cancer genomes. E. Khurana, Y. Fu, M. Gerstein. Yale University, New Haven, CT.

Whole-genome sequencing generally reveals thousands of somatic variants in individual tumors. Identification of key variants responsible for tumor growth and progression is a major challenge. This is especially the case for noncoding variants that are less well understood than coding variants. Discovery of somatic mutations in telomerase reverse transcriptase (TERT) promoter across many different cancer types has shown that regulatory variants may constitute driver events. However, most current methods for identification of cancer driver mutations focus primarily on protein-coding genes. We previously developed a method for identification of candidate noncoding drivers by integrating patterns of selection from 1000 Genomes data with functional annotations from ENCODE (FunSeq). Besides analyzing loss of transcription factor binding sites, our method now identifies creation of new binding sites by somatic mutations in promoters and enhancers of coding genes. We have also integrated function-based FunSeq scoring scheme with patterns of recurrence of somatic variants across hundreds of patients from multiple studies. Tumors exhibit differential expression signatures of thousands of genes. Our unified approach allowed us to explore the fraction of gene expression changes that may be explained by regulatory mutations vs. mutations in coding genes across many different cancer types.


The Human Variation Database (http://gwas.biociscencedbc.jp/cgi-bin/ hvdv/hv_top.cgi) is a repository database set up to achieve continuous and intensive management of Japanese GWAS data and variations identified by NGS and other experimental methods. We have widely called for data submission to aim for data-sharing among researchers. Variations include short/long insertions and deletion and structural variations as well as SNPs/ SNVs. Since vast amount of knowledge about disease-variation relationships are varied in the scientific literature, these data have been extracted by manual curation to organize the disease-related variations and improve our understanding of disease mechanisms and disease-disease relationships. More than 30,000 disease-related variation entries are currently registered with allele/genotype frequencies and statistical data. Pathway data are also provided to facilitate the understanding of epistasis and relationships between similar diseases. Since some disease-related variations are recognized to be different among populations, the reference genomes are built in each population and are registered. The recent improvements of the database are as follows. 1) The scope of the registry was extended to include drug response-related variations and virus resistance-related variations. 2) HLA-DB was constructed because of the characteristics of the HLA region. 3) Links to driver mutations archived in widely used cancer somatic mutation databases were added in order to clarify the relationships between the driver mutations in various cancers and germ-line mutations, although only variations in germ-line cells are targeted in the database. In this presentation, we overview the database structure and discuss the differences of deleterious variations between germ-line cells and somatic cells from the view points of physicochemical properties, functional properties, and evolutionary conservation properties.

Acknowledgement: This work was supported by the contract research fund ‘Database Integration Program’ from the Japan Science and Technology Agency.
1639M Exploring the genome-wide roles of transcription factors and their complexes in chromosome interaction. MJ. Li1,2, L.Y. Wang, PC. Sham1, MO. Zhang1, JS. Li4, JW. Wang1. 1) Biochemistry, The University of Hong Kong, Hong Kong; 2) Statistics Dept, Harvard University, Boston, MA, USA; 3) Psychiatry Dept, The University of Hong Kong, Hong Kong; 4) Molecular and Cell Biology Dept., The University of Texas at Dallas, Dallas, TX, USA.

Huntington’s Disease (HD) is a neurodegenerative disease with the most prominent pathology in the brain. However, human brain tissue is not easily accessible for molecular biology studies and it cannot be isolated from living patients. The widespread pathology of HD indicates that the phenotype is not limited to brain dysfunction. Several studies have shown that understanding alterations in peripheral tissues could be valuable for monitoring disease progression. Considering that transcriptome technologies have successfully been used for biomarker discovery and the study of physiological and pathophysiological mechanisms, it is evident that studying transcriptional changes in peripheral tissue can provide new insights that can lead to the development of new therapies and markers to monitor disease progression. In addition, novel sequencing technologies and the wide availability of data and published articles, create opportunities for novel, more effective ways of data exploitation. Our methodology identifies and prioritizes disease signatures from blood that are also associated with abnormalities in brain, in a robust way by a combination of large scale network analysis, integration of heterogeneous datasets, and incorporation of prior knowledge mined from literature. Our preliminary data show that modules in gene co-expression networks in different human cell types. Existing studies on analyses of TFs and their complexes were only performed at one dimension and not at genome-wide scale. Recently, the unbiased chromosome conformation capture, Hi-C, can detect the genome-wide chromatin interactions, but has restrictions on resolution due to the variation in cell-to-cell chromatin structures and inadequate sequencing depth. In this study, we provide a comprehensive analysis on TFs regulatory pattern within chromosome looping by combining Hi-C and ENCODE ChIP-Seq data from three human cell types (GM12878, H1-HEK293, and K562). We first devised a strategy to map ChIP-Seq peaks of each TF to a normalized 10kb Hi-C contact matrix and construct an interaction matrix for each participant TF. We observed tight correlation for TFs participant activities in high resolution chromosome looping between biological replicates, which indicate the TF activities is more stable than local DNA interactions. To check the enrichment of different chromatin marks and genomic features in the interaction region of each participant TF, we performed enrichment test on several histone modifications marks (H3K4me1, H3K4me2, H3K4me3, H3K27me3, H3K27me3, H3K9Ac, H3K9Ac, H3K36me3, H3K9Ac, H3K9Ac, and H3K27ac), DNA methylation and chromatin accessibility signal. We found chromatin markers and genomic features of each TF is highly correlated between replicates. We also observed that grouped TFs shared consistent patterns of chromatin marks and genomic features, which indicate their similar roles in gene regulation. To explore long range enhancer-promoter interactions in different cell types. Using the combinatorial motifs scanning, we can predict the genome-wide CRMs for specific TF complex with known motifs. Finally, by correlating gene expression profiles of TF and its targets, we can pinpoint the role of detected TF complexes (active or repressive) in controlling cell type specific gene regulation.

1640T Comparing brain and blood gene expression networks in Huntington’s Disease by semantic analysis. E. Mijania, A.J. Heen, W. van Roon, M. Roos. Department of Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands.

Huntington’s Disease (HD) is a neurodegenerative disease with the most prominent pathology in the brain. However, human brain tissue is not easily accessible for molecular biology studies and it cannot be isolated from living patients. The widespread pathology of HD indicates that the phenotype is not limited to brain dysfunction. Several studies have shown that understanding alterations in peripheral tissues could be valuable for monitoring disease progression. Considering that transcriptome technologies have successfully been used for biomarker discovery and the study of physiological and pathophysiological mechanisms, it is evident that studying transcriptional changes in peripheral tissue can provide new insights that can lead to the development of new therapies and markers to monitor disease progression. In addition, novel sequencing technologies and the wide availability of data and published articles, create opportunities for novel, more effective ways of data exploitation. Our methodology identifies and prioritizes disease signatures from blood that are also associated with abnormalities in brain, in a robust way by a combination of large scale network analysis, integration of heterogeneous datasets, and incorporation of prior knowledge mined from literature. Our preliminary data show that modules in gene co-expression networks in different human cell types. Existing studies on analyses of TFs and their complexes were only performed at one dimension and not at genome-wide scale. Recently, the unbiased chromosome conformation capture, Hi-C, can detect the genome-wide chromatin interactions, but has restrictions on resolution due to the variation in cell-to-cell chromatin structures and inadequate sequencing depth. In this study, we provide a comprehensive analysis on TFs regulatory pattern within chromosome looping by combining Hi-C and ENCODE ChIP-Seq data from three human cell types (GM12878, H1-HEK293, and K562). We first devised a strategy to map ChIP-Seq peaks of each TF to a normalized 10kb Hi-C contact matrix and construct an interaction matrix for each participant TF. We observed tight correlation for TFs participant activities in high resolution chromosome looping between biological replicates, which indicate the TF activities is more stable than local DNA interactions. To check the enrichment of different chromatin marks and genomic features in the interaction region of each participant TF, we performed enrichment test on several histone modifications marks (H3K4me1, H3K4me2, H3K4me3, H3K27me3, H3K27me3, H3K9Ac, H3K9Ac, H3K36me3, H3K9Ac, H3K9Ac, and H3K27ac), DNA methylation and chromatin accessibility signal. We found chromatin markers and genomic features of each TF is highly correlated between replicates. We also observed that grouped TFs shared consistent patterns of chromatin marks and genomic features, which indicate their similar roles in gene regulation. To explore long range enhancer-promoter interactions in different cell types. Using the combinatorial motifs scanning, we can predict the genome-wide CRMs for specific TF complex with known motifs. Finally, by correlating gene expression profiles of TF and its targets, we can pinpoint the role of detected TF complexes (active or repressive) in controlling cell type specific gene regulation.

1641S Genetic Risk Prediction and Neurobiological Understanding of Alcoholism. A. Niculescu1,2, D. Levey1, H. Le-Niculescu1, J. Frank1, M. Ayalawi2, N. Jain3, B. Kirlin1, R. Learman4, E. Winiger5, Z. Rodi5, A. Shekh5, N. Schork6, F. Kiefer7, N. Wodarz8, B. Muller-Myhsok9, N. Dahmen9, M. Nothen10, R. Sherva10, L. Farrer10, A. Smith10, H. Kranzler10, M. Rietschel10, J. Gelehrter10, GESGA Consortium. 1) Psychiatry, Indiana Univ Sch Medicine, Indianapolis, IN; 2) Central Institute of Mental Health, Mannheim, Germany; 3) Department of Human Biology, The J. Craig Venter Institute, La Jolla, California, USA; 4) Dept. of Addictive Behavior and Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim / Heidelberg University, Germany; 5) Dept. of Psychiatry, Psychiatry University Medical Center Regensburg, Univ. of Regensburg, Germany; 6) Dept. of Statistical Genetics, Max-Planck-Institute of Psychiatry, Munich, Germany; 7) Dept. of Psychiatry, Univ. of Mainz, Germany; 8) Boston University School of Medicine, Department of Medicine (Biomedical Genetics); 9) Yale University School of Medicine, Department of Psychiatry, Division of Human Genetics; and VA CT Healthcare Center; 10) Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, and Philadelphia VAMC. We have used a translational convergent functional genomics (CFG) approach to discover genes involved in alcoholism, by gene-level integration of genome-wide association study (GWAS) data from a German alcohol dependence cohort with other genetic and gene expression data, from human and animal model studies, similar to our previous work in bipolar disorder and schizophrenia. A panel of all the nominally significant p-value SNPs in the top candidate genes from both the GWAS and a recently completed transcription factor binding site (CRMs) in different human cell types. Existing studies on analyses of TFs and their complexes were only performed at one dimension and not at genome-wide scale. Recently, the unbiased chromosome conformation capture, Hi-C, can detect the genome-wide chromatin interactions, but has restrictions on resolution due to the variation in cell-to-cell chromatin structures and inadequate sequencing depth. In this study, we provide a comprehensive analysis on TFs regulatory pattern within chromosome looping by combining Hi-C and ENCODE ChIP-Seq data from three human cell types (GM12878, H1-HEK293, and K562). We first devised a strategy to map ChIP-Seq peaks of each TF to a normalized 10kb Hi-C contact matrix and construct an interaction matrix for each participant TF. We observed tight correlation for TFs participant activities in high resolution chromosome looping between biological replicates, which indicate the TF activities is more stable than local DNA interactions. To check the enrichment of different chromatin marks and genomic features in the interaction region of each participant TF, we performed enrichment test on several histone modifications marks (H3K4me1, H3K4me2, H3K4me3, H3K27me3, H3K27me3, H3K9Ac, H3K9Ac, H3K36me3, H3K9Ac, H3K9Ac, and H3K27ac), DNA methylation and chromatin accessibility signal. We found chromatin markers and genomic features of each TF is highly correlated between replicates. We also observed that grouped TFs shared consistent patterns of chromatin marks and genomic features, which indicate their similar roles in gene regulation. To explore long range enhancer-promoter interactions in different cell types. Using the combinatorial motifs scanning, we can predict the genome-wide CRMs for specific TF complex with known motifs. Finally, by correlating gene expression profiles of TF and its targets, we can pinpoint the role of detected TF complexes (active or repressive) in controlling cell type specific gene regulation.
1642M

In the setting of a large genomics research center, we constantly have requirements to have up-to-date detailed workflow information as well as efficiency in querying available data in seamless and integrated way to help scientific discoveries. Instead of taking the conventional data warehouse approach to extract data from various sources into a centralized data store, which is both time consuming and hard to co-ordinate with the real-time workflow status, we use distributed database query facility to dynamically query different source databases. The resulting suite of web applications on top of the distributed real-time query layer is being referenced as Analyst Portal. Data modalities searched by Analyst Portal include service order intake, biorepository sample management, SNP-array genotyping, gene array expression data, next generation sequencing, DIoCM neuroimaging, phenotype data from hospital EMR system and from collaborators, catalog of genetic variants (SNP, CNV, indel). For individual modality, Analyst Portal supports three levels of detailed data queries: single subject information lookup, all records in the modality, and a search limited to an uploaded subject list. In case of multiple domains search, Analyst Portal supports cohort selection by combining user defined search terms from different modalities and generating complex distributed SQL query. For example, a query like *get a list of Caucasians, males only, 6-10 years old subjects who have at least 2 ADHD diagnoses over the past 3 years, and on at least one medication such as Adderal, Concerta or Ritalin, and have blood DNA samples in our biorepository, and already genotyped on Illumina or Affymetrix SNP chips, with not NGS samples sequenced yet* can be done real-time easily in Analyst Portal. Among 50 typical data queries scientists and researchers have used during previous research studies, the Analyst Portal has ability to handle 70-80% of them instantly. Here, we review the software design, performance tuning, user experience and areas for future optimization.

1643T
Phenotype terminologies in use for genotype-phenotype databases: A common core for standardization and interoperability. P.N. Robin-son¹, S. Aymé², L. Chanas², A. Hamosh³, A. Rathi⁴, International Consor-tium for Human Phenotype Terminologies. 1) Institute for Medical Genetics, Charité-Universitätsmedizin, Berlin, Germany; 2) INSERM, US14, Paris, France; 3) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD.

The community needs to be provided with terminology standards in order to achieve interoperability between databases intended for clinical research. Phenotypic information is crucial to interpret genomic rearrangements as well as high-throughput sequence data including whole exome and whole genome sequencing. The aim of our work was to develop a core terminology of phenotypes interoperable with all terminologies in current use. Relevant terminologies in use by different communities to describe phenomes were cross-referenced: PhenoDB (2846 terms), London Dysmorphology Data-base (LDDDB; 1318 terms), Orphanet (1243 terms), Human Phenotype Ontology (9855 terms, 22/08/2102), Elements of Morphology (AJMG; 423 terms), ICD10 (1230 terms), as well as medical terminologies in use: UMLS (7,957,179 distinct concept terms), SNOMED CT (>311,000 concepts), Mesh (26,853 concepts) and MedDRA (69,389 concepts). We established a strategy to compare the terminologies to find commonalities and differences, using ONAGUI as a tool to identify exact matches. The non-exact matches were verified manually by an expert. A core-terminology of 2,300 terms was derived and analyzed by a panel of experts (International Consortium for Human Phenotype Terminologies - ICHPT). The resulting consensual terminology will be freely available in a dedicated website (www.ichpt.org) and at the International Rare Disease Research Consortium (IRIDiC) website. Mappings from ICHPT to other terminologies will be given in order to enable interoperability between databases that use the various terminologies to annotate phenotypic data, with the goal of enabling data exchange between all major genetic databases.

1644S
Data exploration through stark visualizations in gene expression profile of down syndrome. J. Ruvalcaba⁵, P. Skinner⁶, D. Weaver⁷, N. Economidou⁸, D.G. Knowles⁹, E.G. Couto³. 1) Translational Medicine, PerkinElmer, Inc., Waltham, MA; 2) Bioinformatics, Integromics, Wauaukee, WI.

Life Scientists are increasingly challenged by huge datasets resulting from High-Throughput Experiments that contain an enormity of information. Making sense of the big data produced is essential to understanding complex biological systems and doing so without wasting unnecessary resources is key. Frequently this problem is addressed using different software solutions for each technology and handling many steps manually or with ad-hoc scripting. This generates problems when comparing results from different technologies in a reliable way and makes it difficult to represent the data for analysis when attempting to answer biological questions of interest. In this work we evaluate if the use of interactive visual exploration of the data can provide the researcher with the ability to better evaluate their data and extract relevant biological conclusions. By taking advantage of visual data representations and the human capacity to immediately recognize patterns in visual data, we demonstrate how this approach accelerates data exploration and helps to reveal information easily missed using the traditional data analysis paradigm. In the traditional approach the data visualization is only used as the end product of the analysis. The initial steps and the exploratory analysis use limited visual aids that are frequently highly summarized visualizations focused on a specific question; thus, important trends are easily missed when not specifically searched. In order to demonstrate how visualizations can help not only in the representation of the final results but also during the entire analysis process we have reanalyzed the expression data from a recent study published in Nature (Letourneau A. et al. Domains of genome-wide gene expression dysregulation in Down’s syndrome. Nature 2014 Apr 17;508(7496):345-50) together with other expression datasets related to Down Syndrome deposited in the GEO database. We show how by using interactive and visual exploration at the different steps of the analysis in combination with the standard tools used during the traditional data analysis we can reveal several interesting features of the data that can be easily missed and ignored when using only the conventional analysis & reporting.

1645M
Evaluating global enrichment of trait-associated variants in epigenomic features. E. Schmidt¹, J. Chen², W. Zhou³, J. Zhang⁴, E. Chen⁵, C. Willer⁶. 1) Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 2) Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI.

Genome wide association studies have identified hundreds of non-protein-coding variants with modest effect sizes, making alterations in transcript or protein expression difficult to detect. This well-associated genetic variation likely plays a role in transcriptional regulation via epigenomic regulatory features, a diverse set of which have been experimentally defined by ENCODE in various cell types. We develop GREGOR (Genomic Regulatory Elements Discovery tool) to address 3 aims: 1) elucidate the important tissue types in which genetic variation impacts transcription for a particular trait, 2) narrow our focus of the regulatory features underlying transcription disrupted by trait-associated variants, and 3) use positional overlap with selected regulatory domains to identify potential functional candidate variants at trait-associated loci.

Taking into account linkage disequilibrium, we evaluate enrichment of 5 sets of metabolic-related trait-associated variants in a range of regulatory features from different comparison tools. We prioritize individual non-coding variants at GWAS SNP with control SNPs based on 3 properties: i) number of variants in high LD, ii) minor allele frequency, and iii) distance to nearest gene. We find enrichment of aggregated sets of trait-specific variants in regulatory enhancer sites in biologically relevant tissue types. Examples include significant enrichment and fold change of variants associated with blood pressure in blood vessel (P=1×10⁻⁶; 1.5), body mass index in oiffactory neurosphere-derived cells (P=4×10⁻⁵; 1.7), coronary artery disease in heart (P=2×10⁻⁷; 1.7), lipid in liver (P=3×10⁻¹⁶; 1.6), and unexpectedly high in 2 diabetes in heart (P=0.4×10⁻⁷; 1.7). We further evaluate the position of index SNPs and their LD proxies related to histone marks and ChIP-seq TF binding sites, as well as functional chromatin states. We prioritize individual non-coding variants at GWAS loci and experimentally test candidate SNPs at 5 GWAS lipid loci using a luciferase reporter to measure transcript activity in human HepG2 cells. In all, our tool systematically employs the wealth of regulatory data available from ENCODE to establish known biological connections with trait-associated variation, find potentially novel biological insights, and guide experimental follow-up to identify functional variants at GWAS loci.
A bioinformatic protocol for the study of rare diseases. F. Tobar Tose, E. Ocampo Toro, P. Hurtado Villa. Department of basic Science for Health, Pontificia Universidad Javeriana Cali, Cali, Valle del Cauca, Colombia.

Introduction: Rare diseases (RDs) involve several genetics and molecular mechanisms, which describe diverse and specific phenotypes and conditions; which difficult the concentration of efforts to research and propose new therapies for the patients. However, there is a relevant feature in these kind of diseases; that is the genetic context, where a single variant in a gene could define the whole context of the disease, for example, in the lysosomal storage diseases. Additionally, there is a very important fact, that is the consequences in the genome; these mean, although several mechanisms are associated to each RDs, these can be generally described as aberrations of the human genome, whose organization is disrupted and therefore the dynamic is changing too. Results and discussion: Accordingly, at the present work we present a bioinformatics workflow for the association of genomic information related with RDs. We improved and applied methods of data mining and nodes theory, considering the phenotypic and genetics descriptions for several RDs. This approach allow us to identify, molecular and genomics associations among several rare diseases. Interestingly, physiological and phenotypic descriptions could be associated to the combination of genomic patterns, and could describe focused of genomic susceptibility. Additionally, it was associated to affection of critical metabolic or physiological processes. For example, we identify several syndromes related to adrenal insufficiency with genomics patterns (Based on density of repetitive elements), that phenotypically are related to hydroxylase deficiency, and the consequences in the lipid metabolisms. The relational network shows this metabolic process as critic node in the genomics associations identified in Rare diseases. Therefore genomics context could define unknown related factors important for the exploration of new genetics, or new targets for disease exploration. Conclusion: Although Rare Disease are described as diverse molecular and genetics anomalies, by bioinformatics approaches could be possible the identification of genomics and metabolic connections, that allows the definition and implementation of new methods for specific target in several RDs.

EnhancerDB: a database of human enhancers and their putative targets. P. Wang, L.Y. Wang, M.Q. Zhang, Q. Lian, J. Wang. The University of Hong Kong, Hong Kong, Hong Kong, Hong Kong.

1647S As a major class of distal cis-regulatory elements, enhancers are crucial to the regulatory system orchestrating gene transcription. Thereby detection of enhancers has been extensively studied and numerous enhancers have been detected by experiments and predicted by computational models. Here we present a database containing ~3,000,000 human enhancers. 3,598 enhancers are experimentally detected and collected from 190 publications, and the rest are predicted based on transcription factor binding motifs, high conservation, histone modification marks or eRNA expression. Meanwhile, we incorporated data of transcription factor binding motifs and ChiP-Seq peaks to annotate the enhancers. In addition, we used available long-range chromatin interaction data, including ChIA-PET and Hi-C in 10 human cell lines to present putative target genes of the enhancers.

A highly efficient and scalable compute platform for massive variant annotation and rapid genome interpretation. J. Warren1, E. Colak1, A. Kiani4, J. Li2, S. Bhattacharya3, N. Bani Asadi1,2, S. Barr3, A. Butler4, G. Nolan4, R. Chen5, W.H. Wong6,7, H.Y.K. Lam2. 1) Department of Engineering, Bina Technologies, Redwood City, CA; 2) Department of Bioinformatics, Bina Technologies, Redwood City, CA; 3) Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford CA; 4) Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford California 94305; 5) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 6) Department of Statistics, Stanford University, Stanford, CA; 7) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

After obtaining the variants from next generation sequencing data, researchers and clinicians still face the undertaking of interpreting the results. Despite the availability of numerous public databases, using this collective information is an arduous task due to inconsistent data, multiple versions and nonstandard formats. Even after aggregating the data and annotating the variants, it remains a laborious exercise to identify the causative variants associated with the disease in question.

To address this challenge we present a highly efficient data pipeline that leverages big data technologies to integrate annotations from a large number of biological databases. The pipeline takes variant callsets, annotates all samples, and indexes the variants to support real-time queries and analytics. Specifically, it uses Hadoop MapReduce to perform extensive precomputation on the annotation data; builds indexes of the annotated variants; versions and manages data in both relational and NoSQL databases; and horizontally scales to increase throughput. With this approach, a 5-node cluster can annotate a typical whole exome sequencing (WES) sample with ~300 million annotation records (>200GB) in under 30 minutes or a whole genome sequencing (WGS) sample under an hour. Annotating a sample and indexing its variants are computationally demanding steps, but these are one-time costs. After the indices are generated, users can concurrently and repeatedly perform ad-hoc queries to identify the variants that meet given criteria. On average, it takes less than one second for any complex query to return the results on a single sample.

Researchers can easily query the samples in a database-centric view or a genome-centric view. In one application we analyzed the WGS data from the DNA of Atlas, the skeletal remains of a 6-inch human found in the Atacama Desert, Chile. The annotation and analysis took less than an hour for >3M total variants and identified 37 functional changing variants located in genes highly associated with dwarfism and skeletal dysplasia. Four of these variants were not found in dbSNP and may be of interest for further investigation. Such filtration greatly reduces the search time and space for researchers. With predefined statistical workflows, researchers can also carry out genetic studies for case-control, trio, cohort, and tumor-normal studies.
1649T Investigating the genetic architecture of pulmonary arterial hyperten-
sion shared with other diseases. L.A. Yancy1,2, V.A. de Jesus Perez3,4, R.T. Zamanian1,2, A.J. Butler5,6, 1) Biomedical Informatics Training Program, Stanford University School of Medicine, Stanford, CA; 2) Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA; 3) Division of Pulmonary and Critical Care Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, CA; 4) Vera Moulton Wall Center for Pulmonary Vascular Disease, Stanford University School of Medicine, Stanford, CA.

Pulmonary arterial hypertension (PAH) is a rare and fatal disease with a median survival of approximately 4 years for those afflicted. Many genome-wide association studies (GWASs) have attempted to determine the variants predictive of PAH, but often due to small sample sizes, studies have not led to highly predictive markers. However, GWASs have been successful in determining variants associated with many other diseases. We have an ongoing effort to curate these variants into a database called VARIants Informing MEDicine (VARIMED). VARIMED contains results from over 17,000 peer-reviewed genetic epidemiology studies (e.g. GWAS), covering over 460,000 variants (in over 18,000 genes) associated with over 6,500 phenotypes. In this study, we leverage VARIMED to determine if PAH exhibits a shared genetic mechanism with other diseases. Finding other diseases that have a significant enrichment of PAH associated variants may provide insight into (novel) variants or genes that play a significant role in PAH (and are otherwise not detectable in a PAH GWAS due to small sample size) or even novel therapies for PAH that can be “borrowed” from diseases with similar genetic architecture. Here, we perform a GWAS on exomes (~88,000 variants) from PAH subjects, rank those variants based on p-value and overlap to our GWAS-derived variants associated with those prioritized variants (~300 genes). Using those “prioritized” genes, we then query VARIMED and identify diseases with a similar genetic architecture (i.e. identify those diseases with a significant enrichment of those genes based on our PAH GWAS data). Through this, we found significant similarities between PAH and a number of immune-related diseases, with the most significant finding being an unusual acquired kidney disease. This work can be extended to investigate the shared genetic architecture of other diseases, and even potentially other diseases that fit within the large spectrum of diseases now with characterized genetic architecture.


The ability to analyze genomic data from differing sequencing methodologies and multiple tissue types in a single cancer patient provides a tantalizing opportunity to explore the capabilities of an integrative genomic study to elucidate potential tumor drivers and mechanisms. Simply because of sample availability, researchers are typically limited to analyses of primary tumor data. Access to data from both RNA-Seq and DNA-Seq datasets from adjacent normal tissue, blood and metastatic tissue samples in a single patient is rare, yet could greatly increase signal and reduce noise in these analyses. Technological advances and cost reductions mean truly integrative genomic analyses, in which mutational signatures can be reviewed in concert with transcriptional profiles across relevant samples from multiple tissues, could be used to understand tumorigenesis and potentially guide intervention. We present here a proof of concept study of a single breast cancer patient where we integratively analyze exome and RNA-Seq data obtained from four sample types - primary tumor, blood, adjacent normal tissue and metastasis. We use GenePool™, a commercial software application for biomarker discovery and validation in large patient cohorts (Station X, San Francisco, CA) to quickly prioritize variants and genes found in the tumor and metastasis by comparison to the blood and the adjacent normal breast tissue. Similarly, we analyze the primary tumor relative to the metastasis to identify differences in mutational signatures and transcriptional profiles. We present the results of this single patient integrative analysis and show how GenePool can extend these types of analyses to cohorts of patients, becoming a useful tool in the understanding and treatment of tumors.

1651M Limitation of multiple testing through the integration of TCF7L2 DNA occupancy and SNAP association data reveals GIP and CEPPE1 as novel type 2 diabetes loci. C.A. Sloan1,2, E.T. Miyasato1, S. Xia1, V. Guy1, J. Sainz2, K.H. Kaestner2, A.D. Wells4, S.F.A. Grant1,2, 1) Division of Human Genetics, The Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Department of Genetics and Institute of Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3) Institute of Biomedicine and Biotechnology of Cantabria (IBBTEC), Spanish National Research Council (CSIC), Santander, Spain; 4) Pathology and Laboratory Medicine, Children’s Hospital of Phila-

The TCF7L2 locus is strongly implicated in the pathogenesis of type 2 diabetes (T2D). We previously mapped the genomic regions bound by TCF7L2 within HCT116, a colon cancer cell line, revealing an unexpected highly significant over-representation of specific GWAS loci categories. Given that only a minority of the predicted genetic component to most complex traits has been identified to date, plus the fact that a number of GWAS-implicated transcription factors have now been shown to exhibit statistically significant preferential binding to loci associated with complex traits, we investigated if restricting association analyses to just the genes uncovered from our Chip-seq approach in order to substantially reduce multiple testing could yield novel T2D loci. When investigating all the known GWAS loci bound within 5kb by TCF7L2 (most likely to be functional), derived from HCT116, apart from the known TCF7L2 locus itself (rs7901695), the coronary artery disease associated T allele of rs46522 within the UBE2Z-GIP-ATPSG1-NAPB locus yielded significant and novel (OR=1.5, p=4.90x10^-12); however, indeed, the occupancy site was approximately 4kb from the transcription start site for GIP in an intergenic region known to be a hub for binding proteins, H3K27Ac histone marks and open chromatin via a DNAse I hypersensitive site. Indeed the regions GIP binds to in the promoter region of GIPR has already been reported in relevant GWAS settings to be associated with BMI and glucose and insulin response metrics. Furthermore, when we analyzed tag-SNPs within genes not previously implicated by GWAS but bound within 5kb by TCF7L2 in HCT116, we observed significant association within the same dataset of the A allele of rs4780476 within CEPPE1 with T2D risk (OR=1.1, p=4.10x10^-12). Furthermore, the TCF7L2 occupancy site was in the immediate CEPPE1 promoter region. This is an equally notable observation, as only one paper has been published to date on this particular product and shows that down-regulation of CEPPE1 expression improves glucose metabolism in vitro in adipocytes. Our limitation of multiple testing is based on biological plausibility, where a GWAS-implicated transcription factor is clearly pointing us to genes that are genetically associated with complex disease more often than expected by chance and thus may also be pointing us to novel genes where their strength of the association was at the level of noise at the genome wide scale.
Accessing ENCODE project data using a REST API and JSON. C.A. Sloan, E.T. Chan1, V.S. Malladi1, L.D. Rowe1, J.S. Strattan1, B.C. Hitz1, N.R. Poduttun1, F. Tanaka1, K. Leamed2, B.T. Lee2, S. Myasato1, M. Simison1, E.L. Hong, W.J. Kent1, J.M. Cherry2, 1) Genetics Department, Stanford University, Palo Alto, CA; 2) School of Engineering, UC Santa Cruz, Santa Cruz, CA.

The Encyclopedia of DNA Elements project (ENCODE) has been producing data for over eight years to investigate DNA and RNA binding proteins, chromatin structure, transcriptional activity and DNA methylation on a variety of human and mouse tissues and cell lines primarily. As the complexity and diversity of the data grows, the tools required to organize, search and access the data in meaningful ways need to be more sophisticated. The ENCODE Data Coordination Center (DCC) has incorporated a representation state transfer application programming interface (REST API) with JSON objects to facilitate the access of ENCODE experimental metadata using a web portal. Meta-data can be accessed and data can be searched for at http://www.encodecc.org/ using the HTTP access commands of get and post. We further expand on the access capability by allowing filtering of the metadata with the use of search urls. This system allows external researchers to write their own interfaces to access, analyze and visualize the ENCODE data. It also facilitates the integration of the ENCODE data with other similar large-scale data sets like Epigenic Roadmap and modENCODE. Here we present our JSON schemas, examples of the REST API and use-cases for the search functions. Our goal is for the genomics community to use the released ENCODE data available through these methods for data mining and integration.

Beyond Flat Files: Creating a web-based data API to simplify parsing and distribution of GTEx data. T. Sullivan, D. DeLuca, K. Hadley, K. Huang, J. Nedziel, A. Segre, E. Gelfand, T. Young, G. Getz, K. Ardlie, the GTEx Consortium. The Broad Institute, Cambridge, MA.

Large-scale genomics projects such as GTEx produce vast amounts of data, and the goals of many of these projects include creating publicly accessible databases of genetic information. Often, to perform meaningful analysis, one must download very large text files, which must then be merged, integrated, and queried. The programming overhead the user faces during these integration and querying steps can be disproportionate to the amount of information sought by the analyst. While web pages can provide some querying capability, we argue that the heart of the solution to this problem lies in providing a web-based API. We outline here the major ways in which the GTEx API development has contributed to the ultimate goals of GTEx data distribution. These benefits fall into three categories: querying, software integration, and insulation. Data querying: The GTEx API provides the user with the ability to ask specific questions. For example, querying with an isoform instantly provides isoform specific expression levels across all GTEx tissues. Integration in any software: A major advantage that a data API has over a website is that the queries can be programmatically integrated into client systems. GTEx data can be accessed within scripts on demand. Non-GTEx web portals can query GTEx data at any time for integration and display within their webpages. Insulation of data storage from data access: The question of which technologies to use to store and query large datasets is non-trivial and often evolving. For example, querying the end user from these issues by providing a domain specific outward face to the data that is intuitive to use. As a result, database technologies could be modified or completely replaced without the API user having to modify existing code.

Thus far the major beneficiary of the GTEx API has been the GTEx Portal itself. Efforts to develop GTEx data visualization tools rely heavily on the GTEx API. The API delivers data efficiently for multi-tissue eQTL data, gene and isoform expression data, and sample metadata. In the future, the availability of public data APIs across large-scale genomics projects will allow researchers to easily integrate data from multiple types of studies and quickly find what measurements have been made about their genes of interest, be it in the context of a variety of normal tissues, cancer types, or disease states.


Association statistics for each locus are important end products from studies that investigate the genomic contribution to disease/trait manifestation. Therefore, since its inception, dbGaP has been a core repository for archive and dissemination of these data. dbGaP currently provides over 3800 data sets for tests of association, linkage and somatic mutation analyses. Over 90 percent of these data sets are constructed with SNP markers, the rest are results for specific genes or whole genome copy number variations. Although these association data cover a variety of phenotypic traits, diseases and pharmacological responses, they have been provided by only a small portion of the studies in dbGaP. The dbGaP team is renewing its call to encourage researchers to upload their association results, including those that do not reach genome wide significance, to dbGaP. These results will be permanently archived and made fully accessible to authorized researchers. After potentially participant-identifiable information being removed, these data are also included to the open-access dbGaP browser and PheGeni web resource (http://www.ncbi.nlm.nih.gov/gap/phegeni), where they are linked to NCBI resources and updated with subsequent genome builds and dbSNP annotations. To facilitate the sharing of these data, we propose a guideline for Minimum Info Required for Association Data (MIRAD). The MIRAD is composed of following data elements: 1) locus identifier (#, geneID, dbVarID); 2) variation summary (sample size, genotype quality and variant counts of each group) within the locus; 3) Statistics (p-value, FDR); and 4) Coding (risk) allele, and effect size. Collection of these minimal elements will aid other researchers to evaluate supporting evidence and independently verify discoveries from other research efforts in a consistent fashion. Association results are strongly encouraged for submission, even in cases where individual level data is not suitably constrained. Broad data sharing allows the broad research community to directly use analysis results for meta-analyses, to increase statistical power, or for the development of hypotheses. This presentation will describe the current association analysis results data archived at dbGaP, and outline and discuss MIRAD.

Genome in a Bottle: So you’ve sequenced a genome, how well did you do? J.M. Zook, H. Panik, M. Salti, Genome in a Bottle Consortium. Biosystems and Biomaterials Division, National Institute of Standards and Technology, Gaithersburg, MD.

To help clinical and research labs understand the trustworthiness of variant calls from Next Generation Sequencing, the Genome in a Bottle Consortium (www.genomeinabottle.org) is developing well-characterized whole human genomes. We recently published our methods to integrate and arbitrate between datasets from multiple sequencing technologies to develop high-confidence SNP, indel, and homozygous reference genotypes for our pilot NIST human genome Reference Material based on Coriell DNA NA12878 (Zook et al, Nat. Biotech. 2014). Since then, we compared our arbitrated multi-dataset calls to orthogonal methods developed by Illumina Platinum Genomes and Real Time Genomics using phased pedigree information to develop high-confidence calls. We found that our arbitrated calls are conservative but have exceptionally few errors and high quality homozygous reference calls, and that phased pedigree methods do particularly well in finding systematic sequencing errors such as homopolymer errors but have some systematic alignment errors that inherit properly. We combine the strengths of each of these datasets to develop the most comprehensive highly accurate benchmark set of SNP, indel, and homozygous reference calls. Genome in a Bottle is also developing similar multi-dataset arbitration and pedigree methods for structural variants. In addition, to enable users of our benchmark to assess sequencing and bioinformatics performance in a uniform manner, the Genome in a Bottle Consortium has developed a specification for performance metrics to be output by benchmarking algorithms using our Reference Materials. Finally, we will describe how labs are assessing their performance using our high-confidence SNP, indel, and homozygous reference genotypes for our pilot NIST human genome Reference Material, and our progress developing future Reference Materials.
1657M Quantity or quality that is the question: integrative genome-wide association. A.M. Mezlini1,2, C. Bacil1, S. Morrissey1, M. Taylor1, D. Malkin1, A. Goldenberg1,2, 1) SickKids Research Institute, Toronto, Canada; 2) Computer Science, University of Toronto, Toronto, Ontario, Canada.

The causal factors behind complex genetic diseases are numerous and heterogeneous. Many studies have identified and functionally validated associated protein-coding variants. Recently, variants modulating gene expression levels and their relation to disease have attracted a lot of attention. This raises the question of how often gene products (proteins) quality versus quantity abnormalities are causing disruption in biological processes in relation to diseases. Both gene products’ quantity and quality might be important for any given disease thus, analyzing both data types simultaneously can provide better models of disease and characterize larger sets of patients. To address this issue we present an integrative method for combining two complimentary data types: exome sequencing data as an indicator of protein quality and gene expression data which is indicative of protein quantity. Note that our approach is very different from eQTL analysis, since we are treating expression data as a predictor of a disease, rather than an outcome. Using variant harmfulness as a prior, we obtain a score per gene per patient and by exploiting gene interaction information such as protein-protein and/or regulatory networks, our method identifies sets of genes whose aberrant states potentially contribute to the disease in a large proportion of the patients. The significance of the gene set is assessed via permutations. Our analysis allows 1) identify a small set of relevant rare and common coding variants; 2) zero-in on potential regulatory aberrations. In the regulatory case, the “drivers” can be retrieved in a subsequent targeted analysis of regulatory SNVs, CNVs and epigenetic regions in a cost-effective manner for the cohort of patients. Varying the proportion of the regulation-attributed vs coding variant-attributed genome aberration in a simulated setting we show that gene-base variant aggregation methods such as CAST, Calpha, RWAS fail in scenarios where less than 60% of the causal variants are coding and that our method is able to retrieve the disease mechanisms even in scenarios where as few as 20% of causes are coding variants. Our method is on par or outperforms all others in terms of precision and recall in all considered scenarios. We applied our methodology to a cohort of medulloblastoma patients and uncovered significant genetic heterogeneity across subtypes of medulloblastoma. Our results confirm and extend previously reported gene-disease associations.

1658T Cross-species genome and epigenome visualization on WashU Epi-Genome Browser. X. Zhou1, R.F. Lowdon1, D. Li1, I. Smirnov1, Y. Cheng2, P.A.F. Madden4, R. Hardison3, J. Costello2, T. Wang1, 1) Department of Genetics, Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO; 2) Brain Tumor Research Center, Department of Neurosurgery, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, CA; 3) Department of Genetics, School of Medicine, Stanford University, Stanford, CA; 4) Department of Psychiatry, Washington University in St. Louis, St. Louis, MO; 5) Department of Biochemistry & Molecular Biology, Pennsylvania State University, University Park, PA.

Integrative comparison of genomes and epigenomes across mammalian species is of critical importance to human genetics research. While genome sequence comparison has pinpointed conserved elements in the human genome, patterns and rules of epigenomic conservation are not yet well defined. The same genome encodes multiple epigenomes. Epigenomes are cell type-specific, change with age, and respond to environment. Epigenetic variation and dynamics may be the foundations of great biological complexity across species and between individuals. For instance, human-mouse conserved genomic regions may have different epigenetic modification patterns, which may explain human-specific traits that are otherwise difficult to explain using the genome alone. Such an analysis requires representing each species’ epigenomic information in the context of genome sequence comparison. It thus poses a great challenge to data access and visualization. We address this challenge through a major innovation in the WashU Epi-Genome Browser (http://epigenomewidget.wustl.edu/browser/). We developed a new mechanism that allows investigators to compare genomes of two species by visualizing their genome alignments at different resolution, smoothly transitioning from gapped-assignment at single-base resolution, to syntenic blocks and rearranged genomes at megabase-scale. In the context of this alignment, vast epigenomic data from both species can be displayed and directly compared. Our Browser breaks the “one-reference-genome-only” status quo of genome browser tools, and contributes a powerful tool for the research community to access the cross-species genomic resources (i.e., Roadmap Epigenomics, ENCODE, modENCODE) in comparative genomics and epigenomics framework.
1660M
Genetic predictive modeling of diabetes based on circulating glyceric measures. The Long Life Family Study (LLFS). A.T. Kraja, M.K. Wojtczynski, J. H. Lee, I. Miljkovic, B. Thyagarajan, S. J. Lin, I. B. Borecki, J. T. Perls, K. Christensen, A. Newman, P. An, M.A. Province. Division of Statistical Genomics, Department Of Genetics, Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, MO, USA; 2) Department of Epidemiology, the Sergievsky Center, Columbia University Medical Center, New York, NY, USA; 3) Graduate School of Public Health, Department of Epidemiology, Center for Aging and Population Health, University of Pittsburgh, Pittsburgh, PA, USA; 4) Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA; 5) Division of Geriatrics, Department of Medicine, Boston University Medical Center, Boston, MA, USA; 6) Danish Aging Research Center, Epidemiology, University of Southern Denmark, Odense, Denmark; 7) Department of Clinical Biochemistry and Pharmacology and Department of Clinical Genetics, Odense University Hospital, Odense, Denmark.

The LLFS is a cohort study of longevous families recruited from 4 centers Boston, New York, Pittsburgh and Odense, Denmark. The LLFS family members have lower prevalence (7.1%) of type 2 diabetes (T2D) than the US European-ancestry population (10.2%) 20 years and older. We compare genetic associations of literature validated loci for glucose, insulin, insulin resistance and diabetes in LLFS versus top ranked results of LLFS. We hypothesize that this approach may identify genes that associate with insulin and glucose metabolism that do not operate under disease conditions. We conducted an additive genetic model genome-wide association scan (GWAS) for fasting glucose (FG), insulin (FI), HOMA-IR, HOMA-B, and elevated hemoglobin A1c (HbA1c) (N=4,250). The Illumina Omni 2.5 chip was used and ~9.2M (MAF > 0.01, r²>0.3) variants were imputed. From public databases, GWAS variants unique per gene with p-values at p ≤ 10⁻⁵ were selected (n=279). Predictive models via Graphical Trees and Support Vector Machine for T2D/pre-diabetes were constructed. Analysis of this list of candidate genes using GeneGO software revealed involvement in "Protein folding and maturation insulin processing" pathway (log₁₀(p)-FDR=10.4) and "Diabetes" (log₁₀(p)-FDR=7.8). The best LLFS SNPs of the 279 candidate genes identified by LLFS five false discovery rate (FDR) with other top ranked variants present in LLFS. Best LLFS SNPs of 279 candidate genes showed respectively a Range, Mean, Median and SD log₁₀(p) for FG 0.3-6.7, 2.2, 2.1, 0.0, FI 0.3-5.6, 2.0, 2.0, 0.8; HOMA-IR 0.4-5.0, 2.0, 2.0, 0.8; HOMA-B 0.2-5.0, 2.0, 2.0, 0.8; and HbA1c 0.4-5.0, 2.0, 2.0, 0.8; compared to LLFS best SNPs (p<0.00001) for FG 5.0-7.8, 5.4, 5.2, 0.5; FI 5.0-7.7, 5.6, 5.4, 0.6; HOMA-IR 5.0-7.6, 5.5, 5.3, 0.5; HOMA-B 5.0-6.3, 5.4, 5.3, 0.3; and HbA1c 5.0-6.6, 5.5, 5.4, 0.5. Between LLFS own top ranked candidates and literature’s top candidate gene list, a few overlap (LLFS own Genes/New, FG 43/37, FI 35/34, HOMA-IR 41/40, HOMA-B 33/33 and HbA1c 35/34). The new LLFS findings may point to loci important for normal glycemic metabolic condition on age. Among new findings rs78375027 (chr6:19304147, -log₁₀(p) =7.7, MAF=3.5%, n=3,943 near LOC100506885 associates significantly with FI. This locus has been previously associated with myocardial infarction (rs9460319, -log₁₀(p)=3.5), hypertension (rs11752516, -log₁₀(p)=5.1), LV mass (rs10498996, -log₁₀(p)=5.7) and stroke (rs2955438, -log₁₀(p)=3.8).

1661T

70% of diseases are preventable, and common complex diseases and conditions like Heart Disease, Cancer, Stroke, Diabetes, Arthritis, and Obesity account for 7 in 10 deaths each year in the US. Changes in behavior can dramatically lower patient risk of developing diseases, even if one has a genetic predisposition. The first wave of consumer-focused genomics took fragmented approaches that focused strictly on genomic data, or health assessment data, and didn’t provide the full view of a person’s health, and ultimately did not lead to patient behavior change at high rates. We’ve since learned that genomic data for common complex diseases is only part of the puzzle, and should be used in concert with all of a patient’s risk factors to get the most personalized and complete view of a patient’s health available. How then, with the care of a physician, can genomic data best be delivered to an individual, alongside comprehensive health data, with clear ways to turn insights into enduring action? At BaseHealth, we have created a platform that leverages big data analysis to synthesize many data sets (genomic data, lab tests, lifestyle factors, family history, and other medical information) to output a comprehensive health assessment unique to each patient. Through multidimensional data modeling and meta-data analysis driven by published and curated genetic data and scientific data, health databases, and metadata analysis, we can capture complex relationships between diseases and risk factors and give physicians a better understanding of their patients’ unique risks for keycommon complex diseases. The platform covers more than 40 common complex diseases that combine both modifiable and nonmodifiable risk factors and engages patients and their physicians in a highly collaborative way. With analysis provided by our platform, physicians can make meaningful recommendations that lead to patient behavior change. In this presentation, we will show how statistical modeling can meaningfully increasing a patient’s chance to prevent a high-risk disease. Furthermore, we’ll examine what happens when physicians use a comprehensive approach that combines all data types that have an effect on health to engage their patients to improve health and lower disease risk over time. In addition, we will share results some of the pilot programs from among the 50 physicians and 250 patients that are currently using Genophen.
Next generation sequencing data has become a prolific tool in genetic research. As such, tools supporting that interpretation must provide an intuitive and powerful user experience that integrates research data with public resources and tools such as a next-generation genome browser that can illustrate genomic context from a web-based browser with public data. We introduce Golden Helix GenomeBrowse 2, a free visualization tool for all common NGS file formats such as BAM, VCF, BED, WIg, FAI, SAM, binary, and BigWig while also hosting multiple parallel and biologically meaningful views at the same time. In addition to the genome browser, we provide a suite of tools for filtering and analyzing data, as well as a suite of tools for annotating data. These tools can be used to identify variations in the genome and to classify and quantify the impact of these variations. For example, we can use these tools to identify variations in the genome that are associated with human diseases and to classify and quantify the impact of these variations. In addition, we can use these tools to identify variations in the genome that are associated with mouse and vertebrate diseases and to classify and quantify the impact of these variations. For example, we can use these tools to identify variations in the genome that are associated with craniofacial abnormalities and to classify and quantify the impact of these variations. In addition, we can use these tools to identify variations in the genome that are associated with craniofacial abnormalities and to classify and quantify the impact of these variations.
1666M

Sequencing of 50 rhesus macaques facilitates identification of new genetic models of human disease. G. Fawcett1, D. Rio Deiros1, R.A. Harris2, M. Raveendran1, B. Ferguson2, Z. Johnson3, S. Kanthaswamy4, E.J. Vallender5, N.H. Kallrin6,7, R.W. Wiseman7,8,9, D.M. Muzny1, R.A. Gibbs1, J. Rogers1. 1) Dept Mol & Human Gen, Baylor College Med- HGSC, Houston, TX; 2) Dept. Mol. & human Gen, Baylor College Med, Houston, TX; 3) Division of Neurosciences, Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR; 4) Yerkes National Primate Research Center, Emory University, Atlanta, GA; 5) Molecular Anthropology Lab, Department of Anthropology, UC Davis, Davis, CA; 6) California National Primate Research Center, UC Davis, Davis, CA; 7) Department of Environmental Toxicology, UC Davis, Davis, CA; 8) New England Primate Research Center, Harvard Medical School, Southborough, MA; 9) Dept. of Psychology and Psychiatry, Univ. of Wisconsin-Madison, Madison, WI; 10) Waisman Laboratory for Brain Imaging and Behavior, Univ. of Wisconsin-Madison, Madison, WI; 11) HealthEmotions Research Institute, University of Wisconsin-Madison, Madison, WI; 12) Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, WI; 13) Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, WI.

Rhesus macaques (Macaca mulatta) are the most commonly used non-human primate model of disease, and are particularly important as models of psychiatric, neurological and infectious diseases. Therefore, characterizing genetic variation within research colonies of these monkeys will advance the study of a variety of different diseases. We sequenced the exomes of 50 unrelated rhesus macaques from 5 NIH-funded primate research centers. Single nucleotide variants (SNVs) are defined as variants with a minimum of 4 reads covering each position with a minimum quality score of 30, whereas validated SNPs also have both alleles observed in at least two animals. Prior studies suggest rhesus macaques have greater genetic diversity overall than humans. Yuan, et al (2012) reported that while rhesus macaques exhibited ~3 SNPs/Kb, the majority of these variants were in intergenic regions, and that coding regions were similarly polymorphic in rhesus macaques and in humans. Estimates of human SNP density in predicted exonic regions (Bainbridge, et al 2011) show 1.45 SNPs/Kb in exons. Our similar to the human data with an exonic SNP density of 1.15 SNPs/Kb. However, our SNP density in off-target intronic regions is significantly greater in macaques (4.8 SNPs/Kb). Our results are consistent with Yuan et al’s expectation of greater non-coding but similar coding variation density in macaques relative to humans. This result is interesting since coding variants are likely to be under strict negative selective pressure. We also sought to identify specific new potential models of disease. Among the 50 exomes, we identified: 133 transcription stops gained, 17 transcription stops lost, 225 splice sites changed, and 11,768 non-synonymous SNPs. We observed variants in genes from many different ontological categories (DAVID analysis) including neurologically relevant and immune system genes. For example, there was one non-synonymous variant in CR2, four in DNAH11, one in ATP6AP2 and two in FGFR3. CR2 is important for immune system function, and DNAH11 and ATP6AP2 are neurologically functional genes, while FGFR3 is associated with a variety of morphological disorders and cancer. Our results indicate that this species carries functionally significant mutations in many genes relevant to human disease risk.

1667T

Identifying mouse models related to human disease. S. Rockwood1, S.M. Bello2, J.T. Eppig2, J.E. Richardson2, M. Sasner2, C.L. Smith2, L. Donahue1. 1) Genetic Resource Science, The Jackson Laboratory, Bar Harbor, ME; 2) Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, ME.

Access to model organisms is of paramount importance to advancing the research objectives of the bio-medical research community. To facilitate this access, the Mouse Repository at The Jackson Laboratory (JAX) has served as a centralized resource for the distribution, cryopreservation and development of high health status mouse models for over 50 years. Hundreds of new strains are added each year to one of the largest collections of characterized mouse strains available. The increasing magnitude and diversity of current day mouse resources present a challenge to researchers who attempt to identify and obtain mouse models most appropriate for their work. In addition to newly designed search interfaces at the JAX website (www.jax.org), other web-based search resources are available to assist with this task. The Mouse Genome Informatics group at JAX maintains two of the most useful: the International Mouse Strain Registry (IMSR) and the Human/Mouse: Disease Connection site (HMDC). The IMSR (www.findmice.org) is a searchable online database of mouse strains and mutant ES cell lines available worldwide. The HMDC (www.diseasemodel.org) provides a seamless human-to-mouse data traversal, enabling clinical and translational researchers to take advantage of the wealth of data and annotations from mouse models; as well as allowing mouse researchers to connect their findings directly to genetic associations reported in human disease. This poster provides several examples illustrating how these resources can be used to quickly identify and obtain mouse models related to human disease research applications (also see related posters by Smith CL et al and Eppig JT et al). Donating a strain to the Jax Mouse Repository is an easy way to fulfill the NIH’s requirements for sharing mice. Researchers wishing to have strains considered for inclusion in the Repository are encouraged to submit their strains at: www.jax.org/donate-a-mouse The JAX Mouse Repository is supported by NIH, The Howard Hughes Medical Institute, The Ellison Medical Foundation and several private charitable foundations.

1668S


Drosophila melanogaster (fruit fly) is a historically long-established model organism and its availability encompasses almost every field in modern biology, such as genetics, development and neurobiology. Moreover, D. melanogaster can provide many human disease models including diabetes and neurological disorders. Recent progress of genome-wide association study (GWAS) in humans has elucidated thousandish disease-associated genetic elements. Nevertheless, model organisms are still indispensable for functional validation and characterization of those candidate disease genes. In this study, we present a genome-wide functional gene network for D. melanogaster, FlyNet, and demonstrate its utility in pathway interpretation and functional characterization of the candidate genes from human GWAS, facilitating genetic dissection of human diseases in D. melanogaster.
1669M

One of the hallmark benefits of RNA-seq over traditional gene expression quantification methods is the ability to capture the sequences of exon-exon junctions, which are indicative of the transcript isoform structure present in the samples. Beyond qualitatively measuring the presence of alternative isoforms, methods exist to infer isoform abundance. However, assessing the accuracy of isoform quantification methods remains extremely challenging due to a lack of ground truth. Orthogonal technologies such as qPCR or NanoString have been used for validation but any approach that targets only portions of each isoform is limited in its applicability to validate the range of alternative isoforms produced by RNA-seq. Here we evaluate a panel of algorithms for the quantification of isoforms, including Flux, Cufflinks, eXpress, RSEM and Sailfish in combination with multi-isoform validation experiments from the Genotype Type Expression (GTEX) project. As orthogonal technologies we are employing long fragment, long read Illumina RNA-seq as well as deep 1-2 kb read PacBio sequencing. We have generated 66 TruSeq libraries from a panel of 13 different GTEx tissues with an average fragment length of 400 base pairs. These libraries were sequenced to 100 million 250bp paired-end reads. The tissue panel consists of adipose, artery, cerebellum, frontal cortex, lymphocytes, heart, lung, muscle, pancreas, testes, thyroid and whole blood. The dataset to be evaluated consists of the standard GTEx sequencing runs (2x76bp fragment TruSeq) derived from the same samples, as well as a subsampled version of the evaluation set in which the read lengths have been truncated to 76bp to simulate standard runs. The opportunity provided by the multi-tissue GTEX sample bank is critical to this evaluation considering that many isoforms are expressed in a tissue specific manner. These data will provide insight into the relative performance of the tools as well as into the process of isoform specific expression in general.

1671S

Dendritic cell (DC) linages coordinate immune system activity through functional specialization. Irf4, a transcription factor, is required for CD11b+ DC lineage development from bone marrow stem cells and has been implicated in multiple inflammatory diseases such as asthma. The epigenetic landscape of immune cell development, and in particular, Irf4, in inflammation and inflammatory diseases remains largely unexplored partly due to the difficulty of using highly purified, and typically, limited populations of cells in ChIP-seq (chromatin immunoprecipitation followed by sequencing) assays.

A dedicated Circuit (DC) ChIP input control - was developed using limited amounts of K562 cells. ChIP-seq* on the Ion Proton™, followed by analysis using MACS2 - of the transcription factor CTCF and histone modification marks (H3Kme3-methylation, H3K27-acetylation) strongly correlates with ENCODE datasets - with confidence rates of 69% (CTCF), 73% (H3K27ac), 89% (H3Kme3). Furthermore, CTCF Proton™ data show significant enrichment for insulator domains at 41%; the moderate presence of DNA methylation concomitantly with H3K27me3 - a signature for polycomb repression - is attested to by the ENCODE-Broad ChromHMM. Also, 55% and 35% of the acetylation marks enrich for the enhancers and promoters respectively. These results are based on chromatin derived from 1 million(M) cells, making it viable for generating data from limited number of primary cells. This is in contrast to the 10M cells recommended by the ENCODE consortium.

This methodology was used to compare Irf4 genomic binding sites generated from flow-sorted population of 1, 3, 5, and 25M CD11b+ lineage murine DCs. We observed comparable Irf4 ChIP-seq results from 5 versus 25M cells (three fold) and more, as low as 5M cells can be used to acquire high quality (FDR: 10^-19) ChIP-seq data. We identified genomic Irf4 binding sites proximal to genes, whose activity is consistent with CD11b+ DC lineage activity and/or known to contribute to inflammatory diseases. We then examined Irf4 functional regulation of the identified gene targets, we performed RNA-seq analysis with CD11b+ DCs and a related lineage, CD103+ DCs. Integrating expression analysis with ChIP-seq indicates a unique CD11b+ DC gene expression program concordant with Irf4 loci association in comparison to CD103+ DCs. *For Research Use Only.

1672M
GEM.app: using hadoop to empower the revolution of large-scale collaborative analysis and data-sharing in the genomic age. M. Gonzalez, R. Schule, S. Zuchner, J. Ziegle, S. Zechner, Human Genetics and Genomics, University of Miami Miller School of Medicine, Miami, FL.

Next-generation sequencing technologies have revolutionized the human genetics research community by enabling large scale genome-wide association studies and collaborative networks adopting GEM.app for analysis, we observe a radical reduction in time to discovery and increasingly complex questions. As we see consortia powered with powerful distributed IT technologies with bioinformatics will lead to efficient and collaborative analysis of large-scale genomic data. The management and interpretation of such data is vital to continue the progress of human genomics research. To this end, we have developed, GEnomes Management Application (GEM.app), in order to make large-scale genomics analysis and data-sharing accessible to users of all technical backgrounds. GEM.app uses the combination of a highly distributed computing cluster and a user friendly graphical web interface to provide users the capabilities to run real time analyses of the ~6,000 exomes and ~900 genomes publicly currently uploaded in GEM.app. In addition, we have developed tools within GEM.app to facilitate the creation of collaborative data-sharing groups, where researchers are in full control to share data-sets instantly. Within these collaborative data-sharing groups, researchers from around the world can contribute data that can be concurrently analyzed by all users in the data-sharing network. GEM.app utilizes the latest variant annotations (conservation, amino acid substitution predictions, ClinVar, EVS, etc.) to allow flexible and powerful analysis options. We provide five major analysis modules: 1) Variants within Families - basic filtering of variants, 2) Genes Across Families - filtering genes are enriched, 3) All Genes - filtering all genes, 4) Allele Sharing - filtering for alleles that are enriched, 4) Cancer Analysis - quickly analyze tumor/normal pairs, and 5) PathFinder - hypothesis driven analyses using Gene Ontology and various pathway databases to identify pathways enriched in functional annotations. Our first GEM.app users, including DDHD1, DDHD2, GABA2, CYP2U1, RTN2, B4GALNT1, BICD2, FBXO38, ANKRD11, PNPLA6, SARM1, DNAJC5, etc. Combining powerful distributed IT technologies with bioinformatics will lead to efficient tools to explore the complex functional landscape as well as ongoing and collaborative networks adopting GEM.app for analysis, we observe a radical reduction in time to discovery and increasingly complex questions resolved.
1673T
A Data Driven Approach to Precision Medicine. P. Lum. Ayasdi, Menlo Park, CA.

The pharmaceutical industry is currently in the midst of a large data problem: there is far too much data to be analyzed in order to promote a viable compound, and there are too many different types of data to fuse in order to get started. There is an urgent need for software that can analyze clinical and genomic data and derive valuable insights in one place. Ayasdi’s application, Ayasdi Cure makes it easy for biologists, chemists, and physicians to consume and derive value from complex data by revealing patterns and subgroups in a visualization-based user interface. This product is built on the Ayasdi Platform which leverages hundreds of machine learning algorithms through the framework of Topological Data Analysis (TDA) to deliver operational insights from data in minutes. The product contains specific features that are best suited to pharmaceutical discovery in research including biological pathways, a gene variant viewer, and a compound visualization tool. However, this application represents a larger theme in the pharmaceutical industry: the movement towards data-driven approaches to precision medicine. With Ayasdi Cure, it is easy to incorporate patient information with chemical information to create more effective and targeted treatments, truly turning data into therapies. For example, a top 5 pharma company used Ayasdi Cure to salvage a chemical compound during clinical trials which was proven to be effective for a specific sub-population. This talk will focus on the trends in data-driven approaches to drug discovery, the types of data ecosystems that need to exist for effective analyses, and a few case studies that illustrate how Ayasdi Cure has helped pharmaceutical companies derive insights from complex data.

1674S
Improved Small RNA Library Preparation Workflows for Next-Generation Sequencing. S. Shore1, N. Paul2, M. Salcedo1, G. Zon1, C. Olsen2, K. Qaadri1, 2) TriLink BioTechnologies, San Diego, CA; 2) Biomatters, Inc. Newark, NJ.

MicroRNAs (miRNAs) are a group of small noncoding RNAs that regulate transcriptional and post-transcriptional gene expression. They regulate gene expression by binding to the 3'-untranslated region (3'UTR) of mRNAs, which leads to mRNA degradation or destabilization. Studies have demonstrated the effect of miRNAs on methylation machinery and protein expression in disease progression via epigenetic mechanisms. With the improvement of sequencing technologies, miRNAs are increasingly being characterized using deep sequencing. One of the challenges of deep sequencing small RNAs (small RNA-Seq) is in the sample preparation workflow. This workflow involves the ligation of fixed sequences (adaptors) onto the 5' and 3' ends of the starting RNA library. These adaptors are prone to mismapping with one another without a segment of the library in between, resulting in adaptor-dimer formation. Most library workflows use affinity capture to remove these adaptor dimers, which is inefficient due to the close size similarity between adapter dimers and adapter-tagged small RNA libraries. Alternatively, a gel purification step can be used, which in consequence, can depurate low-abundance sequences from the starting library. Furthermore, adapter-dimers can predominate when input library concentrations are low, thus suppressing formation of the adapter-tagged library. In this poster, we present a novel approach to small RNA library sample preparation using chemically modified adapters to disfavor adapter-adapter ligation while allowing for efficient joining of adapters onto the 5' and 3' ends of the library. Using this technology, we demonstrate improvements in specificity and yield for small RNA library preparation.

1675M
Enhanced fetal aneuploidy detection using hardware accelerated alignment. M. Sykes1, C. Roddey2, M. Ruehle2, R. McMillen2, P. Whitley2, 1) Sequenom Inc., San Diego, CA; 2) Edico Genome Inc., La Jolla, CA.

Noninvasive prenatal testing can be performed by massively parallel sequencing of DNA from maternal plasma. This method has been shown effective in the detection of fetal aneuploidies of chromosomes 13, 18, 21 and the sex chromosomes. Accurate classification of these aneuploidies requires, in part, alignment of sequencing reads to the human genome, calculation of chromosome fractions based on these alignments and calculation of z-scores for each chromosome based on these fractions. The success of these steps relies upon the choice of aligner and algorithm used to determine the chromosome fractions.

Here we present reclassification of a dataset of 1269 samples previously analyzed using bowtie 2 as the aligner. In this study alignments are generated by the DRAGEN processor, a hardware-accelerated sequencing analysis system developed by Edico Genome. We report systematic differences between the two aligners but equivalent performance in terms of chromosome fraction variability and thus chromosome quantification.

Both the bowtie 2 and DRAGEN based analyses successfully identified all known T13, T18 and T21 cases in the dataset. The sensitivity and specificity were both > 99.9% in each classification. At the same time the DRAGEN system provides speed increases of greater than thirty-fold relative to bowtie 2 running with 6 threads on a 3.5 GHz Xeon CPU, allowing a single computer to replace the efforts of a small cluster.

These results demonstrate that the classification algorithm for fetal aneuploidy is robust and resistant to localized changes in the alignment profile. Furthermore the DRAGEN system provides equivalent performance to bowtie 2 with a significant increase in speed.

1676T
Platform comparison between Ion Proton and Illumina HiSeq 2500 on a 759-gene disease panel across 248 samples. A.V. Uzilov1,2, H. Shah1,2, M. Mahajan1,2, D. Starcevic1,2, R. Sebra1,2, R. Chen1,2, E. Schadt1,2, 1) Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY.

Targeted panels allow making next-generation sequencing libraries enriched for genomic regions of interest, providing more cost-effective sequencing for the needed depth. Commercial panels and custom panel design services exist for both Ion Proton and Illumina HiSeq instruments. However, little publically available comparative data exists to guide choice of appropriate platform for large sequencing projects. We sequenced blood-extracted DNA from 248 individuals on Ion Proton and Illumina HiSeq 2500 using a custom targeted panel covering 4413 genomic regions and full exons from 759 genes selected for their genetic disease relevance. Agilent’s SureSelect hybridization-capture chemistry (covering 2.8 Mb) was used for Illumina and amplicon-based AmpliSeq (covering 4.4 Mb) was used for Proton. We look at coverage uniformity and variant concordance to compare these two platform versions in the same dataset. We compare the two chemistry/platform combinations both in terms of sample volume and panel size. 18.1% of variant calls are unique to the Illumina panel, 6.6% are unique to the Proton panel, 1.0% are conflicting calls at the same site, and 74.5% are concordant (median percent across all samples). We observe that Proton concordance depends on sequencing depth at depths seen in our study (median 118X for Illumina and 100X for Proton); As Proton depth increases, concordance increases and “unique in Illumina” calls decrease, suggesting that raising Proton depth leads it to make more calls found by Illumina. However, there is no improved concordance at higher Illumina coverage, suggesting we are operating at or above optimal depth there. To find optimal depth for Proton, we sequenced NA12878 on two Proton instruments to high depth (one barcode per 318-type chip, 972X and 766X median coverage). We show how subsampling reads impacts variant calling accuracy versus similar high-depth replicate runs on Illumina (323X and 289X median coverage). At highest depths, sensitivity for GIAB gold-standard set of known variants was 96.6% for both Illumina or Illumina and 89.0% and 89.4% for the two Proton replicates. We describe factors influencing accuracy, concordance, and callability and discuss if they can be controlled by adjusting depth or variant caller settings or whether they are a property of the panel or sequencer. Our findings will be useful to investigators when selecting technology appropriate for their NGS study.
1677S  
Sparse sufficient dimension reduction and matrix subset selection methods for big image data analysis in cancer. N. Lin, J. Jiang, S. Guo, N. Xiong. University of Texas School of Public Health, Houston, TX.  
With the advancements in modern biomedical engineering, biomedical imaging has become one of the most reliable tools for disease diagnostics and the evaluation of the progress of different treatments. A key issue for image analysis is high dimension data reduction and feature selection. The most widely used dimension reduction and feature selection methods have two serious limitations. One limitation is that these methods often use unsupervised dimension reduction method, which selects dimensions without taking information on class labels or disease status into account. The reduced data will then lose important class information that is hidden in the original data. The second limitation is that the algorithms for searching features are unable to deal with large number of features. To overcome these limitations, we introduce concepts of sufficient dimension reduction (SDR) and coordinate hypothesis which project the original high dimensional data to very low dimensional space while preserving all information on clinical outcome information, and formulate feature selection problem as a subset selection problem for matrices. We develop matrix approximation theories and error estimation methods as powerful tools for optimal image feature matrix column selection. We develop a randomized algorithm that is designed to select image features with provable guarantees that image reconstruction by selected image features can reach prespecified high accuracy. To further reduce dimensions, we extend one dimensional principal component analysis to two dimensional principle component analysis (2DFPCA) and use 2DFPCA scores to represent image data. The proposed method is applied to 186 kidney cancer histology images from different individuals. Of these 186 images, 121 of them are confirmed to have kidney cancer and the rest of them are served as the control in our study. The randomized based sparse SDR method was applied to 2DFPCA score to select features. We used support vector machine (SVM) as classifier which took selected features as input. By 5-fold cross validation, the average classification accuracy in the test dataset was 89.3% and in the training dataset was 96.4%. If we use only 2DFPCA score of the original image data as features and the same SVM as the classifier, we only can reach average accuracy of 65% and 95% in the training and test sets, respectively. This strongly demonstrates that the proposed method substantially outperforms other methods.

1678M  
A Genomics Analysis Pipeline for Cloud Computing. R.J. Mashi1, K. Ye1, N. Nutter1, D.C. Koboldt1, D.E. Larson1, K. Chen2, L. Ding1. 1) The Genome Institute, School of Medicine, Washington University, St. Louis, MO; 2) Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX.  
As medicine continues to enter the era of personalized medicine, the amount of raw genomics data is expected to increase substantially. The traditional analysis model, wherein data is downloaded from secured internet sites, e.g., data coordination centers, to local storage for processing, may easily result in an accumulation of unprocessed data if data transfer rates, storage capacities, and computing rates cannot keep pace. Individual investigators or those working in small collaborative groups may face additional logistical challenges in terms of cost, compute resources, data storage and security, and computer management expertise. One possibility for addressing these problems is to use cloud computing, which provides access to large amounts of storage and computation as a service. Technologies that enable cloud computing to be performed on large scale data storage, i.e., those that "bring the tools to the cloud," are therefore beneficial. In this study, we describe an easy-to-use web interface for configuring a genomics analysis pipeline featuring software packages that are among the state-of-the-art in variant calling tools for next-generation sequencing methods, namely, VarScan, BreakDancer, and Pindel. This combination of tools allows users to analyze for germline and somatic single-nucleotide polymorphisms (SNPs) and insertions and deletions (indels), copy-number variations, and structural variations that include translocations and inversions. The configured computing jobs can be deployed either on the physical compute cluster or on cloud/virtual compute clusters. The interface utilizes the StarCluster cluster-computing toolkit for interacting with Amazon's Elastic Compute Cloud. The software is available through the Turnkey Variant Analysis Project (TVAP) web site at http://tvap.genome.wustl.edu. Additional information is available at the ASGH 2014 Annual Meeting workshop "SeqTools to Demistify the Cloud and Genomics Analysis for Researchers Seeking Ways to Analyze High-Throughput DNA Sequencing Data." This work is supported by the National Human Genome Research Institute grant U01HG005157 (to L.D.).

1679T  
As the amount of sequencing data that is available in the public domain continues to increase, and the cost of a whole human genome sequence declines, the desire to perform large-scale analyses has grown considerably. By combining large public data sets with one's own data, the statistical power of a study can be increased in a cost effective, streamlined manner. However, the availability of a compute, storage and data-sharing platform that can facilitate these types of "big data" analyses is relatively limited. The options that are available for these studies are typically lacking in one key aspect or another such that identifying an optimal solution often involves some level of pulling together disparate, sub-optimal components. Here we describe a singular environment that includes high performance computing, data storage, data sharing and access to high value public data sets. Annai ShareSeqTM Genomic Resource is a bioinformatics and data management platform that provides a highly scalable, secure and computationally powerful solution for genomic researchers, in particular, those interested in cancer related data. We will describe the various aspects of ShareSeq as well as present examples of how this resource can be used for data analysis. A workflow manager can be used to configure an analysis pipeline by taking advantage of a variety of analysis tools that are resident in ShareSeq. Computing resources can be accessed in an on-demand manner by initiating a virtual machine that can be scaled appropriately, depending on the analysis job size. Other robust tools can be used to transfer sequencing files, create a storage repository to house them, and later, leverage metadata that has been automatically indexed, for flexible querying of the collection of files. These capabilities, in combination with easy access to large data sets, create an invaluable resource for researchers and clinicians that are using genomic data.

1680S  
Validation of a Series of Genomic StripAssays® to Salivary DNA Collection Using the DNA•SAL™ Device. P.D. Slowey1, C. Oberkanins2, A. Berndt1, G.A. Thomas1. 1) Oasis Diagnostics Corporation, Vancouver, WA US; 2) ViennaLab Diagnostics Vienna, Austria.  
PCR hybridization based StripAssays® [ViennaLab Diagnostics] provide a rapid convenient method for research and diagnostic detection of mutations associated with cancer, genetic disorders, genetic predispositions and pharmacogenomics. The assays consist of four steps: gDNA isolation, multiplex PCR, hybridization, and detection. Using DNA isolated from blood, conventional StripAssays® detect a combination of wild type and mutant probes for the identification of a variety of genetic disorders and predispositions including Gaucher disease, thalassemias, CVD and others. While blood is a common specimen matrix, invasive collection limits participant compliance particularly with small children and its utility for remote point of care applications. Due to a number of factors including the non-invasive nature of collection, ease of use and subject compliance, saliva has become an increasingly important alternative to blood for nucleic acid based research and diagnostics. While saliva is simple to collect, the proportion of epithelial cells (buccal,) and non-epithelial cells (leukocytes, mesenchymal cells) varies between individuals and can confound analysis and restrict adoption of saliva based testing. In the present study, the performance of gDNA extracted from saliva collected using the DNA•SAL™ Salivary DNA Collection Device [Oasis Diagnostics®, Vancouver WA US] and isolated using the Mini•SAL™ DNA Isolation Kit [Oasis] was evaluated in 5 different commercially available PCR assays from ViennaLab: α and β thalassemia (4-130, 4-160), Gaucher disease (4-260), CVD T (4-360) and CVD A (4-370) chosen from the ViennaLab website as tests with the highest level of complexity and sensitivity to DNA quality and quantity. For each test, 5 ul of DNA was isolated at a concentration of 2-19 ng/µL using the Mini•SAL™ protocol was used in the StripAssay®. Bands on the StripAssay® test strips were inspected visually and using ViennaLab’s own proprietary software package [Evaluator™]. All assays were run in triplicate directly comparable to results obtained using DNA isolated from blood samples. The data indicate that gDNA extracted from saliva via the Mini•SAL™ DNA Isolation Kit following collection using the DNA•SAL™ device is a robust alternative to blood DNA when used with the ViennaLab StripAssay®. According to this study, this device is a simple-to-use system with the potential for truly realizing point of care genetic testing.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1681M Direct to PCR Genomic Analysis Using Saliva Derived Samples. G.A. Thomas1, C. Oberkanins2, A. Berndt3, P.D. Slowey1. 1) Oasis Diagnostics Corporation, Vancouver, WA; 2) Viennalab Diagnostics, Vienna, Austria. Saliva has become an increasingly important sample matrix as an alternative to blood for nucleic acid based research and diagnostics utilizing PCR, hybridization, microarrays and next-generation sequencing due to the non-invasive nature of the collection, ease of use and participant compliance. While saliva may be readily collected, the proportion of epithelial cells (buccal, leukocytes and non-human nucleic acid from microbial sources varies between individuals and can confound analysis. Recently, the utility of DNA extracted from saliva has been validated as a robust alternative to blood for use in PCR based strip assays. The StripAssay® technology [ViennaLab] consisting of four steps: gDNA isolation, multiplex PCR, hybridization and detection provides a rapid convenient method for research and diagnostic detection of mutations associated with cancer, genetic disorders and pharmacogenomics. In the study, DNA extracted from saliva using the DNA•SAL™ Salivary DNA Collection Device [Oasis Diagnostics®] and isolated using the companion Mini•SAL™ isolation kit [Oasis] performed identically to blood DNA in 5 different multi-probe StripAssays®. To test the suitability of DNA•SAL™ collected saliva samples for direct PCR analysis without DNA isolation, the performance of independently collected saliva samples in two genetic disorder tests from Viennalab [FY-PTH-MTHFR (4-260) and Hemochromatosis (4-210)], was evaluated as a function of input sample dilution. Both StripAssays® examine multiple wildtype and mutant probes (4-260: target V G1691A, prothrombin G20210A and MTHFR C677T. 4-210: HFE H63D, S65C and C282Y). Normal test results were obtained for all samples when simply diluted with water within the range of 1:5 to 1:100. Half of undiluted samples failed presumably due to inhibition of PCR activity, while all samples failed at a dilution of 1:500 presumably due to excessive dilution of the target. The data demonstrates that saliva collected using the DNA•SAL™ Device when diluted within the range of 1:5 to 1:100 can be used directly in PCR assays without substantial signal loss. The elimination of the requirement for DNA isolation when using the DNA•SAL™ collection tool in tandem with PCR hybridization StripAssays® greatly facilitates the potential for truly realizing remote point of care genetic testing.


Detailed information on spatial patterns of gene expression in complex tissues is needed to enable new insights into the molecular differences between normal and disease states. Although in situ hybridization has excellent resolution, it is limited in multiplexing. Conversely, microarrays and RNA-seq multiplex well but are not designed to provide spatial information. Laser capture microdissection enables highly multiplexed analysis of well-defined sites in a tissue section but is relatively slow and laborious. We developed a scalable assay system that is both spatially resolved and enables high-throughput multiplex overcomes limitations of the methods listed above. The assay involves binding of probes to target RNAs in the sample, and encoding the probes by spatially-addressed delivery of sequence-encoded adapter molecules. The sequences of the resulting assay products and their frequency provide both abundance and location data for each target RNA in the sample. We performed the spatial addressing using microfluidic devices that we designed to define a rectangular array of assay sites. Assay products were collected and sequenced in bulk using an Illumina next-generation DNA sequencer. We used this system to make the expression of 134 target RNA sequences from 69 genes across an array of 256 positions on formalin-fixed paraffin-embedded (FFPE) pancreas and liver samples. We constructed virtual images of the FFPE sections based on the data revealing local regions of high gene expression that were well correlated with immunofluorescence data. For example, we mapped regions of high insulin mRNA abundance to islets of Langerhans in pancreas by correlation with insulin protein abundance determined by immunofluorescence. The technology is flexible. We have used the devices to analyze both individual and pooled mRNA samples for a variety of diseases. The technology currently achieves an effective pixel resolution of ~50 × 50 microns by using disposable microfluidic devices instead of complex imaging instrumentation. Although the resolution of our devices is currently lower than that of optical imaging, a 0.5X coverage of 400 samples is achieved by performing 400 different samples to be carried out at many sites in parallel. We are currently developing methods to assay proteins as well as RNAs using this novel technology, and we envision our assays enabling high-dimensional analysis of gene expression in both normal and disease samples.


The analysis of the metagenome of human microbiomes has highlighted the fundamental role of the microbiome in human health. Though the metagenomics field has just recently emerged, there is a paradigm shift from the understanding of the microbiome composition in terms of taxonomical classification towards a functional understanding of the microbiome (the metatranscriptome). Such shift is motivated by the recent findings that functional analysis of the metatranscriptome better addresses the underlying complex molecular interplay which affects the emergence of different phenotypes and diseases. This paradigm shift means that future microbiome studies will be complemented with a metatranscriptome analysis. Human microbial metatranscriptome sequencing poses significant challenges. Specifically, prokaryote non-ribosomal RNA typically constitutes a very small fraction of the total RNA, where ribosomal RNA (rRNA) and contaminating host RNA are the overwhelming majority of transcripts (up to 99%). Despite the urgent need to avoid sequence rRNA, the existing technologies to deplete all rRNA using subtractive hybridization are far from perfect. Here we describe a novel technology that allows specific capture of the active and informative transcriptome of all species of bacteria and archaea while removing both the host RNA and rRNA in a single step. Unlike subtractive hybridization where the RNA sequence must be known a priori to generate RNA probes, our strategy is universal. All prokaryotic and archaeal mRNAs and small RNAs will be captured. On the other hand, processed transcripts such as mature ribosomal and transfer RNAs (rRNA, tRNA) from all organisms and mRNA from eukaryotes will be discarded. As a proof of principle for metatranscriptomics analysis, we demonstrated that mRNA from E. coli could be enriched 10 fold from a mixture of total human RNA and total E. coli RNA while maintaining the relative representation of each prokaryotic transcript.

1684M Extremely low-coverage whole genome sequencing in South Asians captures population genomics information. N. Rustagi1, A. Zhuo2, S. Wang3, N. Ramesh1, W.S. Watkins4, D. Muzny4, R.A. Gibbs1, L.B. Jorde4, F. Yu1, J. Xing3,4. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Rutgers, the State University of New Jersey, Department of Genetics, Piscataway, NJ; 3) Rutgers, the State University of New Jersey, Human Genetics Institute of New Jersey, Piscataway, NJ; 4) University of Utah School of Medicine, Eccles Institute of Human Genetics, Salt Lake City, UT.

The cost of whole genome sequencing (WGS) has been reduced considerably in recent years due to advances in next-generation sequencing technologies (NGS), but many aspects of large-scale cohort WGS studies are still daunting. Challenges include the requirements for quick turn-around times and high variant calling accuracy. We have focused on sensitive variant site discovery, robust association study replication and comprehensive population genomics characterization. The ability to identify common variants (minor allele frequency (MAF) > 5.0%) is a central aim of such studies (Li Y. et al., 2011).

To develop a cost-effective path to variant discovery, we evaluated extremely low-coverage whole genome sequencing (WGS) by sequencing 185 Indian individuals followed by imputation using BEAGLE with reference panels from the HapMap project and the 1000 Genomes project. We applied two variant discovery pipelines, SNPTools and GATK, and generated a consensus call set. Comparisons with previously published SNPs and sequencing results show the consensus set has an over 90% sensitivity for identifying variants with a MAF> 5%, with a false discovery rate (FDR) <15%. Imputation provides further improvement in sensitivity for both common and rare variants. Restricting standard population genomics analysis to shared regions with high overall coverage, shows a high level of consistency with past results. This data set will enable in-depth understanding of genomic diversity and detailed inference of demographic history in South Asia.

We further characterized the expected sensitivity and accuracy of SNP calling in extremely low-coverage sequencing studies using simulations. We down-sampled from the 1000 Genomes data to reflect coverage of 0.25x and 0.75x and found ~ 500 samples with 0.75x coverage are sufficient to recover > 60% common SNPs (MAF > 5%) with an FDR <3%. For the dataset with 0.25x coverage, >40% SNPs are recovered from ~500 samples with MAF> 20% and FDR <3%. Only 350 samples with 0.75x coverage will yield 80% true SNPs, with MAF >= 20%. Overall, we demonstrate that extremely low-coverage WGS with imputation can be a useful study design for variant discovery with a dramatically reduced cost, even in populations without available reference data.
Advancing Clinical Diagnostics Using Whole Exome Sequencing. D.1685T

1687T Unique Haplotype structure determination in human genome using Single Molecule, Real-Time (SMRT) sequencing of targeted full-length fosmids. K. Eng1, R. Hall2, H. Hon1, C. Patel1, D. Geraghty2, S. Rana2, 1) Pacific Biosciences, Menlo Park, CA; 2) Fred Hutchinson Cancer Research Center, Seattle, WA.

Determination of unique individual haplotypes is an essential first step toward understanding how identical genotypes having different phases lead to different biological interpretations of function, phenotype, and disease. Genome-wide methods for identifying individual genetic variation have been limited in their ability to acquire phased, extended, and complete genomic sequences that are long enough to assemble haplotypes with high confidence. We explore a recombininge approach for isolation and sequencing of a tiling of targeted fosmids to capture interesting regions from human genome. Each individual fosmid contains large genomic fragments (~35 kb) that are sequenced with long-read SMRT® technology to generate contiguous long reads. These long reads can be easily de novo assembled for targeted haplotype resolution within an individual’s genomes. The P5-C3 chemistry for SMRT sequencing generated contiguous, full-length fosmid sequences of 30 to 40 kb in a single read, allowing assembly of resolved haplotypes with minimal data processing. The phase preserved in fosmid clones spanned at least two heterozygous variant loci, providing the essential detail of precise haplotype structures. We show complete assembly of haplotypes for various targeted loci, including the complex haplotypes of the KIR locus (~150 to 200 kb) and conserved extended haplotypes (CEHs) of the MHC region. This method is easily applicable to other regions of the human genome, as well as other genomes.

1688T Examination of the Performance of Whole Genome Amplified DNA across Multiple Capture Methodologies and Sequencing Platforms. B. Hicks1,2, C. Dagnall1,2, M. Malasky1,2, H. Lee1,2, R. Eggensperger1, J. Mitchell1,2, W. Luo1,2, X. Zhang1,2, C. Chung1,2, M. Yeager1,2, K. Jones1,2

1) Cancer Genomics Research Laboratory, Leidos Biomedical Research, DCEG, NCI, Bethesda, MD; 2) Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), Rockville, MD, USA.

While input requirements for targeted, whole exome and whole genome sequencing have been driven down considerably by the advent of new and improved technologies, often whole genome amplified DNA (wgaDNA) remains the only source of critical and informative subjects. Additionally, the advent of single cell technologies demands use of wgaDNA in downstream analysis. We performed whole exome and targeted sequencing on a series of matched genomic DNA (gDNA) and wgaDNA samples, utilizing different capture techniques, including the hybridization based Roche NimbleGen SeqCap EZ Human Exome v3.0, the amplification based ThermoFisher Ion AmpliSeq Exome, and two custom Ion AmpliSeq panels (47 amplicons and 856 amplicons). The captured libraries were sequenced on the Illumina HiSeq 2500, the ThermoFisher Ion Proton, and the Ion Torrent PGM, respectively. Data filtering, alignment and variant calling was done per internal standard protocols. We reviewed overall coverage statistics, as well as positional coverage variance in areas of high GC content, centromeric and telomeric regions. Sensitivity and specificity of the different approaches was determined via use of the NIST reference genotype data from the well characterized National Array 2012B GeneMap1000 dataset. We compared sequence coverage statistics (mean coverage for Illumina platform and uniformity for Ion platforms) show sequencing coverage is comparable between gDNA and matched wgaDNA samples. Duplication rates for the NimbleGen/Illumina samples were slightly (0.5%) higher for the wgaDNA source, but well within acceptable limits. All approaches showed some bias towards qDNA in coverage of high GC content areas and telomeric regions, with the amplification based capture methods being most susceptible to differential coverage in these specific areas. SNV concordance for the NimbleGen/HiSeq qDNA and wgaDNA as compared to the high confidence genotypes from GiaB was 99.7% and 99.8% respectively, with indel concordance rates of 98.3% and 98.2%. SNV concordance for the AmpliSeq Exome/Proton gDNA and wgaDNA was 98.9% and 99.1%, with indel concordance rates of 90.7% and 90.9%. False negative rates were higher when utilizing AmpliSeq Exome, but only a slight increase (1.8% for SNVs and 5.5% for indels) when comparing wgaDNA to qDNA. These results show an ability to utilize wgaDNA for all sequencing platforms, with some caution on use of amplification based capture techniques.
1690M  Lower Cost, Higher Throughput Library Preparation with the Echo liquid handler® and the NuGEN Ovation® Single Cell RNA-Seq System. J.D. Lesnick 1, J.D. Heath 2. 1) Labcyte Inc, Sunnyvale, CA, USA; 2) NuGEN Technologies, San Carlos, CA, USA.

The rapid evolution of next-generation sequencing (NGS) technologies is accelerating our knowledge of gene expression, regulation and pathway complexities in mammalian cells. Transcriptome analysis with NGS offers increased transcript coverage to enable the detection of rare transcripts, novel alternative splice isoforms and the measurement of transcript abundance. However, traditional library preparation methods for NGS are often not amenable to transcriptome analysis. Traditional methods carry a requirement for a large amount of total RNA to yield sufficient mRNA to analyze which can be unobtainable or cost prohibitive in most experiments. The NuGEN Ovation Single Cell RNA-seq System is a highly sensitive and complete library preparation procedure for whole-transcriptome sequencing that requires total RNA from samples as small as a single cell or 10 pico-grams. In this study, the Ovation Single Cell RNA-Seq System was validated by examining the resulting library complexity, reproducibility, and evenness of transcript coverage. Furthermore, a majority of the sample and reagent transfers were automated using the Echo acoustic liquid handling technology. Echo liquid handlers transfer a wide range of fluids without contact of tips or recalcification between fluid types. The industry leading accuracy and precision of Echo liquid handlers at microliter and nanoliter volumes in combination with the NuGEN Ovation Single Cell RNA-Seq System increases library preparation throughput while reducing the costs to enable a broader application of transcriptome analysis with NGS.


The HiSeq X Ten was designed to unlock the potential of population sequencing, delivering the world’s first $1000 human genome at 30x coverage. By using flow cells with billions of patterned nanowells at fixed locations, we have significantly increased the density of clusters that could be achieved by random clustering. To support clustering on patterned flow cells we developed a new cluster chemistry called Exclusion Amplification (ExAmp). Here we describe an improved ExAmp v2 formulation that supports simultaneous seeding and amplification of TruSeq PCR-free libraries, the gold-standard for genome sequencing due to more uniform coverage of regions that are typically underrepresented in amplified libraries (high GC-rich regions, promoters and repetitive regions). This new formulation continues to support amplification of TruSeq Nano PCR libraries and also offers a significant improvement in genomic coverage uniformity. We presently are showing improved coverage uniformity for challenging regions of the genome whilst maintaining high quality sequence at genes and exons. Furthermore, by utilizing the Platinum genome data sets from the CEPH pedigree, we have validated our data and found excellent sensitivity and specificity of SNP calling with the improved ExAmp clustering.


Targeted resequencing of exonic regions (subsets to whole exome) using hybridization-based capture has been used as a key discovery tool for understanding many biological problems including disease mechanisms of Mendelian diseases and cancer. In addition, discoveries from next generation sequencing (NGS) efforts are increasingly being used to design targeted panels for clinical research and diagnoses. However, as targeted sequencing moves to the clinical setting, very few automation solutions exist that minimize process variability and reagent costs for low and mid-tier labs. Furthermore, labs are asked to perform multiple tests with limited amount of material. To address these needs, we have tested our the Apollo 324 NGS Library Prep System to enrich for exomes from NGS libraries prepared from low input material. The Apollo 324 is a flexible automation system that has been shown previously to process up to 8 or 32 samples in a single run. The number of samples to run in a given day is matched to the amount of reagents/sample to eliminate reagent waste. Genomic DNA and cDNA library preparation kits for the Apollo 324 robot can process samples for whole genome sequencing, RNA-seq and Chip-Seq experiments. Now, we are extending the utility of automation to include the ability to perform targeted capture. Commercially available target enrichment kits (SureSelect, Nimblegen) were processed manually according to manufacturer’s instructions and compared with the automated system at different input levels of gDNA from HapMap and tumor samples. Data will be presented from exome-enriched samples from manual and automated workflows, sequenced using a HiSeq system and compared for their ability to identify variants.


Background: The throughput of Next-Generation Sequencers has increased dramatically since their advent. However, the read length remains relatively short among sequencers which are able to generate large number of reads, for example, the 150nt maximum officially supported by Illumina on HiSeq 2500. Such read length is generally sufficient for alignment to the genome, but longer read length, even to some small extent, is needed for de novo assembly of genomes and transcriptomes. Methods: The observed error rate of HiSeq 2500 Rapid Run is generally below 1% at the very end of 150nt reads, therefore further extending the read length might yield slightly decreased but still acceptable read quality as long as the sequencing reaction is optimized accordingly. Pair-End Rapid Runs of 215+230nt and 215+215nt reads were performed on HiSeq 2500 with slight modification of the standard sequencing procedure. Small amount of PhiX control library was spiked-in to each lane to estimate the actual sequencing error profile in addition to the Illumina Quality Score. Size-selected libraries of narrow size range were used to generate clusters of homogeneous size and the clustering density was reduced from the standard value to help improve the signal-to-noise ratio and read quality. Result: By controlling the raw cluster density around 600K/mm², the overall error rate of 215nt read pairs could be kept below 1%, while the per cycle error rate at the very end of these reads remained below 2%. The Illumina Q30 bases of each lane could be maintained over 80%. The bias in the base call composition became more noticeable beyond the official read length of 150nt, especially when the cluster density is relatively high. Very low cluster density could lead to reduction in the base composition bias, although not much improvement in the base call accuracy. In conclusion, we have optimized the sequencing condition to extend the HiSeq 2500 read length from 150nt by 43% to 215nt. At the expense of reduced cluster density, the quality of the extended reads is mostly maintained.
1693M
Genome-wide transcriptome enrichment sequencing for research and clinical applications. H. Doddapaneni1, J. Hu1, H. Chao1, X. Liu1, S. White1, K. Walker1, C.J. Buhay1, M. Wang1, M. Bellair1, L. Wang1, K.R. Covington1, A. Roy2,3, P. Sumazin1, S.E. Plon1,3,4, D.W. Parsons1,3,4, D.A. Wheeler1, E.A. Boerwinkle1,5, D.M. Muzny1, R.A. Gibbs1. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX; 3) Department of Pediatrics, Baylor College of Medicine and Texas Children's Hospital, Houston, TX; 4) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 5) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX.

Transcriptome sequencing (RNA-Seq) together with whole exome sequencing (WES) offers an integrated informative dataset for clinical diagnosis and characterization of human transcriptome, especially for characterization of cancer etiology and outcomes. For example, detection of pathogenic gene fusion events is important to accurately characterize childhood sarcomas where morphologic characteristics or WES is not sufficient. However, current RNA-Seq methods rely heavily on very high quality RNA samples, which are often not available in routine clinical setting. Towards the goal of incorporating RNA capture methods in a clinical setting, we have developed an RNA-Seq capture protocol combining the BCM-HGSC strand-specific RNA-Seq protocol with our CLIA certified WES protocol. Inclusion of a capture step enables RNA-Seq to become RNA quality independent as it provides probes to capture the targeted transcripts. Libraries were prepared using Universal Human Reference RNA and co-captured (4-plex) using VCRome 2.1 capture reagents and sequenced on HiSeq 2500. On average, 66 million reads were generated per sample and these data were compared to standard RNA-Seq data generated for the same sample. Capture libraries had a significantly higher mapping rate (95.1% vs 77.5%), higher exonic rate (93.4% vs 77%) and lower intergenic rate (0.2% vs 4.2%) compared to RNA-Seq libraries. Correlation of expression values (in FPKM) between the replicate capture libraries was excellent (R² = 0.9882). Capture libraries detected 18,903 protein coding genes compared to 18,611 genes found in RNA-Seq libraries. There were 756 genes unique to capture libraries (FPKM-0.22) and 289 genes unique to RNA-Seq libraries (FPKM-2.8). The low FPKM values observed for the novel genes found in capture libraries suggest ‘enrichment’ of low expressed transcripts due to presence of their probes. Further, use of a genome-wide capture probe such as VCRome allows for capture and validation of novel fusion transcripts. We are validating the protocol on samples with 13 known fusion transcripts found in heptoblastoma and pediatric sarcoma samples that will then serve as test controls. Concurrently, we are also testing RNA of different types (RIN numbers, input amounts) and from FFPE. We anticipate that our RNA-Seq capture protocol will provide robustness and through high throughput process automation, will result in fast turnaround times to have practical utility in both research and clinical settings.

1694T
Whole Genome Sequencing on DNA extracted from Saliva: a systematic evaluation of SNV, CNV and Structural Variant Calling. S. Germer1, A.K. Emde1, A. Abyankar1, B. Correa1, D.M. Oschwald1, T. Hu-Selig1, P. Yurttas Beim1, R.B. Darnell1. 1) New York Genome Center, New York, NY 10013; 2) Celmatix, 1 Little West 12th St. New York, NY 10014.

As whole genome sequencing (WGS) becomes increasingly affordable, DNA repositories are being established for a variety of research purposes. These repositories often rely on blood specimens, which presents several practical hurdles. The ability to source DNA from saliva would reduce costs and increase flexibility for many researchers. Given sample procurement is a common bottleneck in genomics studies, this development would have significant implications. Previously, it has been reported that saliva yields DNA of sufficient quality to reliably make single nucleotide variant (SNV) calls in WGS datasets. Here we address whether larger variants, including copy number and structural variants (CNVs and SVs) can also be reliably detected from saliva-extracted DNA. We evaluated WGS data (~30x coverage, Illumina HiSeq) generated from matched blood and saliva (DNA Genotek, Oragene) samples selected to represent a spectrum of bacterial contamination levels (avg~10%). Saliva DNA showed both a reduced alignment rate to the human genome (0.74-0.93) compared to blood sample DNA (>0.95) and a corresponding reduction in mean genome-wide coverage. A pre-sequencing qPCR assay predicted bacterial contamination rates that correlated with the alignment rates for the human genome (n=0.94). SNV calls from both sets of samples (GATK best practices pipeline) showed similar rates of discordant calls between replicate saliva and blood samples (~0.02% at 99% VQSR tranche). Similarly, DNA from saliva and blood displayed very similar Mendelian inheritance errors (0.01%–0.05%). To investigate the effect of replication on SV calls in saliva, we performed a dual analysis (a) replicate analysis of saliva-blood SV calls from GenomeSTRiP. Across matched saliva-blood pairs we did not identify any somatic SVs that could be confirmed by replicates or any somatic CNVs in GenomeSTRiP calls exhibited greater variability, but the variability between blood and saliva was comparable to the variability between replicates and largely due to differences in genome coverage. Therefore, the discordant calls likely represent a lack of sensitivity/specificity in the calls made by GenomeSTRiP. Our study suggests that DNA from saliva can yield comparable WGS results to DNA from blood across a range of variant sizes and types, especially when coupled to pre-sequencing estimation of bacterial contamination to allow for comparable mean genome coverage.

1695S

Whole-genome amplification (WGA) is commonly used upstream of next-generation sequencing and microarray protocols when available genetic material is severely limiting, such as in single cell work. However, WGA can suffer from sequence-specific bias, allelic drop out, and non-uniform coverage, hampering data interpretation. To quantitatively assess bias introduced by WGA, we used Droplet Digital™ PCR (ddPCR™), targeting multiple sites across the genome. Limited samples were prepared using standard WGA protocols. The use of droplets to improve WGA output will be discussed.
1696M

Whole exome sequencing (WES) and target panel sequencing (TPS) are currently the most popular applications in the NGS community. Although ThruPLEX-FD has already been approved for WES and TPS applications in the laboratories of current customers, Rubicon wants to make its internal protocols available publicly for the convenience of all its customers. Here, we describe experimental results answering the following questions: Can ThruPLEX-FD be easily used for enrichment with Agilent XT2, Roche Seq Cap EZ for WES and TPS? (protocol establishment) - How do ThruPLEX-FD enrichment metrics compare with the standard Roche (Kapa) and Agilent libraries? - What are the advantages of using ThruPLEX-FD? Library synthesis: Three inputs were used to build ThruPLEX-FD libraries with Covaris sheared 200bp average human gDNA, respectively 500pg, 10ng and 50ng. Recommended low input libraries for Roche (10ng) and Agilent XT2 (100ng) were prepared with the same gDNA and enriched according to the vendor protocols. Capture: ThruPLEX-FD libraries were pooled with the Agilent library for hybridization (different barcode systems), however hybridized separately from the Kapa library (identical barcode system). Vendor-specified reagents and protocols were used except for addition of ThruPLEX-FD compatible blocking oligonucleotides IDT xGen® Universal Blocking Oligo - TS-p5 and xGen® Universal Blocking Oligo - TS-p7 (6nt), spiked into ThruPLEX-FD hybridizations at the concentrations suggested by the vendor. Sequencing: Pools of the enriched libraries were sequenced on Illumina MiSeq, and the reads from each library were down-sampled to 3.9M read-pairs. Reads that contained the same phasing block were also combined as a function of phasing block overlap. For all pairs that are on the same phasing block, the chance that a pair is phased correctly, again as a function of distance is plotted. From these results it can be concluded that phasing yield is in Mb-scale (>80% SNPs sharing phasing block) and accuracy of 99.9% and 99% extends to 100kb and 200kb pairwise SNP distances, respectively. We demonstrate utility of this imputation-free approach by accurate and complete phasing of de novo and compound heterozygous SNPs. We anticipate the scalable, rapid, and cost-effective workflow could enable haplotype resolution to become routine in human genome sequencing.

1698S
Contiguity Preserving Transposition Sequencing (CPT-seq): Haplotype-resolved sequencing and assembly. J. Fisher1, C. Turk1, L. Christiansen1, F. Zhang1, E. Kostem2, S. Amini2, M. Ronaghi1, A. Axley2, K. Shendure2, K. Gunderson3, F. Steemers2. 1) Advanced Research Group, Illumina, Inc, San Diego, CA, 92122, USA; 2) Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. Haplotype-resolved genotyping enables medical interpretation of genetically relevant variant data, deep inferences regarding population history, and the non-invasive prediction of fetal genotypes. We describe a novel library preparation and haplotyping approach that enables haplotype analysis with a high percentage of short sequencing reads as little as 2-4 lanes of HiSeq over the standard VCF. The method employs a Nextera TruSeq transposition approach that preserves the contiguity and ordering information of the target DNA sequence while inserting common primer, and optionally index sequences at 100-200 bp intervals. Subsequently, the contiguously linked-libraries are processed through the assay while maintaining this contiguity information. Finally, the individual short libraries of each DNA molecule are resolved through sequencing. The method effectively generates long-strobed reads that are on average ~30-50kb in length with ~5-15% coverage. Phased strobed reads can be as long as the physical length of input DNA. Additionally, we have greatly streamlined the haplotyping protocol. The total assay time is <3 h with ~30 min hands-on-time and the workflow can be integrated with robotic or microfluidic processing. As a proof-of-concept, we apply this method to phase over 95% of heterozygous variants from a HapMap trio into long, accurate haplotype blocks with N50 block sizes ranging from 1.4-2.3 Mb. We evaluated phasing yield and accuracy of the current method by plotting the probability that heterozygous SNP pairs are phased correctly as a function of distance between them. For all pairs that are on the same phasing block, the chance that a pair is phased correctly, again as a function of distance is plotted. From these results it can be concluded that phasing yield is in Mb-scale (>80% SNPs sharing phasing block) and accuracy of 99.9% and 99% extends to 100kb and 200kb pairwise SNP distances, respectively. We demonstrate utility of this imputation-free approach by accurate and complete phasing of de novo and compound heterozygous SNPs. We anticipate the scalable, rapid, and cost-effective workflow could enable haplotype resolution to become routine in human genome sequencing.

1697T
Identify enhancer elements at genome-wide scale using MIT-seq. X. Wu, L. Su, D. Wells. Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Frederick, MD.

Enhancer elements are important for regulation of gene expressions. However, enhancers are not easy to identify and they are far away from genes. Retrovirus integrates into the host genome as part of their life cycle and leaves a permanent tag at the location of integration. Different retroviruses use different mechanisms for target selection. Recently, we have demonstrated that murine leukemia virus (MLV) integration site clusters correlate highly with enhancer elements in the human genome. MLV retroviral vector has been widely used for gene delivery and is highly efficient to transduce a broad range of cells. Using linker-mediated PCR, we can easily identify millions of MLV integration tag by sequencing (MIT-seq). The MIT-seq tags are analyzed similarly as ChiP-seq tags to identify peaks and clusters, making this a new approach to identify enhancer elements in a variety of cells. (Funded by FNLCR Contract HHSN261200800001E).

1699M
A Method for Selectively Enriching Microbial DNA from Contaminating Vertebrate Host DNA. E. Yigit1, G. R. Feehery1, F. J. Stewart1, S. O. Oyola4, Y. Wei Lim2, B. W. Langhorst1, V. T. Schmidt2, E. T. Dimalanta1, A. L. Amaral-Zettler4, T. Davis1, M. A. Quail2, S. Pradhan1. 1) New England Biolabs, Ipswich, MA; 2) Wellcome Trust Sanger Institute, Cambridge, UK; 3) Department of Geobiology, Brown University, Providence, RI, USA; 4) The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA, USA; 5) Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, USA; 6) Department of Geobiology, Brown University, Providence, RI, USA.

Recent discoveries have implicated the human microbiome as playing a role in certain physical conditions and disease states, and these advances have opened up the potential for development of microbiome-based diagnostics and therapeutic tools. The majority of microbiome DNA studies to date have employed 16S analysis, but these provide very little information regarding function. In contrast, sequencing of the total DNA of a microbiome sample provides a broader range of information including genes, variants, polymorphisms, and putative functional information. However, many samples, including those derived from vertebrate skin, bodily cavities, and body fluids, contain both host and microbial DNA. Since a single human cell contains approximately 1,000 times more DNA than a single bacterial cell, even low-level human cell contamination can substantially complicate the analysis of a sample. In some cases, as low as 1% of sequencing reads may pertain to the microbes of interest and a large percentage of sequencing reads must be discarded, making such experiments impractical. To address this issue, we developed a method to enrich for microbial DNA using methyl-CpG binding domain (MBD) to separate methylated host DNA from microbial DNA. Importantly, microbial diversity and relative abundance is maintained after enrichment. This simple magnetic bead-based method was used to remove human or fish host DNA from bacterial and protistan DNA. We describe the enrichment of DNA samples from human blood, a mock malaria-infected blood sample, human cytosin fibrosis sputum, and a black molly fish, followed by next generation sequencing on multiple platforms. Sequence reads aligning to host genomes were reduced approximately 10-fold, while the percentage of sequence reads corresponding to microbial sequences increased approximately 10-fold. This new method for microbiome sequence analysis holds promise for use with a variety of sample types, enabling enrichment while accurately reflecting the diversity of the original sample.

The recent progress in single-cell analysis has unlocked clinical interests of women's health and infertility research like endometrial pathologies, implantation failure and etiology of endometriosis. Although woman endometrium is extensively studied, the most essential players - individual endometrial cells, have gained little attention. The objective of my study is to take deeper view on endometrial functioning at individual cellular level by providing gene expression based phenotyping and defining the functionally active major cell populations and their molecular interactions. This goal will be achieved by using clinical endometrial biopsies and combining the most advanced cell- and molecular-biology techniques, like scrutinized biopsy handling from collection to treatment, multiple-color FACS analysis with single-cell sorting option, and highly multiplex single-cell transcriptome analysis by RNA-sequencing. Voluntary healthy fertile women donate the endometrial biopsies during the proliferative and mid-secretory, e.g. receptive phase of the natural menstrual cycle. The samples are cryopreserved in proper media and the thawed tissues are enzymatically disaggregated with the cell suspensions containing the epithelial- (CDH1+), stromal- (CD13+), and the rest repertoire of living cells. During the sample treatment, minimum manipulation times at +4°C are used to retain the natural gene expression profile. Cells are labelled by combination of surface marker antibodies for single-cell FACS sorting directly into 96-well plate lysis buffer and are converted to 48-plex illumina libraries using Single-cell Tagged Reverse Transcription (STRT) protocol. After cell sorting and cDNA synthesis, qPCR-based QC assay is used to evaluate each well/cell's sorting accuracy and cDNA yield before pooling and further library preparation. FACSArial single-cell sorting accuracy and cDNA turn-around success rate is 90-100%. This multidisciplinary study have faced us to different challenges like (i) proper biopsy cryopreservation and treatment procedure to prepare stable single-cell solution in minimum time, (ii) single-cell QC after FACS and cDNA synthesis, (iii) capture methods where random shearing is deployed, it is not possible for capturing a wide range target region sizes. In contrast to hybridization methods, Amplicon based sequencing B-cell antibodies and T-cell receptors is gaining in popularity due to recent throughput and read length improvements in next-generation sequencing (NGS) technologies. However, the structural and sequence complexities of antibody genes have made reliable targeting approaches challenging. We have developed and optimized a method for accurate sequencing of the immune gene repertoires of B-cells and T-cells. In contrast to previous studies, our method generates full-length sequences of B-cell antibody and T-cell receptor genes. This allows for exhaustive somatic mutation profiling across complete V, D and J segments, full isotype information analysis (IgM, IgD, IgG, IgA and IgE), and the possibility for synthesis and expression of complete antibody chains for downstream immunological assays. By introduction of a unique barcode ID into every captured mRNA molecule, all PCR copies of each mRNA fragment can be collapsed into a single consensus sequence, making the assay extremely accurate by resolving PCR bias and sequencing errors, as well as allowing quantitative digital molecule counting. The assay can work with as low as sub-nanogram levels of input total RNA and is the first method that allows targeted amplification of all possible antibody heavy and light chains (IGH+IGK+IGL) and T-cell receptors (TCRA+TCRB) in a single reaction simultaneously.


Immune sequencing, which allows for the study of complex immunological diseases by sequencing B-cell antibodies and T-cell receptors, is gaining in popularity due to recent throughput and read length improvements in next-generation sequencing (NGS) technologies. However, the structural and sequence complexities of antibody genes have made reliable targeting approaches challenging. We have developed and optimized a method for accurate sequencing of the immune gene repertoires of B-cells and T-cells. In contrast to previous studies, our method generates full-length sequences of B-cell antibody and T-cell receptor genes. This allows for exhaustive somatic mutation profiling across complete V, D and J segments, full isotype information analysis (IgM, IgD, IgG, IgA and IgE), and the possibility for synthesis and expression of complete antibody chains for downstream immunological assays. By introduction of a unique barcode ID into every captured mRNA molecule, all PCR copies of each mRNA fragment can be collapsed into a single consensus sequence, making the assay extremely accurate by resolving PCR bias and sequencing errors, as well as allowing quantitative digital molecule counting. The assay can work with as low as sub-nanogram levels of input total RNA and is the first method that allows targeted amplification of all possible antibody heavy and light chains (IGH+IGK+IGL) and T-cell receptors (TCRA+TCRB) in a single reaction simultaneously.

1703T Optimized DNA extraction and repair improves the yield and quality of sequencing libraries derived from FFPE samples. L. Chen1, E. Rudder2, G. Dunn1, T. Evans1, L. Ettwiller1, 1) New England Biolabs, Inc. Ipswich, MA 01938; 2) Covaris, Inc. Woburn, MA 01801.

Cancer biopsy samples are routinely formalin fixed and embedded in paraffin in order to preserve the morphological features of suspected tumor samples. While these formalin fixed, paraffin embedded (FFPE) samples are valuable for archival histological data, the formalin treatment is detrimental to the quality of nucleic acids. Indeed, DNA isolated from archived tissue is often of poor quality and damaged due to fragmentation, oxidation, deamination, and protein-DNA crosslinks. Thus, it remains challenging to 1) extract DNA from FFPE samples and 2) obtain high quality DNA for NGS library preparation. Here we successfully coupled Covaris® truXTRAC™ to extract FFPE DNA with PreCR®-B Repair Mix to repair DNA damage. Covaris truXTRAC use highly controlled acoustic energy to efficiently remove paraffin from FFPE cores, sections, and slides enabling efficient tissue rehydration, tissue denaturation, crosslink reversal, and nucleic acid release. The PreCR®-B Repair Mix is an enzyme cocktail formulated to repair damaged template DNA and blunt ends in one step. PreCR® is active on a broad range of DNA damage, including modified bases, nicks and gaps, and a variety of blocking moieties at the 3’ end of DNA. We found that DNA repair pre-treatment of extracted FFPE DNA samples consistently increased library yield up to 2-fold in the four samples tested. Furthermore, downstream data analysis of the sequencing reads demonstrated a remarkable reduction of sequencing mismatches. In conclusion, the incorporation of PreCR®-B in library preparation workflows remarkably improves the quantity and quality of NGS libraries from FFPE samples resulting in more robust library preparation and more reliable base calling.
1704S
Targeted enrichment of forensically relevant STRs for improved human DNA profiling. M.R. Nandini, S.R. Gadi pattly, A. Sarkar. Lab of DNA Fingerprinting Services, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, Telangana, India. Human DNA profiling is employed worldwide for unambiguous identification of individuals in mass fatality incidents, criminal cases and also to establish parentage by genotyping repeat length polymorphisms in 16-18 highly polymorphic short tandem repeats (STRs) or microsatellites. DNA typing involves a multiplex PCR-based assay to generate STR profiles from even trace amounts of available biological source in a reasonably short duration. In spite of the availability of various commercial STR kits for DNA typing, there is a need for development of newer technologies to overcome limitations posed by challenging and recalcitrant forensic samples like those containing excess non-human DNA and PCR inhibitors, which often leads to allele/locus dropouts, peak imbalances and loss of signal during STR profiling. In the present study, a sequence-specific DNA enrichment strategy has been developed and evaluated to selectively ‘capture’ and enrich the forensically relevant human STR regions from highly compromised samples. The strategy involved simultaneous enrichment of the 18 widely used STR loci (which includes the 13 core CODIS loci) with the aid of two sequence-specific biotinylated oligonucleotide probes targeting each locus. The enriched fragments were captured using streptavidin-coated magnetic beads, eluted and subsequently subjected to multiplex amplification using commercially available STR kits followed by fragment analysis. Experiments conducted with simulated forensic samples revealed that prior enrichment of STR regions greatly facilitated the generation of complete and/or improved DNA profiles even in the presence of 5-fold excess of non-human DNA contaminants and 3-4 fold of excess PCR inhibitors (than those normally tolerated by thermostable DNA polymerases incorporated into the commercial STR kits). The pre-PCR enrichment technique provides definite improvement over the existing methodologies for DNA (STR) profiling from challenging forensic samples, and is a promising approach for increasing the DNA testing based human identification from skeletal remains in missing persons and mass disaster victim identification programmes.

1706T
Microbial detection employing Affymetrix Axiom® Genotyping Solution. M. Shapero1, S. Gardner2, M. Mühmann3, K. McLaughlin2, C. Jaing4, L. Bellon2, T. Slezak2,1. Affymetrix, Inc., 3420 Central Expressway, Santa Clara, CA 95051; 2 Lawrence Livermore National Laboratories, Livermore, CA 94551. The human body contains a diverse set of microbes including eukaryotes, archaea, bacteria, and viruses. These microbes play a key role in both human health and disease. As sequences of microbial genomes continue to expand, microarrays are well-positioned to capitalize on this evolving information content. Here we present an overview of Affymetrix® Genotyping Solution with regard to the design and testing of a microbial detection array (MDA) to evaluate differentiation of several model organisms at the species and strain level. A pilot array was designed leveraging organisms where samples and full genome sequences are available for both the primary target species and relevant near-neighbors. Multiple probe designs that iterate on probe length and tolerance to position-specific mismatches were evaluated for six families, which include Bacillaceae, Burkholderiaceae, Enterobacteriaceae, Francisellaceae, Thermotogaceae, and Paramyxoviridae. Using pure and mixed samples of defined origin as the hybridization targets, probes were scored using a composite likelihood maximization algorithm for their ability to detect perfect match sequences as well as identify novel, previously uncharacterized species/stains within a known family. Probe- and system-level performance characteristics will be presented with regard to sensitivity and specificity.

In summary, new Affymetrix® applications in broad microbial detection further extend the platform’s capabilities. The scalable sample throughput enabled by 24-96- , and 384-array layouts coupled with laboratory automation will allow processing of tens of thousands of samples per month with minimal manual intervention and is consistent with the need for cost-effective examination of the microbiome in the context of bioterrorism, environmental monitoring, food safety, and human and animal health.

1707S
Design of a biobanking genotype array optimised for Chinese populations. R.G. Walters1, I.Y. Millwood2, Y. Lu3, H. Lin4, J. Brodsky5, M.A. Ansari6, Y. Zhan4, W.W. Kretzschmar4, N. CaR7, R. Bowden2, J. Flint8, P. Donnelly9, T. Webster9, G. Li,10 J. Schmidt10, L. Li10, R. Peto1, R. Collins10, Z. Chen10, 1) Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), University of Oxford, UK; 2) Affymetrix, Inc., Santa Clara CA, USA; 3) BGI-Shenzhen, Shenzhen, China; 4) Department of Statistics, University of Oxford, UK; 5) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 6) School of Public Health, Peking University Health Science Center, Beijing, China; 7) Chinese Academy of Medical Sciences, Beijing, China. The China Kadoorie Biobank is a prospective study of 515,000 adults recruited from 10 geographically diverse areas of China. Preliminary genotyping of 82,494 subjects (after QC), conducted using a panel of 384 SNPs selected on the basis of previous genome-wide or candidate association studies as well as under clinical endpoints (e.g. stroke, diabetes, coronary artery disease) and intermediate phenotypes (adiposity, blood pressure, cholesterol), revealed substantial population stratification. Common SNPs varied in minor allele frequency (MAF) between different recruitment centres by as much as 10-fold, and 293 out of 335 SNPs displayed MAF differences between regions that were significant after Bonferroni correction. Even though data were available for only 216 unlinked common SNPs, principal component analysis nevertheless revealed a strong North-South structure that scale and pattern of the differences across China strongly suggest that they are due to substantial admixture between northern and southern populations with distinct ancestries.

In part to address the challenges that admixed population structure presents for the establishment of genetic associations with disease and intermediate phenotypes, we have designed a new Affymetrix Axiom® genotyping array optimised specifically for Chinese populations, using data from diverse sources: 1,000 genomes data from CHB/CHS subjects; high coverage sequence data for 156 of the same subjects and exome data from 1,746 subjects mainly from China (and also from BGI), allele frequencies from the CEU, YRI and CHS as well as 1,800 Taiwan Biobank subjects; and allele frequencies from the CONVERGE consortium, derived from sequence data for 9,000 subjects recruited at 59 hospitals across China. Together, these provided improved MAF estimates for known SNPs and indels, identified novel variation not previously reported in Chinese, and enabled efficient selection of array content to maximise genome-wide coverage. The array was designed following a strategy similar to that for the UK Biobank array, and includes: 142,454 variants selected on the basis of linkage disequilibrium, including 1,222 SNPs that lie on or within known or suspected functional regulatory elements, 24,142 previously reported GWAS hits; 81,522 variants not present in 1,000 genomes CHB/CHS data; and a GWAS grid of 479,385 additional variants. With imputation, the array provides coverage of 93% of 1,000 genomes variants with MAF>0.05, and 87% of variants with MAF>0.01.
1708M


Most extracted RNA samples contain ribosomal RNA and messenger RNA with only a low percentage of microRNA (miRNA). In order to measure miRNA more precisely and achieve better quality miRNA sequencing data, it is desirable to remove the larger unwanted ribosomal RNA as well as other species of RNAs. This abstract describes the use of Beckman Coulter’s SPRIselect reagent kit for miRNA enrichment. The Solid Phase Reverse Immobilization (SPRI) magnetic bead method is a nucleic acid purification procedure that does not require centrifugation or filtration steps. This simple, rapid, automation friendly and produces high quality samples of miRNA for use in downstream applications. Enrichment of miRNA from total RNA can be achieved in three binding steps. Step 1: Removal of large ribosomal RNA fragments. Step 2: Removal of intermediate RNA fragments. Step 3: Capturing of miRNAs and small RNAs. The micro-particles with bound miRNA are thoroughly washed with molecular biology-grade ethanol. In the elution step, the purified miRNA is easily recovered from the micro-particles using nuclease free water, which provides maximum flexibility for downstream applications. The poster presents that SPRIselect is a superior miRNA enrichment method. 1. It provides flexibility to separate different fragment sizes. 2. It shows higher miRNA binding efficiency. 3. It enables users to keep all fractions of RNA species during miRNA enrichment. Beckman Coulter, the stylized logo, SPRI, SPRIselect are registered trademarks of Beckman Coulter, Inc. All other trademarks are the property of their respective owners.

1709T

Standardizing High-Throughput Sequencing of Extracellular RNA from Human Plasma. Y.E. Wang1, K.M. Danielsen2, R. Rubio1, S. Das2. 1Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Boston, MA; 2Cardiovascular Institute, Beth Israel Deaconess Medical Center, Boston, MA.

Background. Extracellular vesicles have been shown to regulate cellular signaling by transmitting RNA materials such as mRNA, microRNA and snRNA. This phenomenon implies extracellular RNA (exRNA) may partially reflect cellular content within the human body and may show disease specific variation. As such, exRNA profiles have great potential as disease biomarker. While high-throughput sequencing technology offers a potentially sensitive means to characterize and quantify exRNA, there is a lack of a understanding of the efficacy and reliability of commercial kits for extracellular RNA sequencing. In this study, we aim to optimize protocols for sequencing extracellular RNA from human samples. Methods. Blood plasma samples were obtained from one healthy individual at a single time point. The samples were frozen and thawed to some overlaps. This suggests miRNA species detection is confounded by the sequencing protocol used. Conclusion. In this study, sample treated with proteinase K treatment preceding Exiqon purification; c) addition of carrier glycoblue during proteinase K treatment/Exiqon purification. Sequencing libraries were generated using random hexamer based, SMART technology based, and smRNA specific library preparation kits. The resulting libraries were sequenced on a HiSeq 2000 on a Single Read 50 bp flowcell. The sequence data was mapped against the transcriptome using BWA and analyzed using miRDeep2 algorithm to quantify miRNA. Results. Approximately 10 million raw reads were obtained for the majority of the samples but with variable mapping rates across library preparation protocols. The variation is less across different sample treatment process within a protocol. The NEB protocol resulted in the highest number of miRNA species identified in all three conditions. Proteinase K and glycoblue treatment appear to enhance the number of identifiable miRNA species. We then compared the sets of miRNA species derived from different library preparation and found large differences between them with some overlaps. This suggests miRNA species detection is confounded by the sequencing protocol used. Conclusion. In this study, sample treated with proteinase K treatment prior to vesicle lysis and prepared by smRNA specific protocol resulted in the best outcome. However, with sample preparation and sequencing protocol differing strongly impact on extracellular RNA miRNA quantification and detection, additional studies is necessary to optimize the process.

1710S


RNA sequencing (RNAseq) is gradually replacing micro-array technologies as the method of choice for transcriptional profiling studies. Greater resolution and lower processing costs are two important factors that are driving this trend; however technical limitations exist that have prevented more widespread adoption. RNA seq from whole blood represents a particular challenge in that ~50% of the transcriptome is encoding by globin. This issue of globin overabundance reduces detection of informative transcripts and increases processing and storage costs. Recently, several methods have been developed to address the problem of excessive globin; however no comprehensive comparison has been reported. To this end, we undertook a systematic and rigorous evaluation of three promising approaches; insert dependent adapter cleavage (InDAC), globin removal prior to sequencing, and ultra-deep sequencing and report our findings here. Both InDAC and Globin removal reduced globin reads by ~50% to 2% compared to untreated controls. Off target effects were few and linearity was maintained in the samples that had globin removed prior to sequencing. InDAC resulted in more detectable off-target sequences and slight compression of dynamic range, however all three approaches were largely compared when evaluated in fold change space.

1711M


Recent advances in NGS library construction and a digital microfluidics platform has enabled us to significantly lower the sample input requirements into TruSeq Stranded mRNA preparation workflow. We find that the reduced reaction volume of our digital microfluidics platform has greatly improved the kinetics of each enzymatic step. In addition to improved kinetics, the encapsulated fluidic system facilitates improved sample recovery and yield of each step in the process. As a result of these efficiencies, it is possible to prepare high quality RNA-Seq libraries even when starting with only nanogram levels of total RNA. We will show detailed sequencing results generated from libraries made with inputs of 100, 10, and 1 ng, or less, of total RNA using this new technology. Even at these levels, we observe a large dynamic range of gene expression, and see results that are typically observed with 100 ng of total RNA utilizing the standard bench protocol. Furthermore, the improved sensitivity of the system is demonstrated by detecting equivalent number of genes in 10 ng of RNA processed on the microfluidics platform versus 100 ng of RNA processed with the standard bench protocol. The digital microfluidics platform enables researchers to access sample types with limited RNA recovery such as RNA extracted from small numbers of isolated cells or fine needle aspirates (FNA) resulting in a greater understanding of gene expression at the cellular level. Finally, this microfluidics technology is very easy-to-use and automated.

1712T


In RNA sequencing, duplicates or reads that map to the same position are discarded but in RNA sequencing (RNA-Seq), these reads can represent highly expressed genes. The issue of duplicates in RNA-Seq is even more pronounced in low input or degraded samples. Higher percentages of duplicates in very low input and degraded samples are routinely observed in RNA-Seq using standard bioinformatics tools such as picard but the source of duplicates is commonly misunderstood. Under normal assay conditions and with recommended input levels, three different RNA-seq assays give different apparent numbers of duplicates on the same UHRR and Brain samples: <20% duplicates for targeted RNA seq, <10% duplicates for miRNA-seq and <5% duplicates for Total RNA-seq. These differences are not necessarily due to PCR artifacts but occur as a result of the differences in complexity between the coding regions, the miRNA, and the total RNA of a cell. When we measure true PCR duplicates using a molecular bar coding approach, it becomes clear that there are in fact much lower numbers (much less than 1%) of potential PCR duplicates in standard RNA-Seq prep. However, we find that when reducing DNA input amounts for any of these three assays to 1ng or less, we observe dramatic increases in percentage of duplicates. This value then becomes an important metric for overall efficacy of the experiment.
1713S

Protein coding genes constitute approximately 1% of the human genome but harbor most of the disease-associated variants. By focusing on only the protein-coding regions of the human genome, scientists are now able to more efficiently detect these variants. Targeted re-sequencing enables highly sensitive and comprehensive detection of variants and provides insights into the biology behind a given phenotype. With recent major advances in this technology, combined with a better understanding of biological pathways, NGS is now being considered for use in clinical research.

To efficiently and comprehensively catalogue variants within a sample, we developed a new set of SureSelect target enrichment research panels optimized for analysis of disease-associated regions. 1) Our Latest Clinical Research Exome targets the full human exome for broad research applications. It is specifically designed to achieve excellent coverage of disease-associated content, especially those relevant to constitutional disease research. It provides deep and comprehensive coverage of genomic content derived from highly curated databases including CCDS, RefSeq, GENCODE, Vega, MirBase, UCSC known genes, Human Gene Mutation Database (HGMD®), ClinVar and Online Mendelian Inheritance in Man (OMIM™); 2) The Focused Clinical Research Exome narrows the capture regions to only include ~5000 targets with strong association to disease and annotated within HGMD®, OMIM™ and ClinVar; 3) The Inherited Disease Panel further narrows the focus to cover only a set of ~2700 genes within OMIM™ that are highly relevant targets for Mendelian disease research. These three exome options allow researchers to maximize efficiency depending on their application space and sequencing platform. Here we demonstrate capture efficiency for these exome designs and describe sequencing coverage statistics as well as sensitivity and concordance for various sequencing workflows. Additionally we show low-input SureSelect workflows that enable greatly reduced hybridization and turn-around times. Further, we provide data analysis and visualization using Agilent’s SureCall software.

1714M

Identifying patterns of gene expression and SNP alleles related to a pathway or disease can help identify specific genes of interest. Fixed content panels are a good first approach, but there are frequently additional genes pertinent to the pathway. To this end, we are developing customizable TaqMan® assay panels as flexible research tools for pathway and disease related genes and markers. The panels contain assays for a suggested list of target genes and SNP markers that are customizable to omit and add assays of interest. The researcher can choose between three platforms: TaqMan® OpenArray®, TaqMan® Array Card and TaqMan® Array Plate. The initial panels are gene expression assay sets covering 15 popular research areas in human, mouse and rat: Alzheimer’s disease, Cardiotoxicity, Hepatotoxicity, Nephrotoxicity, Neurotoxicity, Inflammation, Diabetes, Cell cycle, Apoptosis, Hypertension, Angiogenesis, Glycosylation, Parkinson’s disease, Cardiovascular disease and p53 Signaling. Each panel contains up to 92 markers as a core assay set, and some panels contain additional assays in the “Off Your Array” list that may be used to substitute core assays. Additional assays of interest may be added, pre-designed or custom assays. The repertoire of panels will be expanded to include additional gene expression panels and SNP genotyping panels. Flexible Content TaqMan® Pathway Panels enable a better selection of assays for the immediate research interest.

1715T
Design and implementation of a transplantation-targeted whole genome genotyping array. A. Shaked1, Y.R. Li2, W. Asselbergs4, S. Balsa1, K. Karczewski3, K. Birdwell6, W. Oetting7, P. de Bakker1, R. Jacobson3, D. MacArthur2, Y. Lu11, G. Lord11, A.K. Israni1, B.J. Keating2, 1) Dept of Transplantation, University of Pennsylvania, Philadelphia, PA, USA; 2) The Center for Applied Genomics, The Children’s Hospital of Philadelphia, Philadelphia, PA, USA; 3) Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 4) Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht, Utrecht, Netherlands; 5) Yale Medical School; 6) Harvard Medical School and Broad Institute, Massachusetts General Hospital, MA, USA; 7) School of Medicine, Vanderbilt University; 8) College of Pharmacy, University of Minnesota; 9) Epidemiology and Bioinformatics at UMC Utrecht; 10) Dept of Nephrology, Transplantation and Internal Medicine, Guy’s and St Thomas’ Hospital and King’s College London, UK; 11) Affymetrix, Santa Clara, California.

More than 559,000 solid organ transplantations (kidney, liver, heart and lung) have been performed in the United States since 1988, with nearly 30,000 occurring annually in the US. Despite significant improvements in transplant outcomes, chronic organ rejection affects up to 60 percent of patients representing a major cause of morbidity and mortality following allogeneic transplantation. Whole genome sequencing projects have identified millions of common and rare genetic polymorphisms across human populations, which are thought to contribute to genetically-associated risk factors for transplant and immunosuppression-related adverse event, as well as serve as a source of genetic disparity between donors and recipients that can lead to both acute and chronic organ rejection. Array-based genotyping technologies permit significant flexibility in choosing the scope and density of SNPs to examine for disease or phenotype-specific genome-wide association studies. We report and discuss the design, implementation, and preliminary results from thousands of samples typed on a state-of-art high-content, high-throughput commercial platform composed of nearly 800,000 markers with coverage of a number of genomic features of interest in particular to the transplant community: 1) Exonic and Loss of Function (LoF) variants (274K) from >40K human exons. 2) Imputation panel variants (459K) with all available EUR ancestry data from the 1000 genome project and ~90K African ancestry panels SNPs to boost coverage and improve fine-mapping. We added significant content (55K) for UTRs, pharmacogenomic markers (7.5K) from PharmaADME.org, genome-wide and the candidates used in the literature related to immunosuppression therapy response, and copy number variant probes (16.4K) from the Wellcome Trust and UKbiobank (n=2.4K) and analysis of 68K GWAS samples from CHOP. 5) Expression QTL markers (17.4 K SNPs from NCBI/NIH GTEX eQTL database; 6) Dense coverage across the MHC and KIR regions; and 7) Cardiovascular & metabolic disease related content for analysis of new onset of diabetes after tx (NODAT) & Cardiac allograft vasculopathy (CAV) by collision of SNPs from >600 PubMed manuscripts. This genotyping array offers rich information for the transplant community utilizing state-of-the-art array design technology and data from large-scale resequencing projects and publicly available functional and gene expression datasets.

1716S
Performance of seven mutation pathogenicity prediction methods in the classification of missense variants of the CYP1B1 gene. G. Chavarría-Soley, Biology, University of Costa Rica, San José, Costa Rica, San José, Costa Rica.

Nonsynonymous single nucleotide polymorphisms (SNPs) in the coding regions of genes can lead to aminoacid changes and potentially affect protein function and, therefore, susceptibility to disease. Several computational methods have been developed for the classification of SNPs according to their predicted effect on protein function and resulting pathogenic potential. In this study, we evaluated the performance of seven commonly used pathogenicity prediction methods available on the Internet (SIFT, nsSNPAnalyzer, Panther, pMut, PolyPhen, PhD-SNP, and SNAP). In order to test them, nonsynonymous SNPs in the CYP1B1 gene-which codes for the cytochrome P450 1B1 enzyme-were selected. A total of 129 missense variants in CYP1B1 were identified in the literature, from which 87 could be classified as pathogenic or neutral according to criteria such as segregation with disease phenotype, effect on protein function, among others. The algorithms showed significant variation in the assignment of the variants to three categories (non-neutral, neutral, no prediction), with a low 37% prediction rate for Panther. Pairwise concordance between methods in the classification of the same variants pathogenic or neutral varied between 37% and 94%. The accuracy in the prediction of the pathogenicity of the variants was higher than 68% with all methods except pMut (47%). The highest false positive and false negative rates were found for SIFT and pMut, respectively. Taking into account the rate of pathogenicity (positive) and false negative rates, the method with the overall best performance in the present study was nsSNPAnalyzer, closely followed by SIFT, Polyphen and SNAP.
1717M
A Comprehensive IT System to Support GTEx Biospecimen Collection Operations, P. Guan1, C. Shive2, L. Qi3, D. Tabor4, P. Hanharan5, S. Wu1, K. Umm7, V. Santanharam7, P. Kigomy9, J. McLean9, J. Vaugh9, H.M. Moore1, on behalf of GTEx consortium.
1) National Cancer Institute, Bethesda, MD; 2) Leidos Biomedical Research Inc., Frederick, MD.

The NIH Common Fund's Genotype-Tissue Expression (GTEx) project aims to study gene expression and regulation across multiple human tissues (30+ tissue types) from approximately 900 healthy normal donors. It is expected to provide valuable insights into gene regulation and its tissue specificity, to identify correlation between genetic variations and variations in gene expression, and to find a variety of disease genes for early diagnosis. Results are expected to help understand inherited susceptibility to diseases. To meet the challenge of GTEx requirements for collecting and tracking high quality biospecimen samples, a custom-built software system, the Comprehensive Data Resource (CDR) was developed to support sample collection, workflow, clinical data entry, case management, and review and curation of study data. CDR is built with a combination of technologies from Grails, Oracle, Groovy, jQuery, and Apache Solr. CDR provides secure data access based on pre-defined user roles and privileges. Personally Identifiable Information (PII) and Protected Health Information (PHI) are restricted to a limited data set (LDS) and to authorized users through dynamic content redaction. Intuitive graphic user interfaces for Biospecimen Source Sites streamline data entry work flow using intuitive SOPs for sample collection and processing. Automated data check and business rule validation confirm data integrity and SOP adherence simultaneously. Web service API's allow a Pathology Resource Center to access digital imaging data housed remotely at a Comprehensive Biospecimen Resource (CBR). API's connect to the CBR's LIMS system for real-time sharing of the identified data. CDR is connected to the Laboratory Data Analysis and Coordinating Center (LDACC) at the Broad Institute through a private API before the final release into dbGaP. CDR's analytics and reporting module supports data analysis and aggregation, report generation and real-time operational data snapshots. CDR is a distributed web-based bioinformatics system that has supported GTEx biospecimen operations from an earlier pilot phase to the current scale-up stage. CDR manages and maintains multi-dimensional data models around each donor case (average 500+ data elements per case). An efficient case management tool capable of connecting to various remote informatics systems. CDR could be adapted by the broader research community to standardize and streamline biobanking operations.

1717T
Clinical phenotype-based gene prioritization using semantic similarity and the Human Phenotype Ontology. A. Masino1, E. Dechene2, M. Dullik3, A. Wilkens4, N. Spinnere5, J. Krautz6, J. Pennington4, P. White6, 1) Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA; 2) Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA; 4) Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA; 5) Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 6) Department of Pediatrics, Cincinnati Children's Hospital and Medical Center, Cincinnati, OH; 7) Department of Biomedical Informatics, University of Cincinnati, College of Medicine, Cincinnati, OH.

Exome sequencing is a promising method for diagnosing patients with a complex phenotype. However, variant interpretation relative to patient phenotype can be challenging in some scenarios, particularly clinical assessment of rare complex phenotypes. Each patient’s sequence reveals many possibly pathogenic variants that must be individually assessed. Variant impact to gene function alone is insufficient to determine clinical relevance. Clear association between a gene harboring a damaging variant and patient phenotype is also required for accurate clinical diagnosis. We expect that automated gene prioritization can assist this process by ranking a given set of genes relative to patient phenotype. The model orders genes by the semantic similarity between Human Phenotype Ontology terms associated with a gene and those describing the patient. Model validation was performed for 33 Mendelian diseases with 100 simulated baseline patients per disease. Each patient was assigned phenotypic features with probability based on published feature penetrance. To measure robustness, we modeled likely clinical conditions by adding noise, i.e. terms unrelated to the disease, and imprecise phenotypic features. Our results suggest that clinical diagnostic precision of automated gene prioritization can substantially improve the ability to prioritize genes, and by extension variants, for clinical care. Our preliminary conclusions are confirmed by our expectation that automated gene prioritization based on phenotypic descriptors may increase accuracy and decrease effort for predicting genomic variant significance.

1719S
The Orphanet Rare Diseases Ontology (ORDO) : a reference tool integrating clinical and genetic data. A.M. Rath1, A. Oiry1, C. Gonthier2, L. Chanas1, H. Parkinson2, J. Pennington4, P. Vasant2, W. Kaechele3, M. Hanauer1, B. Ubertó1, A. Olry1, C. Gonthier1, S. Aymé1, 1) US14, INSEMR, Paris, Paris, France; 2) EMBL-EBI, Hinxton, UK.

The growing complexity of genetic knowledge, the scattered phenotype data, the large number of distinct rare diseases, as well as the multiplicity of medical terminologies in use impose the need for a reference tool in which all these data are integrated in a normalized fashion in order to render them accessible for health information systems and for research. Since 1997, Orphamnet maintains the Orphanet Rare Diseases Ontology (ORDO) based on the available medical literature and on expert advice. This data is manually curated, comprised of a nosology (classification of RD), relationships (genes-diseases, epimediological data, orphan drugs) and cross-references with other terminologies (MeSH, SNOMED CT, UMLS), databases (OMIM) or classifications (ICD10) in use. Genes are cross-referenced with other scientific databases (HGNC, OMIM, UniProt, Genatlas, Reactome, ensembl, IUPHAR). These data are already freely available for download via the OrphaData platform (www.orphadata.org). However, it is necessary to make data available for researchers in a machine-readable format, ready to be integrated to any technical environment (obo and owl formats). In order to achieve this, Orphanet set up a collaboration with the European Bioinformatics Institute (EBI, Hinxton, UK) to order to produce the Ontology of Rare Diseases, which provides a robust and consistent modeling of data and their semantic relationships, as well as interoperability standards with other scientific resources in use in both research and in public health. The Orphanet Rare Diseases Ontology is available on BioPortal (http://biocenter.bioontology.org/ontologies/ORDO) and OrphaData and is updated monthly.

1720M
Comprehensive Transcriptome Analysis Reveals that Nonsense-Mediated mRNA Decay Is Not Globally Suppressed in Lung Adenocarcinomas. L. Hu1, P. Zhang. State Key Lab of MedicalGenomics, SIBS, Shanghai, China.

The nonsense-mediated mRNA decay (NMD) pathway ensures the rapid degradation of mRNAs containing premature translation termination codons (PTC). Prior research has shown that genes containing NMD targets tend to have high expression levels and low expression noise. Thus, we believe that cancer efficiently uses the NMD mechanism.

1721T
Combined use of mutant loxP sites, JT15 and JTZ17, is a useful approach for sophisticated genome engineering. K.C. Chen1, P.L. Chen2,3,4, 1) Graduate Institute of Medical Genomics and Proteomics, National Taiwan University College of Medicine, Taipei, Taiwan; 2) Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan; 3) Research Center for Developmental Biology and Regenerative Medicine, National Taiwan University, Taipei, Taiwan; 4) Research Center for Developmental Biology and Regenerative Medicine, National Taiwan University College of Medicine, Taipei, Taiwan.

Cre-mediated site-specific integrative recombination system is a useful tool for genome engineering. Integrative recombination with a pair of mutant loxP sites is a developed and successfully strategy to increase the diversity and possible combinations. Using several mutant loxP sites combination system for manipulation of multiple genes on knock-out or knock-in experiments can improve complex genetic studies. In 2003, Pedra et al. (Gene- sis, 2003, 36:162-167) reported a couple of mutant loxP combination, JT15 and JTZ17, which yielded the most efficient integration in E. coli. However, their recombinacn efficiency with wild-type loxP or other mutant loxP sites remains unclear. In order to address this issue, we constructed a series of neoeycin selection cassette flanked with the wild-type or mutant loxP sites to test the frequency of site-specific integration. Our results showed that the frequency of targeted recombination between JT15 and JTZ17 was 74%. We also observed that the recombination between JT15 and the JT15:JTZ17 mutant loxP sequence was very low. There was no recombination between JT15 and the JT15:JTZ17 mutant loxP sequence. Most importantly, JT15, JTZ17 and JT15:JTZ17 could not recombine with wild-type loxP sites. These results confirm our conclusion and indicate that the JT15 and JTZ17 is a useful approach in sophisticated genome engineering strategies.
Identification of rare causal variants in sequence-based studies: methods and applications to VPS13B, a gene involved in Cohen syndrome and autism. L. Ionita-Laza1, M. Caparini1, S. De Rubeis2, K. McCallum1, J. Buxbaum2. 1 Biostatistics, Columbia University, New York, NY; 2 Memorial Sloan-Kettering Cancer Center; 3 Icahn School of Medicine at Mount Sinai.

We propose a general framework to perform gene/region based analysis of sequence data by regressing a functional response on one or multiple scalar predictors. Next generation sequencing technologies make it possible to uncover genetic information from millions of variants. Since the observed sequenced variants are very close in their genetic positions, we can consider them to be realizations of random continuous functions. Therefore, instead of analyzing multiple individual genetic variants per subject, we can estimate the underlying continuous function and treat it as a functional response in a regression model. Smoothing splines are used to fit these functional responses by maximizing the penalized likelihood. Covariates can also be incorporated in the analysis to control for confounding, including qualitative and quantitative predictors. By utilizing a connection between penalized spline regression and linear mixed models, we are able to fit our model using standard linear mixed models statistical packages. To illustrate our approach, we conduct simulation studies and apply our proposed methodology to sequencing data from the Dallas Heart Study.


We developed likelihood-based methods for analyzing data from case-control studies, if data from case participants is sequenced to a greater depth than data from controls, the difference in haplotype frequency can be calculated under the assumption that each individual’s estimated haplotype frequency is its true haplotype frequency. We performed extensive simulations (at several alpha levels) to assess the type I error rates under various null hypothesis and the power of the test under various alternative hypothesis.

A GENERALIZED SIMILARITY U TEST WITH APPLICATION TO MULTIPLE-TRAIT SEQUENCING STUDIES UNDER TRAIT-DEPENDENT SAMPLING. Z. Z. Tang1, R. Y. Lin2. 1) Vanderbilt University, Nashville, TN; 2) University of North Carolina, Chapel Hill, NC.

We propose a general framework for analyzing binary traits under the trait-dependent sampling design. Trait-dependent sampling design has been adopted in many sequencing studies to enrich causal signal and increase power of the association tests. In the National Heart, Lung, and Blood Institute Exome Sequencing Project (NHBLI ESP), subjects with the highest or lowest values of body mass index, low-density lipoprotein, or blood pressure were selected for whole-exome sequencing. However, such study design produces challenges for genetic association analysis. We provide a valid and efficient maximum likelihood approach for analyzing binary traits under the trait-dependent sampling. Under this framework, we produce the commonly used gene-based association tests including burden tests, variable-threshold tests and variance-component tests. We compare our methods with the naive methods, namely the standard logistic regression methods. We demonstrate through extensive simulations that our methods preserve type I errors, whereas naive methods can yield severely inflated type I errors. For a particular trait of interest, our approach properly combines the association results from all studies with measurements of that trait. This meta-analysis is substantially more powerful than the analysis of a single study. By contrast, meta-analysis of naive methods can be less powerful than the analysis of a single study. The usefulness of the proposed methods is further illustrated with data from NHBLI ESP.

Functional regression for genetic association studies. O. Vsevolozhskaya1, D. Zaykin2, L. Qing3. 1) Epidemiology & Biostatistics, Michigan State University, East Lansing, MI; 2) Biostatistics Branch, NIEHS, Research Triangle Park, NC.

We propose a general framework to perform gene/region based analysis of sequence data by regressing a functional response on one or multiple scalar predictors. Next generation sequencing technologies make it possible to uncover genetic information from millions of variants. Since the observed sequenced variants are very close in their genetic positions, we can consider them to be realizations of random continuous functions. Therefore, instead of analyzing multiple individual genetic variants per subject, we can estimate the underlying continuous function and treat it as a functional response in a regression model. Smoothing splines are used to fit these functional responses by maximizing the penalized likelihood. Covariates can also be incorporated in the analysis to control for confounding, including qualitative and quantitative predictors. By utilizing a connection between penalized spline regression and linear mixed models, we are able to fit our model using standard linear mixed models statistical packages. To illustrate our approach, we conduct simulation studies and apply our proposed methodology to sequencing data from the Dallas Heart Study.

1724T Bootstrap Tests of Association For NextGen Sequence Data That Allow for Systematic Differences in Read Depth between Cases and Controls. G.A. Satten†, Y. Hur†, H.R. Johnston‡, P. Liao†, Y. Jiang‡, A.S. Aller‡. 1) Centers for Disease Control and Prevention, Atlanta, GA; 2) Dept of Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 3) Dept of Biostatistics and Bioinformatics, Duke University, Durham, NC; 4) Duke Clinical Research Institute, Duke University, Durham, NC.

We develop likelihood-based methods for analyzing data from case-control and trio studies that directly uses data on reads without first making intermediate genotype calls. When the location of polymorphic loci is known, we show these likelihood approaches have appropriate size and good power compared with methods that use called genotypes. When the locations of polymorphic loci are not known in advance, we develop screening methods to screen out loci that are expected to be genotyped on read data alone. We use a bootstrap approach to estimate which of the loci that screen in are truly polymorphic. Using these estimates, we then construct bootstrap tests for association that properly account for screening and provide power gains over other single study designs. Under this framework, we produce the commonly used gene-based association tests including burden tests, variable-threshold tests and variance-component tests. We compare our methods with the naive methods, namely the standard logistic regression methods. We demonstrate through extensive simulations that our methods preserve type I errors, whereas naive methods can yield severely inflated type I errors. For a particular trait of interest, our approach properly combines the association results from all studies with measurements of that trait. This meta-analysis is substantially more powerful than the analysis of a single study. By contrast, meta-analysis of naive methods can be less powerful than the analysis of a single study. The usefulness of the proposed methods is further illustrated with data from NHBLI ESP.

1726M Binary Trait Analysis in Sequencing Studies under Trait-Dependent Sampling. Z.Z. Tang†, D.Y. Lin‡. 1) Vanderbilt University, Nashville, TN; 2) University of North Carolina, Chapel Hill, NC.

We propose a general framework to perform gene/region based analysis of sequence data by regressing a functional response on one or multiple scalar predictors. Next generation sequencing technologies make it possible to uncover genetic information from millions of variants. Since the observed sequenced variants are very close in their genetic positions, we can consider them to be realizations of random continuous functions. Therefore, instead of analyzing multiple individual genetic variants per subject, we can estimate the underlying continuous function and treat it as a functional response in a regression model. Smoothing splines are used to fit these functional responses by maximizing the penalized likelihood. Covariates can also be incorporated in the analysis to control for confounding, including qualitative and quantitative predictors. By utilizing a connection between penalized spline regression and linear mixed models, we are able to fit our model using standard linear mixed models statistical packages. To illustrate our approach, we conduct simulation studies and apply our proposed methodology to sequencing data from the Dallas Heart Study.

A Generalized Similarity U Test with Application to Multiple-Trait Sequencing Association Study. C. Wei1,2, O. Lu1. 1) Department of Biostatistics and Epidemiology, University of North Texas, Fort Worth, TX; 2) Department of Biostatistics and Epidemiology and Biostatistics, Michigan State University, East Lansing, MI.

Sequencing-based studies are emerging as a major tool for genetic association studies of complex diseases. It also poses great challenge to the traditional statistical methods (e.g., single variant analysis) due to the high-dimensionality of the data and low frequency of the genetic variants. Joint test has been shown to be more suitable for sequencing studies, by jointly testing multiplex variants to increase the power and reduce the dimensionality. Meanwhile, there are growing needs for statistical methods that are distribution free and can handle multiple phenotypes. In this paper, we proposed a generalized similarity U test, referred to as GSU. GSU rst summarizes the genetic information and multiple traits into the genetic similarity and trait similarity, and then combines the two similarities in the framework of weighted U statistic. We derived the asymptotic distribution of GSU under null hypothesis as to recursively calculate the significance level. We also studied the asymptotic behavior of GSU under alternative hypotheses using simulations. We then applied the calculation for study design. To evaluate the performance of GSU, we conducted extensive simulation studies and compare it with the existing methods. Through simulation, we found GSU had advantage over existing methods in terms of power comparison and robustness to trait distribution. Moreover, GSU is computationally more efficient than the existing methods. Finally, we applied GSU to the multiple traits analysis of Dallas Heart Study and identified joint association of 4 genes with 5 metabolic related traits.
1728S

High-throughput sequencing technologies have facilitated the identification of large numbers of single-nucleotide variations (SNVs), many of which have already been proven to be associated with diseases or other complex traits. Several large sequencing studies, such as, the 1000 Genomes Project, the UK10K project, or the NHLBI-Exome Sequencing Project, have consistently reported a large proportion of private SNVs, that is, variants that are unique to a family or even a single individual. The role that private SNVs play in diseases and other traits is currently poorly understood — which is largely due to the fact that it is statistically very challenging to consider private SNVs in association testing. While it is generally impossible to use single-marker tests for private SNVs, burden tests are potentially able to deal with private SNVs, but only if the number of private SNVs occurring in a region is correlated with the trait under consideration. Moreover, burden tests have a disadvantage if deleterious and protective SNVs occur together in the same region. Non-burden tests like the popular SNP-set (Sequence) Kernel Association Test (SKAT) are typically utilizing correlations between SNVs — a strategy that is not applicable to private SNVs either, since singular events are generally uncorrelated. We propose the Position-Dependent Kernel Association Test (PODKAT), which is designed for detecting associations of very rare and private SNVs with the trait under consideration even if the burden scores are not correlated with the trait. PODKAT assumes that, the closer two SNVs are on the genome, the more likely they have similar effects on the trait under consideration. This assumption is fulfilled as long as deleterious, neutral, and protective variants are grouped sufficiently within the same region. This contribution focuses on the use of PODKAT for large whole-genome studies. On the one hand, we will discuss issues related to data handling, computational complexity, and statistical significance. On the other hand, we will present results obtained for UK10K whole-genome SNP cohorts that underline the potential of considering private and very rare SNVs in genome-wide association studies.

1729M
A non-threshold region-specific method for detecting rare variants. D.P. Chen1, A.R. Hsieh2, C.S.J. Fann1. 1) Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; 2) Graduate Institute of Biostatistics, China Medical University, Taichung, Taiwan.

Rare variants have a proven role in some complex diseases. Many statistical methods proposed for the detection of rare variants associated with diseases have some limitations, such as the threshold of rare variants, and the direction of effects. Accordingly, we developed a region-specific method that do not use the threshold for defining rare variants and take the directions of effects into account. Our method also considers the linkage disequilibrium (LD) within the region, and can handle common and rare variants simultaneously. Our region-specific method used the concept of weighting variants according to their minor allele frequencies and odds ratios (OR) to combine effects of common and rare variants on disease occurrence into a single score, and provided a test statistic in assessing the significance of the score. To evaluate the performance of our method, we simulated extensively under different effect sizes according to Basu and Pan (2011). We found that the power of our method increased as the effect sizes increased. The type I error of our method was controlled well in spite of the simultaneous variations. Moreover, we compared our proposed method to several currently available methods, including kernel-based adaptive cluster (KBAC) and Sequence Kernel Association Test (SKAT). We found our method can generate comparable or better power in simulations. Results from our method showed a 15% increase in power comparing with SKAT (61% vs 47%) under small OR and lower LD, and 36% increase in power comparing with KBAC (98% vs 62%) while variants have different directions. However, our method performs well in 2-direction setting, but moderate in independent-variant scenario. We conclude that our proposed method can be used as a complementary tool with others to assist the dissection of the etiology of complex diseases.

1730T
Evaluating the calibration and power of three gene-based association tests for the X chromosome. C. Ma, M. Boehnke, S. Lee. Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI.

While genome-wide association studies (GWAS) have identified thousands of trait-associated genetic variants, the proportion of findings on the X chromosome lags behind those on the autosomal chromosomes. Existing X chromosome analysis methods focus on single marker association analysis. For analysis of rare variants (minor allele frequency < 0.5%), gene-based tests where multiple markers are analyzed jointly as a unit can be more powerful than single marker tests. To date, there are no gene-based tests designed to analyze the X chromosome. Using simulated case-control and quantitative trait (QT) studies, we evaluate the calibration and power of three gene-based tests for the X chromosome: burden, Sequence Kernel Association Test (SKAT), and optimal unified SKAT (SKAT-O). Specifically, we evaluate the impact of different ratios of males and females in cases and controls, and different coding of males alleles with X-inactivation (coding minor alleles as X=2) and without (X=1).

For case-control studies, all three tests are well-calibrated or slightly anti-conservative for all scenarios evaluated. As previously shown, power of the three tests depends on the underlying genetic architecture of the genomic region analyzed; burden is most powerful for multiple causal variants with the same direction of effect, SKAT is most powerful for causal variants with opposite directions of effect, and SKAT-O is generally powerful. For variants simulated assuming X-inactivation, coding male minor alleles as X=2 is slightly more powerful; for variants simulated assuming no X-inactivation, coding male minor alleles as X=1 is slightly more powerful. However, the power loss for misspecifying the generally unknown model is small. Different ratios of males and females in cases and controls have little effect on power. For QT studies, burden and SKAT are well-calibrated, while SKAT-O can be slightly anti-conservative across all scenarios. Power comparisons between tests for QTs are very similar to those for binary traits. We demonstrate that these three gene-based tests are well-calibrated and powerful for both binary and quantitative trait data, and can be directly applied to analyze rare variants on the X chromosome.
1731S
Exploiting correlation of genetic effects in rare variant association studies. M.A. Rivas1, M. Pirinen2, L. Moutsianas3, C. Spencer4, K. Banasik1, D. van Heel5, K. Hunt4, P. Soininen6, A.J. Kangas7, M. Alaperto4,5,6, M.J. Daly8, F. Karpe9, P. Donnelly1,10, M.I. McCarthy1,2,11,12.
1) Wellcome Trust Centre for Human Genetics Research, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK; 2) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland; 3) Barts and The London School of Medicine and Dentistry, Queen Mary University of London; 4) Computational Medicine, Institute of Health Sciences, University of Oulu, FI; 5) NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio, FI; 6) Computational Medicine, School of Social and Community Medicine and the Medical Research Council, Integrative Epidemiology Unit, University of Bristol, UK; 7) Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, US; 8) Broad Institute of MIT and Harvard, Cambridge, MA, USA; 9) Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, UK; 10) Department of Statistics, University of Oxford, Oxford, UK.

High-dimensional data present exciting opportunities in enhancing our understanding of the relationship between genetic variation and multiple diseases or health related quantitative traits. A first step towards utilizing high-dimensional phenotypes in genetic studies is to understand how their genetic components are related. A second step is to apply this knowledge in association studies.

We present novel statistical approaches that leverage estimates of the correlation of genetic effects for rare variant association studies that enable association testing (via extension of the C-alpha framework) and effects estimation in cross-disorder and cross-phenotype study designs. The intuition is simple: 1) for some groups of genetic variants, such as protein truncating variants, annotation suggests high correlation of effects; we want to exploit this by building models which incorporate information on the correlation between the effects of these variants on a single phenotype; 2) for any causal genetic variant that is pleiotropic, incorporation of the correlation of genetic effects for rare variant association studies that may improve power to detect association.

Simulations suggest that our approach improves power to detect association compared to standard univariate approaches. First, we consider the study of rare variants and multiple continuous traits. We apply the tests to coding rare variant data and plasma NMR metabolite levels in 4,522 healthy individuals from the population based Oxford Biobank study and identify novel genetic signals that were not identified with univariate aggregate methods. As an example, we identify multivariate association (p=1×10^-6) of protein-altering variants in PAH (phenylalanine hydroxylase) to the amino-acid trait profile (univariate min(p) = 4.5×10^-4). Second, we consider the study of rare variants, multiple diseases, and shared controls. We apply the tests to coding sequence data from 25 GWAS risk genes in 41,911 UK residents of white European origin, comprising 24,892 subjects with six autoimmune disease phenotypes and 17,019 controls and identify new association of protein truncating variants at TNFAIP3. Our framework is computationally efficient, making the analysis of large rare variant association study designs practical.

1732M
Integrated statistical model of genetic variation reveals new insights into the genetics of autism. X. He1, K. Roeder1,2, B. Devlin1,2, M.J. Daly1,4, J.D. Buxbaum3,4,5, Autism Sequencing Consortium. 1) Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh, PA; 2) Department of Statistics, Carnegie Mellon University, Pittsburgh, PA; 3) Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA; 4) Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA; 5) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA; 6) Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY; 7) Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY; 8) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 9) Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY.

Analysis of de novo mutations has been effective in identifying susceptibility genes of developmental disorders such as autism. Recurrent de novo loss-of-function (LoF) mutations on the same gene generally provides strong evidence for the gene’s involvement in risk. Recently we developed a statistical model, named TADA, that extends this approach by combining information on de novo LoF mutations, de novo missense mutations and inherited variants for the same gene (He et al, PLoS Genetics, 2013). In the current work, we improve the TADA model by refining the weights mapping onto different types of genetic variation. Under the new model, de novo LoF mutations contribute most to a gene’s evidence, followed by de novo missense mutations, transmitted LoF mutations and transmitted missense mutations. We achieve this weighting strategy by using different prior distributions for the relative risks of different types of genetic variation. Using parameters derived from whole exome sequencing (WES) data of autism, we show through simulations that TADA is almost twice more powerful than using only de novo LoF mutations. Using TADA we analyzed a large WES dataset of autism, consisting of more than 2,000 parent-child trios and 7,000 case-control samples. The data supported an estimate of more than 1,000 genes involved in risk for autism. We next used TADA to predict 33 high-confidence autism risk genes (FDR < 0.1) and 107 genes at a relaxed threshold (FDR < 0.3). We developed two strategies for estimating the relative risks of different types of mutations. Under the new model, de novo mutations contribute most to a gene’s evidence, followed by de novo missense mutations, transmitted LoF mutations and transmitted missense mutations. Other enrichment analyses identify important biological processes involved in autism (De Rubeis for the Autism Sequencing Consortium). Finally, a newly developed statistical method named DAWN was used to analyze the gene networks underlying autism. DAWN identified several subnetworks that are enriched with genes of high TADA scores and gene pairs of high co-expression in brain transcriptome data.
1733T

Meta-Analysis of rare variants association studies with multiple correlated traits. X. Wang, X. Zhao, C. Huang. P.O. Box 413, Zilber School of Public Health, University of Wisconsin-Milwaukee. Milwaukee, WI.

For many diseases such as asthma, attention deficit hyperactivity disorder, or hypertension, genetic association studies are often conducted to test multiple correlated traits which are routinely measured and thought to be more proximal to the biological etiology of the clinical disorder. Testing multiple correlated traits can identify quantitative trait loci (QTL) shared between correlated traits in addition to trait-specific QTL. Meta-analysis as a cost-efficient powerful tool for combining distinct genome-wide association studies (GWAS) has been widely used to detect common variants for single trait in GWAS. However, existing meta-analysis methods may not be optimal or even not be applicable for detecting rare variants in sequencing data with multiple correlated traits due to allelic heterogeneity, extreme rarity of individual variants, and correlation among traits. We propose a general statistical framework for meta-analysis of gene-based rare variants association studies with multiple correlated traits. Due to low minor frequencies of rare variants, it is often unstable or not feasible to estimate regression coefficients of rare variants in a multi-marker regression model. Thus, we developed a score test to test the effect of an optimally weighted combination of variants in a gene in each study for a single trait, and perform meta-analysis for each single trait with Lee’s meta-analysis method and obtain the p value for each gene for all the studies. Then, we use the Brown’s method to combine p values for the multiple correlated traits and obtain the p value of the meta-analysis with the multiple correlated traits. There are three key advantages of the proposed method: 1) through aggregating score statistics, it circumvents the estimation of the regression coefficients of rare variants; 2) it is computationally efficient even for whole-genome analysis since p values can be calculated analytically; 3) it can decrease the false-positive rate by incorporating the extra information provided by the cross trait covariance. To investigate the performance of the proposed method, we conducted extensive simulation studies for type I error rates and power comparison. The multiple correlated traits meta-analysis method maintained appropriate type I error rates in all the simulation settings. Further simulation studies are ongoing to compare power of the proposed method with the single trait meta-analysis method proposed by Lee et al.

1735M

A fast and powerful test of independent assortment with implications for the analysis of ‘big data’. V. Hager*, W.C.L. Stewart1,2,3. 1) Nationwide Children’s Hospital, 700 Children’s Drive Columbus, Ohio 43205; 2) The Ohio State University Department of Pediatrics, 700 Children’s Drive Columbus, Ohio 43205; 3) The Ohio State University Department of Statistics, 1958 Neil Avenue, 404 Cockins Hall Columbus, OH 43210.

Quantitative methods for the analysis of genetic sequences on related individuals continue to place a high premium on computational and statistical efficiency. As such, the most commonly used methods tend to incorporate as much information as possible as quickly as possible. However, much less attention is paid to the perhaps more important issue of statistical significance. For example, although most modern workstations can easily handle the computations associated with a large-scale, genome-wide linkage scan, the p-value for such a scan is either accurate but time consuming, or fast but inaccurate. To address this issue, we developed a test of independent assortment (TIA) that permits the accurate and rapid computation of p-values for large-scale genome-wide linkage scans. Furthermore, we show that our proposed test (1) has a limiting distribution under the null hypothesis of no linkage; (2) yields an approximate 300x speed-up over existing methods like MERLIN; and, (3) has the same statistical power as the widely used Kong & Cox Iod2. Overall, these attractive features should facilitate the analysis of ‘big data’ genetics and genomics through the meta-analysis of linkage, association, and gene expression data.

1736T

Generation of sequence-based data for pedigrees segregating Mendelian or Complex traits. B. Li, G.T. Wang, S.M. Leal. Center for Statistical Genetics, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

There is great interest in analyzing next generation sequence data that have been generated for pedigrees. However, unlike rare variant association methods for population-based data there are limited rare variant methods which have been developed to analyze pedigree data. One of the limitations in developing methods to analyze pedigree data is the availability of software to generate realistic pedigree sequence data with the variants generated either conditionally or unconditionally on qualitative or quantitative phenotypes. Rare variant pedigree simulated data is necessary to evaluate type I and II errors and also to compare the power of different methods. Therefore, we have developed RarePedSim (rare-variant pedigree-based simulator), a program to simulate pedigree-based gene-level genotype and phenotype data for complex and Mendelian trait rare variant studies given any user-specified pedigree structures. For complex traits rare variant association studies, RarePedSim can generate genotypes of founders based on site-specific variant information and non-founders according to principles of segregation. Site-specific variant information including allele frequencies, positions and functionalities can be derived from simulated sequences under realistic demographic and genetic models or obtained from real sequence data. RarePedSim can generate phenotypes based on a wide range of genetic etiology for both qualitative and quantitative complex traits, including logistic model of odds ratios, population attributable risk model and linear mean-shift model. For Mendelian disease rare variant analysis RarePedSim generates genotypes for individuals within the family given their phenotypes and mode of inheritance, e.g. dominant and recessive. Allelic heterogeneity between and within families and locus heterogeneity can also be simulated. Currently, RarePedSim is the only available program that can generate both genotypes and phenotypes for gene-based rare variants regardless of pedigree structure. The data generated by RarePedSim are in standard Linkage file format, ready to be used for a variety of purposes including evaluating type I error and power, for association methods including mixed model tests and parametric and non-parametric linkage analysis.
Poster: Statistical Genetics and Genetic Epidemiology

1737S
GenLib: an R package for the analysis of genealogical data. M.-H. Roy-Gagnon1, S. H. Gauvin2, J.-F. Lefèvre3, C. Moreau4, E.-M. Laviole5, D. Labuda6, H. Vézina7. 1) Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, ON, Canada; 2) CHU Sainte-Justine Research Center, Montreal, QC, Canada; 3) Department of Social and Preventive Medicine, Université de Montréal, Montreal, QC, Canada; 4) BALSAC Project, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada; 5) Department of Pediatrics, Université de Montréal, Montreal, QC, Canada.

Founder populations play an important role in the study of genetic diseases. However, their advantages often include access to detailed genealogical records. These genealogical data provide unique information for researchers in evolutionary and population genetics, demography and genetic epidemiology. However, analyzing large genealogical datasets require specialized methods and software. The GenLib software was originally developed to study the large genealogies of the French Canadian population of Quebec, Canada. These genealogies are accessible through the BALSAC database, which contains over 3 million records covering the whole province of Quebec over four centuries. Using this resource, extended pedigrees of up to 17 generations can be constructed from a sample of present-day individuals. We have implemented GenLib as a package in the R environment for statistical computing and graphics, thus allowing optimal flexibility for users. GenLib is designed to manage genealogical data, for example extracting a part of a genealogy or selecting specific individuals. Functions to describe genealogies using relevant summary measures, such as genealogical completeness and generational depth, are also available. In addition, GenLib can compute measures of relatedness (kinship and inbreeding) and the genetic contribution of founders. Finally, functions for generating pedigrees are also available in GenLib. We illustrate the use of GenLib with a sample of 140 individuals from regional populations of Quebec previously described in Roy-Gagnon et al., 2011. Ascending genealogies were reconstructed for these individuals using BALSAC, yielding a large pedigree of 41,523 individuals. With GenLib, we provide a more detailed description of these genealogical data in terms of completeness, genetic contribution of founders, relatedness and inbreeding, further illustrating the regional differences reported in the data. We also produced a genealogy for each founder based on the whole genealogy to estimate the probability of sharing, identical by descent, chromosomal segments of varying lengths introduced at varying frequencies by different number of founders. In conclusion, the R package GenLib provides the necessary flexibility to analyze genealogical data, allowing a more efficient and easier integration of different types of data and analytical methods and making it ideal for future developments.

1739T
Assessing mitochondrial DNA variation and copy number in lymphocytes of 2,077 Sardinians using tailored sequencing analysis tools. J. Ding1, C. Sidorenko1, T. J. Butler2, M. K. Wing3, O. Maireles1, Y. Qian1, F. Busonero1, R. Nagaraja1, F. Cucca2, G. R. Abecasis2, D. Schlesinger3. 1) Laboratory of Genetics, National Institute on Aging, Baltimore, MD; 2) Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 3) Institute of Biomedical Sciences, Consiglio Nazionale delle Ricerche, Monserrato, Cagliari, Italy.

To quantify mitochondrial DNA (mtDNA) variation, genotype calling and analytic programs developed for nuclear DNA must be modified, because mtDNA is linear and contains two separate chromosomes. We have developed an algorithm that jointly calling genotypes across multiple individuals, based on linkage disequilibrium (LD) patterns and familial transmission of mtDNA. As expected, mothers and their children share essentially all homoplasies but a lesser proportion of heteroplasies. Overall, heteroplasmy increases with age, but on average only ~1 heteroplasmy reaches the 4% level between ages 20 and 80. The total extent of accumulation of variants, however, remains hard to assess because heteroplasmy levels could be similar to or even lower than sequencing error rates; but resolving power can be increased in several ways, including deep sequencing, pooling, and using genealogical data.

We have also made a sequence-based estimate of mtDNA copy number based on the observed ratio of sequence coverage between mtDNA and autosomal DNA. The average in the cohort studied was 112 copies ±25 per cell, a number in agreement with estimates from qPCR assays. The computational methods thus facilitate comprehensive mtDNA analysis from whole-genome sequencing data. The algorithms can also be applied to quantitate sequence variation in other instances of high DNA copy number, such as cancer.

1740S
Improving effective sample size using extrapolated log p-values. B. Engelhardt1, X. Guo2, S. Mukherjee2, J. Tung2. 1) Biostatistics & Bioinformatics, Duke University, Durham, NC; 2) Department of Statistical Science, Duke University, Durham, NC; 3) Evolutionary Anthropology, Duke University, Durham, NC.

While the per-sample cost of genomic data collection has plummeted in recent years, population-based genomic studies, counterintuitively, remain extremely expensive. This expense arises because attempts to maximize power have relied almost exclusively on increases in sample size. An alternative, more cost-effective approach is to improve the statistical power associated with a given sample through development of more sensitive statistics. This approach is particularly attractive in the context of association mapping, or identifying genetic variants that are associated with complex traits. In particular, a large number of genetic variants may have an effect on a given complex trait, and effect sizes across these associated variants are thought to have a double exponential distribution (i.e., a few large magnitude effects and many effect sizes close to zero). However, association mapping, both in genome-wide association studies (GWAS) and quantitative trait locus (QTL) studies, continues to struggle with their ability to detect associated variants with low effect size because of the limited number of samples in each study and the large number of associations tested. We present a statistical approach to extrapolate association significance (extrapolated log p-values, or ELPS) in a given study using subsampling methods. We validate this approach by quantifying gains in statistical power using simulated trait values and real genotype information from 4,646 individuals in the Wellcome Trust Case Control Consortium (WTCCC) data set. Our simulations show substantial gains in statistical power to detect true associations for data sets with limited sample size using ELPS. We identify this increase in effective sample size for association mapping in three diverse data sets: case-control mapping of Crohn’s disease, and quantitative trait mapping of gene expression levels (eQTL mapping) in HapMap phase 3 control offspring and nuclear DNA samples. We demonstrate that this increase in power is due to the increased sensitivity of ELPS; the theoretical justification for this approach is elegant in the context of classical statistics, ELPS enables the detection of small effects in studies where the sample size has prevented these associations from reaching statistical significance. ELPS improves the power to detect associations for a given sample through development of more sensitive statistics and phenotypes for which existing methods are ineffective due to insufficient power. Broadly, ELPS maximizes the power to draw statistical signal from expensive samples.
1741M

Alternative peak calling methods on Hi-C data accommodating the whole spectrum of dispersion. Z. Xu, Y. Li. Department of Genetics and Department of Bioinformatics, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Advancement in chromatic conformation capture and next generation sequencing technologies are enabling genomewide investigation of dynamic chromatic interaction. To analyze Hi-C data, Ay et al (2014) developed a peak calling method (Fit-Hi-C) that assigns statistical confidence estimates to mid-range intra-chromosomal contacts based on a varying-coefficient Binomial model with the success rate as a smoothing function of the distance between the midpoints of a fragment pair. Ay et al (2014)’s model assumes that Hi-C data are equal-dispersed given the distance between the midpoints of a fragment pair. We found that the dispersion (given the distance) for real Hi-C data within different domains, either before or after outlier/peak filtering, can be (1) under-dispersed, (2) equi-dispersed, (3) over-dispersed or (4) dispersed differently for different distances. To address the dispersion issue in real-hi-C data, we consider alternative peak calling methods: (1) varying-coefficient Poisson model, (2) varying-coefficient Negative Binomial model, (3) varying-coefficient COM-Poisson model and (4) nonparametric standardization method. We benchmark these methods together with Fit-Hi-C using the number of interactions between known enhancers and transcription start sites given the same number of total peaks called. We found that each method has its own advantages. For example, we found that varying-coefficient Negative Binomial model performs better than varying-coefficient Poisson model for the domains showing over-dispersion based on Jin et al (2013)’s Hi-C data. This motivates us to consider a combined two-step peak-calling method, where the level of dispersion is quantified in the first step to guide the method of choice in the second step.

1742T

Haplotype based fine mapping algorithms using meta-analysis summary results. J. Zheng1,2, S. Rodriguez1,2, J. White1, C. Giambartolomé2, D. Zabaneh1, I.N. Day3, UCLEB Consortium. 1) Bristol Genetic Epidemiology Laboratory, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom; 2) University College London Genetics Institute, Department of Genetics, Environment and Evolution, London, United Kingdom; 3) Department of Primary Care & Population Health, University College London, Royal Free Campus, London, United Kingdom; 4) Centre for Clinical Pharmacology, University College London, London, United Kingdom; 5) MRC Integrative Epidemiology Unit, School of Social and Community Medicine, Bristol, United Kingdom.

Meta-analysis of multiple genome-wide association studies (GWAS) is becoming a popular method to increase statistical power and reduce false positive findings. As a follow-up procedure to meta-analysis, detecting secondary association signals at loci are necessary to determining variants with true causality. However, individual-level data is often not pooled, making fine mapping of associated loci via conditional analysis time consuming and impractical. Here, we present two statistical methods, Sequential Sentinel SNP Regional Association Plot (SSS-RAP) (Zheng et al., 2013) and E-method HAPloType-based Regional Association analysis Program (HAPRAP) (Zheng et al., under review). Both methods enable fine mapping using group-level summary statistics and reference haplotype information. SSS-RAP detects SNPs with independent effects conditional on the top associated signal. HAPRAP extends multiple regression and conditional analysis to meta-analysis levels. We demonstrate by a performance comparison that SSS-RAP is as accurate as conditional analysis and ten model selection methods in individual-level. We compare HAPRAP with existing methods using simulated data, BWHHS cohort and 1000 Genomes. The results of HAPRAP are highly consistent with multiple regression results using individual-level data. Moreover, HAPRAP consistently outperforms GCTA joint and conditional analysis (Yang et al., 2012) across different linkage disequilibrium (LD) correlations (r²) and reference panel sample sizes. The Web-based interfaces of SSS-RAP and HAPRAP are available online: http://apps.biocompute.org.uk/sssrapt/sssrapt.cgi and http://apps.biocompute.org.uk/haprap.

1743S

PedBLLIMP: A Linear Predictor based Approach to Impute Genotypes in Pedigrees. W. Chen, D. Schaid. Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN., USA.

Recently, Wen and Stephens proposed a linear predictor, called BLIMP, that uses conditional multivariate normal moments to impute genotypes with accuracy similar to current state-of-the-art methods. The key insight is that it regularized the estimated covariance matrix based on a model from population genetics. We extended multivariate moments to impute genotypes in pedigrees. Our proposed method, PedBLLIMP, utilizes both the linkage disequilibrium (LD) information estimated from external panel data and the pedigree structure or identity by descent (IBD) information. By stacking the genotype matrix of single nucleotide polymorphisms (SNPs) for all individuals in a pedigree into a vector G, we showed that the mean of G is essentially the same as the mean of genotypes for any individual which can be estimated from external panel data. The covariance of G is a Kronecker product of two matrices. The first is the covariance matrix between SNPs within the same individual and the second is the correlation matrix between individuals in the pedigree. The first covariance matrix can be estimated from external panel data similarly as in BLIMP. The second matrix can be estimated in two ways. For global correlation, it is twice the kinship matrix which can be estimated from the pedigree structure. When genotype is available for pedigrees, methods, local correlation of a specific locus or region between two individuals can be estimated based on inferred IBD in each region. Of auto from MORGAN package is used to sample inheritance vectors and IBD is estimated accordingly. With the estimated mean and covariance of G, the best linear predictor is applied to estimate the expected dosage of untyped markers. The proposed method was evaluated on a pedigree design where some individuals were genotyped with dense markers and the rest with sparse markers. We found that incorporating the pedigree/IBD information can improve imputation accuracy compared to BLIMP. Because rare variants usually have low LD with other SNPs, incorporating pedigree/IBD information largely improved imputation accuracy for rare variants. We also compared PedBLLIMP with IMPUTE2 and GIGI. Results show that when sparse markers are in a certain density range, our method can outperform other methods. In our experimental data, we found that the PRECAL and GIGI model may be useful for imputing small to medium sized pedigrees where some individuals are densely genotyped and other individuals are sparsely genotyped.

1744M

Detecting maternal-offspring gene interactions using linear mixed effect models: The Quantitative-MFG Test. M.M. Creek1, E. Sobel2, J.S. Sinsheimer1,2,3, I.N. Day3, UCLEB Consortium. 1) Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA; 2) Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA; 3) Department of Bioinformatics, University of California, Los Angeles, CA.

The genetic etiology of many complex diseases remains to be discovered. Maternal and offspring gene interactions or maternal-fetal genotype (MFG) interactions may provide a new avenue for investigating the effects of maternal genes on the offspring’s phenotype and have been shown to be involved in complex diseases such as schizophrenia and autism. Existing methods to investigate the effects of MFG incompatibility on quantitative traits using retrospective likelihoods are limited to case-parent trios. Presently, there are no methods to investigate the effects of MFG incompatibility on quantitative traits that are suitable for both small and large families. We propose the Quantitative-MFG Test, a prospective approach using a linear mixed effects model, which takes into account familial correlations by partitioning the variance. This approach can handle varying pedigree sizes, general and specific scenarios of MFG incompatibility, and the inclusion of covariates. We validate the model under two familiar maternal-offspring gene interaction scenarios, RHD and NIMA incompatibility. We estimate type I error rates and power by simulation and investigate the effect of risk allele frequency on power. The type I error rates for both scenarios are approximately 0.05 (SE=0.005). When the expected variance explained by a specified effect size is 1%, the power is over 0.8 for RHD incompatibility and is over 0.7 for NIMA incompatibility when testing for any genetic effect, a NIMA effect, or an offspring effect with sample sizes of 500 nuclear families. Simulations demonstrate that parameters and variances are well estimated and that this method works with arbitrary sets of families. Additionally, we find that treating the nuclear families within an extended pedigree as independent results in inflated false positive rates, supporting the need of a test for association of MFG incompatibility and quantitative traits that utilizes all available family data.
1745T On the null distribution of Bayes factors. Y. Guan. Pediatrics and Molecular & Human Genetics, Baylor College of Medicine, Houston, TX.

The p-value is the blood-love of many theoreticians and practitioners alike. A Bayesian procedure often faces an inconvenient demand of producing p-values. When a data analysis involves multiple testing, such as in a genome-wide association study, the significant threshold of the p-value is exceedingly small, owing to Bonferroni correction, and it requires a prohibitively large number of permutations to produce a p-value. Thus, characterizing the null distribution of Bayes factors, which enables one to compute p-values without the need of permutation, is of practical importance. On the other hand, it is beneficial to understand how priors affect Bayes factors under the null, which may in turn shed light on prior specification. We investigated how priors affect the null distribution of Bayes factors for Bayesian linear regression (of SpS covariates) under the Normal-Inverse-Gamma prior specification. For a finite sample, the null distribution of $2 \log(\text{Bayes factor})|S$ under the independent prior is approximately a mean-shifted linear combination of $\chi^2_n$, whose distribution function can be evaluated using an existing polynomial algorithm, and hence the corresponding p-value of a Bayes factor can be computed efficiently, without the need of permutation. The independent prior favors null for a large sample size; Zellner's SpS-prior does not--asymptotically, $2 \log(\text{Bayes factor})|S \sim g/(1+g) \chi^2_p + p \log(1+g) + O(1)$ under the SpS-prior.

1746S Mixed model with correction for case-control ascertainment increases power in multiple sclerosis association study. T. Hayeck1, N. Zaitlen2, P. Loh3, B. Vihjalmsdottir1, S. Samuela4, A. Gusev5, J. Yang2, G. Chen6, M. Goddard7, P. Visscher7, N. Patterson8, A. Price1. 1) Harvard School of Public Health, Boston, MA; 2) University of California, San Francisco, CA; 3) University of Queensland, Brisbane, Australia; 4) University of Melbourne, Melbourne, Australia; 5) Broad Institute, Cambridge, MA.

We introduce a Liability Threshold Mixed Linear Model (LTMLM) association statistic for ascertained case-control studies that increases power vs. existing mixed model methods, with well-controlled false-positive rate. Recent work has shown that existing mixed model methods suffer a loss in power under case-control ascertainment (Yang et al. 2014 Nat Genet), but no solution has been proposed. Here, we solve this problem using a chi-square score statistic computed from posterior mean liabilities (PML) under the liability threshold model. Each individual's PML is conditional not only on that individual's case-control status, but also on every individual's case-control status and on the genetic relationship matrix obtained from the data. For example, disease cases with higher genetic relationships to other disease cases are assigned a larger PML than disease cases with lower genetic relationships to other disease cases. The PML are estimated using a multivariate Gibbs sampler, with the liability-scale phenotypic covariance matrix based on the genetic relationship matrix (GRM) and a heritability parameter estimated via Haseman-Elston regression on case-control phenotypes followed by transformation to liability scale (Lee et al. 2011 AJHG), Leave One Chromosome Out (LOC0) analysis was used to calculate the GRMs, avoiding the decrease in power that arises from including candidate SNPs in the GRM (Yang et al. 2014 Nat Genet). The Gibbs sampler does not iterate over SNPs, thus overall running time is comparable to existing mixed model methods. In simulations of unrelated individuals, the LTMLM statistic was correctly calibrated and achieved higher power than existing mixed model methods in all scenarios tested, with the magnitude of the improvement depending on sample size and severity of case-control ascertainment.

1747M GARFIELD - GWAS Analysis of Regulatory or Functional Information Enrichment with LD correction. V. Iotchkova1,2, M. Gehs1, G. Ritchie1,2, E. Birmeyer1, L. Dunham1, N. Soranzo3, UK10K Consortium Cohorts Group. 1) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1HH, United Kingdom; 2) The EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SD, United Kingdom.

Genome-Wide Association Studies (GWAS) have been increasingly fruitful in discovering genotype-phenotype associations. The mechanisms underlying these associations, however, are still largely unknown as only a small fraction of SNPs directly affects gene expression. The correlation of functional consequences of non-coding variants has been greatly enhanced by large-scale efforts to identify regulatory genomic regions (e.g ENCODE). However, robust methods are still lacking to systematically evaluate the functional relevance of variants to gene expression in diseases or quantitative traits. Here we propose a novel approach that leverages GWAS findings with regulatory or functional annotations to find features relevant to a phenotype of interest. We perform greedy pruning of GWAS SNPs (LD r2>0.1) and then annotate them based on functional information overlap. We then systematically compared enrichment levels between associations from whole genome sequencing (WGS from the UK10K project) and sparser genotyping datasets (e.g. HapMap). When applied to 4 main lipid sub-phenotypes, the method found similarly large enrichment values as well as larger number of significant enrichments in the WGS data, highlighting a greater sensitivity of WGS data. Finally, we developed new software to facilitate the application of our method by the research community. Overall, it is expected that a more accurate classification of enrichment patterns might lead to biological insights and help prioritise variants for follow-up studies.

1748T Comparison of machine-learning methodologies to prioritize genetic variants based on functional data. S.A. Gagliano1,2,3, R. Ravil1,2,3, M.R. Barnes4, M.E. Weale2, J. Knight1,2,3, Schizophrenia Working Group of the Psychiatric Genomics Consortium. 1) Centre for Addiction and Mental Health, Toronto, Canada; 2) Institute of Medical Science, University of Toronto, 610 University Ave., Toronto, Ontario, Canada; 3) Department of Medical & Molecular Genetics, King’s College London, London, London, United Kingdom; 4) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1HS, United Kingdom; 5) Department of Medical & Molecular Genetics, University of Cambridge, Cambridge, United Kingdom.

Incorporating a variety of functional characteristics as predictors, such as DNase I hypersensitive sites, transcription factor binding sites, and histone modifications, we developed a methodology using elastic net, a regularized regression, to prioritize genetic variants that are truly associated with complex disease (Gagliano et al. PLoS ONE 2014). We compared the results from our elastic net method to two other published methods: Ritchie et al. (Nat Methods 2014) and Kircher et al. (Nat Genetics 2014). These models use different methodologies (modified random forest and a support vector machine, respectively) and different predictors with regard to the quantity, type and coding of the functional characteristics. We tested which methodology performed the best for prioritizing sub-genome-wide-significant variants (5E-6p<1E-6) by investigating which one assigned higher scores to the sub-genome-wide-significant variants from the first round of a schizophrenia genome-wide association study (GWAS) that come up as associated in the second round, and lower scores to those variants that do not come up as associated. We obtained prediction values of the sub-genome-wide-significant variants in the Psychiatric Genomics Consortium (PGC1, Schizophrenia GWAS Consor- tium, 2011) for the three methodologies. A quantile-quantile plot was created using the p-values from the larger second round of the schizophrenia GWAS (PGC1) (corrected) stratified into the PGS from the sub-genome-wide-significant variants from the first round belonging to the top and bottom quartiles with regard to their prediction value. Elastic net outperforms the other two methodologies when it comes to identifying truly associated genetic variants in the PGS. It is worth mentioning that we are exploring the use of the different algorithms, for instance, elastic net compared to random forest, while holding constant the functional characteristics used in order to discriminate performance differences as being attributed to the methodology and/or the functional predictors.
1749S
A mixed model methodology to correct technical artifacts and enable meta-analysis of sequence based association studies. C. Murphy1, P. Syrris2, P. Lambiase2, D. Speed4, V. Plagnol1, UCL-exomes Consortium.
1) Genetics Institute, University College London, London, United Kingdom; 2) Institute of Cardiovascular Science, University College London, United Kingdom; 3) Heart Hospital 16-18 Westmoreland Street London W1G 8PH.

High throughput DNA sequencing technologies, either whole exome (WES) or whole genome (WGS) sequencing are revolutionizing the diagnosis and novel gene discovery for rare disorders. As the field transitions from the early discovery for Mendelian and near Mendelian diseases to more complex and oligo-genic diseases, there is substantial benefit in being able to combine data across studies, performing the type of meta-analysis for cases and controls that have proven to be so successful for genome-wide association studies (GWAS). However, WGS and WES are substantially more affected by sequencing errors and technical artifacts than genome-wide genotyping arrays. As a consequence, meta-analysis of sequence based association studies are often dominated by spurious associations, which result in technical limitations. Here, we show that it is possible to take advantage of the type of mixed models developed initially to control for population structure in GWAS studies, and apply these ideas to control for technical artifacts. We developed computational improvements to optimize the estimation of the mixed model parameters to enable genome-wide testing in a realistic time (from days to minutes). These techniques are being implemented as part of the LDAK package (www.ldak.org). Using a dataset of 2,800 WES (UCL-exomes), which aggregates a diverse set of studies, we demonstrate that substantial reduction in the association statistic inflation can be achieved by applying these novel analytical techniques, both for single variant and gene based tests, while preserving the sensitivity of the test. We focus on several cardiovascular diseases (ARVC and sudden cardiac death) to illustrate the ability of these novel methods to produce more interpretable results. Our methodology provides a general framework to facilitate current and future meta-analysis of sequence based association studies.

1750M
1) Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA; 2) Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA 23298, USA; 3) Human Genetics Center and Division of Biostatistics, The University of Texas School of Public Health, Houston, TX 77030, USA; 4) Department of Biostatistics, Virginia Commonwealth University School of Medicine, Richmond, VA, 23298, USA; 5) Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA; 6) Department of Biomedical Engineering, Tulane University, New Orleans, LA 70118, USA; 7) Department of Epidemiology and Biostatistics, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA.

High-throughput DNA sequencing technologies have been generating datasets of massive rare and common genetic variants. In a typical sequence dataset, a short genomic region may contain sequences of many single-nucleotide polymorphisms (SNPs). For example, in the sequence data of Mexican pedigrees from GAW18, 68% of all the 11,933 non-overlapped 100-kb regions on odd chromosomes contain more than 500 SNPs. It is a formidable challenge to identify a high-dimensional SNP set with sparse functional variants, where the set size m is close to or larger than the sample size n. Existent prominent association methods appear suitable to identify dense low dimension SNP sets, i.e., m/n<10%, and >20% of set wise SNPs are functional loci. Such methods are ineffective and insufficient to identify sparse high-dimensional SNP sets due to perceived limitations. First, they may be severely confounded by inherent population structure and cryptic relatedness when applied to high-dimensional sets of rare and/or common variants. Second, they suffer unnecessarily large degrees of freedom due to massive neutral variants. Third, they model mean genetic effects only and ignore inherent heterogeneity. In this report, we propose an efficient method for identifying sparse high-dimensional sets from sequence data with arbitrary population structure and cryptic relatedness. In this method, we first perform a heterogenetic feature selection to search for SNPs that best explain phenotypic data; we then link the selected variants with the confounders by a hetero heterogenetic mixed-effect model and perform likelihood ratio test for the genetic effects. In both feature screening and formal association test, we model population structure as fixed effect and model cryptic relatedness as a random effect in respective heterogenetic mixed-effect models. Under extensive simulations, the proposed method properly controlled type I error rates and appeared more powerful than existing prominent sequence association methods (e.g., SKAT, famSKAT and fastLMM). The proposed method is computationally efficient, integrates heterogeneity and diverse genetic effects, and simultaneously adjusts for arbitrary covariates and cryptic relatedness. Application to the deep sequence data of Mexican American pedigrees from GAW18 demonstrated practical utility of the proposed method.

1751T
Estimation of prognostic marker genes by public microarray data in patients with ovarian cancer epithelial. S. Yang, J. Kim. Biomedical Sciences Dept, Seoul National University College of Medicine 103 Daehakno, Jongnogu, Seoul 110-799, KOREA.

Lymphatic invasion is regarded as a predictor of aggressiveness of ovarian cancer. However, lymphatic invasion do not accurately represent 5 years survival. To diagnosis and treatment of ovarian cancer, we analyzed the differentially expressed genes between 5 years survival group and 5 years death group of lymphatic invasion in serous ovarian epithelial cancer with DNA microarray. Data from 63 ovarian cancer patients with lymphatic invasion and 35 ovarian cancer patients without lymphatic invasion from TCGA data were analyzed. Among these 98 patients, 16 patients were survived 5 years or more. DEGs identified by Bioconductor R package. Functional analysis of genes that were analyzed with DAVID web tool. We found 20 DEGs (P value<0.001) from 5 years survival 64 patients and 5 years death 8 patients without lymphatic invasion. Also we found 55 DEGs (P value<0.01) from 5 years survival 55 patients and 5 years death 8 patients with lymphatic invasion. Pathway analysis showed that the lymphatic invasion related DEGs were related with starch and sucrose metabolism pathways while lymphatic invasion related DEGs were related with renal cell carcinoma pathways. These DEGs for ovarian cancer results in high grade serous ovarian cancer patient survival, particularly in 5 years death and survival patients with and without lymphatic invasion. These findings may have implications for the diagnosis and treatment of the ovarian cancer.
A novel meta-analysis approach for genome-wide association studies with sex-specific effects. E. Kang¹, J. Joo², N. Furlotte¹, E. Kostem¹, B. Han¹,²,₄,₅, E. Eskin¹,², J. Joo², N. Furlotte¹, E. Kostem¹, B. Han¹,²,₄,₅, E. Eskin¹,². 1) Department of Computer Science, UCLA, Los Angeles, CA; 2) Interdepartmental Program in Bioinformatics, UCLA, Los Angeles, CA; 3) Division of Genetics, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA; 4) Division of Rheumatology, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA; 5) Partners Center for Personalized Genetic Medicine, Harvard Medical School, Boston, MA; 6) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA; 7) Department of Human Genetics, UCLA, Los Angeles, CA.

Over the past several years, a significant amount of loci implicated in genome wide association studies show differences in effect sizes for males and females. The prevalence of sex-specific effects both motivates strategies for discovery of sex-specific effects as well as raises questions of how to analyze association studies consisting of both males and females. The traditional approach to discover sex-specific effects is to analyze each sex separately and the traditional approach to analyze association studies consisting of both males and females is to include sex as a covariate in the statistical model when performing association analysis. Unfortunately, these approaches may lead to a loss in power when sex-specific effects are present at the loci being tested. In this paper, we present a novel meta-analytic approach for the analysis of genome-wide association studies consisting of both males and females. In our approach, males and females are analyzed separately and the results are combined using a random effects meta-analysis approach allowing for differences in effect sizes between sexes. We show that by analyzing males and females separately, our method reduces the overall variance in each study leading to an increase in statistical power. Through simulations and application of our method to the Northern Finland Birth Cohort data, we show that our method has increased power over the traditional approaches for discovering associated loci with and without sex-specific effects while controlling for false positives.

Mapping of novel regulatory influences on genes encoding subunits of the L-type calcium channel, using digital measurement of allelic skew. N. Kamitaki¹,², S. Ghosh¹,², K. Egan¹,²,³, S.A. McCarroll¹,². 1) Department of Genetics, Harvard Medical School, Boston, MA; 2) Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA; 3) Harvard Stem Cell Institute, Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA; 4) Howard Hughes Medical Institute, Cambridge, MA.

Most of the SNP haplotypes implicated by genome-wide association studies (GWAS) do not include protein-coding variants, suggesting that the functional variant(s) has a regulatory function. One of the more proximal phenotypes to genotype is gene expression, but precise quantification of gene-regulatory effects in brain has both biological and technical challenges - in particular, genetic background and sample-to-sample variability from varying mixtures of different cell types in a complex tissue. Allele-specific expression can be used to control for environmental and trans-acting genetic influences, but techniques for measuring allelic-specific expression often exhibit amplification bias or are insufficiently sensitive to detect modest, quantitative effects on gene expression. We developed a set of new molecular and statistical methods for making precise digital measurements of allelic-regional specific-expression and using such measurements to map gene-expression effects. These methods include: 1) A novel molecular method for measuring allelic-specific expression with high precision, by digital counting of alleles in RNA. 2) A novel molecular method for inferring the chromosomal phase of candidate regulatory variants with transcribed sequence variants; this method is fast and scalable, and we have validated it at genomic distances up to 200 kb. 3) Companion statistical methods for combining the results of independent assays with genome-variation data to map regulatory effects on gene expression. We applied these methods to post mortem brain samples from 5 brain regions in 105 individuals, focusing initially on genes encoding subunits of the L-type calcium channel, now implicated in genetic studies of schizophrenia and bipolar disorder. Using this approach, we were able to map novel regulatory effects that have not been identified in expression QTL studies of the same genes. For example, we find multiple LD-independent effects on the expression of CACNA1C; both effects replicated in an independent series of post mortem brain samples. Significant regulatory effects were brain-region-specific while another showed a regulatory effect consistent in magnitude across brain regions. We have introduced several variants into embryonic stem cells using the Crispr/Cas9 system and are currently differentiating these into neurons to evaluate effects on expression.
A Candidate Pathway Approach Identifies Many Genetic Variants Associated with Colon Cancer Risk and Survival. N. Shariat-Fard1, M.L. Slattery2, Q. Liu1, C. Friedman1, B.J. Caan1, J.D. Potter5,6,8, Y. Yasui1. 1) School of Public Health, University of Alberta, Edmonton, Alberta, Canada; 2) Department of Internal Medicine, University of Utah Health Sciences Center, Salt Lake City, Utah, USA; 3) Division of Research, Kaiser Permanente Medical Care Program, Oakland, California, USA; 4) Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; 5) Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA; 6) Centre for Public Health Research, Massey University, Wellington, New Zealand.

Genetic association studies have traditionally focused on effects of individual genetic markers, usually single nucleotide polymorphisms (SNPs), on disease of phenotype. A novel genome-wide approach more suited to detect causal relationships and underlying molecular mechanisms, however, considers multiple genes implicated in the same pathway and further interacts with environmental exposures. This analysis aimed at assessing the interaction effects of the angiogenesis gene-pathway with three main lifestyle exposures (dietary protein intake, smoking, and alcohol consumption) on colon cancer risk and survival. Logic regression, which identifies Boolean combinations of predictor variables, was used to assess gene-pathway effects, followed by assessments of gene-environment interactions (GEIs) using logistic regression for risk and Cox proportional-hazards models for survival. We analyzed data of 1,541 colon cancer cases and 1,934 controls from the Diet, Activity and Lifestyle as a Risk Factor for Colon Cancer Study conducted at three centers in the United States. The study selected a total of 257 SNPs in 34 genes in the angiogenesis candidate pathway based on standard pathway maps and experimental evidence. We found five statistically significant GEIs for colon cancer risk and three GEIs for colon cancer survival with increasing levels of all three environmental exposures. For risk: FLT1 rs2387632 OR rs9513070 and high animal protein intake (ORINT = 1.69, 95% CI (1.03, 2.76)); KDR rs6838752 and heavy alcohol consumption (ORINT = 1.33, 95% CI (1.02, 1.72)); and BMP4 rs17563 and ≥ 20-pack-years smoking (HRRINT ≥ 1.45, 95% CI (1.03, 2.02)). For survival: TNF rs1800630 and high animal protein intake (interaction hazard ratio HRRINT = 1.80, 95% CI (1.13, 2.86)); BMP1 rs12357490 OR rs2160978 and ≥ 20-pack-years smoking (HRRINT ≤ 1.78, 95% CI (1.03, 3.09)); and BMP2 rs12477602 and heavy alcohol consumption (HRRINT = 1.78, 95% CI (1.66, 39.03)). Our study shows that GEI effects on colon cancer risk and survival can be identified by adopting a comprehensive candidate pathway approach that emphasizes the biologic hypothesis in the selection of the pathways genes and environmental exposures and carries that logic through to the analysis.
1759M
Analysis of pleiotropy at a fine genomic scale. D.J. Balding, D. Speed.
Inst Genetics, Univ College London, London, United Kingdom.

Bivariate SNP-based heritability analysis is being widely applied to measure the overlap in genetic architecture between pairs of diseases. It is especially popular for psychiatric diseases, as better understanding of the concordance or discordance between different traits can inform nosology and improve prediction of disease progression. Given SNP-level data for a pair of diseases, the approach implemented in the software GCTA returns an estimate of , the genome-wide average correlation between the SNP effect sizes for each trait: >0 indicates that risk alleles for the first disease tend to also increase risk for the second, and vice-versa. A recent study considered five psychiatric disorders (schizophrenia, bipolar disorder, major depressive disorder, autism spectrum disorders (ASD) and attention deficit hyperactivity disorder (ADHD)), and found significant concordance (p>0) for five of the ten pairings; the strongest concordance was reported for schizophrenia and bipolar disorder, in line with previous knowledge of the overlap between these two conditions. A major limitation of this approach is that represents concordance averaged across the entire genome, which means that similarity in one genomic region can be cancelled out by dissimilarity in another. Furthermore, a significantly non-zero estimate of provides no indication of which genomic regions are concordant or discordant, knowledge of which would assist in identifying shared genes and pathways. We propose a tool for measuring concordance at a fine genomic scale, at the level of individual genes or short genomic regions. Using computational optimization, our algorithm is extremely fast, able to analyse genome-wide data for many thousands of individuals in a matter of minutes. As proof of principle, our method identified the overlap between Rheumatoid Arthritis and Type 1 Diabetes, with a substantial and statistically significant correlation coefficient. In the study of psychiatric disease, the major subtypes of Inflammatory Bowel Disease, Crohn's Disease and Ulcerative Colitis, it identifies a gene where risk alleles for the former are protective for the latter. We re-examine the study of five pairs of diseases in children, and find significant genetically-local concordance even where none exists genome-wide.

1760T
Development of efficient polygenic risk scores for personalized medicine: methodological concepts and examples. K. Fischer1, K. Läll1, R. Mägi1, A. Morris1–2, A. Metspalu1. 1) Estonian Genome Center, University of Tartu, Tartu, Estonia; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, UK.

Polygenic risk scores have great potential in predictive personalized medicine, even in cases where each of the individual variants identified by a genome-wide association study has a relatively weak effect on the disease of interest. There are, however, several methodological issues that need to be considered while constructing such scores. First, the genetic variants to be used in such scores, as well as their relative weights, are often chosen on the basis of a discovery study. This may lead to overestimation of the real effect of the score, due to selection bias towards variants that have an effect that is overestimated by chance (“winners curse”). Another potential source of biases is present when the outcome is binary and the discovery study (or a replication study) uses simple logistic regression to estimate the score, the genome-wide average correlation between the SNP allele counts to a polygenic score. Due to non-collapsibility of the odds ratios, such estimates are biased due to non-collapsibility of the odds ratios, such estimates are biased. Crohn's Disease and Ulcerative Colitis, it identifies a gene where risk alleles for the former are protective for the latter. We re-examine the study of five pairs of diseases in children, and find significant genetically-local concordance even where none exists genome-wide.

1761S
A Proper and Efficient Approach to Integrative Analysis of Sequencing and GWAS Data for Rare Variant Associations. Y.J. Hu1, Y. Li1,2, P. Auer1, D. Y. Lin2. 1) Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 2) Biostatistics, University of North Carolina, Chapel Hill, NC; 3) Genetics, University of North Carolina, Chapel Hill, NC; 4) Joseph J. Zilber School of Public Health, University of Wisconsin, Milwaukee, WI.

In the large cohorts typically used for genome-wide association studies (GWAS), it is not economically feasible to sequence all cohort members. A cost-effective strategy is to sequence subjects with extreme values of quantitative traits or those with specific diseases. By imputing the sequencing data from the GWAS data, we can obtain increased power to detect for sequencing, one can dramatically increase the number of subjects with information on rare variants. However, treating the imputed rare variants as observed quantities in downstream association analysis may inflate the type I error, especially when the sequenced subjects are not a random subset of the whole cohort. Although this problem can be alleviated by restricting the analysis to the variants that are accurately imputed, a large number of rare variants will be excluded with this strategy. In this article, we show how to properly account for the uncertainties in the imputation of rare variants. We consider all commonly used gene-level association tests, including the burden test, variable threshold (VT) test, and sequence-kernel association test (SKAT), all of which are based on the score statistic for assessing the effects of individual variants on the trait of interest. We show that the score statistic based on the observed genotypes for sequenced subjects and the imputed genotypes for non-sequenced subjects is unbiased. We construct a robust variance estimator that reflects the true variability of the score statistic regardless of the sampling scheme and imputation quality, such that the corresponding association tests always have the correct type I error. Through extensive simulation studies, we demonstrate that the proposed tests are substantially more powerful than the use of accurately imputed variants only or the use of sequencing data alone. An application to the data from the Women's Health Initiative (WHI) is provided. The relevant software is freely available.

1762M
Genetic Studies of Functional Quantitative Trait with both GWAS and Next-Generation Sequencing Data. D. Lee1, C. Hans1, G. Bell1, D. Aguilarr1, B. Cade2, J. Below2, M. Xiong2. 1) Biostatistics, University of Texas School of Public Health, Houston, TX; 2) Human Genetic Center, University of Texas School of Public Health, Houston, TX; 3) Department of Human Genetics, University of Chicago, Chicago, IL. 4) Cardiology, Baylor College of Medicine, Houston, TX; 5) Division of Sleep Medicine, Harvard medical school, Boston, MA.

Traditional quantitative genetics has primarily studied traits with cross sectional data. However, in the biologic world, many quantitative traits change over time. These quantitative traits are repeatedly measured as functions of time or complete curves. The traditional methods for such genetic studies summarize the time or space varying traits as means across the time or space. The traditional interval and quantitative trait locus (QTL) analysis for functional quantitative traits has the correct type I error, especially when the sequenced subjects are not a random subset of the whole cohort. Although this problem can be alleviated by restricting the analysis to the variants that are accurately imputed, a large number of rare variants will be excluded with this strategy. In this article, we show how to properly account for the uncertainties in the imputation of rare variants. We consider all commonly used gene-level association tests, including the burden test, variable threshold (VT) test, and sequence-kernel association test (SKAT), all of which are based on the score statistic for assessing the effects of individual variants on the trait of interest. We show that the score statistic based on the observed genotypes for sequenced subjects and the imputed genotypes for non-sequenced subjects is unbiased. We construct a robust variance estimator that reflects the true variability of the score statistic regardless of the sampling scheme and imputation quality, such that the corresponding association tests always have the correct type I error. Through extensive simulation studies, we demonstrate that the proposed tests are substantially more powerful than the use of accurately imputed variants only or the use of sequencing data alone. An application to the data from the Women's Health Initiative (WHI) is provided. The relevant software is freely available.
1763T
Genotype risk score may mislead physiological interpretation of quantitative trait associations. N. Wang1,2,3, Y. Shu1,2,3, H. Allayee1,2,3, A. Xiang4, T. Buchanand5, R. Watanabe1,2,3 1) Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, CA; 2) Physiology and Biophysics, Keck School of Medicine of USC, Los Angeles, CA; 3) USC Diabetes & Obesity Research Institute, Keck School of Medicine of USC, Los Angeles, CA; 4) Department of Research and Evaluation, Kaiser Permanent Southern California, Pasadena, CA; 5) Department of Medicine, Division of Endocrinology, Keck School of Medicine of USC, Los Angeles, CA.

The genotype risk score (GRS) has been used as a measure of total genetic risk and a predictor of complex disease risk. Better understanding of the physiology of complex disease can be achieved by studying disease-related quantitative traits (QTs). GRS has recently been increasingly used to explore association with disease-related QTs, although its validity has not been examined. We tested 38 type 2 diabetes (T2D) risk SNPs for association with 17 T2D-related QTs in the BetaGene study, which consists of Mexican American families of probands with or without previous gestational diabetes. We compared association by single-SNP, GRS, and stepwise forward selection (FS), to assess which method provides clear physiological interpretation. We tested each SNP for univariate association with Bonferroni correction for 646 tests (38 SNPs×17 QTs), which is overly conservative due to correlation among QTs. GRS was constructed by summing risk alleles without weighting and testing for association. FS picked SNPs with a threshold of P ≤ 0.1, followed by omnibus likelihood ratio test for association. All tests were adjusted for age, sex and familial correlation. 727 subjects with complete data, to ensure fair comparison, were included.

Univariate testing identified KCNQ1 rs2237892 (P=0.030) and MTNR1B rs10830963 (P=0.041) associated with disposition index (DI) and GCKR rs780094 with triglyceride (P=0.036). The proportion of variation explained ranged from 1.1-8.5% (median=3.8%). GRS was only associated with 3 QTs ranging from 0.2-7.6% (median=1.4%). GRF showed significant association with 16 of 17 QTs, with 3-11 SNPs entering into final models. The proportion of variation explained ranged from 1.1-8.5% (median=3.8%). The two strongest associations were for AIR (P=4.6×10^{-5}; 2.5% variation explained), acute insulin response (AIR), P=4.6×10^{-5}; 2.7% variation explained), and glucose effectiveness (P=0.022; 0.68% variation explained). GRS showed significant association with 16 of 17 QTs, with 3-11 SNPs entering into final models. The proportion of variation explained ranged from 1.1-8.5% (median=3.8%). The two strongest associations were for AIR (P=2.7×10^{-7}; 11 SNPs accounting for 8.5% of variation) and DI (P=9.9×10^{-3}; 8 SNPs accounting for 7.6% of variation). In summary, GRS was only associated with 3 of 17 QTs, while FS was associated with 16. GRS does not identify which loci contribute to the association with a given QT. FS identifies the subset of SNPs contributing to the association, providing detailed information facilitating physiological interpretation. We conclude GRS may lead to misleading physiologic interpretation of genotype-phenotype relationships due to lack of detailed information.

1764S
Development and application of a population based statistical framework addressing the n=1 problem in human genetics. A.B. Wilfert1, M. Lek2,3, J.N. Constantino4, W.A. Gahl4,5,7, E. Flynn2,7, E. Valkanas2,7, D.G. MacArthur2,3, D.F. Conrad1 1) Genetics, Washington University in St. Louis, St. Louis, MO; 2) Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA; 3) Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA; 4) Psychiatry, Washington University in St. Louis, St. Louis, MO; 5) NIH Undiagnosed Diseases Program, NIH, Bethesda, MD; 6) Office of the clinical director, NHGRI, NIH, Bethesda, MD; 7) Medical Genetics Branch, NHGRI, NIH, Bethesda, MD.

The diagnosis of rare, idiopathic diseases is emerging as a primary application of medical genome sequencing, and has motivated high-profile research efforts such as the NIH’s Undiagnosed Diseases Program (UDP). However, the application of standard tools from genetic epidemiology for many of these cases is frustrated by a combination of small sample sizes, genetic heterogeneity and gene by environment interactions. In response, we have developed an inference framework that uses massive control population sequencing datasets to measure the disease potential of a genotype or diplotype in a sample size as small as a single individual, what we refer to as the "n=1" problem. Our framework is built on 3 principles that are advances on historically applied pathogenicity prediction algorithms. (1) Like others, we integrate multiple prediction algorithms to estimate the probability that a variant damages protein function. (2) Most prediction methods that we incorporate are generic in that they use the same model to annotate variation in all genes and are naive to ploidy. We believe the best predictors will be gene-specific and genetic in nature and have taken steps towards this by integrating information on gene physiology using model-based measures of haploinsufficiency for all genes in the genome. (3) We use simulation, mutation rate modeling, and exome sequencing data from over 15,000 individuals to create a null model of "normal" genetic variation and estimate the probability of sampling a genotype or diplotype of a given pathogenicity score from healthy individuals. Our models have unmatched ability to discriminate putative disease mutations in the Human Gene Mutation Database (HGMD) from variants identified by population re-sequencing (AUC of 91% and 97%, respectively). We simulate over 30 million cases of the n=1 problem by spiking HGMD variants into 1092 healthy exomes from the 1000 Genomes Project. For 30% of all HGMD variants our gene-specific and genetic model always ranks them into the top 10 most pathogenic variants in an exome. The results of our modeling and spike-in analyses replicate when replacing HGMD with ClinVar data. We discuss the weaknesses of our approach and our attempts to identify areas for improvement. Finally, we apply our framework to a number of real-life n=1 cases, including a pair of twins from a consanguineous marriage concordant for Autism and several hundred cases from the NIH’s Undiagnosed Diseases Program.
1765M
Meta-analysis on polygenic effects. J.H. Zhao, J.A. Luan, S.J. Sharp. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, United Kingdom.

Background: It has been shown recently that whole genome data can facilitate estimation of genetic contributions to a variety of traits via a mixed model framework (Yang et al. Nat Genet 2010; 42:565-9; Zhao & Luan. J Prob Stat 2012, doi 10.1155/2012.485174). Our aim was to use this approach to investigate the impact of meta-analysis using study specific summary statistics such as (the genomic) heritabilities and/or polygenic/residual variance components as compared to those estimates from individual level data. Methods: We analysed body mass index from population-based and family-based cohorts, EPIC-Norfolk, Fenland, and Framingham studies involving Affymetrix 500K and Affymetrix 6.0 SNPs, involving ~10,000 individuals in total. We obtained genomic relationship matrices, which were then used in mixed models as implemented in the computer program GCTA and alternative procedures in PLINK. R. We compared results from meta-analysis using summary statistics and from individual level data. We also conducted simulation experiments. Results: Although the participants in the three cohorts were all European descent, there was variability in their genomic heritability estimates, with point heritability estimates being ~20%, 30% and ~50%, respectively. In the case of Framingham data, although the familial and genomic relationship matrices did not necessarily take equal values numerically they yielded comparable estimates. When polygenic effects are involved, meta-analysis using either heritability or variance components generally agreed with those from individual level data but with larger standard errors, which were likely to be larger than those assuming independence between inter-cohort participants in other scenarios not involving polygenic effects. The availability of whole-genome data seamlessly allows for meta-analysis of population-based and family-based data under the polygenic model. We have made the relevant implementations available in the R package. Conclusions: This highlights the utility of whole genome data in heritability estimation as well as availability of individual level data so as to account for correlation among all individuals. Further work will be to assess the uncertainty due to sampling of the genome as with other aspects of association analysis in the presence of polygenic effects.

1767S

1) Harvard School of Public Health, Boston, MA; 2) The University of Queensland, Brisbane, Queensland, Australia; 3) Massachusetts General Hospital, MA, USA; 4) Broad Institute, Cambridge, MA, USA; 5) Brigham & Women’s Hospital, Boston, MA, USA; 6) Harvard Medical School, Boston, MA, USA; 7) Keck School of Medicine, USC, Los Angeles, CA, USA; 8) Schizophrenia Working Group of the Psychiatric Genomics Consortium; 9) Icahn School of Medicine at Mount Sinai, New York, NY, USA; 10) UCSF, San Francisco, CA, USA; 11) UCLA, CA Los Angeles USA; 12) The University of Melbourne, Parkville, Victoria, Australia.

Polygenic risk scores are a promising approach for predicting disease risk, with prediction accuracy expected to improve substantially as sample sizes increase (Chatterjee et al. 2013 Nat Genet; Dudbridge et al. 2013 PLoS Genet). The standard approach involves pruning variants in linkage disequilibrium (LD) and applying a P-value threshold to association statistics, but this discards information and may limit predictive accuracy. We developed a Bayesian polygenic risk score, LDpred (https://bitbucket.org/bjarni_vilhjalmsson/lodpred), that uses LD from a reference panel to estimate posterior mean causal effect sizes from GWAS summary statistics, producing optimal (best linear unbiased prediction) polygenic risk scores when model assumptions hold. These estimates have a closed-form solution under an infinitesimal model for any number of SNPs, and can be drawn from a mean-zero normal distribution. Under a non-infiniteesimal prior where only a fraction of markers are causal, we approximate them via Markov chain Monte Carlo (MCMC). In simulations using real genotypes, our LDpred method is more accurate than existing methods (particularly at large sample sizes), well-calibrated, and computationally efficient. We applied LDpred to seven WTCCC diseases and observed substantial improvements in prediction R² (4-5%) for type-1 diabetes it improved from 30.4% to 32.8% and for schizophrenia it improved from 24% to 26.7%). LDpred is an efficient method for quantitative genetic and genomic data, although the familial and genomic relationship matrices did not necessarily take equal values numerically they yielded comparable estimates. When polygenic effects are involved, meta-analysis using either heritability or variance components generally agreed with those from individual level data but with larger standard errors, which were likely to be larger than those assuming independence between inter-cohort participants in other scenarios not involving polygenic effects. The availability of whole-genome data seamlessly allows for meta-analysis of population-based and family-based data under the polygenic model. We have made the relevant implementations available in the R/gap package. Conclusion: The work supports the utility of whole genome data in heritability estimation as well as availability of individual level data so as to account for correlation among all individuals. Further work will be to assess the uncertainty due to sampling of the genome as with other aspects of association analysis in the presence of polygenic effects.

1766T
Estimation of causal effects distribution from genome-wide association studies. L. Zhang1, Y.F. Pei1, Y. Lin1. 1) School of public health, Soochow university, Suzhou, JiangSu, China; 2) University of Shanghai for Science and Technology, Shanghai, China.

Estimation of the distribution of causal SNP effects and their heritability from genome-wide association study has been a research interest as it could explain in part the mystery of missing heritability. In this study, we propose a novel statistical method for such estimation. Specifically, we study the full range of GWAS summary results and link observed p-values and unobserved effect sizes by (non-central) chi-square distributions. By modeling the observed full set of GWAS p-values into a multinomial event, we build likelihood function in terms of causal SNP effects. We present both parametric and non-parametric forms of maximal likelihood estimation. The simulation studies showed that the proposed method had the ability to accurately estimate the number of causal SNPs and their effect sizes. As a real application, we analyzed a publicly available GWAS summary dataset that was released by the GIANT consortium, for height trait. Our analyses showed that there was a total of over 6,000 SNPs that might be associated with height, and explained ~40% heritability, where the number was much larger than that was previously estimated. As a conclusion, our proposed method has the potential to estimate common SNP-based genetic basis of complex traits from large-scale GWAS meta-analyses results.

1768M
Mediation Analysis of Integrated Genetic and Genomic Data in the Presence of Missing Data. R. Barfield, X. Lin. Harvard University, Boston, MA.

It is of increasing interest to analyze integrated different types of genetic and genomic data. Mediation analysis provides a useful tool for analysis of integrated genetic and genomic data to understand disease causing mechanisms. In genetic and genomic studies, SNP data, such as GWAS or sequencing data, are often collected on all individuals enrolled in a study. The genomic data, such as gene expressions and DNA methylations are often collected in a subset of study subjects. When the two types of data are combined to perform an functional analysis, the individuals who only had the SNP data collected are typically ignored. These individuals however can still provide useful information to the analysis. We propose a mediation analysis method using all the data by leveraging the information from the individuals with only the SNP data. We show using all available data, we gained more efficient estimators of the direct effects of SNPs and the indirect effects of SNPs mediated through gene expressions/DNA methylations on a phenotype with varying level of missingness. As the proportion of individuals with missing data increased, we gain an increase in the efficiency of our estimates. We also show power gain in detecting genetic and genomic associations using all available data. We applied our method to data from the MESA cohort using DNA methylation data and SNP data with obstructive sleep apnea as the outcome.

Genome-wide association studies have found thousands of variants robustly associated with complex traits. However, for most of them the underlying biological mechanisms are not well understood. The enrichment of gene expression associated variants (eQTLs) among disease-associated variants indicates that a substantial component of trait variability is likely to be determined through the regulation of the transcriptome. We propose a method termed PrediXcan that directly tests this hypothesis by correlating genetically predicted gene expression levels with the phenotype. For this purpose, we developed prediction models of gene expression using data from multiple eQTL studies (GTEx pilot data, GENUVADIS/1KGenomes, Framingham). Among advantages of PrediXcan are that 1) actual gene expression levels are not needed since they are calculated directly from the genetic data, 2) mechanism of the effect is directly built into the test making biological interpretation straightforward, 3) unlike actual differential expression studies, disease status cannot affect in silico levels thus ruling out reverse causality, 4) it tells whether up or down regulation drives the risk, 5) multiple testing burden is greatly reduced. Using these models we computed the in-silico expression levels of 15200 genes in six disease studies from the Wellcome Trust Case Control Consortium (WTCCC). For example, in Rheumatoid Arthritis 23 HLA region genes were differentially expressed (genome-wide significant threshold p<3.3e-6). Outside of the HLA region, RSBN1 (p=2.4e-6) and PSME1 (p=1.5e-7) were of interest. We found many positive control genes as well as many novel ones with replication efforts under way. In conclusion, we propose a powerful gene based association test that directly tests the hypothesis that genetic variants affect disease status through the regulation of genes. Application to the WTCCC data identified many novel genome-wide significant genes.

Addressing Potential Bias in Heritability and Coheritability Estimates within Ancestrally Homogeneous Populations. J. Liu1,2,3, T.J. Hoffmann1,2, J.S. Witte1,2,3,4 1) Department of Epidemiology & Biostatistics, UCSF, San Francisco, CA; 2) Institute for Human Genetics, UCSF, San Francisco, CA; 3) Diller Family Cancer Center, UCSF, San Francisco, CA; 4) Department of Urology, UCSF, San Francisco, CA.

A substantial heritability of complex traits can be explained by considering all SNPs across the genome simultaneously using linear mixed-effects models with an estimated genetic relationship matrix (GRM). This approach has been extended to assess the coheritability—or potential shared genetic basis—among different traits, including psychiatric disorders, inflammatory bowel disease, and major chronic obstructive pulmonary disease-related traits. However, the estimated GRM used in such (co)heritability analyses depends on sample-based allele frequency estimates across the genome. Any systematic differences in allele frequency estimates among the study samples can lead to a biased GRM and incorrect (co)heritability estimates. For example, when calculating the coheritability of a trait across different ancestral populations, differences in their allele frequencies can over- or under-estimate the relatedness between ancestrally similar individuals, resulting in more exclusions of samples from the analysis, and biased GRM and (co)heritability. As shown previously, one can address this issue with (co)heritability among ancestrally distinct or admixed populations by calculating the GRM by using individual-specific (or population-specific) SNP allele frequencies based on the individual ancestry. However, little if any consideration has been given to the potential impact of this issue on (co)heritability estimation in relatively homogeneous populations. It is increasingly common to combine samples from different studies with similar ancestral backgrounds for (co)heritability analysis. Here, samples from different studies may typed on different genotype platforms or have distinct cryptic relatedness. As with ancestrally distinct or admixed populations, ignoring these differences can bias the GRM and in turn biased (co)heritability estimates. We highlight this issue with two examples: 1) the bias in estimating lung cancer heritability within a European population composed of samples from three studies; and 2) the bias in estimating the coheritability of lung-breast cancer within European population in which the traits are measured on different samples. In conclusion, we show that the allele frequency adjustment approach mentioned above is able to correct the bias efficiently. Our findings emphasize the need to control for the bias due to allele frequency differences between groups of samples even when they appear to be from a relatively homogeneous population.

Statistics for genetic association in the presence of covariates - genome scanning considerations. H. Lin1, E. Feingold1,2, Y. Lin1. 1) Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA.

A number of different statistics are available for genetic association analysis in the presence of covariates. In the context of a genome-wide association study, hundreds of thousands to millions of SNPs are tested, and whatever covariate model we specify is likely to be imperfect. In addition, the results of the study often focus on the list of SNPs ordered according to the statistics rather than on certain p-value cutoffs. Therefore, it is important to investigate the behavior of extreme values of the statistics rather than the behavior of the expected values. Gail et al. (2008) discussed this issue and proposed “detection probability” and “proportion positive” to measure the success (power) of a genomic study when ranked lists are the primary outcome. In theory, the ranked lists can be dominated by SNPs with misfit models rather than by true positive results. We are conducting a comprehensive comparative study to investigate the behavior of different association statistics that model covariates. We evaluate the statistics from the perspective of which statistics can provide robust ranked lists of “top hits.” These are not necessarily the same statistics that have the highest power in a conventional single-test context.
A new prognostic model to predict renal outcome in autosomal dominant polycystic kidney disease (ADPKD). E. Cornec-Le Gall1, MP. Audrèze1, M. Trouillas4, M. Morinière5, M. Voisin5, M. Borini5, P. Mooser5, A. Zalloua5, J. Kwan1, M. Li1, J. Deng1, P. Sham1. 1) Department of Nephrology, Centre Hospitalier Universitaire de Bretagne, Saint-Nazaire, France; 2) Department of Nephrology, Centre Hospitalier Universitaire de Rennes; 3) Department of Nephrology, Centre Hospitalier de Quimper; 4) Department of Nephrology, Centre Hospitalier de Pontivy; 5) Centre de Dialyse, Association des urémiques de Bretagne, Brest; 10) Société Brestoise du Rein, Brest; 11) Department of Nephrology, Centre Hospitalier de Saint Nazaire; 12) Department of Nephrology, Centre Hospitalier de Vannes; 17) Centre associatif de Dialyse, ECHO, Nantes; 18) Etablissement Français du Sang, Brest.

ADPKD is marked by a high clinical variability, especially in regards of age at end-stage renal disease (ESRD). As we are shifting to the end of targeted therapies in ADPKD, one needs to target the patients who are more likely to develop ESRD and thus to benefit from those therapies. In that setting, we conducted a cross-sectional study in a population of 1130 patients (525 males) from all the nephrology centers (n=17) of a single area, Brittany, France. We applied the least-squares method to evaluate the association between several clinical, genetic, and molecular genetic data (mutation of PKD2, non truncating mutation (NTM) of PKD1 or truncating mutation (TM) of PKD1) using Kaplan-Meier curves and univariate followed by multivariate Cox regression analyses and built a score weighting each significant factor according to the HR obtained. Internal validation was assessed using non parametric bootstrapping with replacement. Discrimination of the model was evaluated using the c-statistic, which represents area under time-dependant ROC curves. Median age at inclusion was 54.3 yrs [54-53-54]. After multivariate analysis, 4 factors remained significantly associated with renal survival : Gender (HR=1.15, 0.95 CI=1.25-2.93), Hypertension onset before age 35 (HR=2.24, 0.95 CI=1.76-2.86), first urological complication before age 35 (at least 1 amongst gross hematuria, flank pains related to cyst or cyst infection) (HR=2.1, 0.95 CI=1.62-2.73), and genetic status (PKD2 vs NTM of PKD1 HR=1.93, 0.95CI=1.27-2.93, vs TM of PKD1 HR=4.88, 0.95CI=3.39-7.02). We thus defined a score ranging from 0 to 9 (male gender:1 pt, hypertension onset before age 35: 2 pts, first urological complication before age 35: 2 pts, PKD2 mutation: 0 pt, NTM of PKD1: 2 pts and TM of PKD1: 4 pts). The accuracy of the PRO-PKD score applied to the sample was high, with a c-statistic of 0.863 at 65 yrs. To facilitate ease-of-use, we defined 3 risk categories: low risk, intermediate and High-Risk of early progression to ESRD with median ages at ESRD of respectively 73.6, 57.7 and 48.5 yrs (HR: Intermediate Risk=3.99, High-Risk=13). Predicted probabilities of ESRD at age 60 were 16.5%, 57.5%, and 91.2%, for the Low, Intermediate, and High Risk Group, and accounted for 75% of all observed ESRD. The PRO-PKD score is simple, accurate and easy-to-use prognostic score that might be of interest in the selection of the patients who should be included in the ADPKD clinical trials and be a valuable tool for personalized medicine in ADPKD.
standing the patterns of recurrence in transmission genetics. That support a maternal origin bias of recurrent de novo mutations. Our shows that the predictions of our model are strongly consistent with data on observation of an affected child led to unexpected predictions for recur-
ing the mean and variance of the proportion of mutant gametes based on the observation of an affected child led to unexpected predictions for recurrence risk. Consistent with empirical data, our results that Unrecognized ASE can reduce estimates of genotype/phenotype relations.

1777M Parent of origin and recurrence risk bias: probabilistic modeling explains the broken symmetry of transmission genetics. C. Shaw 1,2, J. L. Campbell 1, P. Olsson 2,3. 1) Dept Molecular & Human Gen, Baylor Coll Medicine, Houston, TX; 2) Mathematics Department, Trinity University, San Antonio, TX.

Most de novo human mutations are observed to arise in fathers, and increasing paternal age strongly correlates with increased risk of paternally transmitted mutations. Strikingly, de novo mutations in X-linked recessive disease show strongly elevated familial recurrence rates, and in male offspring these mutations must be transmitted from fathers. We are intrigued by these contrasting patterns of mutation transmission and their consequences for human genetics. We developed a comprehensive stochastic process model of mutation in the human germ line, closely modeling the sex-specific demography between males and females and considering the effect of increasing paternal age. We extended our prior work on exact sampling formulas for multiple-gene linkage analyses to develop new formulas for the mean and variance of the proportion of mutant gametes in each paternal age category. Our method has several advantages: it allows for updating the mean and variance of the proportion of mutant gametes based on observation of an affected child led to unexpected predictions for recurrence risk that implicate parent of origin as a central variable in the analysis of recurrence risk. Consistent with empirical data, our results that more transmitted mutations arise in fathers, and that this trend increases with a father's age. Surprisingly, the variance in the proportion of mutant gametes is lower in fathers and decreases with age, so that knowledge of a transmitted mutation has small effect on updating the expected proportion of mutant gametes in the male offspring. Quite different is the case for the female germ line, which arrests after clonal expansion in early develop-
ment, variance in the proportion of mutants is much higher, and observation of a transmitted mutation in an affected child dramatically increases the risk of recurrence in another pregnancy. Our model demonstrated that parent of origin determined in families with recurrent genetic disease shows that the predictions of our model are strongly consistent with data that support a maternal origin bias of recurrent de novo mutations. Our findings explain the broken symmetry of transmission genetics.

1778T What are genome-wide association studies detecting? Our experience predicting cystic fibrosis-related diabetes onset. D. Soave 1,2, M. Miller 1,2, K. Keenan 1,2, U. Durante 1,3, L. Summ 1,2, J. Rommens 1,2, L. Straub 1,2. 1) Program in Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Division of Biostatistics, Dalhousie School of Public Health, University of Toronto, Toronto, Ontario, Canada; 3) Program in Physiology and Experimental Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; 4) Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada; 5) Department of Statistical Sciences, University of Toronto, Toronto, Ontario, Canada; 6) Program in Genetics and Genomic Biology, Genome Institute for New England, Hospital for Sick Children, Toronto, Ontario, Canada; 7) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada.

It is assumed that genome-wide association studies are identifying genetic determinants of the phenotype under study. However, we were reminded that this may not always be the case when a recent genome-wide association study for Cystic Fibrosis (CF)-related diabetes (CFRD) identified associated SNPs in SLC2A9 (HR=1.34, at rs7512462). SLC2A9 is an apical epithelial cell transporter expressed in human pancreas, but is unlikely to have a direct role in the endocrine compartment as its protein payload is delivered to the bloodstream and surrounding cells without a specialized ductal sys-
tem. We recently showed that rs7512462 accounts for >10% variability in early exocrine pancreatic damage in children with severe CF, and Mendelian randomization provided evidence that this early exocrine pancreatic damage causes later-onset of CFRD. These findings may explain how the SLC2A9 epithelial transporter contributes to CFRD. Despite only an indirect associa-
tion between SLC22A9 and CFRD risk, we hypothesized that SLC22A9 may have a differential role with prediction and validation of CFRD risk in CFRD predictive models using 126 participants from the Canadian CF Gene Modifier Study on whom longitudinal measures of immunoreactive trypsinogen (IRT; the early exocrine pancreatic damage biomarker) and rs7512462 genotype were available. Cox proportional hazards models with 10-fold cross-validation were used to generate risk scores for each individual and these risk scores were used to assess the CFRD predictive ability for various ages using time-dependent receiver operating characteristic curves and the corresponding area under the curve (AUC). Using risk scores calculated from estimated IRT at birth and its rate of decline to predict CFRD by age 12, 16, 20 and 24y, our cross-validated AUC's were 65%, 64% and 78%, respectively. With the addition of rs7512462 genotype into the risk score calculation, the predictive accuracy improved, including 72% to 73%, 74% and 80%, respectively. Although we are cognizant that genome-
wide significant SNPs require follow-up to identify the underlying causal variant that induced the association, less consideration is paid to identifying the underlying phenotype to which the variant is more proximally associated. Our findings suggest that once we better model the pancreatic phenotypic complexity, the addition of rs7512462 genotype can significantly improve predictive accuracy despite having only a weak and indirect effect on CFRD.
1780M Statistical method for analyzing allele-specific expression across individuals for multiple statuses. F. Lee, X. Wen, M. Boehnke. Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI.

It is well known that genetic variants in cis-regulatory regions can have large effects on the expression level of the target genes. Such effects can be detected by eQTL mapping and allele-specific expression (ASE) analysis. In this study, we are interested in examining the concordance between eQTL and ASE analysis, especially when the regulatory variants show state or tissue-specific effects. To this end, we develop a novel statistical method to perform ASE analysis using RNA-seq data across multiple individuals. One of the unique features of our method is that we require genuine ASE signals to be consistent across individuals, while accounting for potential heterogeneity among individuals. Also, using Bayesian model averaging, we effectively phase uncertainty into account. We apply this method to the RNA-seq data from the Exome Aggregation Consortium (ExAC) project. In particular, we attempt to confirm strong (tissue-specific) eQTL signals using ASE analysis, and we also evaluate the consistency of effect size from eQTL mapping and ASE analysis.

1781T A New Approach to Finding Association with Complex, Longitudinal Phenotypes using Population Data. A.M. Musolf1, D. Londono1, A.O. Nato Jr.2, P. Vuijtser3, J. Brandon1, J.A. Herring3, C.A. Wise5,6, H. Zou5, M. Jin1, L. Yu3, S.J. Finch8,9, P. Bovel1, M. Bochud2, T.C. Matise3, D. Gordon1. 1) Department of Genetics, Rutgers University; 2) Statistical Genetics Lab, Division of Medical Genetics, University of Washington, Seattle, WA 98195, USA; 3) Swiss Institute of Bioinformatics, 1014 Lausanne, Switzerland; 4) Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, Dallas, TX 75219, USA; 5) Department of Orthopedic Surgery, Texas Scottish Rite Hospital for Children, Dallas, TX 75219, USA; 6) Department of Orthopedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX 75219, USA; 7) Shanghai Institute of Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China; 8) ShanghaiBio China, Pudong, Shanghai 201203, China; 9) Center of Alcohol Studies, Rutgers University, Piscataway, NJ 08854, USA; 10) Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11790, USA; 11) Ministry of Health, P.O. Box 52, Mont Fleuri, Republic of Seychelles.

Previously, we detailed a new method for testing association between longitudinal phenotypes and causal genotypes. The method uses growth mixture models to determine longitudinal trajectory curves. The Bayesian posterior probability (BPP) of belonging to a specific curve was then used as a quantitative phenotype in association analyses. To identify association for multiple SNPs, we did not perform association analyses on individual SNPs; instead the genome was sliced into blocks of 50 SNPs. A significance value was obtained on each block via the program TDT-HET. This method displayed greater than 80% empirical power in most simulations scenarios. This method exclusively used family-based data. Here, we extend the method to population-based data sets. The method maintains many ideas from the family-based method. However, the program SumStat is used to acquire significance levels on each block, instead of TDT-HET. Multiple scenarios are tested including four causal variants located within a single locus and eight causal variants spread between two loci on different chromosomes. Reduced models using environmental covariates were also considered. Our data set was highly stratified to ensure robustness in the presence of population stratification. To correct for population stratification, ancestry fractions from the program ADMIXTURE are regressed on the BPPs and the residuals are used as the phenotype for association analyses. Our method also utilized three distinct data sets, which represent a discovery data set and two confirmatory data sets. The final p-values of the association value was obtained on each block via the program TDT-HET. This method exclusively used family-based data. Here, we extend the method to population-based data sets. The method maintains many ideas from the family-based method. However, the program SumStat is used to acquire significance levels on each block, instead of TDT-HET. Multiple scenarios are tested including four causal variants located within a single locus and eight causal variants spread between two loci on different chromosomes. Reduced models using environmental covariates were also considered. Our data set was highly stratified to ensure robustness in the presence of population stratification. To correct for population stratification, ancestry fractions from the program ADMIXTURE are regressed on the BPPs and the residuals are used as the phenotype for association analyses. Our method also utilized three distinct data sets, which represent a discovery data set and two confirmatory data sets. The final p-values of the association

1782S Multiple testing procedures for GWAS with high-dimensional phenotypes. C.B. Peterson1, M. Bogomolov2, Y. Benjamini3, C. Sabatti1. 1) Department of Health Research and Policy, Stanford University, Stanford, CA; 2) Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Technion City, Haifa, Israel; 3) Department of Statistics and Operations Research, Tel Aviv University, Tel Aviv, Israel.

We develop statistical methodology which have power and reduce errors in identifying genetic variants that are relevant to multivariate phenotypes such as imaging features, actigraphy measures, or gene expression. Since both the predictor and response variables are high dimensional, this represents a massive multiple testing problem.

A standard approach to handle this multiplicity is to control the false discovery rate across the entire set of hypotheses representing the association of each SNP to each phenotype. Unfortunately, this method does not control the false discovery rate for SNPs or the average false discovery rate for selected SNPs. This is a serious drawback for gene mapping studies, where the identification of SNPs with functional effects is one of the primary goals. In addition, it fails to account for the genetic architecture of traits. In this type of study, most genetic variants will not affect any of the phenotypes of interest, but those with functional effects are likely to influence multiple traits. In applying error control to the pooled set of hypotheses across all SNPs, we are reducing power to detect the effects of functional SNPs, while increasing errors for SNPs that have no influence on any traits.

In this work, we employ a two-step selection procedure recently proposed by Benjamini and Bogomolov that allows us to both gain control of the appropriate error rates and improve power by taking advantage of the genetic structure of the problem. Here we explore the application of this strategy in the context of GWAS. Specifically, we define families of hypotheses by SNP, where each family consists of hypotheses on the association of a particular SNP to each of the phenotypes under study. In step (1), we perform selection of SNPs that affect one or more phenotypes using, for instance, a standard statistic for which SNPs were obtained from the North Finland Birth Cohort (NFBC) study.


In studies of case-parent triads, information may be available about occurrence of the condition in the parents. Typically parental phenotypes are ignored, but including that information in analyses may increase power to detect genetic association for autosomal variants. We propose a method of using parental phenotypes that assesses association independently of the usual case-parent-based association test. Our proposed method enjoys many advantages of case-parents designs: it is robust to population stratification and allows testing of maternally-mediated genetics effects as well as inherited-gene effects. The parental information can enable a natural internal replication for findings based on offspring and their parents, and it can be used to improve power. We develop composite tests that combine evidence from this parent-phenotype-based test with the traditional log-linear transmission-based test. A likelihood-based approach builds information from these two sources into a single coherent model: The model can impose equality of parental and offspring relative risks and permit testing of that equality. This model also allows us to use the expectation-maximization algorithm when some parental genotypes are missing. We evaluate the proposed method through non-centrality parameter calculations and simulation studies. We show that incorporation of parental phenotype data often improves power to detect a genetic contribution to the phenotype. As illustration, we apply our method to a nuclear-family-based study of young-onset breast cancer.
1784T Efficient multiple imputation for missing phenotype using genomewide DNA methylation data. W. Guan1, W. Wu2, Y. Li3, J. Pankow2, E.W. DeNora3, J. Breasler, M. Fornage5, M.L. Groves2, T. Mosley4, C. Hicks2, E. Boerwinkle5. 1) Division of Biostatistics, Univ of Minnesota, Minneapolis, MN; 2) Department of Genetics, University of North Carolina, Chapel Hill, NC 27516; 3) Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27516; 4) Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN 55455; 5) Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX 77030; 6) Department of Medicine, School of Medicine, University of Mississippi, Jackson, MS 39216.

DNA methylation is a widely studied epigenetic mechanism and alterations in methylation patterns may be involved in the development of common diseases. Unlike inherited changes in genetic sequence, variation in site-specific methylation varies by tissue, developmental stage, disease status, and may be impacted by aging and exposure to environmental factors such as diet or smoking. While these wide-range correlations pose analytical challenges, including reverse causality and confounding by non-genetic factors in epigenome-wide association studies (EWASs), it brings opportunities to infer missing phenotype values using rich methylation data. We propose a multiple imputation method using large-scale methylation data, such as that provided by the Illumina HumanMethylation450 (HM450) BeadChip. Based on established methylation-phenotype association at multiple CpG sites, plausible values for missing data in phenotype variables can be drawn multiple times. The imputed datasets can then be analyzed separately using standard complete-case analysis, with the point estimates and standard errors adjusted for missingness at each site.

1786M Confounded by Ancestry? Considerations for Ancestry Adjustments in Genetic Association Tests. E.R. Martin1, I. Tunc1, Z. Liu, M.A. Schmidt1, C.D. Bustamante2, C.W. Beecham1. 1) Hussman Institute for Human Genomics, Dr. John T. Macdonald Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 2) Department of Genetics, Stanford University, Palo Alto, CA.

It is well known that ancestry differences may lead to confounding in genetic association tests. It has now become standard practice to adjust for principal components that capture global ancestry of samples, but this is often done without assessment of whether they are true confounders. Adjusting for covariates that are not true confounders can lead to loss of power; thus, judiciously assessing whether ancestry is a confounder prior to analysis is likely to improve power in association tests. In addition, there have been recent reports considering whether it is preferable to adjust for local or global ancestry, but the conclusions depend on the model for population structure and trait/disease model. Here, we review the definition of statistical confounding in general and the properties of covariate adjustment (with and without confounding) in relation to statistical power. We then examine theoretical expectations to illustrate when we might expect confounding and the consequence of ancestry adjustment in genetic association tests in the context of various genetic models for population structure and admixture. Theoretical calculations were verified and generalized to more complex situations using simulations. We showed that in a single admixed population, when testing the trait locus itself there is no confounding by ancestry, and in this scenario, adjusting unnecessarily for ancestry could lead to a substantial decrease in power. In stratified admixed populations when the trait values or disease prevalence differ between strata, global ancestry is a confounder and must be adjusted for to avoid inflated false positive rate. When testing for global ancestry captures the effect and there is no additional benefit adjusting for local ancestry. The one scenario that we examined where we observed a local ancestry effect in addition to a global ancestry effect was testing a marker in a focal rather than global ancestry locus. The population structure, dynamics and genetic model are usually unknown, we propose a variable selection strategy for assessing whether ancestry adjustment is necessary. Our simulations show that this strategy maintains correct Type I error, has good power compared to the true model, and has greater power than arbitrarily adjusting for ancestry covariates. This data-informed strategy provides a viable alternative to ubiquitous adjustment that preserves power while protecting against spurious results.
1788S

Effective genetic risk prediction using mixed models. D. Golan, S. Rosset. Statistics, Tel-Aviv University, Tel-Aviv, Israel.

Despite identifying thousands of genetic variants associated with dozens of high-impact diseases using genome-wide association studies, our ability to accurately predict the disease status using genetic data remains disappointingly low for many highly heritable diseases. One leading theory is that highly heritable diseases for which prediction is difficult (e.g., bipolar disorder, type-2 diabetes and hypertension) are typically driven by many common variants with small effects. The effects of most variants are too small to reach significance, but cumulatively these variants account for a considerable portion of the disease burden. This theory suggests that using only those variants which were found to be significantly associated with the disease is too conservative, and that a more permissive inclusion criterion should be used when selecting variants for the purpose of prediction. However, there is an obvious trade-off: when the inclusion criterion is stringent, a considerable part of the signal is left out, but when using a more permissive inclusion criterion, the number of estimated parameters (effect sizes) increases dramatically, resulting in overfitting problems and leading to poor performance. We propose a novel statistical approach for genetic risk score inference (GeRSI), based on mixed-effects models. Variants with strong associations are included in the model as fixed effects, just as in standard approaches. Additionally, variants with no significant association are not discarded, but instead are treated as random effects, thus circumventing the need to estimate their respective effects and considerably reducing the number of parameters in the model. By doing so, GeRSI takes advantage of all variants for better prediction, without losing accuracy due to estimating a huge number of parameters. We demonstrate the superiority of GeRSI to state-of-the-art methods in extensive simulation. When applying GeRSI to seven phenotypes from the WTCCC study, we confirm that the use of random effects is most beneficial for diseases that are known to be highly polygenic: hypertension (HT) and bipolar disorder (BD). For HT, there are no significant associations in the WTCCC data. The best existing model yields an AUC of 54%, while GeRSI improves it to 59%. For BD, using GeRSI improves the AUC from 55% to 62%. For individuals ranked at the top 10% of BD risk predictions, using GeRSI substantially increases the BD relative risk from 1.4 to 2.5.

1789M

Allele-specific DNase I hypersensitive sites exhibit H3K27ac enrichment in GM12878. J.M. Peralta1, M. Almeida1, L.J. Abraham1, E. Moses1, J. Biangero1. 1) Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX; 2) Centre for Genetic Origins of Health and Disease, University of Western Australia, WA.

Open chromatin accessibility is often assessed by the presence of DNase I hypersensitive sites (DHSs). It correlates with cis-located elements that regulate the expression of a gene, like promoters and transcription binding sites. The acetylation of Lysine 27 of histone H3 (H3K27ac) is a modification highly correlated with cis-located elements that regulate the expression of a gene, like promoters and transcription binding sites. The acetylation of Lysine 27 of histone H3 (H3K27ac) is a modification strongly correlated with the active state of enhancers to promote transcription and gene expression. Non-coding variants that affect the binding of proteins to cis-regulatory elements can change the local sensitivity to DNase I digestion, leading to allele-specific DHSs that can serve as an indication of variants with a potentially functional effect on the regulation of neighboring genes. These allele-specific DHSs can be detected, in heterozygotes, as a significant deviation from the expectation of equal allele depth coverage at SNP sites within DHSs. We refer to the significance of this effect, in the -log10[p] scale and derived from a likelihood ratio based test, as the functional potential (FP) of a SNP. Using mapped short sequencing reads (reads) from DNase-seq of five pooled GM12878 lymphoblastoid cell line replicates (R1-R5) for NA12878, released by the ENCODE Project, we examined the allele-specific chromatin accessibility at 2423308 known heterozygous SNPs sites, based on Illumina’s Platinum WGS of NA12878. FP estimates were obtained from independent tests for 482371 (99%) SNPs using the read depths from each of the two SNP alleles from GM12878 pooled DNase-seq reads. A stringent Bonferroni threshold (-log10[n]≥6 ) was used to classify our FP estimates as two groups, FP≥6 (n=161, 0.33%) and FP<6 (n=48076, 99.67%). Then we determined the FP overlap (n=108) within H3K27ac signals. Relaxation of the cutoff threshold for the classification of FP estimates by 2-log10 units lead to an even better enrichment signal (p=6.76×10^-50). These results support a role for allele-specific DHSs in the regulation of gene expression and suggest that the FP of SNPs can be used to quickly classify potentially functional non-coding variants.

1790T

Genetic modifiers in TGFβ pathway affect disease severity in Duchenne Muscular Dystrophy. J. Punetha1-2, H. Gordish-Dressman1-2, L. Bello1, A. Kesari1, M. Girir1, E.P. Hoffman1-2. Cooperative International Neuromuscular Research Group. 1) Department of Integrative Systems Biology, The George Washington University School of Medicine, Washington DC, USA; 2) Center for Genetic Medicine Research, Children’s National Medical Center, Washington DC, USA.

Duchenne muscular dystrophy (DMD [MIM 310200]), the most common monogenic disease in boys, shows marked heterogeneity in disease onset and progression. This inter-patient phenotypic variability observed in DMD can be attributed to the presence of genetic modifiers. Genetic modifiers of DMD (SPP1 rs28357094, LTBP4 rs10880) have been associated with loss of ambulation and lower grip strength in patients. Both known DMD genetic modifiers lie within the transforming growth factor-beta (TGF-β) signaling pathway, which is known to be dysregulated in DMD patients, thus increasing its significance in disease pathophysiology. We used a TGF-β pathway driven approach to help investigate potential genetic modifiers and their interactions in a large natural history cohort of DMD patients (CINRG) followed longitudinally for 5-8 years. Key single nucleotide polymorphisms (SNP’s) in TGF-β pathway (SPP1, LTBP4, TGFBR2, and IBSP) genes were genotyped and tested for associations with outcome measures. Outcome measures chosen were those that could predict disease progression before loss of ambulation occurred i.e. time to run/walk, grip strength. In accordance with our previous studies, dominant SPP1 rs28357094 genotype showed significant association with grip strength and appeared to be dependent on steroid use. We identified a novel genetic modifier in DMD; patients with the recessive IBSP rs2616262 T/T genotype showed a significant association with disease progression in the time to run/walk (TTRW) test. IBSP genotype also showed significant association with grip strength in the non-ambulatory patients (p=0.04, n=111). Preliminary data for SNP interaction studies between SPP1 and IBSP polymorphisms show evidence of interaction in the top 10% of the two loci. Providing sufficient statistical confidence of genetic modifier effects in different patient cohorts has been problematic due to differences in steroid treatment regimens, ethnicity, outcome measures, age of study population, and study sites tested. In our cohort, there were differences in functional outcomes between different ethnic groups as well as different standards of care provided. Validation of SNP association studies requires selection of a homogenous population to reduce confounding factors and hidden stratification. Further analyses of multi-locus modifier approaches will help identify the key players in transition from successful to unsuccessful remodeling in muscular dystrophy.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1791S
Linkage disequilibrium clustering can improve power of weighted-sum-type multi-marker tests for genetic association analysis. Y. Yoo1,2, S. Kim1, L. Sun1,2, S. Bulik3,5 1) Mathematics Education, Seoul National University, Seoul, South Korea; 2) Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, South Korea; 3) Division of Biostatistics, Dalca Lana School of Public Health, University of Toronto, Toronto, Canada; 4) Department of Statistical Science, University of Toronto, Toronto, Canada; 5) Prosserman Centre for Health Research, The Lunenburg-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Canada.

Background For gene-based analysis of genome-wide association studies, a multi-SNP regression analysis using common and low frequency SNPs can be performed for a gene unit and global hypothesis tests can be constructed. Methods constructed as quadratic forms are especially effective in the presence of both deleterious and protective causal SNPs. On the other hand, weighted-sum type tests benefit from signals of both causal SNPs and are deleterious. The power of both tests suffers in the presence of many neutral SNPs relative to causal SNPs. In this study, we investigate how clustering of SNPs can improve power of a hybrid quadratic-linear combination test using weights incorporating the correlations between SNPs. Methods The Wald statistic is defined as a quadratic sum of beta coefficients with covariance matrix weights. The LC statistic is linear sum using weights incorporating variances and covariances between SNPs. We propose the MLC statistic defined as a quadratic sum of bi-cluster linear combinations, and apply a spectral clustering algorithm to cluster closely correlated SNPs. The coding for base and risk alleles within each cluster is assigned using a pairwise LD coding method so that the number of pairs of SNPs with positive value of linkage disequilibrium measure r is as many as possible in each cluster. We compared the power of Wald, LC, MLC tests with numerical calculation of asymptotic power over different disease model scenarios. For comparison over realistic correlation patterns, we selected 1000 genes, and specified genotype-phenotype models using the structure of SNPs from HapMap data for these genes. Three types of disease model scenarios were examined: A: one causal SNP, B: two causal SNPs in different clusters both with deleterious effects, C: two causal SNPs in different clusters with one deleterious effect and one protective effect. Results and Conclusions For a one causal SNP scenario, the number of pairs of SNPs that can capture the true effect is usually higher than the other tests under all three disease models, and the power of the MLC test is usually higher than the other tests under all three disease models. For the scenarios where the XCI is random, we compared our method to other state-of-the-art packages (e.g. hapLOH, BAFsegmentation) to identify aberrations and improve detection performance in mixtures with aberrations in less than <5% of the sample. We focus our comparisons on a novel data set, constructed from series of mixtures of DNA from cell lines derived from tumor and paired normal tissue of a lung (T: HTB-175, N: CRL5949) and breast (T: CRL-2343, N: CRL-2363) cancer. Methods such as hapLOH that capture haplotype information alone excel over those that simply model the magnitude of BAFs. However, the more subtle inflections in signal at the interpretable threshold can be represented. In several instances are detectable by our method, we incorporate a more use of the available data. Our flexible framework can be utilized in a variety of contexts, including testing specific regions of a priori interest for AI.

1792M
A unified analysis approach for X-chromosome that accounts for random, skewed and escaping of X-chromosome inactivation. J. Wang, R. Yu, S. Shete. Dept Biostatistics, UT MD Anderson Cancer Ctr, Houston, TX.

X-chromosome inactivation (XCI) on female X-chromosome loci states that in females, one X chromosome is randomly inactivated during early embryonic development. The XCI process is random, skewed and escaping. Due to the random XCI, about half of the cells have one allele inactive while in 5-10% of the cells, the XCI event is random and one allele is active. Our framework allows for a probabilistic characterization of each bi-marker subset and each bi-marker subset with 1 of the 2 copies of X chromosome inactive while the other is active. We developed a novel data set, constructed from series of mixtures of DNA from cell lines derived from tumor and paired normal tissue of a lung (T: HTB-175, N: CRL-5949) and breast (T: CRL-2343, N: CRL-2363) cancer. Methods such as hapLOH that capture haplotype information alone excel over those that simply model the magnitude of BAFs. However, the more subtle inflections in signal at the interpretable threshold can be represented. In several instances are detectable by our method, we incorporate a more use of the available data. Our flexible framework can be utilized in a variety of contexts, including testing specific regions of a priori interest for AI.
1795M
Quality and accuracy assessment for NGS data analysis and interpretation. J. Li, M. Mohiyuddin, A. Kian, M. Camerio, P. Jiang1, S. Tabrizi2, S. Saffranter3, I. Shlyakhter4, N. Assadi5, P. Sabeti2, W.H. Wong5, H.Y.K. Lam1. 1) Department of Bioinformatics, Bina Technologies, Redwood City, CA; 2) Department of Engineering, Bina Technologies, Redwood City, CA; 3) Program in Medical and Population Genetics, the Broad Institute, Cambridge, MA; 4) Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA; 5) Department of Statistics, Stanford University, Stanford, CA; 6) Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA.

Next-generation sequencing (NGS) has become the mainstay for characterizing genetic variations for both targeted genomic regions and whole genome. The quickly decreasing cost at increasing coverage is enabling numerous possible applications of large-scale studies, ranging from discovering disease-causing mutations and pathways to identifying cancer somatic alterations.

To conduct meaningful research and generate correct interpretation based on the identified variations, quality assessment (QA) of the data becomes extremely important, especially in clinical settings. Accumulated cohort study results, in turn, provide learning resources for differentiating true variations from errors. Here we present a comprehensive, user-friendly quality control pipeline for NGS secondary and tertiary analysis. This pipeline, synergized with the secondary analysis framework, provides across-the-board quality measurements addressing potential quality issues residing in NGS reads, such as sequence quality, length, and adaptor detection; in alignment, such as coverage statistics, mapping distribution, and insert estimation; and in variant calling, such as known variant ratio, heterozygous/homozygous ratio, and transition/transversion ratio.

For secondary analysis, to ensure the relevant biological interpretations are not caused by low quality or conflicting data, our pipeline provides a novel scheme for quality assurance by learning from accumulated homogeneous samples. A wide range of quality control is performed on the identified variations from the new samples, such as calculating missing data rate and heterozygosity rate, detecting duplicated or related individuals, checking for ethnicity, gender and blood type, identifying mitochondrial haplotype, and computing Hardy-Weinberg log p-value.

Visualization methods such as clustering heatmap, VQS/LOD score plots, and a novel method we have developed - QC-Radar, are used in QC report for users to easily identify potential problems at a glance. We have applied our pipeline to a large number of samples, including recently to 20 Lassa fever patient genomes of a cohort in which we successfully identified the optimal VQSR LOD score threshold to maximize the sensitivity of the high-confidence variant call set.

1796T
A comprehensive survey of genetic variation in 20,769 subjects from the Harvard Cohorts. S. Lindstrom1, S. Loosier2, C. Chen1, H. Huang2, J. Huang2, A. Chen3, H. Choi4, G. Curhan5, I. De Vivo6,7, C. Fuchs8, F. Hui9, K. Khera1, F. Pasquier10, E. Rimm11, S. Twerdeger12, D. Hunter11,6, P. Kraft11. 1) Harvard School of Public Health, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) Massachusetts Eye and Ear Infirmary, Boston, MA; 4) Massachusetts General Hospital, Boston, MA; 5) Boston University School of Medicine, Boston, MA; 6) Brigham and Women’s Hospital, Boston, MA; 7) Dana-Farber Cancer Institute, Boston, MA.

The Nurses’ Health Study (NHS), Nurses’ Health Study II (NHSII), Health Professionals Follow Up Study (HPFS) and the Physicians Health Study (PHS) have collected detailed longitudinal data on a wide range of exposures and traits for more than 310,000 study participants over the last 35 years. Across the four cohorts, more than 90,000 study participants donated a blood sample between 1982 and 1999 allowing for prospective analysis of biomarkers and genetic variation. Over the last eight years, 20,769 subjects have been genotyped as part of genome-wide association studies (GWAS) of 12 primary traits within the cohorts. However, these studies have utilized 6 different GWAS platforms, making it difficult to conduct analyses of secondary phenotypes by borrow controls from other studies. To allow for secondary analysis in this data, we have created new datasets merged by platform and imputed them to a common reference panel, the 1,000 Genomes Phase I release. We describe the methodology behind the data merging and imputation and present new imputation quality statistics and association results from the 1 release. We describe the methodology behind the data merging and imputation quality statistics and association results from the 1 release.

1797S

Judiciously selecting carriers of target genetic variants is crucial for the success of a target deep sequencing study. Extreme phenotype sampling (EPS) has been widely employed to identify complex trait genes. EPS can enrich rare genetic causals variants compared to naïve random sampling (NRS). An intuitive alternative is extreme residual sampling (ERS) - selecting individuals with extreme phenotypic residuals after adjusting for some covariates. However, it remains unclear when local genetic covariates within target genomic regions should be adjusted for to enrich target genetic variants at sampling stage. Heteroscedasticity is ubiquitous in real data but largely ignored in gene mapping for complex traits, especially that of admixed subjects. People pay most of attention on the phenotypic mean instead of phenotypic variance. In this article, we first formulated the powers of ERS with adjusting for the local genetic covariate and ERS to include carriers of the target variant under heteroscedasticity model. In terms of inclusion power, we specified diverse superiorities of EPS and ERS. Second, given budget restriction, our findings provide genomic researchers with novel sequencing extremes strategy combining EPS and ERS to determine the truncation threshold and design cost-effective target sequencing studies. Last, we apply our novel method in real data to identify genetic variants for complex diseases.

1798M
Use of exome sequencing data for the analysis of population structures, inbreeding, and familial linkage. V. Pedergnana1,2, A. Belkadi1, A. Abhyankar1, O. Vincent1, Y. Han1, B. Boisson2, J.L. Casanova1,2, L. Abel1,2, 1) Laboratory of Human Genetics of Infectious Disease, INSERM U1163, University Paris Descartes. Imagine Institute, Paris, France; 2) Wellcome Trust Centre for Human Genetics, Oxford, United Kingdom; 3) New York Genome Center, New York, NY, USA; 4) St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, the Rockefeller University, New York, NY, USA; 5) Howard Hughes Medical Institute, New York, NY, USA.

Numerous methods have been proposed to analyze whole exome sequencing (WES) data in order to discover potential causal variants in Mendelian disorders and in more complex traits. These methods could benefit from additional information such as linkage studies in the study of Mendelian diseases. Population stratification could also be an issue in the analysis of WES data when focusing on complex traits. Both linkage and population structure analyses are classically conducted through genome-wide (GW) SNP arrays. Here, we compared the information yielded by WES data to that provided by SNP array data in terms of analyses usually performed by SNP array data such as principal component analyses (PCA), linkage studies, and homozygosity rate estimation. We analyzed 123 subjects originating from six world regions, including North Africa and Middle East which are regions poorly covered by public database and presenting a high consanguinity rate. A number of quality control (QC) filters were tested and applied to the WES data. Compared to results obtained with SNP array data, we found that WES data provided accurate prediction of population substructure and led to highly reliable estimation of homozygosity rates (correlation > 0.94 with the estimations provided by SNP array). Linkage analyses showed that the linkage information provided by WES data was on average 53% lower than the one provided by SNP array at the GW level, but 58% higher in the coding regions. In conclusion, WES data could be used after appropriate QC filters to perform both analysis and adjust for population substructure, to estimate homozygosity rates, and to perform linkage analyses at least in coding regions.
decrease heterogeneity. Molecular analyses including high-throughput genotyping to assess potential genetic heterogeneity. For instance the study was to recruit affected sibpairs. A total of 785 affected sisters accepted to provide a blood sample, 68% of them were able to provide mammograms. 2012. A total of 1,711 cases and 2,829 sisters and controls were included. Of the study design and of the available resources. Index cases (and their unaffected controls) were recruited. In total, 785 affected sibpairs participated in this study, including 306 affected sisters in cases of breast cancer, 300 unaffected controls and 189 unaffected siblings. SIM, we generated combined independent individual samples and family data. On the other hand, it is difficult to examine the confounding effect on the estimation of phenotype mean effect. We are now extensively analyzing this issue in different settings: sample data contained with two generations, and one-way analysis of variance to determine whether the confounding effect of CR was affected by environmental risks in cohort studies. Tohoku University, Sendai, Japan. The study showed that the probability of survival during gestation ranged from 98.4% at week 20 to 79.0% during the first 24 hours after birth. In the subgroup of patients who survived the first day after birth, only one death was recorded, so by the end of week 96 of follow-up, the group had a survival probability of 91.7% average. Conclusion: It is of importance to extend the network of surveillance and monitoring of birth to other health institutions in the city of Cali, for the purpose of establishing a surveillance system, and that allows for an adequate sampling that provides an accurate overview of the current state of epidemiology, health care and survival of patients with birth defects in the city of Cali.
1802T Assessing the potential impact of low participation in DNA buccal swab collection on the validity of effect estimates. M.M. Jenkins1, J. Reethuis1, H. Razzaghi1, A. Herrling1, M.L. Gallagher1, M.A. Honein1. The National Birth Defects Prevention Study. 1) Centers for Disease Control and Prevention, Atlanta, GA; 2) Oak Ridge Institute for Science and Education, Oak Ridge, TN; 3) The University of North Carolina, Chapel Hill, NC.

Low participation and potential selection bias for specimen collection in gene-environment interaction studies threaten validity. To better understand if effect estimates obtained from participants who submitted DNA buccal swab samples will be impacted by selection bias, we examined the associations between self-reported pregestational diabetes mellitus (PGDM) and four birth defects among families who did and did not submit samples, stratified by maternal race-ethnicity (non-Hispanic white, non-Hispanic black, and Hispanic). Following a telephone interview in a multi-site, population-based, case-control study (National Birth Defects Prevention Study), mothers with estimated dates of delivery between 1997 and 2009 were asked to collect and submit buccal swab samples. Samples were submitted for 52% of interviewed mothers, their infants, or both. For each birth defect, maternal race-ethnic-stratified ratio of odds ratios (ROR) and corresponding 95% confidence intervals (CI) were calculated using logistic regression, comparing the ORs for the PGDM/birth defect association of families who did not submit samples to the ORs for the PGDM/birth defect association of families who did submit samples, while adjusting for maternal age at delivery. Analyses included control families with live born infants who had no major structural birth defects (N=8775), families with pregnancies affected by limb reduction defects (N=958), neural tube defects (N=1615), orofacial clefts (N=3503), or septic heart defects (N=5762). 1% of control mothers reported PGDM versus up to 7% of case mothers. Submitting samples was not associated with reporting PGDM among non-Hispanic black and Hispanic control families but non-Hispanic white control families who submitted samples reported PGDM less often. Statistically significant differences in the effect estimates for the PGDM/birth defect association between families who did or did not submit samples were observed among non-Hispanic black families whose child had an orofacial cleft (ROR=7.58, 95% CI 1.02-56.49, P=0.048) and non-Hispanic white families whose child had a septal heart defect (ROR=0.32, 95% CI 0.13-0.76, P=0.010). We found no evidence to support different effect estimates of PGDM and birth defects in the 10 other analyses among these three race-ethnicities. Further analyses will be completed with other exposures and birth defects.


The Million Veteran Program (MVP) has been proposed to create a large cohort of veterans, one-milllion strong to leverage the excellent electronic medical record resource and combine it with genomic data and self-reported survey data to allow for both original and replication studies on diseases and conditions prevalent in Veterans. With mail-based recruitment, face-to-face enrollment at approximately 50 VA medical centers and broad consent for research, MVP has enrolled over 275,000 Veterans to date. Future models of online enrollment are currently under development. A centralized resource for computation and data analysis called the Genomic Information System for Integrative Science (GenSiS), is under development to facilitate data analysis in a secure environment. Approximately 200,000 samples are currently being genotyped on a customized Affymetrix Axiom chip and smaller subsets of samples are undergoing whole exome and genome sequencing, all through contracted services. The initial focus of analysis will be on mental health conditions such as PTSD, schizophrenia and bipolar disorder, other conditions predominant in the MVP cohort, as well as the exceptionally aged. Approved researchers will access data within the GenSiS environment. Results of the recruitment strategies, overview of demographics and diseases prevalent in the cohort, and updates on the informatics infrastructure as well as genetic analysis will be presented.


It is well known that genotype imputation can boost the power of genome-wide association studies by imputing genotypes not observed in the sample. It facilitates the identification of susceptibility loci for complex diseases. Over the past decade, many imputation algorithms and software have been developed to impute genotype for samples from various populations. In this study, we investigated the genotype imputation accuracy of IMPUTE2 and MaCH-Admix using 148 subjects from multi-ethnic populations (African, Asian, European and Hispanic). The subjects were genotyped in Illumina HumanOmniExpress plus exome array and at the same time whole genome sequenced at high coverage (30X). Instead of masking the genotypes in some specific loci, we directly assessed imputation accuracy by comparing the imputed genotype with the real genotype from sequencing data. In order to address the potential imputation bias, imputation accuracy is based on the concordance for heterogeneous genotype. We found that IMPUTE2 is more computationally efficient than MaCH-Admix. For common variants (MAF>0.01), IMPUTE2 and MaCH-Admix have similar imputation accuracy. However, MaCH-Admix has relatively higher imputation accuracy than IMPUTE2 for the rare variant (MAF<0.01). Both IMPUTE2 and MaCH-Admix show better imputation accuracy in common variants than rare variants. The imputation results also provide valuable information for post-imputation quality control. Optimal threshold for selecting imputed SNPs can be determined for one fixed imputation accuracy. In addition, we also imputed HLA classical alleles and compared with HLA loci called from sequencing data. The imputation accuracy across ethnicities correlated with increased posterior probability of imputed HLA type. In general, we demonstrated that genotype and HLA classical alleles can be accurately imputed for multiple ethnicities using current imputation methods via comparing sequencing data. IMPUTE2 and MaCH-Admix have similar accuracy performance, and IMPUTE2 has better computation efficiency.


It is widely known that population substructure can distort the results of genetic association analysis, especially when rare variants are present such as in next-generation sequencing (NGS) studies. Linear mixed effects model (LMM) were proposed before for adjusting population substructure while testing typically rare variant as a fixed effect. To improve the power of testing multiple rare variants while adjusting for population substructure, one can jointly test them as a random effect in conjunction with the random effect from population substructure. This brings challenge to the current testing methods that are computationally intensive and hard to scale to NGS. We propose a powerful and scalable strategy for testing zero variance component in presence of multiple variance components in LMM. Our approach combines the recent development of exact (restrictive) likelihood ratio test with single variance component and a strategy for reducing LMM with multiple variance components to that with a single one. The performance of the strategy is demonstrated on both synthetic and real data.
1806S  E

Estimating base-calling error rates in next-generation sequencing data using overlapping read pairs. Y. Y. Lo, S. Zöller et al., The BRIDGES Condonum. 1) Biostatistics, University of Michigan, Ann Arbor, MI, 2) Psychiatry, University of Michigan, Ann Arbor, MI.

Genotype calling from short read sequencing data requires accurate estimates of base errors in the reads. Such errors can occur in multiple steps of the sequencing experiment, particularly during the PCR and sequencing steps. Base errors from all steps appear as mismatches to the reference genome on aligned reads, hence the error models typically aggregate errors from all sources. Besides, base information from non-independent sequence reads are often discarded or wrongly incorporated into the base-calling error model intended for independently sequenced reads. Disentangling the sources of base-calling errors would allow more efficient use of all sequencing reads.

Here we aim to distinguish between base-calling errors generated by PCR and by the sequencing machine (machine error), using overlapping read pairs (ORPs) from short PCR fragments in paired-end sequencing. ORPs replicate PCR errors in the overlapping region, while each read in the pair is independently sequenced. Therefore, per genomic position, discordant bases from an ORP are the result of at least one machine error but not PCR error, regardless of the true genotype. On the other hand, if an ORP reports the same base, but different from the reference allele, either the underlying base is a true variant, or the ORP has at least one PCR error or two machine errors. Based on these observations, we derive maximum-likelihood estimators for PCR error and machine error using the counts of discordant and concordant bases from ORPs. We calculate these estimates, as functions of read cycle and base quality score, using an expectation-maximization algorithm.

We apply our method to a subset of data from a whole-genome sequencing study (median 8x coverage, 806 individuals) with 40-60% ORPs. From these ORPs, we estimate that base-calling errors generated by the sequencing machine are approximately twice as common as PCR errors. We quantify the contribution of read cycle and base quality score to these base-calling error estimates. We evaluate the mean and variance of these predicted error rates across individuals. Understanding the contribution of both error sources to base-calling errors allows for more efficient use of ORPs, hence better modeling of errors and improved accuracy in variant calling.

1807M  O


A plethora of linear mixed model based procedures have been developed recently to control Type I error rates in genetic association studies of complex traits by simultaneously accounting for both population and pedigree structure among sample individuals. The general approach taken by these methods is to compute a genome-wide genetic similarity matrix to be included as part of the covariance structure of the phenotype when fitting the model. These mixed model based approaches control genomic inflation genomewide; however, it has recently been shown that they may not adequately control for false positives at SNPs that are unusually differentiated between ancestral populations (Price et al., 2010). In reality, the differentiation in allele frequencies between ancestral populations varies greatly across the genome, and proper calibration of test statistics from these methods should be a concern even at SNPs that are not unusually differentiated. Through the use of a larger number of samples with related individuals, we demonstrate that adjusting for a genetic similarity matrix calculated genome-wide controls Type I error well at SNPs with an average amount of differentiation, but it also leads to a systematic inflation or deflation of test statistics for SNPs that are more or less differentiated. By including ancestry representative principal components as fixed effect covariates in the mean model in addition to the genetic similarity matrix included in the covariance structure, we are able to correct for this issue and obtain well-calibrated test statistics genomewide. Additionally, we demonstrate that in contrast to the original ancestry adjusted genetic similarity matrix may increase power to detect true associations with traits that are correlated with ancestral background. We apply our testing procedure to a variety of blood phenotypes measured on African American and Hispanic women in the Women's Health Initiative and the WHI SHARe (WHI-SHARe) study, and we see increased significance over existing methods at SNPs associated with white blood cell count.

1808T  E


Overcoming systematic miscalibration of linear mixed model test statistics in genetic association studies of complex traits by simultaneously accounting for both population and pedigree structure among sample individuals. The general approach taken by these methods is to compute a genome-wide genetic similarity matrix to be included as part of the covariance structure of the phenotype when fitting the model. These mixed model based approaches control genomic inflation genomewide; however, it has recently been shown that they may not adequately control for false positives at SNPs that are unusually differentiated between ancestral populations (Price et al., 2010). In reality, the differentiation in allele frequencies between ancestral populations varies greatly across the genome, and proper calibration of test statistics from these methods should be a concern even at SNPs that are not unusually differentiated. Through the use of a larger number of samples with related individuals, we demonstrate that adjusting for a genetic similarity matrix calculated genome-wide controls Type I error well at SNPs with an average amount of differentiation, but it also leads to a systematic inflation or deflation of test statistics for SNPs that are more or less differentiated. By including ancestry representative principal components as fixed effect covariates in the mean model in addition to the genetic similarity matrix included in the covariance structure, we are able to correct for this issue and obtain well-calibrated test statistics genomewide. Additionally, we demonstrate that in contrast to the original ancestry adjusted genetic similarity matrix may increase power to detect true associations with traits that are correlated with ancestral background. We apply our testing procedure to a variety of blood phenotypes measured on African American and Hispanic women in the Women's Health Initiative and the WHI SHARe (WHI-SHARe) study, and we see increased significance over existing methods at SNPs associated with white blood cell count.

Evaluating the current standard approximation of probabilistic genotype estimation in the ‘sub-genus’ genotype approach. This approximation can be easily avoided with simple resampling techniques. We recommend that uncertainty and bias correction be considered in all future analyses of uncertain genotype data.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Assessing the power of the Affymetrix Axiom® CEU array for studying rare and low-frequency variants in a European population sample. B. Schormair1,2,3, E. Tilch1,2, E. Dzemara5, B. Czamara4, D. Czamara4, B. Müller-Myhsok5, B. Müller-Myhsok5, A. Albrechtsen1, W. Zheng4, T.L. Edwards1, A. Mottke1,2, J.C. M. Williams2, 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Genetics, Dartmouth College, Hanover, NH; 3) Center for Research on Genomics and Global Health, NIH, Bethesda, MD; 4) Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, TN; 5) Shanghai Cancer Institute, Shanghai, China.

The number of effectively independent tests performed in genome-wide association studies and the corresponding genome-wide significance level varies by population. Therefore, a common p-value threshold may be inappropriate. To assess this, we estimated the number of independent SNPs for all Phase 3 HapMap samples using the LD pruning function in PLINK. We also used an autocorrelation-based approach to verify the HapMap findings, and tested it on 1000 Genomes data to estimate the number of independent tests in whole genome sequences. The number of independent tests is consistent with autocorrelation corresponding to LD pruning at r²=0.3. Using the number of independent tests from both methods, we calculated appropriate population-specific thresholds by Bonferroni correction. African populations had the most stringent thresholds (e.g., 1.49x10^-7 for YRI, r²=0.3), followed by European/Indian populations (3.12x10^-7 for CEU, r²=0.3), and East Asian populations (3.75x10^-7 for JPT, r²=0.3).

Applying our methods to a previously published GWAS of melanoma, we identified two additional SNPs as significant when data were LD-pruned at r²=0.3, and five more when pruned at r²=0.1. These SNPs all mapped to genes previously associated with melanoma; two SNPs were not significant at 5x10^-8, likely representing type 2 errors. Applying our method to a Chinese breast cancer GWAS (r²=0.3) yielded 50 additional significant SNPs, 19 of which were in or near genes previously associated with the phenotype. Our methods indicate that the conventional correction threshold is not appropriate for many association studies, especially in populations that were recently founded.

1810M

Assessing the power of the Affymetrix Axiom® CEU array for studying rare and low-frequency variants in a European population sample. B. Schormair1,2,3, E. Tilch1,2, E. Dzemara5, B. Czamara4, D. Czamara4, B. Müller-Myhsok5, B. Müller-Myhsok5, A. Albrechtsen1, W. Zheng4, T.L. Edwards1, A. Mottke1,2, J.C. M. Williams2, 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Genetics, Dartmouth College, Hanover, NH; 3) Center for Research on Genomics and Global Health, NIH, Bethesda, MD; 4) Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, TN; 5) Shanghai Cancer Institute, Shanghai, China.

The number of effectively independent tests performed in genome-wide association studies and the corresponding genome-wide significance level varies by population. Therefore, a common p-value threshold may be inappropriate. To assess this, we estimated the number of independent SNPs for all Phase 3 HapMap samples using the LD pruning function in PLINK. We also used an autocorrelation-based approach to verify the HapMap findings, and tested it on 1000 Genomes data to estimate the number of independent tests in whole genome sequences. The number of independent tests is consistent with autocorrelation corresponding to LD pruning at r²=0.3. Using the number of independent tests from both methods, we calculated appropriate population-specific thresholds by Bonferroni correction. African populations had the most stringent thresholds (e.g., 1.49x10^-7 for YRI, r²=0.3), followed by European/Indian populations (3.12x10^-7 for CEU, r²=0.3), and East Asian populations (3.75x10^-7 for JPT, r²=0.3).

Applying our methods to a previously published GWAS of melanoma, we identified two additional SNPs as significant when data were LD-pruned at r²=0.3, and five more when pruned at r²=0.1. These SNPs all mapped to genes previously associated with melanoma; two SNPs were not significant at 5x10^-8, likely representing type 2 errors. Applying our method to a Chinese breast cancer GWAS (r²=0.3) yielded 50 additional significant SNPs, 19 of which were in or near genes previously associated with the phenotype. Our methods indicate that the conventional correction threshold is not appropriate for many association studies, especially in populations that were recently founded.

1811T

Assessing the power of the Affymetrix Axiom® CEU array for studying rare and low-frequency variants in a European population sample. B. Schormair1,2,3, E. Tilch1,2, E. Dzemara5, B. Czamara4, D. Czamara4, B. Müller-Myhsok5, B. Müller-Myhsok5, A. Albrechtsen1, W. Zheng4, T.L. Edwards1, A. Mottke1,2, J.C. M. Williams2, 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Genetics, Dartmouth College, Hanover, NH; 3) Center for Research on Genomics and Global Health, NIH, Bethesda, MD; 4) Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, TN; 5) Shanghai Cancer Institute, Shanghai, China.

The number of effectively independent tests performed in genome-wide association studies and the corresponding genome-wide significance level varies by population. Therefore, a common p-value threshold may be inappropriate. To assess this, we estimated the number of independent SNPs for all Phase 3 HapMap samples using the LD pruning function in PLINK. We also used an autocorrelation-based approach to verify the HapMap findings, and tested it on 1000 Genomes data to estimate the number of independent tests in whole genome sequences. The number of independent tests is consistent with autocorrelation corresponding to LD pruning at r²=0.3. Using the number of independent tests from both methods, we calculated appropriate population-specific thresholds by Bonferroni correction. African populations had the most stringent thresholds (e.g., 1.49x10^-7 for YRI, r²=0.3), followed by European/Indian populations (3.12x10^-7 for CEU, r²=0.3), and East Asian populations (3.75x10^-7 for JPT, r²=0.3).

Applying our methods to a previously published GWAS of melanoma, we identified two additional SNPs as significant when data were LD-pruned at r²=0.3, and five more when pruned at r²=0.1. These SNPs all mapped to genes previously associated with melanoma; two SNPs were not significant at 5x10^-8, likely representing type 2 errors. Applying our method to a Chinese breast cancer GWAS (r²=0.3) yielded 50 additional significant SNPs, 19 of which were in or near genes previously associated with the phenotype. Our methods indicate that the conventional correction threshold is not appropriate for many association studies, especially in populations that were recently founded.
1813M  
The Power Comparison of the Haplotype-based Collapsing Tests and the Variant-based Collapsing Tests for Detecting Rare Variants in Pedigrees. W. Guo, Y. Yao. Unit of Statistical Genetics, NIH, Bethesda, MD.  
Abstract Background: Both common and rare genetic variants have been shown to contribute to the etiology of complex diseases. Recent genome-wide association studies (GWAS) have successfully investigated how common variants contribute to the genetic factors associated with common human diseases. However, understanding the impact of rare variants, which are abundant in the human population (one in every 17 bases), remains challenging. A number of statistical tests have been developed to analyze collapsed rare variants identified by association tests. Here, we propose a haplotype-based approach. This work inspired by an existing statistical framework of the pedigree disequilibrium test (PDT), which uses genetic data to assess the effects of variants in general pedigrees. We aim to compare the performance between the haplotype-based approach and the rare variant-based approach for detecting rare causal variants in pedigrees. Results: Extensive simulations in the sequencing setting were carried out to evaluate and compare the haplotype-based approach with the rare variant methods that drew on a more conventional collapsing strategy. As assessed through a variety of scenarios, the haplotype-based pedigree tests had enhanced statistical power compared with the rare variants based pedigree tests when the disease of interest was mainly caused by rare haplotypes (with multiple rare alleles), and vice versa when disease was caused by rare variants acting independently. For most of other situations when disease was caused both by haplotypes with multiple rare alleles and by rare variants with similar effects, these two approaches provided similar power in testing for association. Conclusions: The haplotype-based approach was designed to assess the role of rare and potentially causal haplotypes. The proposed rare variants-based pedigree tests were designed to assess the role of rare and potentially causal variants. This study clearly documented the situations under which either method performs better than the other. All tests have been implemented in a software, which was submitted to the Comprehensive R Archive Network (CRAN) for use as a computer program named rHAPDT.

Currently many large GWAS consortia are expanding to simultaneously examine the joint role of DNA methylation. In this paper we propose a two-step model where in the first step we test for the joint genetic and epigenetic effect on certain phenotype and in the second step we evaluate the potential causal relationship between these variables. We develop statistical approaches that work at gene level for both genotype and methylation, allowing for common units between these two different data types. Powerful kernel machine framework was used in which pair-wise similarity in the trait values between individuals are compared to pairwise similarity in methylation and genotype values for a particular gene, with correspondence suggestive of association. Similarity in methylation and genotype is found by constructing an optimally weighted average of the similarities in methylation and genotype. For a significant gene, we continue to a second step which uses mediation analysis procedure which enables elucidation of the manner in which the different data types work, or do not work, together. We demonstrate through simulations and real data applications that our proposed testing approach often improves power to detect trait associated genes, while protecting type I error. Additionally, the approach can be easily applied to analysis of rare variants and sequencing studies.


Stratification may affect genetic association studies in admixed populations. In Brazil, no study has thoroughly assessed the accuracy of different baseline populations used in controlling stratification. We used simulations to compare the accuracy of 3 baseline populations (African and European from 1000 Genomes Browser and Amerindian from Northern Brazil). A stationary population of 30,000 individuals was evolved for 10 generations with starting SNP frequencies calculated from the original baseline. Migration was allowed between populations at varying rates and with 20% fixation rate/generation. For each individual, 3 variables tracked ancestry backgroud. Each scenario was repeated 100 times to account for simulation variability. Data was analyzed with the Admixture software set to obtain 3 clusters, both considering as baseline population the same one utilized for simulation and using a different baseline population. Correlation coefficients and an extension of Kaplan-Meier curves were employed for comparisons. Fst was calculated to compare the degree of admixture considering 0.15-0.25 as great, 0.05-0.15 as moderate and 0-0.05 as small genetic divergence. The simulation process was implemented in Python with the library SimuPOP, and statistical analyzes were performed with the R environment. Our results indicate that the baseline population composed of individuais of African ancestry (YRI, LWK, ASW), European ancestry (FIN, GBR, IBS, TSI) and Amerindian (Northern Brazil) performed best in comparison to other baseline populations. Median correlation coefficients for this baseline population ranged from 0.97, 0.95 and 0.90 at migration rates of 0.005 (Fst= 0.45), 0.01 (Fst=0.34) and 0.02 (Fst=0.20) respectively. These results were robust when the other populations were used to start the simulation. The difference between the correlations from the 3 clusters using this baseline population presented a narrower interval as compared to other baseline populations tested. As expected, increased migration rates of 0.03 (Fst= 0.11) and 0.04 (Fst=0.06) reduced median correlation coefficients to 0.78 and 0.79, respectively. These results were also supported by Kaplan-Meier curves with a similar tendency. Our results point that one of the baseline population set is superior to other 2 tested and the accuracy decreases as admixture increases. Further studies are necessary to address the impact in a scenario of an association study with stratification.
Genotyping of the UK Biobank resource, a large extensively phenotyped population collection. D. Petkova1, S. Murphy2, C. Bycroft1, C. Freeman1, P. Donnelly1, 1) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 2) UK Biobank, Units 1-4 Spectrum Way, Adwood, Stockport, Cheshire, SK3 0SA, UK.

The UK Biobank project is a large prospective cohort study of 500,000 individuals aged 40-69 from the UK. With its large sample size and wealth of phenotypic information, the UK Biobank is a powerful resource to investigate how genetic and environmental factors influence a wide range of complex diseases. Samples of the participants’ DNA are being genotyped at the Affymetrix Research Service Laboratory on the custom designed UK Biobank Affymetrix Axiom® array with over 800,000 SNPs, including content to provide good genome-wide imputation coverage for variants with frequency > 1%, genome-wide content such as eQTLs, missense and loss of function variants, and content for cancers, metabolic and neurological disorders (including Alzheimer’s). To introduce the UK Biobank resource, we describe results from quality control and analysis of the first 100,000 samples. In addition to standard QC methods, we discuss strategies for quality control of a very large dataset genotyped in multiple batches, with a large proportion of rare variants. We also characterize various aspects of the UK Biobank dataset: population structure, relatedness, gender misidentification and other sample statistics. The extension of the resource to include genome-wide genetic information on each individual is currently underway.

Spurious cryptic relatedness can be induced by population substructure, population admixture and sequencing batch effects. D. Zhang, S. Li, G. T. Wang, S. M. Leal. Center for Statistical Genetics, Baylor College of Medicine, Houston, TX.

It is important to identify cryptically related individuals in population-based association studies, since inclusion of related individuals can increase type I & II errors. Mixed models have been proposed to analyze data when related individuals are included in the study, but these methods can be computationally intensive and do not always sufficiently control type I & II errors. Another option is to only include one individual from a relative pair/group in the analysis. In population-based studies cryptically related individuals should be identified when performing data quality control. However, caution should be used, since population substructure/admixture and sequence data batch effects can cause detection of spurious relatedness. In order to investigate the problem we systematically evaluated the relatedness of 1,092 samples in the 1000 Genomes project and 2,300 African-American (AA) subjects from the NHLBI-Exome Sequencing project (ESP) via two published methods for kinship inference: (i) the PLINK algorithm which is based on identical-by-descent (IBD) statistic under the assumption of homogeneous population, and (ii) the KING-robust algorithm which uses an estimate of the genome-wide average heterozygosity across individuals to compute an estimator of kinship coefficient. For the 1,000 Genomes project we analyzed the data by country of origin, e.g. Japan, continent of origin e.g. Asia and sequencing batch by continent of origin and for the ESP data we analyzed the data for all AA individuals and separately by sequencing batch. We identified spurious cryptically related pairs of samples due to population substructure/admixture and sequencing batch effects with both methods, but the problem was more extreme for PLINK. For example, an excess of 3rd degree relatives were observed for AA individuals, when sequencing batch was not considered and also for 1000 Genomes project data when population substructure was introduced by analyzing the data by continent. Additionally, the kinship coefficients varied depending on how the analysis was performed, e.g. accounting for sequencing batch or not, caused reclassification of individuals, e.g from 1st degree to 2nd degree relatives. In addition to presenting the results of these analyses and showing the severity of the biases in the kinship coefficients, we also demonstrate strategies which can be used to evaluate if individuals are truly cryptically related or spurious relationships have been identified.

Posters: Statistical Genetics and Genetic Epidemiology

1821S Correcting for population stratification in secondary genetic association studies using subsamples. M. C. Babron1,2, S. Benhamou1,2, E. Génin3, R. Kazma2, 1) Inserm U946, Variabilité Génétique et Maladies Humaines, Paris, France; 2) Université Paris-Diderot, Sorbonne Paris-Cité, UMR5-946, Paris, France; 3) Service de Biostatistique et d’Épidémiologie, Gustave Roussy, Villejuif, France; 4) Inserm U1078, Génétique, Génomique Clinique et Épidémiologie, Hôpital Saint-Louis, Paris, France; 5) Centre National de Génotypage, Institut de Génomique, CEA, Evry, France.

Population stratification (PS) is a potential cause of false positive results in genome-wide association studies, when cases and controls are drawn from a population comprising multiple groups with different disease prevalences. To correct for PS, the inclusion of several principal components (PCs) of genome-wide genotypes as covariates has been shown to control for the inflation of association statistics and has become standard procedure. After testing for genetic association with a primary phenotype of interest, secondary hypotheses are often tested using a fraction of the initial sample for which the secondary phenotype is available. In practice, correction for subsample stratification is done using the PCs that were calculated for the full initial sample. This approach makes the assumption that the subsample has a population structure similar to that of the initial sample, used to calculate the PCs. However, this might not always be the case. In fact, the strategy that uses PCs calculated on the initial sample to correct for PS, when testing secondary hypotheses on a subsample, has not yet been evaluated. Here, we assess the robustness of PS correction using PCs calculated with the full initial sample or a subsample that comprises a distribution of subpopulations that is different from the initial population. First, the collection of 5,811 individual genome-wide genotypes (illumina 317k) from 13 European countries is considered as the initial case-control sample. Then, to simulate subsamples with a different population structure, we randomly draw sets of cases and controls with different proportions from extremes. These extremes match up with the North-East and South-West extremes of the space defined by the first two PCs calculated on the initial sample. Simulating 100 replicates of a subsample with 75% of cases and 25% of controls belonging to the North-East quadrant (the rest belonging to the South-West quadrant), the Q-Q plots show a strong type 1 error from with as low as 2 PCs calculated on the initial sample (θPC = 1.3513) without PS correction satisfactorily with as low as 2 PCs calculated on the initial sample (θPC = 1.0019) or on the subsample (θPC = 1.0014). However, the use of PCs calculated on the subsample corrects PS better than PCs calculated on the initial sample for specific variants in genes which are strongly stratified along this axis (LCT). Further ongoing simulations will assess these approaches in different situations of subpopulation structure.


In genome-wide association studies (GWAS), population stratification and relatedness can result in spurious association findings. Principal component analysis (PCA) has been widely used to adjust for population structure. However, PCA does not work well in the presence of familial or cryptic relatedness in addition to population stratification. Linear mixed models have been proposed as an alternative method for correcting for population stratification and relatedness for continuous phenotypes. However, the normality assumption required by linear mixed models is violated for binary phenotypes in case-control studies. We propose to use logistic mixed models to control for both population stratification and familial or cryptic relatedness in GWAS for case-control data. We develop a computationally efficient score test for association analysis using logistic mixed models. Our simulation studies show that our method performs well in controlling type I error rates in the presence of population stratification and relatedness for case-control data. We illustrate our method in a real data example.

Genetic data collected from families with disease affected members are more likely to enrich disease causal variants. Population-based association tests, when applied to family data, need to control for both population structure and family relatedness. Principal components analysis (PCA), which has been widely used to control for population structure in unrelated samples, does not work well in capturing the population structure in family data. We have previously developed a method called LASER, which can estimate individual ancestry in a reference principal component space based on a set of unrelated reference individuals. The LASER method is robust to the presence of family relatedness in the study sample. Based on ancestry coordinates estimated from LASER, we propose to control for population structure and family relatedness using mixed model and generalized estimation equation (GEE) approaches. We adjust the population structure by regressing on ancestry coordinates and account for family relatedness using the kinship matrix to specify the covariance structure in mixed models and GEEs. Using both simulations and empirical data from the Framingham Heart GWAS Studies, we show that this approach performs well for both continuous traits and binary traits. When genome-wide genotypes are available, our approach outperforms the standard PCA method. Furthermore, because LASER can accurately estimate individual ancestry using extremely low coverage sequencing data, our approach can be applied to targeted sequencing or exome sequencing studies where genome-wide data are not available.

Optimal strategies for studying singletons associated with quantitative traits. S. Rashkin, G. Jun, G. Abecasis. Center for Statistical Genetics, University of Michigan, Ann Arbor, MI.

As the focus genetic association studies shift from common variants to rare variants, it is important to determine the optimal strategies for conducting studies of very rare variants. A common argument is that, in order to detect rare variants with high power, it is necessary to deep sequence samples. However, some studies have shown that low-pass sequencing can also detect many rare variants while enabling larger sample sizes. We aim to identify sequencing strategies that maximize power (for a fixed sequencing effort) for studies that explore the role of rare variants in human disease. We specifically focused on variants that appear in a single sequenced sample as they are the most difficult type of variant to identify; designs with sufficient power to detect associations with disease for singletons should also work well for more common variants. Our previous work on binary traits indicates that power to detect association is maximized at 10x coverage. Here, we extend this analysis to quantitative traits.

We assessed the power to detect singletons using a simulated multi-sample caller using different read depths, sample sizes, sequencing error rates, and false positive rates. The power to detect singletons increases as coverage increases, reaching saturation at coverage about 15-20x. We validated these simulations by down-sampling deeply sequenced exome samples and assessing our ability to detect previously called singletons for different sample sizes and read depths.

We performed simulations to assess the power to detect association between the count of singletons in a gene and a quantitative trait, using a simple linear regression framework. We estimated association study power across a range of values for read depth (2-50x), sample size (100-5000), population frequency of singletons (0.001-0.5%), baseline trait value, sequencing error rates (0.001-0.01), and false positive rate for detecting singletons (0.0001-0.05). Preliminary results indicate a similar trend as seen in binary traits: power to detect an association between singletons and a quantitative trait seems to be maximized around 10x and coverage should only be increased beyond this threshold if it does not require a decrease in sample size.

Factors affecting relative telomere length measurements by quantitative PCR. C.L. Dagnall1,2, B.D. Hicks1,2, A.A. Hutchinson1,2, S.M. Gadalla1, L. Mirabell1, B.J. Baille1, J.D. Figueroa1, C. Bodelon1, J. Liu1, J. Hofmann1, S.A. Savage1, M. Yeager1,2, 1) Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI, National Institutes of Health (NIH), Bethesda, MD; 2) Cancer Genomics Research Laboratory, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD.

As variation in telomere length have been associated with risk of cancer, heart disease, and various other illnesses it is important to understand the pre-analytic and analytic factors affecting relative telomere length measured by quantitative PCR (qPCR). We have conducted multiple studies comparing various DNA extraction methods, methods of removing potential residual PCR inhibitors, and sample storage conditions as factors that may influence relative telomere length measurements obtained from qPCR assays which lead to discrepancies within study results. We have confirmed that the DNA extraction method affects the measurement of telomere length defined by the T/S ratio (the concentration of telomeric DNA relative to the concentration of single copy gene sequence). We studied 40 individuals had DNA extracted from the same buffy coat source specimen by three extraction techniques. The dynamic ranges of the T/S ratios from QiaAmp and QIAasympy-extracted DNA samples were lower (0.39-0.87 and 0.29-0.74, respectively) than the Promega ReliaPrep-extracted DNA samples (0.51-1.46). We observed only moderate correlation between paired individuals and different extraction methods (R-squared ranges from 0.32 to 0.46). In addition, we evaluated the differences in telomere length measurements due to inter-plate variability and show that analysis techniques effectively standardize results across assay plates, significantly reducing the average %CV of replicates (from 17.4% to 8.2% in a data set of 170 replicates from 3 plates all extracted using the same extraction technique). Due to these large differences in dynamic range and correlation between extraction methods, it is advisable to only measure relative telomere length when samples have been extracted using the same technique. Eliminating the factors which cause variation in relative telomere length measurements increases the reliability of results when investigating associations between telomere length and various diseases. Results from inhibitor removal methods and sample storage conditions will also be presented.
1827S
Identifying relative pairs within large datasets. Z. Zeng1, D.E. Weeks1,2, W. Chen2, N. Mukhopadhyay1, E. Feingold1,2. 1) Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 2) Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA; 3) Division of Pediatric Pulmonary Medicine, Allergy and Immunology, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA.

When a large number of people from a small geographical region are sequenced (e.g., by a regional healthcare system), it is of interest to identify individuals within the datasets who are closely related. The first step in this process is to identify regions of identity-by-descent (IBD) between pairs of individuals. Shared IBD segments between individuals can be inferred by identity-by-status of streaks of SNPs using dense marker data. Existing methods for detecting IBD segments, such as implemented in BEAGLE, GERMLINE and PLINK, are likelihood-based. Some of these methods are extremely time-consuming (BEAGLE), and some require extra assumptions, such as independence of SNPs (PLINK) or correctly phased genotypes as input (GERMLINE). To strike a balance between computational time, accuracy, and extra requirements on the data, we propose an empirical method for detecting IBD segments and indentifying recombination events in close relative pairs. Our method first determines confident IBD status in small windows and then fills in the unsure gaps between the confident ones to generate the intact IBD segments. We explored combinations of different strategies (large sliding window vs. small fixed window; reference panel vs. no reference panel; windows based on physical distance vs. those based on a fixed number of SNPs) and developed a new algorithm that is computationally efficient and does not require knowledge of putative relationships.

Our algorithm can be used as a building block for detecting relationships large datasets of putatively unrelated individuals or for testing relationships in pedigree datasets. We applied our algorithm on a group of grandparent-grandchild pairs and evaluated its performance by comparing with a relationship-aware algorithm we are developing as well as on artificially synthesized IBD data. This research was funded under NIH (grant number: R01 38979).

1828M
Leveraging Family Structure for the Analysis of Rare Variants in Known Cancer Genes from WES of African American Hereditary Prostate Cancer. C.D. Ford1,2, M.S. DeRycke1,2, S. Middha1,2, J. Karyadi1, D. Schaid1, S.N. Thibodeau1, W.B. Isaacs2, E.A. Ostrander3, J. Stanford4, K.A. Cooney5, J.E. Bailey-Wilson5, J.D. Carpenters6. On behalf of the International Consortium for Prostate Cancer Genetics (ICPCG), 1) Computational and Statistical Genomics Branch, National Human Genome Research Institute/National Institutes of Health, Baltimore, MD; 2) Department of Health Science Research, Mayo Clinic, Rochester, MN; 3) Cancer Genetics Branch, National Human Genome Research Institute/Institute of Clinical Systems Improvement, UnitedHealthcare, MN; 4) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 5) Johns Hopkins University School of Medicine, Johns Hopkins Hospital, Baltimore, MD; 6) Public Health Sciences Division, Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, WA; 7) University of Michigan Comprehensive Cancer Center, Ann Arbor, MI; 8) Integrated Cancer Genomics Division, Translational Genomics Research Institute (TGen), Phoenix, AZ.

Prostate cancer (PRCA) is the second leading cause of cancer death in North American men and it disproportionately affects African American (AA) men, who have higher incidence and mortality rates compared to men without known African ancestry. Although several studies of hereditary PRCA exist resulting in a few significant and reproducible loci associated with an increased risk of PRCA across different study populations (e.g., HOXB13), disentangling the environmental and genetic contributions towards the persistent health disparity in AA with hereditary PRCA remains elusive. As part of the international effort to address this disparity, the African American Hereditary Prostate Cancer Study (AAHPC) was developed as a national collaboration to explore the role of genetics in the causation of hereditary PRCA in AA. AAHPC is in partnership with the International Consortium for Prostate Cancer Genetics (ICPCG), which conducts collaborative studies of PRCA genetics in multiplex families. As part of an ICPCG sequencing study of 539 affected individuals from 366 PRCA pedigrees, we performed whole exome sequencing on 16 AAHPC affected men from 12 pedigrees. In the combined ICPCG AA cohort (N=22), there were 60/141 rare and/or novel coding variants (minor allele frequency, MAF < 5%) in 15 known cancer-associated genes, 35 of which were non-synonymous. Post-variant calling AAHPC quality control was implemented using Golden Helix SVS 8 software with filters set for removal of variants with Read Depth < 10, Quality Score < 20, Quality Score:Read Depth Ratio < 0.5, Call Rate < 0.75. Variants were additionally filtered by MAF threshold of 5%. After QC, 207,741 variants remained for further analysis. In these analyses, we focused on the same aforementioned cancer causing genes. Two AAHPC sequenced families had > 1 affected members (3 per family). Under the dominant model, all affecteds in Family 1 shared 2 variants in two of these genes (25 shared by 2 affecteds). All affecteds in Family 2 shared 9 variants in two different known cancer genes (29 shared by 2 affecteds). Of all 65 shared rare variants in these 2 families, 40 are shared by other ICPCG AA PRCA-affected men. Additional QC is underway to validate these variants and bioinformatic analyses are being used to predict effects of the variants in an effort to unravel the complex genetic heterogeneity of hereditary PRCA in AA.

1829T
Detection of meiotic breakpoints in families using dense genotyping data. N. Mukhopadhyay1, F. Begum1, Z. Zeng1, E. Feingold1,2. 1) Oral Biology, Univ Pittsburgh/Sch Dental Med, Pittsburgh, PA; 2) Dept of Epide- miology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; 3) Dept. of Biostatistics, School of Public Health, Pittsburgh, PA; 4) Dept. of Human Genetics, School of Public Health, Pittsburgh, PA.

Meiotic recombination is important for proper segregation of chromosomes; errors in recombination may result in chromosomal abnormalities and non-disjunction. Both the total number and the pattern of recombination events are known to vary genome-wide and from person to person. Using genome-wide genotype data to detect locations of recombination in pro-bands is the first necessary tool to study recombination. Earlier methods, e.g. CRIMAP, used linkage-style modeling on three-generation families and sparse microsatellite markers to detect recombination events. More recently, methods that look for “streaks” of SNPs showing IBD status on dense GWAS SNP data have been used to score recombination locations in sibs/blocks. We have now developed a new SNP streak method and software to score recombination locations in pedigree types not previously handled, such as half-sibling pedigrees, and pedigrees with one or more ungenotyped individuals. Our method can be used to substantially increase recombination loci in whole sib-pairs for performing GWAS on recombination phenotypes. We discuss ways to handle inherent problems such as uninformative mating combinations, genotyping errors and missing genotypes. Recombination scoring results on real family study datasets will be presented.
1830S
PIX-LRT: A parent-informed test for SNPs on the X chromosome using case-parent triads. C.R. Weinberg1, A. Wise1,2, M. Shi1, 1) Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, Select a Country; 2) Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA.

The X-chromosome is generally understudied in association studies, in part because of the limitations of methodological options. For nuclear-family-based association studies, most current methods extend the transmission disequilibrium test (TDT) to the X-chromosome. We present a new method to study association in case-parent triads: the parent-informed likelihood ratio test (PIX-LRT). Our method takes advantage of parental genotype information and the sex of the affected offspring to increase statistical power to detect an effect. Under a parent exchangeability assumption for the X, if case-parent triads are complete, the parents of affected offspring provide an independent replication sample for findings based on transmissions to the affected offspring. For each offspring sex we combine the parent-level and the offspring-level information to form a likelihood ratio test statistic; we then combine the two to form a composite test statistic. Our method can estimate relative risks under different modes of inheritance or a more general co-dominant model. In triads with missing parental genotypes, the method accounts for missingness with the Expectation-Maximization algorithm. We calculate non-centrality parameters to demonstrate the power gain and robustness of our method compared to alternative methods. We apply PIX-LRT to publicly-available data from an international consortium of genotyped families affected by the birth defect oral cleft and find a strong, internally-replicated signal for a SNP marker related to oral cleft with or without cleft palate.

1831M
Simulation Analysis to Assess Linkage Results of Class III Malocclusion and Human Chromosome 11. L.K AlOthman, L. Mkhuphadhay1, L. Otero2, P.M. Cruz3, J. Tumer4, L.A. Morford2, S. Oliveria4, J.K. Hartsfield4, J.P.5, M. Govil5, 1) Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA; 2) Centro de Investigaciones Odontológicas, Pontificia Universidad Javeriana, Bogota, Colombia; 3) Private Orthodontic Practice, Brasilia, Brazil; 4) Private Orthodontic Practice North Carolina, USA; 5) Division of Orthodontics, University of Kentucky College of Dentistry, Lexington, KY; 6) Hereditary Genomics/Genomics Core Laboratory, Center for Oral Health Research, University of Kentucky College of Dentistry, Lexington, KY; 7) Department of Genetics and Morphology, Universidade De Brasilia, Brasilia, Brazil.

A simulation study was conducted to evaluate multipoint linkage results for chromosome 11 genetic data from two multi-generational family-based cohorts located in South America with a high prevalence of class III malocclusion. Based on previous data implicating a region of chromosome 11 in the Class III phenotype, 4 single nucleotide polymorphisms (SNPs; rs666723, rs758169, rs1368719 and rs12416856) were genotyped within 17 Brazilian families (178 individuals) and 15 Colombian families (248 individuals). The families in each cohort varied greatly in family size, number of affected individuals. Class III affected and unaffected individuals were diagnosed based on cephalometric measurements, models, photographs and/or oral examination. MMLS multipoint HLODs maximized over different levels of heterogeneity and two genetic models (reduced penetrance dominant and recessive) were generated using SimWalk2. To estimate the empirical significance of these multipoint HLODs, 1000 replicates of unlinked genotype data based on real data pedigree structures, affection status and pattern of inheritance were generated and the simulations were performed using SLINK and SIMULATE respectively. These replicates were then analyzed using SimWalk2 with the original maximizing mode of inheritance (MOI). Power was estimated similarly for each cohort by generating 1000 replicates of pedigree data linked to the SNPs with the highest HLOD. The corresponding cohort-specific MOI was used for the power simulation genetic parameters. For the Brazilian cohort: (a) the empirical HLOD for α=0.05 was 0.51, (b) empirical p-value for HLOD=1.84 was <0.001, and (c) power for suggestive linkage (HLOD=2) was 72%. For the Colombian cohort: (a) the empirical HLOD for α=0.05 was 0.28, and (b) an empirical p-value 0.023 for HLOD=0.51 was observed. These simulation results support potential linkage on chromosome 11. The current linkage results are for the MMLS maximized (by definition) over two genetic models; work is underway to compute causal effects for the inter-allelic genetic heterogeneity, i.e. due to complexity of the Colombian cohort pedigrees, power calculations are pending.

1832T
Rare Variant Association Test for Nuclear Families. Z. He1, N. Krumm2, G. Wang3, B. O’Road3, E. Eichler4, S. Leal5, Simons Simplex Sequencing Consortium. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; Center for Statistical Genetics, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Genome Sciences, University of Washington, Seattle, WA; 3) Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR.

The development of next-generation sequencing (NGS) technologies, has led to many population-based complex trait association studies of rare variants. While these studies are prone to type II errors due to power limitation of population stratification. Approaches employing trio (two parents and an affected proband) data such as the transmission disequilibrium test [TDT (Spielman et al. 1993)] can avoid the problems of population stratification. Recently, for rare variants the TDT was extended to include one affected and one unaffected parent [He et al. (2014)]. It was shown that RV-TDT controls type I error and also has considerable power to detect associations. The TDT analyses only employ information on an affected offspring and their parents. When there are additional siblings, it is advantageous to include them in analysis since they also provide association information. We extended the RV-TDT to analyze all types of independent nuclear families (NF) with at least one affected offspring (RV-NF) and test for RV associations using four commonly used methods: CMC (Li & Leal, 2008), WSS (Madsen & Browning 2009), BRV (Morris & Zeggini, 2010) and VT (Price et al, 2010). We demonstrated that for all RV-NF test type I error is well controlled even when there is a high level of population stratification or admixture. The power of the RV-NF test was evaluated using a number of disease models and nuclear pedigree configurations. The RV-NF test is particularly powerful when RVs are enriched in families.

For example using the CMC version of the RV-TDT and RV-NF the power was evaluated by generating variant data for a 1.500bp gene for which the causal RVs [minor allele frequency (MAF)<1%] have an odds ratio of 2.0. The power to detect and identify one affected child and one unaffected child was 65% for 1,000 nuclear families with one affected child and an unaffected child; and 0.65 for 1,000 nuclear families with two affected children. In order to illustrate the application of the RV-NF method, the exome data from 600 autism spectrum disorder nuclear families with one affected child and one unaffected child were analyzed. RV associations with autism were found for several genes. Given the problem of adequately controlling for population stratification and admixture in RV association studies, the capability of analyzing all types of nuclear families and the growing number of nuclear family studies with NGS data, the RV-NF method is extremely beneficial to elucidate the involvement of RVs in the etiology of complex traits.

1833S
The collapsed haplotype pattern method for linkage analysis of next-generation sequencing data. G. T. Wang, D. Zhang, B. Li, H. Dai, S. M. Leal. Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Traditionally, linkage analysis was used to map Mendelian diseases and genes with well-defined regions were sequenced to identify causal variants. Recent advances in next generation sequencing (NGS) make it possible to directly sequence genomes and exomes of individuals with Mendelian diseases and identify causal mutations by filtering variants in an affected individual from the family members. An example is the joint allelic and non-allelic variant frequency, e.g. >0.1% in variant databases. Linkage analysis of SNP data are sometimes used in conjunction with NGS to increase the success of identifying the causal variant. With the reduction in cost of NGS, DNA samples from entire families can be sequenced and linkage analysis can be performed directly using NGS data. Inspired by "burden" tests which are used for complex trait rare variant association studies, we developed the collapsed haplotype pattern (CHP) method to generate markers from sequencing data for linkage analysis. To demonstrate the power of the CHP method compared to analyzing individual variants, we analyzed and performed empirical power calculations using the all-allelic architecture for several known non-syndromic hearing loss genes, i.e. GJB2, SLCO2A4, MYO7A & MYH6. Power analysis demonstrated that the CHP method is substantially more powerful than analyzing individual SNVs in the presence of inter-allelic familial heterogeneity, i.e. families have different pathological variants within a gene or intra-familial heterogeneity e.g. compound heterozygotes. Specifically for an autosomal recessive model with allelic heterogeneity and locus heterogeneity, i.e. 40% of the 12 families carry a different allele, the CHP method to achieve a power of 90% for the SLCO2A4 gene, while analyzing individual SNVs requires >50 families to achieve the same power at a genome-wide significance level of α=0.05. Under the commonly practiced filtering approaches used on sequencing data, the CHP method is powered to take advantage of the involvement of a gene in Mendelian disease etiology. Additionally because it incorporates inheritance information and penetrance models it is less likely than filtering to exclude causal variants in the presents of phenocopies and/or reduced penetrance. We applied the CHP method in parallel to filtering methods to take full advantage of the power of NGS in families. The CHP method is incorporated in the SEQLinkage software which is freely available http://www.bioinformatics.org/seqlink/.
1834M
Identifying rare variants in linkage regions through pedigree-based conditional linkage analysis. C.W. Bartlett1, S. Buyske2, S.L. Wolock1, L.M. Brzustowicz1. 1) Battelle Ctr Mathematical Med, The Research Institute at Nationwide Children's Hospital, Columbus, OH; 2) Department of Statistics, Rutgers University, NJ; 3) Department of Genetics, Rutgers University, NJ.

Linkage analysis is sensitive for detection of rare mutations, even when multiple rare variants in that region are causal in the population, as long as the same single variant segregates with any single family. However, linkage analysis has poor resolution, implicating large genomic regions with many candidate genes and numerous candidate mutations within each candidate gene. Here we develop a method, based on a variation on likelihood ratio testing, to assess if a set of candidate variants in a linkage region underlie or produce the observed linkage. We show that our conditional linkage test can be used to determine which variants are responsible for an observed linkage signal. Our hypothesis that pedigrees contained information useful for conditional linkage was first validated using both and 1m model probability calculations to develop exact analytical demonstrations and to inform downstream modeling choices. Our approach involves calculating a pedigree likelihood stratifying subjects by liability classes determined by a candidate variant “propensity” score. When variants contributing to the propensity score are primarily responsible for the linkage signal, the linkage signal drops once liability classes based on that score are incorporated, demonstrating that the tested variants are required to observe the linkage signal. We use lasso logistic regression of all loci in the gene against affection status to generate subject-level propensity scores and apply a decision rule to create liability classes, though in our simulations we did not observe ambiguous case that required more than 3 liability classes. Using an integrated (i.e., model-averaged) LOD score, we have derived an empirical null distribution for the variability in the integrated LOD with inclusion of liability classes relative to baseline, to allow us to define p-values for significance testing. Power was estimated through simulation over a range of genotyping conditions with varying numbers of causal and non-causal SNPs within a linked region. In general, power increases as the absolute number of non-causal candidate variants decreases and the ratio of causal variant to total test variant numbers increase. In combination with least bioinformatic filtering of variants contributing to the propensity score, our pedigree method for conditional linkage analysis may provide statistical evidence for a specific gene, or gene set, to help reduce the testing in model organisms and systems.

1835T
A general framework for group-wise transmission/disequilibrium tests for identifying rare variant associations. R. Chen1, W. Chen2, X. Zhan3, B. Li1. 1) Molecular Physiology and Biophysics, Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Pediatrics, University of Pittsburgh, Pittsburgh, PA; 3) Department of Clinical Science, University of Texas Southwestern Medical Center, Dallas, TX.

A major focus of current sequencing studies for human genetics is to study rare variants with the aim of determining the role they may play in disease. However, the power of detecting associated rare variants, controlling for population stratification is particularly challenging for rare variants. Transmission/disequilibrium tests (TDT) based on family designs have been proved to be immune to population stratification and admixture, and therefore provide an effective approach to rare variant association studies to eliminate spurious associations. Moreover, group-wise strategies, originally proposed for population-based studies, have been extended to family designs to collectively analyze multiple rare variants as a single unit. In this study we describe a general framework for group-wise TDT (gTDT) and develop an efficient tool that accounts for relatedness among family members. As ignoring the between-family component may lead to power loss, we further adopted a two-stage test to enhance power: at the first stage, we construct a variance-component score test using the between-family component to identify top hits; at the second stage, we follow up these top hits by constructing the variance-component score test on the independent within-family component. Our method maintains power when a region harbors variants acting in a manner that is not fully penetrant. Further illustrate the approach using sequencing data from the SardiNIA Medical Sequencing Discovery Project.

1836S
Rare Variant Association Analysis of Quantitative Traits in Pedigrees of Arbitrary Size and Structure. Y. Jiang1, K. Connelly2, M. Epstein3. 1) Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA; 2) Department of Human Genetics, Emory University, Atlanta, GA.

Family-based rare-variants sequencing studies have attractive characteristics: they can be more powerful compared to case-control and population-based studies due to increased genetic load and can further enable the implementation of rare-variant association tests that, by design, are robust to confounding due to population stratification. However, current methods for family-based studies mainly focus on nuclear families or parent-offspring trios while ignoring the information provided by other types of relatives. With this in mind, we propose a powerful rare-variant test for analysis of quantitative traits in extended pedigrees that can accommodate any family size, and is robust to population stratification. For each non-founder, our method partitions rare variants in a region of interest into two components: a between-family component (sensitive to population stratification but less sensitive to genotyping error) based on information from all ancestors as well as an orthogonal within-family component (which is robust to population stratification). We then test for association between the robust within-family component and the trait of interest using a variance-component score test that accounts for relatedness among family members. As ignoring the between-family component may lead to power loss, we further adopted a two-stage test to enhance power: at the first stage, we construct a variance-component score test using the between-family component to identify top hits; at the second stage, we follow up these top hits by constructing the variance-component score test on the independent within-family component. Our method maintains power when a region harbors variants acting in a manner that is not fully penetrant. Further illustrate the approach using sequencing data from the SardiNIA Medical Sequencing Discovery Project.

1837M
Dissecting the Genetic Architecture of Longevity with Millions of Individuals. J. Kaplanis1, B. Marks1, M. Gershovits2, M. Sheikhi2, A. Price1, D. MacArthur4, Y. Eritch1. 1) Whitehead Institute for Biomedical Research, Cambridge, MA; 2) Harvard School for Public Health, Boston, MA; 3) Broad Institute of Harvard and MIT, Cambridge MA; 4) Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA.

Understanding the genetic architecture of complex traits is one of the top missions of human genetics. Emerging lines of studies have highlighted the entangled etiologies of these traits, which can include epistasis, parent-of-origin effects, sex and age interactions, and environmental risk factors. To conduct robust genetic epidemiological analysis, statistical models require sampling substantial amount of data from large families. However, the recruitment of large cohorts of extended kinships is both logistically challenging and cost-prohibitive. Here, we present a big data strategy to address this challenge: harnessing existing, free, and massive Web 2.0 social network resources to trace the aggregation of complex traits in extremely large families. We collected millions of public profiles from Geni.com, the world’s largest genealogy-driven social network. Using this information, we constructed a single pedigree of 13 million individuals spanning many generations up to the 15th century. In addition, Natural Language Processing was used to convert genealogical information into birth and death locations to obtain a proxy for environmental factors. We obtained multiple of phenotypes from this resource including longevity, fertility, and migration patterns. This dataset provides a wide range of kinships for familial aggregation studies. We compiled this data in a new resource called FamiLinx which will be free available for the entire scientific community. We will show the utility of FamiLinx for heritability estimates across distant relatives to disentangle analysis of epistasis, parent-of-origin, and shared environments.
Adaptive Combination of P-values for Family-based Association Testing with Sequence Data. W. Lin, Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan.

Family-based study design will play a key role in identifying rare causal variants, because rare causal variants can be enriched in families with multiple affected subjects. Furthermore, different from population-based studies, family studies are robust to any bias induced by population substructure. It is well known that rare causal variants are difficult to detect with single-locus tests. Therefore, burden tests and non-burden tests have been developed to combine signals of multiple variants in a chromosomal region or a functional unit. This inevitably incorporates some neutral variants into the test statistics, which can dilute the power of statistical methods. To guard against the noise caused by neutral variants, we here propose an ‘adaptive combination of P-values method’ (abbreviated as ‘ADA’). This method combines per-site P-values of variants that are more likely to be causal. Variants with large P-values (which are more likely to be neutral variants) are discarded from the combined statistic. In addition to performing extensive simulation studies, we applied these tests to the Genetic Analysis Workshop 17 data sets, where real sequence data were generated according to the 1000 Genomes Project. We show that this ADA method outperforms the burden tests and non-burden tests when the percentage of causal variants is small. ADA is recommended for its potential to guard against the noise induced by neutral variants.

Exome sequencing in an isolated population reveals multiple rare variants affecting both high-density lipoprotein cholesterol and the levels of certain blood metabolites. E.M. van Leeuwen1, A. Demirkan1,2, N. Amin1, A. Verhoeven1, A. Meissner1, R.W.W. Brouwer1, W.F.J. van Ijcken1, A. Isaacs1, T. Hankemeier6,8, K. Willems van Dijk6,7, C.M. van Duijn1. 1 Epidemiology, Erasmus MC, Rotterdam, Zuid-Holland, Netherlands; 2 Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands; 3 Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands; 4 Center for Biomics, Erasmus MC, Rotterdam, the Netherlands; 5 Leiden Academic Center for Drug Research, Division of Analytical Biosciences, Leiden University, Leiden, The Netherlands; 6 The Netherlands Metabolomics Centre, Leiden University, Leiden, the Netherlands; 7 Department of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands.

Finding rare variants implicated in complex diseases has proven to be difficult. Large family-based studies in isolated populations yield increased power for identifying rare coding variants among others due to underlying founder effects. We explored this in the ERF study, a genetic isolate from the Netherlands. We tested if rare variants coding for high density lipoprotein cholesterol (HDL-C) also influence other metabolic traits which are also known to be under strong genetic control (30-80%). Participants of the ERF study include 3000 related individuals selected on the basis of genealogy. Exome sequencing was conducted in 1300 individuals. HDL-C was determined by the enzymatic method, whereas 400 metabolites (glycerophospholipids, triglycerides, amino-acids, acyl-carnitines, small metabolic compounds) were determined by NMR or MS. We tested around 550,000 coding variants within the exomes of 1300 samples for associations with the HDL-C level, followed by segregation analysis in families. We found both common (minor allele frequency (MAF) > 0.1) and rare variants (MAF < 0.02) to be associated with HDL-C. The common variants are all located within the CETP region, a region which is known to be associated with HDL-C level. We found nine rare variants to be significant (p-value < 2.57 x 10^-6). Bonferroni corrected for the number of unique genes tested) after adjusting for multiple testing and family structure. Each of the rare variants segregated within families. Of note is that all carriers of the rare variants have a (extremely) high HDL-C, and thus the rare variant may be of interest for therapeutic development. Each rare variant was also associated with several metabolites (p-value < 3.09 x 10^-5) which may help us understand the mechanisms involved. This study shows that combining next-generation sequencing with metabolomics within large family studies can help us unravelling the process from variant to biological processes influencing a clinical measurement. The fact that we found variants associated with high HDL-C levels opens up opportunities to translational research for medication.

Efficient gene-gene interaction test for discordant sib pairs in genome-wide association studies. R. Chung1, P. Song2, Y. Wang2, P. Lin2, W. Tsai1. 1) National Health Research Institue, Zhunan, Taiwan; 2) National Tsing-Hua University, Hsinchu, Taiwan.

Genome-wide association studies (GWAS) have been a popular strategy to identify single nucleotide polymorphisms (SNPs) associated with complex diseases. Gene-gene interactions may play an important role in complex diseases. Therefore, many methods have been developed for gene-gene interactions for GWAS. However, most of the methods are developed for case-control studies, and only a few methods are available for family-based interaction analysis for complex diseases. Moreover, family-based interaction methods are computationally intensive, which are not applicable to genome-wide interaction studies, which test all possible pairs of SNPs across the genome. We propose an efficient family-based gene-gene interaction test, which compares the difference in log odds ratios for a pair of SNPs between cases and controls. Extended from the sib TDT (S-TDT), cases and controls are defined as genotypes in affected and unaffected sibs in discordant sib pairs (DSP), respectively. A multivariate hypergeometric distribution is used to calculate the variance and covariance for the test statistic. We used simulations to demonstrate that the proposed test has correct type I error rates under different scenarios, such as different sample sizes and minor allele frequencies. We also performed power studies to evaluate the power for the proposed test with other family-based interaction methods under different scenarios. The results suggested that the proposed test is a valid test and has comparable power to the existing tests. We also used simulations to show that the proposed test can run 20 times faster than the regression-based interaction test. Finally, we applied the test to a family GWAS study for hypertension and several promising SNP-SNP interactions were identified for the disease.
1842S
Robust and powerful family test for rare variant association, K. Lin1, 2, S. Zöllner1, 2. 1) Biostatistics, University of Michigan, Ann Arbor, MI; 2) Psychiatry, University of Michigan, Ann Arbor, MI.

Modern sequencing technology allows investigating the impact of rare variant on complex disease. However, in a conventional case-control design, large samples are needed to achieve sufficient power for testing association of suspected loci to complex disease. In such large samples, population structure can cause spurious signals. One approach to overcome low power and population stratification is family-based study design since the count of risk variants can be enriched in a family containing multiple affected members with the same genetic background. Traditional methods such as the TDT and the FBAT are not very powerful when applied to rare variants. Thus, there is a need to develop a method to directly incorporate the advantage of family-based design to discover association between rare variants and disease phenotypes.

We propose a novel test for a sample of families. In each family, we first determine the chromosomes shared IBD/non-IBD at the locus of interest for all family members. Conditional on the sharing status, we evaluate the number of risk alleles shared IBD among affected family members. Since risk variants are more likely to be shared among affected members, we test for an excess of shared rare variants among affected members by comparing the observed count of shared variants to the expected count of shared variants under the null. This expectation can be calculated conditional on the founders' haplotypes. The proposed test is robust to population stratification since each family is a matched unit. We evaluate the power of this approach analytically as well as with computer simulations using a general model for the effect size of rare risk variants. We compare our method with existing methods and present several novel kernels for rare variant association in the presence of population structure.

Considering different models from rare to common variants, we show that the proposed approach is especially powerful for rare variants as compared to the conventional case-control design.

1843M

Technological advances in high-throughput sequencing platforms have made it possible to extend genome-wide association studies to rare variants by exome, whole-genome, and targeted sequencing. Custom genotyping arrays such as the ExomeChip have been utilized for their low cost in genotyping rare variants. Family data can enrich rare variant association analyses, and allow the use of well characterized pedigreed data that were collected previously in linkage and association studies. Population structure (such as in admixed populations) presents great challenges to statistical methods that analyze datasets with rare variants. The standard approach of including principal components of ancestry as covariates in the regression of a trait may not be sufficient to eliminate the effect of population structure on the rare variant association.

We extend the sequence kernel association test (SKAT) to family studies, and present several novel kernels for rare variant association in the presence of population structure: 1) weighted linear kernel (to confirm the existing results from SKAT); 2) distance kernel that adjusts pedigree structure; 3) distance kernel that adjusts pedigree and local population structure; and 4) distance kernel that adjusts pedigree and local population structure. In addition to including principal components of ancestry as covariates in the linear regression model, we standardize the distance between any pair of individuals at a gene by dividing their global or local distance across the genome. Extensive computer simulations show that our distance kernels can effectively eliminate population stratification as well as increase the power of association test. Finally we apply our methods to a data set of rare variants genotyped in African American and Hispanic populations.

1844T

Discovering regions of pairs of genomes that are identical by descent (IBD) is an important part of many genetic analyses. Given a large sample of 100K people, it is prohibitive to be able to test every possible pair of IBD segments with a haplotype model. However, RefinedIBD is not computationally efficient, even though it uses the GERM-LINE algorithm to discover the putative IBD segments. TIMBER is an algorithm that decides if a putative IBD segment discovered by GERM-LINE has a sufficient level of evidence for IBD to be retained. TIMBER uses GERM-LINE-discovered putative IBD segments to filter the very same segments. Within GERM-LINE, the genome is split up into non-overlapping windows, where TIMBER calculates a weight for each of these windows. Each weight provides the relative level of evidence of IBD from that window in a GERM-LINE-discovered putative IBD segment. TIMBER’s main action is to down-weight windows of the genome that show excessively high degree of putative IBD segments. TIMBER is made possible given the large count of putative IBD segments that are discovered in running GERM-LINE on a large data set (over 100K people). We have evidence that TIMBER is a very useful IBD filter from both real and simulated data given a large data set of over 100K people. While TIMBER is not as accurate as a method such as RefinedIBD, it provides a significant improvement running GERM-LINE on its own and it is computationally efficient for a large data set.

1845S
Statistical Tests for Co-Segregation of Genetic Variants with Disease in Pedigrees, D. Schaid1, S. McDowell1, J. Sinnwell2, L. Cannon-Albright2, C. Teerlink1, J. Stanford1, E. Ostrander2, W. Isaacs3, J. Xu4, K. Cooney5, E. Lange6, J. Schleyk7, J. Capiten10, I. Powell11, J. Bailey-Wilson12, O. Cussenot13, G. Cancel-Tassin13, G. Gilles14, L. FitzGerald14, C. Maier15, A. Whittemore16, C.L. Hsieh17, F. Wiklund18, W. Caton19, W. Foulkes20, O. Mandali1, R. Eeles12, Z. Kote-Jarara1, S. Thibodeau3, 1) Dept Hlth Sci Res, Mayo Clinic, Rochester, MN; 2) Dept Internal Medicine, Univ Utah School of Medicine, Salt Lake City, Utah; 3) Fred Hutchinson Cancer Research Center, Seattle, WA; 4) National Human Genome Research Institute, Bethesda, MD; 5) Johns Hopkins Hospital, Department of Urology, Baltimore, MD; 6) Center for Human Genetics, Wake Forest University, Winston-Salem, NC; 7) Departments of Internal Medicine and Urology, University of Michigan Medical School, Ann Arbor, MI; 8) Department of Genetics, University of North Carolina, Chapel Hill, NC; 9) Medical Biochemistry and Genetics, University of Turku, Finland; 10) Integrated Cancer Genomics Division, The Translational Genomics Research Institute, Phoenix, AZ; 11) School of Public Health, University of Minnesota, Minneapolis, MN; 12) Statistical Genetics Section, National Human Genome Research Institute, Bethesda, MD; 13) Service Urology-Batiment Gabriel, Hospital Tenon, Paris, France; 14) Cancer Epidemiology Centre, Cancer Council Victoria, Australia; 15) Department of Urology, University of Ulm, Ulm, Germany; 16) Dept. Health Research and Policy, Stanford Univ, Stanford, CA; 17) Department of Urology, University of Southern California, Los Angeles, CA; 18) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 19) Department of Medicine, University of Utah, Salt Lake City, UT; 20) Dept. Of Oncology and Human Genetics, Montreal General Hospital, Montreal QC, Canada; 21) Department of Genetics, LSU Health Sciences Center, New Orleans, LA; 22) Genetics and Epidemiology, Institute of Cancer Research, Sutton Surrey, UK; 23) Department of lab Medicine/Pathology, Mayo Clinic, Rochester, MN.

Sequencing studies for rare variants in pedigrees often evaluate whether affected pedigree members carry the same variant - an evaluation of co-segregation of variant with disease. This approach, however, requires considerable penetrability, because pedigrees without affected members are not informative. That is, statistical assessment of co-segregation requires comparing the observed number of affected carriers with random expectation, where expectation is based on Mendelian segmentation of variants, conditioned on at least one affected parent per variant.

We developed new statistical methods to test for co-segregation of genetic variants with disease, for studies that genotype only affected pedigree members. Our methods, a simple comparison of observed and expected carrier counts within each family, provided a method and score (QLS) statistic on quasi-likelihood score (QLS) statistics, provide was to perform gene-level tests, important for situations where different pedigrees have different segregating variants, yet all within the same gene. Our gene-level test allows use of weights for different variants, such as weights based on Mendelian penetrability or based on variant function. Furthermore, our QLS method accounts for correlations in the data, such as correlations from related subjects, or correlation among multiple variants (e.g., linkage disequilibrium). We will present simulations to illustrate the power of our approach, as well as results from application to a whole exome scan of familial prostate cancer.
**1846M**

Cross pedigree shared ancestry reveals rare, disease-causing variants in the presence of locus heterogeneity. H. J. Abel, MA. Province. Genetics, Washington University School of Medicine, St. Louis, MO.

Currently, there is great interest in the use of family studies to identify rare variants underlying complex disease. Attempts at fine mapping, however, are often confounded by locus heterogeneity, which can produce noisy and poorly localized linkage signals, as well as by an abundance of rare variants, which frequently segregate with phenotype by chance. Because rare variants shared across pedigrees are likely to be of recent origin, we have developed an approach to leverage identity-by-descent (IBD) sharing between pedigree founders in order to better localize linkage signals in the presence of heterogeneity. Our method relies on the lengths of segments shared identically-by-state (IBS) across pedigrees, using as test statistic a score based on the sum of maximal pairwise shared lengths at each locus, and optimized over all genotyped pedigree members. Use of unphased IBS renders it both computationally efficient, so that pedigree-based permutation tests assessing significance are tractable, and robust to genotyping/sequencing and haplotype-phase switch errors. Further, our approach produces a cross-family metric to facilitate local clustering of families near IBD regions which, in simulations, accurately recovers ancestral relationships and permits stratification by recent shared ancestry. We have evaluated the performance of our method using a combination of coalescent simulation of founder individuals, followed by gene-dropping onto pedigrees. Under a variety of scenarios, with rare causal variants (population MAF<0.01) and modest effect sizes (OR=5-10), our approach achieves 60-90% power, and is able to detect shared ancestral segments harboring rare causal variants when multipoint linkage and rare-variant burden tests fail.

**1847T**

Increasing Power to Detect Rare Variant Associations by Integrating Linkage Data: A Bayesian Approach. S. Lutz, A.L. Peljto, T.E. Fingerlin.

Biostatistics, University of Colorado, Denver, CO.

Large resequencing studies often combine familial cases of disease from linkage studies with sporadic cases to achieve the large sample sizes necessary for detecting low-frequency alleles. In general, these studies include only a single case from each family. As a result, family-based association tests cannot be applied and standard tests for rare variants generally ignore family-specific allele-sharing information that may be available, resulting in an inefficient use of resources and loss of power for this common study design. We propose a Bayesian approach for testing associations between groups of variants and a phenotype that uses family-specific linkage information in the prior to increase power to detect associations with rare variants and the phenotype of interest. Our method directly combines family-specific linkage and rare variant data by giving more weight in the prior to familial cases with evidence for excess allele sharing near the position of the variant being tested. To assess the performance of our method, we simulated genetic data for familial cases, singleton cases, and controls under various genetic models. We then compared our Bayesian approach that incorporates family-specific linkage information to the sequence kernel association test (SKAT) and several burden tests. We show that our method has increased power compared to the standard methods (SKAT and burden tests), which do not incorporate the family-specific allele-sharing information for several underlying genetic models. Our method provides a framework for tests of association with rare variants that are more statistically efficient when family-specific linkage information is available on at least a subset of cases, while leveraging the resources that have been devoted to recruiting, describing, and genotyping familial linkage cohorts.

**1848S**

Impact of screening for precancerous lesions on family-based genetic association tests: an example of colorectal polyps and cancer. S. L. Stenzen1, J.C. Figueiredo2, V.K. Cortessis2, D.C. Thomas1. 1) Dept of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 2) Dept of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, CA.

Estimates of the association between genetic variants and cancer risk may be biased by screening and subsequent removal of detected precancerous lesions. We conducted a simulation study using a discordant sibship case-control design assuming various screening and prevention parameters and determined the association between genetic variation and risk of colorectal cancer (CRC) under four analysis strategies: 1) CRC cases vs. all non-cases with no adjustment for polyp history; 2) same as (1) with adjustment; 3) CRC vs. polyp-free non-cases; 4) CRC or polyp cases vs. neither. We considered genes involved in polyp development (G1), in cancer risk but not polyps (G2), and in progression from polyps to carcinoma (G3). We then applied these strategies to a family-based case-control study of folate-related genes and risk of CRC. In our simulation, estimates of the association between G2 or G3 and CRC comparing CRC cases against all non-cases without adjustment for polyp history were not biased, but estimates were underestimated for G1 if detection of a polyp led to reduction in CRC risk. Analysis strategies 2 through 4 led to substantial biases, with directions varying across strategy and gene type. Type I errors were correct for all approaches, but strategy 1 was the most powerful. In our application, estimates of relative risk differed for selected variants when controls with polyps were excluded or when individuals with polyps were treated as cases. Although estimates of gene-CRC associations may be biased due to detection of precancerous lesions, the degree of bias appears to differ only modestly by analytic strategy in applied studies with population-based recruiting. The impact may be more unpredictable in studies that are not population-based.

**1849M**

TITLE: Familial aggregation of blood pressure in Ramadasia population of north-west Punjab. R. Kumari1, B. Doza2. 1) GHS MARI KAMBO KE, TARANTAN. 2) GURU NANAK DEV UNIVERSITY, AMRITSAR.

In India, the burden of hypertension has increased many folds in recent times due to increase of westernized diets, life styles and the increasing mean age of populations. Therefore, it is essential to understand how blood pressure is influenced by familial factors (both genetic and environmental) and how these contribute to the risk for cardiovascular diseases. The major objective of the study was to describe the genetic heritability and familial household contribution to the phenotypic variation for cardiovascular risk factors especially for SBP and DBP. Three generations i.e. grandparental, parental and offspring generation, aged 7 years and above, from Ramadasia, a scheduled caste population, were selected for the study. The questionnaire grossly included socio-economic lifestyle variables, physiometric measurements and anthropometric measurements. A total of 600 families with 1827 individuals were sampled. Results demonstrated familial aggregation of CVD risk factors. The brother-sister, father-offspring (male/female), mother-offspring (male/female) correlations were all greater than spouse correlations (brother-sister: 34% for SBP, 31% for DBP, 38% for BMI; father-offspring: 33% for SBP, 30% for DBP, 21% for BMI; mother-offspring: 21% for SBP, 20% for DBP, 15% for BMI; mother-offspring: 17% for SBP, 15% for DBP, 23% for BMI; female offspring: 20% for SBP, 18% for DBP, 29% for BMI; spouse: 4% for SBP, 5% for DBP, 9% for BMI). The magnitudes of these correlations for SBP, DBP and BMI were lower in grandparent-offspring. Therefore, the higher estimates of correlations among parent-offspring generations and siblings suggested the genetic closeness. But the correlation between siblings was even larger than correlation between parent-offspring and others. Familial aggregation of blood pressure was observed to be largely due to genes rather than familial environment from the heritability explained by the genetic variation in blood pressure. Therefore, the present study has a great potential for the study of quantitative genetic variations within and between the generations. The estimates of heritabilities through variance-components approach has produced a wide range of heritabilities (10%-98.7%) for selected anthropometric and phsyiometric phenotypes between different combinations of the three generations.
Genetic variants on the X chromosome could potentially play an important role in some complex traits. However, statistical methods for genetic association studies have primarily been developed for variants on the autosomal chromosomes with significantly less attention given to the X chromosome. Mixed linear models (MLMs) have recently emerged as a powerful and effective method of choice for association mapping in the presence of sample structure such as population structure and/or relatedness. Existing MLM approaches are not directly applicable to association testing on the X chromosome. Here, we propose a MLM approach for genetic association testing on the X chromosome in samples with related individuals. In simulation studies with both unrelated and related individuals, we demonstrate that our proposed MLM has the correct type I error for association testing with SNPs on the X chromosome. We also show that existing mixed model approaches are not properly calibrated and have inflated type I error for X-linked traits. We further demonstrate that our proposed MLM method for the X has high power to detect genetic association with X-linked markers.

lincRNA-mRNA transcriptional regulatory co-expression network from RNA-seq data in 624 Sardinian individuals. P. Forabosco1, M. Pala1,2,6, M. Marongiu1, A. Mulas, R. Cusano1, F. Crobu1, F. Reineri2, R. Berutti3,4,5, M.G. Piras5, C. Jones3, D. Schlessinger6, G. Abecasis7, A. Angus1, S. Sanna1, S.B. Montgomery2,7, F. Cucca1,2,6. 1) Research Institute of Genetics and Biomedicine, CNR, Cagliari, Italy; 2) Pathology and Genetics Department, Stanford University, Stanford, CA, USA; 3) CRS4, Advanced Genomic Computing Technology, Pula, Italy; 4) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA; 5) Laboratory of Genetics, National Institute on Aging, Baltimore, MD, USA; 6) Department of Biomedical Science, University of Sassari, Sassari, Italy; 7) co-senior authors.

Long non-coding intergenic RNAs (lincRNAs) are thought to play important roles in regulating gene expression, but the exact function of the majority of them remains unknown. RNA-seq represents an effective tool for high-throughput analysis of lincRNA expression, as it provides a far more precise and accurate measurement of transcript levels than other methods. Bioinformatics approaches might be useful in providing insight into lincRNA functions. In particular, co-expression network analysis of lincRNA and mRNA transcripts enables a systematic and global interpretation of expression data, by assessing co-expression patterns (through correlation matrices), and by identifying functionally relevant modules. Annotation of lincRNAs probable functions can be predicted and linked to biological pathways based on module sharing, i.e. by assigning functions according to the functional enrichment of coding transcripts in the same module. Analysis of co-expression of lincRNAs with mRNA can therefore help prediction of their functional role as a foundation for further mechanism studies. Here we present the construction of a regulatory network comprising both coding and long non-coding transcripts. Gene expression levels have been quantified with RNA-seq from Poly(A)+ PBMC samples in 624 Sardinian individuals. We construct a co-expression network, where nodes are transcripts and edges represent co-expression strength, using the Weighted Gene Co-expression Network Analysis package (WGCNA) implemented in R. We inspect the network in detail in terms of connectivity and density patterns, network hubs, and module identification. Compared to microarray technologies, RNA-seq data results in more interconnections among transcripts (e.g. higher mean connectivity is observed in a subsample of 188 unrelated subjects when using the suggested soft thresholding power in the adjacency transformation in order to meet scale free topology). Using gene significance (a measure correlated with intramodular connectivity) encoding lincRNA status, we identify modules where lincRNAs are most relevant. This study is the first to explore genome-wide lincRNA expression and co-expression with mRNAs using RNA-seq technology.

Random forest for genetic analysis: Integrating the X chromosome. G. Jenkins, J. Biernacka, S. Winham. Mayo Clinic, Rochester, MN. Despite the fact that it is important to play a major role in many Mendelian disorders, the X chromosome is routinely excluded from genome-wide association studies. Data mining methods such as Random Forest (RF) have been proposed to investigate complex genetic models involving many variants. However, for traits associated with sex, inclusion of X chromosome SNPs in RF analysis yields biased results, as we illustrate using simulations and an example from a case-control study of alcohol dependence. We propose three extensions of the RF algorithm to include X chromosome variants. However, for traits associated with sex, inclusion of X chromosome SNPs, based on (1) the principle of X chromosome inactivation (XCI), (2) Beta (TPB) prior that scales with the sparsity in the data. To address the latent biological and technical confounders, we modeled each factor and loading as a mixture of sparse and dense components, where dense vectors capture technical or biological confounding when estimating the sparse co-regulated gene sets. We used simulated data to quantify and compare our methods relative to other biclustering methods. We analyzed gene expression data from the GTEx project, extracting large numbers of sparse groups of co-regulated genes. We discuss the implications of the recovered gene networks in revealing important biological, and potentially disease-related, mechanisms.
1854S

Integrative Metabolomics of Asthma Severity using Bayesian Networks. J. Lasky-Su1, M. McGeeachie1, W. Qiu1, J. Savage1, A. Dahlin1, A. AlGarawi1, D. Croteau-Chonka1, J. Sordillo1, A. Wu1, E. Chen1, D. DeMeo1, A. Litonjua1, C. Clish2, B. Raby1, S. Weiss1. 1) Brigham & Women’s Hosp, Boston, MA; 2) Broad Institute, Cambridge, MA.

Asthma is a heritable disease with both environmental and genetic components. Although a number of molecular determinants have been identified, much remains to be understood about how these variants impact the disease. Metabolites have the distinct advantage of being more proximal markers of disease processes than transcriptional, translational or post-translational changes. Therefore the integration of metabolomics, with genetics, genomics, and epigenetics may more effectively link disease outcomes with genetic determinants. We generated lipidomic pilot data using liquid chromatography tandem mass spectrometry (LC-MS) using plasma samples from 20 Caucasian individuals from the Asthma BioRepository for Integrative Genomic Exploration (Asthma BRIDGE), an open-access repository from 1,435 asthmatics. A total 64 metabolites were generated of which 26 were identifiable. We used two asthma severity phenotypes, a binary measure, “hospitalized overnight for asthma” and a continuous, composite measure, “acute asthma control.” We performed the following analyses for all 64 metabolites: 1) GWAS analyses for each metabolite and all genetic variants; 2) regression analyses for each metabolite and genome-wide gene expression variants; 3) regression analyses for each metabolite and the genome-wide CpG methylation data. We then identified the top variants from each of these analyses including the top SNPs, mRNA transcripts, and CpG sites that were associated with any metabolites. We then applied a Conditional Gaussian Bayesian Network analysis to the top findings from these analyses using the main effect and all SNP interaction terms within each gene pair to generate a gene-gene interaction P-value. Nominally significant findings from GenSalt were evaluated for replication using publically available data from the Multi-Ethnic Study of Atherosclerosis (MESA). Findings from both studies were combined using meta-analysis. Statistical significance was determined by nominal significance in both GenSalt and MESA studies, along with significance after Bonferroni correction to meta-analysis results (α-threshold=0.0025 and 0.00026 for single-gene and gene-gene interaction score, respectively). In single-gene determinants, CACNA1C was associated with long term systolic BP (P=0.0014), while ADRAD1A was associated with long term diastolic BP (P=0.0004). In addition, multiple gene-gene interactions influencing BP phenotypes were identified, including: CACNA1AxADRA1A, CACNA1xCACNA1C, CACNA1CxCACNA1C, CACNA1CxADRA1A, CACNA1CxADRA1D, CACNA1DxADRA1D, CACNA1DxSLC6A2 for both systolic and diastolic BP (all P<1×10^-4); CACNA1AxADRA1D, CACNA1CxADRB2, ADRA1AxADRA1B, ADRA1AxSLC6A2 for systolic BP (all P<1×10^-4); and CACNA1AxCACNA1D, CACNA1AxSLC6A2, CACNA1CxSLC18A1, ADRA1DxADRA1A, ADRA1DxDBH, ADRA1DxSLC18A2 for diastolic BP (all P<1×10^-4); and These findings provide strong evidence for influence of SNS genes and gene-gene interaction.

1855M

Closed-form Wald tests for genome-wide analysis of gene-gene interactions. Z. Yu1, M. Demetrou1, D. Gillen1. 1) Department of Statistics, University of California, Irvine, CA; 2) Department of Neurology, University of California, Irvine, CA.

Despite the successful discovery of hundreds of variants for complex human traits using genome-wide association studies, the degree to which genes jointly affect disease risk is largely unknown. One obstacle toward this goal is that the computational effort required for testing gene-gene interactions is enormous. As a result, numerous computationally efficient tests were recently proposed. However, the validity of these methods often relies on unrealistic assumptions such as additive main effects, main effects at only one SNP, and no linkage disequilibrium. Here we propose to use closed-form Wald tests. The Wald tests are asymptotically equivalent to the corresponding likelihood ratio tests, largely considered to be the gold standard tests but generally too computationally demanding for genome-wide interaction analysis. Simulation studies show that the Wald tests have very similar performance with their computationally intensive counterparts. Applying the proposed tests to a genome-wide study of multiple sclerosis, we identify interactions within the major histocompatibility complex region. In this application, we found that (1) focusing on pairs where both single nucleotide polymorphisms (SNPs) are marginally significant leads to more significant interactions when compared to focusing on pairs where at least one SNP is marginally significant; and (2) parsimonious parameterization of interaction effects might decrease, rather than increase, statistical power.

1856T

Gene based analyses of sympathetic nervous system genes on long term blood pressure: The GenSalt study. C. Li1, J. He1, J.E. Hixson*, D. Gu1, D.C. Rao1, L.C. Shimmin1, J. Huang1, C.G. Gu2, J. Chen1, F. Liu2, J. Li2, T.N. Kelly1. 1) Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA; 2) State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; 3) Department of Medicine, Tulane University School of Medicine, New Orleans, LA; 4) Department of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, TX; 5) Division of Biostatistics, Washington University School of Medicine, St. Louis, MO.

Although genomic factors are known to influence BP, much of the heritability of the BP phenotype remains unexplained. Gene-based analysis of long term BP may increase power to detect genetic factors underlying BP phenotypes. The objective of the current study was to examine the influence of sympathetic nervous system (SNS) genes and gene-gene interactions on long term average BP among 1,881 Han Chinese participants of the family-based Genetic Epidemiology Network of Salt Sensitivity (GenSalt) follow-up study. Nine BP measures were collected at a 2003 baseline examination and at each of two follow-up visits in 2007 and 2011. Single marker associations of 448 tag SNPs in 20 SNS genes and the pairwise interactions of SNPs between each pair of genes were assessed for association with long term BP using mixed models to accommodate the repeated BP measures and GenSalt family structure. For single gene analysis, single-marker analysis P-values for SNPs within each gene were combined using the truncated product method (TPM) to estimate an overall gene-based P-value. TPM was also used to combine P-values at all SNP interaction terms within each gene pair to generate a gene-gene interaction P-value. Nominally significant findings from GenSalt were evaluated for replication using publically available data from the Multi-Ethnic Study of Atherosclerosis (MESA). Findings from both studies were combined using meta-analysis. Statistical significance was determined by nominal significance in both GenSalt and MESA studies, along with significance after Bonferroni correction to meta-analysis results (α-threshold=0.0025 and 0.00026 for single-gene and gene-gene interaction score, respectively). In single-gene determinants, DBH was associated with long term systolic BP (P=0.0014), while ADRA1D was associated with long term diastolic BP (P=0.0004). In addition, multiple gene-gene interactions influencing BP phenotypes were identified, including: CACNA1AxADRA1A, CACNA1xCACNA1C, CACNA1CxCACNA1C, CACNA1CxADRA1A, CACNA1CxADRA1D, CACNA1DxADRA1D, CACNA1DxSLC6A2 for both systolic and diastolic BP (all P<1×10^-4); CACNA1AxADRA1D, CACNA1CxADRB2, ADRA1AxADRA1B, ADRA1AxSLC6A2 for systolic BP (all P<1×10^-4); and CACNA1AxCACNA1D, CACNA1AxSLC6A2, CACNA1CxSLC18A1, ADRA1DxADRA1A, ADRA1DxDBH, ADRA1DxSLC18A2 for diastolic BP (all P<1×10^-4); and These findings provide strong evidence for influence of SNS genes and gene-gene interaction.
Rapid Variance Component Aggregation Test (RVCAT) for evaluating interaction effects of rare-variants. R. Marceau1, W. Lu1, F.-C. Hsu2, J.-Y. Tseng3. 1 Department of Statistics, North Carolina State University, Raleigh, NC; 2 Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC; 3 Bioinformatics Research Center, North Carolina State University, Raleigh, NC.

Studying gene-environment (GxE) and gene-gene (GxG) interactions in genetic association studies allows us to better understand complex disease etiology, uncover missing heritability, and opens the door to personalized medicine. Gene-gene and gene-environment aggregation approaches have been useful alternatives to the more traditional single SNP GxE tests, especially for rare variants. There are two categories of aggregation approaches that are commonly used: fixed effects methods, which collapse at the genotype level (e.g., the burden test), and random effects methods, which collapse at the similarity level (e.g., kernel machine regression and similarly regression). For assessing interaction effects, random effects models have found to be more robust to model misspecifications than fixed effect approaches. That is, these approaches are still valid when the genetic main effect term and/or the GxE term are misspecified. However, because random effects models use variance components (VCs) to capture the effects, such robustness comes at the cost of increased computational burden from having to estimate nuisance VCs for each effect when assessing GxE effect. These nuisance VCs are computationally intensive to calculate when the trait of interest is non-quantitative. In this work, we propose a rapid variance component aggregation test (RVCAT) which sidesteps the need to estimate the nuisance VCs by treating the genetic main effect as a fixed effect term. We show that the RVCAT maintains the validly and power of the traditional VC tests, it is also computationally efficient—for a quantitative trait analysis.

GWAS, we will show how multiple personal exposures can be assessed simultaneously in terms of their association with diseases such as type 2 diabetes (T2D), heart disease risk factors, and mortality in cohorts from the United States National Health Nutrition Examination Survey and the International Study of Macronutrients in Blood Pressure (N-5K-30Y). In these studies, we show how an array of exposures ranging from pollutants, nutrients, and pesticides are associated with these diseases and have effect sizes that are comparable or exceed GWAS findings. We will discuss the hurdles, including biases such as reverse causality, confounding, and challenges of inferring independence in midst of the dense correlation structure of the exposure. We outline and show preliminary findings that combine the exposome with genome to assess how genetic risk for disease may be dependent on environmental exposure-defined phenotypes. We explored the best strategy to detect gene-environment interaction for environmental exposure-defined phenotypes, where the exposure had four ordered levels (unexposed; and low, medium, high exposure). We compared the statistical power and Type 1 error of 4 different tests that are commonly used in this context: 1) gene main effect test (G); 2) genotype-exposure interaction test (GE); 3) case-only gene-environment interaction test (CO); 4) joint test of both the gene main effect and the interaction (GGE). We considered factors that may affect the power and Type 1 error of these tests, including: genetic risk allele frequency; genotype-exposure frequency; genotype-exposure correlation in the population; effect sizes of exposure, gene and gene-environment interaction; and exposure misclassification. Finally, we compared analyses for non-exposure-defined phenotype and exposure-defined phenotype with both exposed and non-exposed controls, or with exposed controls only. When only main effects were present with no interaction effect, the G test showed higher power than the GGE test. When main and interaction effects were present, the CO test showed higher power than the GE test for non-exposure-defined and exposure-defined phenotypes. We note that the power of GE test reduced dramatically when restricting to exposed controls for exposure-defined phenotypes. For non-exposure-defined phenotype, the GGE test showed higher power than the G test (2% increase in power). However, for exposure-defined phenotype, the G test showed higher power than the GGE test (12% increase in power). Restriction to exposed controls improved the power of the G test and the GGE test for exposure-defined phenotype. Based on these results, we recommend the CO test for investigating gene-environment interaction, with a caveat on inflated type 1 error due to G-E correlation, and recommend the exposed controls for detecting genetic effects on exposure-defined phenotypes.
A novel functional data analysis approach to detecting gene by longitudi-

The recent switch to a modern lifestyle has coincided with a rapid increase

A simulation study of gene-by-environment interactions in GWAS

GWAS for a longitudinal trait with non-uniform errors: Recovery of

Many methods have been developed for associating SNPs with disease,

Detecting clusters of disease-associated SNPs. D. Swanson, J. Laramie.

1865T

Biochemical network-driven analysis of genetic control of human metabolome. Y.A. Tsepilov1,2, K. Strach1,2, C. Gieger3, Y.S. Altschuler1,2.
1) Institute of Cytology and Genetics SD RAS, Novosibirsk, Russia; 2) Novosibirsk State University, Novosibirsk, Russia; 3) Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany; 4) Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.

Genome-wide association analysis (GWAS) is one of the most popular methods of identification of alleles that affect complex traits. Recent accumulation of GWAS data ("omics" data) could give a new insight into the functional role of specific changes in the genome. Such data require special statistical methods for their analysis, as they are characteristically high dimensional and contain statistical interactions between individual components, which all reflect biological interactions. Development of such methods is of current importance as the progress of molecular biology techniques marches on. We introduce a new approach to evaluate the impact of genetic variants to "omics" phenotypes with relatively high power and lower modest computational intensities. Our approach is based on the idea that if biological pathways and relationships in data could be reconstructed, we can use the knowledge about biochemical neighbors for chosen trait and include this information into analysis of the data. The sources of biological relationship information can vary: biological pathway databases or reconstructed network from the data. We assessed our approach using real population study data from big German study KORA (n=1,784, 2M SNPs) with measured metabolomic data (151 metabolites). For reconstruction of the pathways we used Gaussian graphical models (GGM) which were used as effective tool for the unbiased reconstruction of metabolic reaction. Previously it was demonstrated that using GGM-driven ratios instead of all ratios leads to comparable power with much less computational expenses. Our approach as well has shown the comparable power with all ratio approach with computational complexity similar to the analysis of original concentrations only (n instead of n^2). Using of biochemically related covariates decreases the phenotype variances and the noise driven by network influences between metabolites with the reduction of induced signals and the improvement of actual associations that helps with further functional annotation.

1866S

Joint Analysis of Genetic Interaction and Imprinting in Family Studies. C.C. Wu1, S. Strieb1. 1) Environmental/Occupational Health, National Cheng Kung University, Tainan, Taiwan; 2) Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

Genomic imprinting is considered a primary epigenetic phenomenon, which results in 2 alleles at a locus being expressed differentially depending on the parental origin of the chromosome. The role of imprinted genes and gene-gene interactions in disease susceptibility is well recognized. Imprinting leads to manifestation of parent-of-origin effects. More than 1000 individual gene-gene variants associated with common diseases have been detected by recent whole genome studies; the efforts to map interacting genes have been less successful. In response, we proposed a method to detect genetic interaction effects in the presence of genomic imprinting. We extended our previous allele-sharing method that accounts for parent-of-origin effects in the affected sib pairs framework. We derived 3 mathematical two-locus models that incorporate parent-of-origin effects: additive, multiplicative, and general models. We further proposed to use a novel test to assess the gene-gene interaction and parent-of-origin effects individually and jointly. The statistical properties and performance of the proposed method will be evaluated using simulated data.

1867M

Genetic gene interactions in admixed populations. E. Zy, D. Hu, L. Fejerman. Department of Medicine, Institute for Human Genetics, Helen Diller Family Comprehensive Cancer Center, UCSF, San Francisco, CA.

Identifying gene-gene interactions is a challenge in human genetics due to the limitations of power after adjusting for the large number of potential interactions assessed. We examine the approach to identifying gene-gene interactions in an admixed population. We model a dichotomous trait (disease vs. no disease) in an admixed population consisting of two ancestral populations. First, we consider how the type of genetic interaction and the range of allele frequencies in the ancestral populations affect the association between genetic ancestry and disease. We demonstrate that there are distinct non-linear patterns of the association between disease risk and genetic ancestry depending on the allele frequency differences between ancestral populations. Furthermore, we consider how, by capitalizing on these patterns, investigators can design studies that enhance the power to detect gene-gene interactions. We illustrate the approach using simulations and a study of breast cancer in a U.S. Latina population.
1869S  
Identification of shared genetic etiology between epidemiologically linked disorders with an application to obesity and osteoarthritis. J. Asimina1, K. Panoutsopoulou1, E. Wheeler2, S. Berndt1, T. Barroso1, E. Zeggini1, the GIANT consortium, the arcGEN consortium.  
1) Human Genetics, Wellcome Trust Sanger Institute, Cambridge, United Kingdom; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom; 3) Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom; 4) US Department of Health and Human Services, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. 
Diseases often co-occur in individuals more likely than expected by chance. Shared genetic etiology may explain this, and the discovery of shared variants may suggest a common biological process. A common approach to a genetic overlap analysis is to select a p-value threshold and compare the signals from the genome-wide association study (GWAS) of each trait. However, p-values do not account for differences in power, whereas Bayes’ factors do, and may be approximated using summary statistics. We use simulation studies to compare the power of frequentist and Bayesian approaches to overlap analyses, and to decide on appropriate thresholds for comparison between the two methods. It is empirically illustrated in single-disease associations that BFs have the advantage over p-values of a decreasing proportion of false positives (PPF) as study size increases. For a log10(BF) threshold L0 of 1.69 (R=type II error cost/type I error cost=2, π0 = P(rno association at SNP)=0.99), the PPF decreases from 7.38×10-4 (N=2,000 each cases/controls) to 3.37×10-4 (N=20,000), while for p-values the PPF fluctuates near the p-value threshold α regardless of study size. Power increases with sample size are more noticeable for the Bayesian approach. For OR=1.2 and MAF 0.1, the power in studies with N=3,000 are 0.413 (L0=1.69) and 0.365 (α=5.5×10-4), while at N=5,000 the respective powers rise to 0.750 and 0.716. In a preliminary overlap analysis of obesity (GIANT consortium) with OA (arcGEN consortium), the number of signals is similar at comparable threshold levels between BFs and p-values. The two sets of results are not always overlapping and additional signals are often in already detected genes. For L0=0.91 (R=type II error cost/type I error cost=12, π0=P(rno association at variant)=0.99), there are 15 additional variants identified by PCCA and 14 by BFP statistical comparison by α levels of 0.003 and 0.004 yield 15 and 28 hits, respectively. The most notable difference is that the Bayesian list contains rs13107325 (in SLC39A8/ZIP8), a variant previously associated with obesity-related phenotypes such as BMI, blood pressure, and HDL cholesterol, and animal studies on this gene have shown that the zinc-ZIP8-MTF1 axis regulates OA pathogenesis. This variant is only detected via p-values if we increase α to 0.005. We are pursuing replication of this finding. Extensive simulations to systematically evaluate power differences between the BF and p-value-based overlap analysis approaches are underway.

1870M  
Multilevel dimensionality reduction algorithms for high-dimensional genetics data. K. Cho1,2, D.R. Gagnon3,2, H. Wu2,4, 1) Division of Aging, Brigham and Women’s Hospital/Harvard Medical School, Boston, MA, USA; 2) Massachusetts Veterans Epidemiology Research and Information Center; VA Boston Healthcare System, Boston, MA, USA; 3) Boston University School of Public Health, Boston, MA, USA; 4) Computer Science and Networking, Wentworth Institute of Technology; Boston, MA, USA. 
With rapid advancements in genotyping and sequencing technologies, there are more genetic data available than one can thoroughly analyze with existing tools and methods. This has become one of the analytical challenges in the initial phases of big scale studies. In whole-genome or genome-wide exploration of novel associations, without a priori knowledge of specific regions or genes of interest, researchers rely on the available techniques that perform association tests based on a way of summarization or assuming independence. In addition, traditional statistical procedures present eminent challenges in using these data, where the number of parameters p is scalably larger than number of observations n. These limitations may compromise the level of false positive results and power to conduct studies with acceptable level of confidence. One of the ways to reduce the initial efficiency burden with high dimensional genetic data is to first perform a data dimensional reduction process and then evaluate the resulting panels of markers. We propose several algorithms using traditional PCA based algorithms, LASSO, and combinations of these to evaluate the impact of type one error. Using the Genetic Analysis Workshop 18 simulated data, we apply the several algorithms on null regions of chromosome 3 with regards to the diastolic blood pressure phenotype. Then we use SKAT as the evaluation tool to perform association analysis using the resulting panels from each algorithm. Two methods were applied to compare type I error rates across different algorithms with respect to the baseline panel. Our preliminary work using 360,717 SNPs containing 1431 gene groups in unrelated samples, shows a substantial reduction of the markers through applying proposed algorithms. Our PCA based method was empirically compared to different algorithms on different gene group sets. Among these, the PCA top SNP and PCA genes algorithms resulted in 106 and 23297 unique SNPs, respectively, at 85% variation. At 95% variation, 127 and 35382 resulting SNPs remained, respectively. We are also implementing linkage disequilibrium PCA based heritability approaches using the family datasets. As the amount and dimensionality of the network of genomics data is only escalating in the coming years, a timely, robust and practical analytical pipeline through multilevel dimensionality reduction techniques provides an efficient approach for the initial screening tool.

1871T  
We present a functional method for fitting simultaneously the effects of scalar predictors, confounding factors and environmental exposures on sequenced allelic variants in a genetic region. These effects are fitted as continuous curves smoothly varying over genetic loci to help researchers shed light on which regions harbor causal variants. In a densely sequenced region, one can plot genotypic values on the y-axis versus variant position on the x-axis and ‘connect the dots’. The resulting continuous genotype curve will change over variants position. By utilizing continuous curves, our method exploits linkage disequilibrium among variants in the ways that otherwise is difficult to achieve. For example, one can regress genotype curves on a set of environmental variables and obtain a subject-specific set of curves for each environmental effect that will also change over variants position. The regression on the genotype curve is produced based on a penalized spline regression and can be linked to linear mixed models which allows utilization of standard statistical software packages. We investigate advantages and disadvantages of our approach by conducting an extensive simulation study. We illustrate the approach with sequencing data from the Dallas Heart Study.
1872S
Multivariate approach for finding gene sets differentially expressed by complex phenotype. E. Drigalenko, H.H.H. Göring. Dept Genetics, Texas Biomedical Research Institute, San Antonio, TX.

Genes almost never act alone in a biological system; they typically work in a cascade of networks. This motivates the consideration of clusters of genes, instead of individual genes, as functional genomic units. Variations of certain phenotypes, including complex diseases and quantitative traits, may be associated with differential expressions of multiple genes instead of a single gene.

Transcriptome sequencing (RNA-seq) technology allows measuring of the expression levels of tens of thousands of genes simultaneously. Expression profiles of multiple genes are correlated and thus are appropriately modeled as mutually correlated variables in a statistical testing framework.

Here, we propose three multivariate methods for testing the association of sets of correlated genes with discrete or continuous phenotypic outcomes: (1) comparison of covariate matrices under different conditions using Box's M test; (2) joint regression of all genes in a pathway; (3) orthogonal decomposition of the variance and the regression of principal components (PC) using Kaiser's criteria for the number of essential PCs. KEGG pathway classes C, B, and A were studied separately. We use programs implemented in R. The last two methods can be used for quantitative traits.

To demonstrate the performance, we apply these methods to publicly available data from the NCBI Gene Expression Omnibus database. For this presentation, we used the gene expression (RNA-seq) data from 21 individuals (Akula et al, 2014; GSE53239) - 10 with bipolar disorder and 11 controls. As an example of gene sets, we have used Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways.

We found that the proposed methods do identify differentially expressed gene pathways. In our example, the PC-based method identified the most pathways. Using it, after the Bonferroni correction for the number of pathways is a class, we found that 27 pathways are significant of total 282 in KEGG pathways. We used programs implemented in R. The last two methods can be used for quantitative traits.

1873M
Estimating and interpreting pairwise genetic correlations between hundreds of quantitative traits from population samples of thousands of individuals. M. Pirinen1, C. Benner1, M.A. Rivas2, T. Lehtimäki3, A.J. Kangas4,5, P. Soininen6,7, M. Ala-Korpela1,5,6, J.G. Eriksson7,8,9, O.T. Raitakari10,11, M.R. Järvelin12,13,14, V. Salomaa15, S. Ripatti1,15,16, 1) Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 3) Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere School of Medicine, Tampere, Finland; 4) Computational Medicine, Institute of Health Sciences, University of Oulu and Oulu University Hospital, Oulu, Finland; 5) NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio, Finland; 6) Computational Medicine, School of Social and Community Medicine and the Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK; 7) Department of Chronic Disease Prevention, National Institute for Health and Welfare, Finland; 8) Department of General Practice and Primary Health Care, University of Helsinki, Finland; 9) Unit of General Practice, Helsinki University Central Hospital, Finland; 10) Department of Clinical Physiology and Nuclear Medicine, University of Turku and Turku University Hospital, Turku, Finland; 11) Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; 12) Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment and Health, School of Public Health, Imperial College London, UK; 13) Institute of Health Sciences, University of Oulu, Finland; 14) Biocenter Oulu, University of Oulu, Finland; 15) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; 16) Hjelt Institute, University of Helsinki, Helsinki, Finland.

Several modern technologies, such as nuclear magnetic resonance and mass spectrometry platforms, produce high-dimensional phenotype data on individuals. A first step towards utilising high-dimensional phenotypes in genetic studies is to understand how their genetic components are related. Recent algorithmic advances in multivariate linear mixed models have enabled variance component estimation for pairs of traits using population samples of individuals and genome-wide panels of SNPs. However, current methods have not been tailored for situations where hundreds of traits are available on the same set of individuals. For such settings, we introduce an algorithm that efficiently decomposes pairwise phenotypic correlations into genetic and environmental components. We illustrate our approach with an application to thousands of pairs of metabolic and anthropometric traits measured on up to 14,000 Finnish individuals. For example, we estimate that the observed phenotypic correlation (-0.41) between triglycerides and HDL cholesterol decomposes into an additive genetic correlation (-0.59, s.e. 0.06) and an environmental correlation (-0.36 s.e. 0.02). We assess the interpretation of genetic correlations as correlations between locus-wise genetic effects by applying multivariate statistical models on the results of a recent genome-wide meta-analysis on lipid levels. Finally, we consider settings where prior information about genetic correlation increases statistical power to identify pleiotropic loci, i.e. loci that contribute to multiple traits.
1874T Systemic genetics of Systemic Sclerosis through protein-protein interaction network-based analysis. J. Hamon1, Y. Allano2, M. Martini2, N. Nez2, 1) INSERM UMR1043, Hôpital Purpan, Toulouse, France; 2) INSERM UMR1016, Hôpital Cochin, Paris, France.

Genome-wide association studies (GWAS) of complex traits have revealed reproducible genetic associations with human complex disorders, including Systemic Sclerosis. However, these studies rely on analysis at the single-nucleotide polymorphisms (SNP) level and, consequently, lack power in presence of allelic heterogeneity (multiple independent risk variants within a given gene/genomic region) and/or for detecting interacting risk variants (with low heritability). This could contribute to the considerable interest in multi-locus analysis techniques as, for instance, knowledge-driven methods that integrate known functional and interaction networks. These approaches vary according to the type of known information (protein-protein network/gene function pathways) that is integrated to the GWAS data; the way GWAS data in genes is used (single-SNP or SNP-set statistical significance); and the model used to test for association between a module/pathway of genes and the phenotype. Here, we present results from various study designs for network-based analysis integrated to our Systemic Sclerosis GWAS data (564 cases and 1,776 controls). We have built a robust two-stage network analysis by randomly splitting our GWAS data into a scan and a replication dataset. In the scan dataset we performed a PPI network-based approach using a dense module search strategy and using different approaches to assess GW. For instance, from the lowest single-SNP P value adjusted (Min) or Bonferroni-adjusted (Bonf) for the number of SNPs within the gene, or by combining all single-SNP P values with the Fisher’s method. We first compared the results obtained under the different strategies in terms of the length (number of genes) and depth of the enriched sub-networks and the characteristics of the genes within the modules. The top (5 and 10%) most enriched modules were further tested for enrichment analysis in the replication dataset. We finally assessed whether the genes from the replicated sub-networks were significantly clustered within KEGG pathways. Our results showed that the gene-wise association P-value modeling has a large impact on the results, i.e., length (number of genes) of sub-networks, characteristics of their genes and replication rate of top most enriched sub-networks. Overall, we found considerable consistency across the results from different strategies: the different strategies tend to select different sets of genes but also different KEGG pathways.

1875S Mixed-model analysis of common variation reveals pathways explaining variance in AMD risk. J. Hall1, M. Pericak-Vance2, W. Scott3, J. Kovach4, S. Schwartz5, A. Agarwal6, M. Brantley7, J. Haines8, W. Bush9. 1) Institute for Computational Biology, Case Western Reserve University, Cleveland, OH; 2) John P. Hussman Institute for Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 3) Department of Ophthalmology and Visual Sciences, Vanderbilt University, Nashville, TN.

Age-related macular degeneration (AMD) is the leading cause of irreversible blindness in the elderly in developed countries and affects approximately 10% of individuals over the age of 80 in many populations. AMD has a large genetic component, with overall heritability estimated to be between 45% & 70%. Numerous AMD loci have been identified and implicate various molecular mechanisms and pathways in AMD pathogenesis. Eight pathways, including angiogenesis, antioxidant activity, apoptosis, complement activation, lipoprotein metabolism, the tricarboxylic acid cycle, were selected for our study based on an extensive literature review. While these pathways have been proposed in literature, the overall extent that each pathway contributes to AMD risk is unknown. In a dataset of 1,154 AMD cases and 668 controls, we used Genome-wide Complex Trait Analysis (GCTA) to conduct mixed-model regression to estimate the proportion of variance on AMD risk explained by all SNPs in each pathway. Genes were assigned to pathways using Gene Ontology terms and SNPs were mapped to these genes and then assessed using GCTA. To extend the analysis beyond coding variation, nearby SNPs within a 50 kilobases region flanking each gene, many likely to be regulatory, were included in the analysis. In the final analysis, we considered 500 kilobases of each gene, based on retinal pigment epithelium (RPE) DNaSe I Hypersensitivity mappings from the ENCODE project. To account for baseline known risk, we tested 19 established AMD risk SNPs and found that they contributed to 13.3% of the variation in risk in our dataset, while the remaining 659,181 SNPs contributed to 36.7%. After adjusting for these 19 risk SNPs, the complement activation and inflammatory response pathways explained a significant P = 5 x 10^-26 and 9.5 x 10^-7, respectively) proportion of additional variance in AMD risk (9.8% and 17.9%, respectively), with other pathways showing no significant effects (0.3% - 4.4%). Our results show that additional variants associated with complement activation and inflammation genes contribute to AMD risk and that these variants are likely in coding and nearby regulatory regions.

1876M A non-parametric approach for detecting gene-gene interactions associated with age-at-onset outcomes. M. Li1, J.C. Gardiner2, N. Breslau2, J.C. Anthony2, Q. Lu2. 1) Section of Biostatistics at Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, AR; 2) Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI.

Cox-regression-based methods have been commonly used for the analyses of survival outcomes, such as age-at-disease-onset. These methods generally assume the hazard functions are proportional among various risk groups. However, such an assumption may not be valid in genetic association studies, especially when complex interactions are involved. In addition, genetic association studies commonly adopt case-control designs. Direct use of Cox regression to case-control data may yield biased estimators and incorrect statistical inference. We propose a non-parametric approach, the weighted Nelson-Aalen (WNA) approach, for detecting genetic variants that are associated with age-dependent outcomes. The proposed approach can be directly applied to prospective cohort studies, and can be easily extended for population-based case-control studies. Moreover, it does not rely on any assumptions of the disease inheritance models, and is able to capture high-order gene-gene interactions. Through simulations, we show the proposed approach outperforms Cox-regression-based methods in various scenarios. We also conduct an empirical study of progression of nicotine dependence by applying the WNA approach to three independent datasets from the Study of Addiction: Genetics and Environment. In the initial dataset, two SNPs, rs6570989 and rs2930357, located in genes GRK2 and CSMD1, are found to be significantly associated with the progression of nicotine dependence (ND). The joint association is further replicated in two independent datasets. Further analysis suggests that these two genes may give rise to and be associated with the progression of ND.

1877T Estimation of Heritability and Association for Quantitative Traits with Repeated Hba1c Measures: Biomarker for Metformin Response. L. Wu1, J. Mellford2, S.Y. Yee3, K.M. Giacomini4, J.S. Witte5. 1) Department of Epidemiology/Biostatistics University of California, San Francisco, CA; 2) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA; 3) Department of Epidemiology/Biostatistics and Urology, University of California, San Francisco, CA.

Assessing heritability is important for understanding the genetic architecture of quantitative traits. Heritability explained by typed SNPs is commonly calculated using linear mixed-effects models with an estimated genetic relationship matrix (e.g., with GCTA). We adopt a linear mixed-effects model to estimate heritability and carry out association study of quantitative traits with repeated measures. We also propose a new estimate of the genetic relationship matrix that only guarantees the average of diagonal elements to be close to one. The marginal model implied by the linear mixed-effects model is used to lower the computational burden in association analyses. Following an initial screen, the top-ranked SNPs can be reevaluated at a time to produce more accurate p-values of their association. In our study, we use hemoglobin A1c as a biomarker for response to anti-diabetic drugs such as metformin. We estimate the heritability of metformin response in a cohort of 500 African American diabetic patients with 1,320 repeated HbA1c measures after metformin treatment. For this data, the genetic relationship matrix obtained from GCTA has diagonal elements ranging from 0.926 to 1.511 (average = 1.004, SD = 0.014); in contrast, our approach gives values all equal to 1. The off-diagonal elements of the genetic relationship matrix are the cosine similarity of mean-centered genotype vectors of two individuals; assuming no inbreeding, this method guarantees that the diagonal elements equal one, which provides a good characterization of additive genetic relationship of the same individual across time. Standard methods can be directly applied to prospective cohort studies, and can be easily extended for population-based case-control studies. Moreover, it does not rely on any assumptions of the disease inheritance models, and is able to capture high-order gene-gene interactions. Through simulations, we show the proposed approach outperforms Cox-regression-based methods in various scenarios. We also conduct an empirical study of progression of nicotine dependence by applying the WNA approach to three independent datasets from the Study of Addiction: Genetics and Environment. In the initial dataset, two SNPs, rs6570989 and rs2930357, located in genes GRK2 and CSMD1, are found to be significantly associated with the progression of nicotine dependence (ND). The joint association is further replicated in two independent datasets. Further analysis suggests that these two genes may give rise to and be associated with the progression of ND.
1878S

Genome-Wide Association and Gene-Gene Interaction Studies to Explore Etiology of Glaucoma and Ocular Hypertension. S. Verma1, A. Verma2, A. Locas2, J. Linnemann3, P. Peisig4, M. Brilliant5, C. McCarthy6, J. Haines7, T. Vrabec8, G. Tromp9, J. Patakha10, C. Chute11, D. Crosslin12, G. Jarvik13, E. Larson8, M. Hayes8, M. Ritchie11. 1) Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA; 2) Marshfield Clinic, Marshfield WI; 3) Essentia Rural Health, Duluth, MN; 4) Case Western University, Cleveland, OH; 5) Geisinger Health System, Danville, PA; 6) Mayo Clinic, Rochester, MN; 7) Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA; 8) Group Health Research Institute, Seattle WA; 9) Northwestern University, Chicago, IL.

Ocular hypertension (OHT) is a leading risk factor for glaucoma. Previous studies have identified interesting genetic and environmental factors for these two traits, which are among the leading causes of blindness worldwide. In the electronic Medchalcal Records and Genomics (eMERGE) network, sites extracted phenotype information on patients with glaucoma and OHT from the EHR who were already genotyped and part of DNA biobanks. We performed GWAS and gene-gene interaction (GxG) studies for both of these traits using data imputed to 1000Genomes. A total of 3,253 and 3,154 unrelated samples of ages 40-90 across five different eMERGE site biorepositories were assembled for glaucoma and OHT studies respectively. We observed novel genome-wide significant variations on Chr1 (rs4147800) associated with glaucoma (p= 1.04×10^-23) and on Chr13 (rs4646227) and Chr16 (rs36120466) associated with OHT (p= 3.92×10^-6 and 4.52×10^-6). To examine epistasis between multiple genetic variants, considering extensive computational demands and multiple testing burden of exhaustive testing, we applied a main effect filter (SNPs with p-value < 0.01 from GWAS) and then performed genome-wide interaction study (GWIS) using regression methods. We found several interesting interactions that remain significant after Bonferroni correction. ABCA4 shows a significant interaction with GAD2 (LRT p= 2.71×10^-11) and also we observed significant interaction (LRT p= 1.85×10^-7) between two genes (PON3 and PON2) in member of paraoxonase family that validates previously published associations between primary open-angle glaucoma and antioxidant enzymes affecting high density lipoprotein (HDL). We also merged these results and performed pathway-enrichment analysis for all results with main effect p<1×10^-8. KEGG Pathway analysis found 10 pathways that are shared for glaucoma and OHT. Notably, ABC transporter pathway contain the most significant genes associated with both Glaucoma and OHT. Similarly serotonin 5-HT2C receptor agonists (LTB4R and AGRP genes) linked to neuroactive ligand-receptor interactions (most significant association in OHT) are associated with both phenotypes and they have also been identified as potential agents for the treatment of both OHT and glaucoma. Our findings indicate novel SNP associations and GxG interactions for both of these traits and demonstrate the relationship between these two phenotypes at a potential molecular level with the guide of pathway analysis.

1879M

Major-Effect Loci for Lipids also Impact Phenotype Variability in the Old Order Amish. J. Yerges-Armstrong, J. O’Connell. Program in Personalized and Genomic Medicine, and Department of Medicine, Division of Endocrinology, Diabetes and Nutrition - University of Maryland School of Medicine, USA. Baltimore, MD 21201.

Association studies for complex traits have traditionally focused on identifying loci that shift the mean of a quantitative trait. Several recent papers describe methods for modeling genetic loci that instead impact trait variability. These genetic pathway variance quantitative trait loci (or vQTL) are of particular interest as they may indicate undetected genetic or environmental interaction. With this in mind, we tested three loci for lipid segregating in the Old Order Amish (OOA) for the presence of vQTL effects. Analyses were conducted on over 1500 participants genotyped on the Illumina Human Exome BeadChip. All analyses were conducted using our mixed models analysis for pedigrees and populations (MMAP) software. Levene’s F-test was implemented in MMAP to allow us to model genotype differences accounting for polygenic effects, age and sex. Both the mean and median F-tests were calculated but only the median test is presented. The first variant tested was the APOC3 null mutation (R19X) associated with markedly lower mean triglycerides (p=2×10^-14). In addition to the large effect on the mean, there was a significant vQTL effect (p=4.6×10^-10). The second variant was the common, promoter variant in CETP (rs3764261), which was highly significant for both mean differences in high-density lipoprotein cholesterol (p=2.6×10^-10) but had only a modest vQTL effect (p=0.02). Due to a founder effect ~10% of Amish are carriers of the APOB 3500Q allele which is associated with ~60mg/dL lower high-density lipoprotein (LDL) cholesterol (p=1.2×10^-7). We observed a significant vQTL (p=1.5×10^-5) for LDL with this variant. We have previously reported on the higher coronary artery calcification (CAC) in 3500Q carriers thought to be caused by high LDL cholesterol. In 896 OOA with CAC measures we continue to see a strong vQTL for LDL cholesterol (p=1.7×10^-6) but not for CAC (p=0.17) despite the much higher mean CAC levels in mutation carriers. To our knowledge, the vQTL effects for R19X and R3500Q are the first reported for lipids. The lack of vQTL effect for CETP may indicate that highly-penetrant, rare variants are more likely to be vQTLs. Interestingly, we only detected a vQTL for APOB with CAC in OOA with CAC < 2.0 mm2 which suggests a different mechanism than CAC on both the genetic and environmental factors. This could be a reflection of R3500Q increasing CAC through LDL and not impacting calcification directly, but additional work with other pleiotropic variants is needed before this generalization can be made.

1880T

An integrative imputation method for multi-omic datasets. D. LIN1,2, J. ZHANG1,2, J. LI1,2, H. DENG2,3,4, Y. WANG1,2,3,4. 1) Biomedical Engineering, Tulane University, New Orleans, LA, 70118, USA; 2) Center of Genomics and Bioinformatics, Tulane University, New Orleans, LA, 70112, USA; 3) Department of Biostatistics and Bioinformatics, Tulane University, New Orleans, LA, 70112, USA; 4) Center for Systems Medicine, Shanghai University for Science and Technology, Shanghai, CHINA.

Despite the increasing significance of integrating diverse types of genomic and epigenetic data (e.g. mRNA, microRNA and DNA methylation), the currently available multi-omics datasets inevitably suffer missing values due to technical limitations and various constraints in an experiment. These missing values will severely hinder the downstream analysis such as the construction of regulation patterns among genetic and epigenetic factors, and the identification of differential expressed biomarkers. Currently, imputation methods for single data type have been well studied and compared. More evidence has shown the biological interconnections among these genomic and epigenetic data. It is significant to integrate these correlated multi-omic datasets for improving the imputation accuracy of missing data. In this study, a novel imputation method was proposed to address this issue by: 1) combining the estimates of missing value from individual omic datasets in a way of model stacking to obtain an optimal estimate; and 2) incorporating prior biological knowledge into the imputation across different datasets. The real glioma cancer datasets from TCGA including 50 patients with mRNA measurements of 5939 genes, expression data of 104 microRNAs and DNA methylation measurements of 5013 genes were included in the analysis. We compared the our method with five single data imputation methods (kNN, LLS, ILLS, BPCA, and SVD) in different noise levels, sample sizes and missing rates, and the performance was evaluated by normalized root mean square error. The results indicated that our integrated method showed the lowest imputation error among all other methods across all situations. The superior performance was more prominent in cases of smaller sample size and higher noise level. We concluded that our proposed imputation method will improve the imputation accuracy by incorporating more omic datasets as well as prior biological information.
Combining allele-specific and population signal boosts power for association mapping of multiple DNA sequence-based cellular traits. N. Kumasaka, D. Gaffney. Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

Association mapping of cellular traits, such as gene expression or epigenetic marking, is a powerful method for uncovering the function of genetic variation. When cellular traits are quantified using high-throughput DNA sequencing, association signals can be detected using two orthogonal sources of information: population level differences between individuals, and allele specific differences within individuals. Here, we develop a unified probabilistic model for combining allele-specific and population level sequencing data that significantly boosts power for association detection in DNA sequencing-based cell traits, and apply this to multiple publicly available population level data sets of RNA-seq, DNase-seq and ChIP-seq data. A couple of similar approaches has already been proposed (Sun 2012; and McVicker et al 2013). However those methods integrate allele-specific information within individuals that are heterozygous at the variant within sequenced feature (Sun 2012) or those for heterozygous at the linked putative causal variant (McVicker 2013). Our approach combines information not only from heterozygotes but also homozygotes across multiple biallelic SNPs within a sequenced feature. It shows a significant improvement over the publicly available approaches, in particular, it significantly improves power for association detection in small sample sizes, with an approximately 7.5% increase in sensitivity at FDR 5% over the next best performing method (Sun 2012) in a sample size of 24 individuals and a 14.3% increase over simple linear regression. In addition, an internal analysis also shows that using information from both heterozygotes and homozygotes gives better power than just using heterozygotes with 4% additional power for the same data set. Our approach also explicitly models important experimental biases including mapping errors, genotype uncertainty and reference mapping bias. We believe our method is a general approach that will significantly improve association detection for any cellular phenotype in which allele-specific signature can be quantified, and will be particularly useful in cases where sample sizes are small.

Explaining missing heritability using Gaussian Process Regression. K.J. Sharp1, W. Wiegerink1, A. Arias-Vasquez2,3, B. Franke2,3, C.A. Albers2,4, H.J. Kappen1. 1) Biophysics, Radboud University, Nijmegen, Netherlands; 2) Human Genetics, Radboud University Medical Centre, Nijmegen, Netherlands; 3) Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands; 4) Molecular Developmental Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University Medical Centre, Nijmegen, Netherlands.

For many traits and common human diseases, causal loci uncovered by genetic association studies account for little of the known heritable variation. ‘Missing heritability’ might lie in the effect of non-additive interactions between multiple loci, but this has been difficult to test using existing parametric approaches. We employed a non-parametric, Bayesian method, based on Gaussian Process Regression, for identifying associated loci in the presence of interactions of arbitrary order. On both simulated and real datasets we demonstrate that the method has considerable power to detect high-order interactions and explain missing heritability.

As a proof of principle, we analysed 46 quantitative yeast phenotypes. Whereas detected pairwise QTL-QTL interactions accounted for little of the variance (a median of 3% per trait), we found that over 70% of the total known missing heritability could be explained using common genetic variants, many without significant marginal effects. Interestingly, the availability of biological replicates significantly improved the power to identify such loci and, hence, to explain variance.

These results already represent a significant advance in approaches to understanding the missing heritability problem with potentially important implications for studies of complex, quantitative traits. Importantly, however, features of the algorithm can be exploited to permit application to datasets incorporating much larger numbers of putative QTLs. In particular, the most computationally expensive steps consist of large numbers of independent computations that conform to the SIMT parallel computation model of Graphics Processing Units (GPUs). We also describe work in progress to develop such an implementation. Using a single Nvidia Tesla M2090 GPU we already achieve two orders of magnitude improvement in wall-clock time per iteration over a serial implementation. This indicates the potential of the approach for application to human GWAS datasets.
Incorporating Functional Information in Tests of Excess De Novo Load. 

S. McGarvey et al. 

Comparison of GWAS results from imputed SNPs and multiple anchor and partner genotyped SNPs in an isolated population, Samoa. R.L. Minster, H. Duhautox, S. Petrovski, L. Hawley, S. Viali, R. De Sandre, D.E. Weeks 1,2, S.T. McGarvey 1. 


Beyond random effects meta-analysis: explaining why effect sizes differ between studies. E. Eskin, Dept Computer Sci, Univ California, Los Angeles, Los Angeles, CA. 

Over the past few years, the aggregation of the results of many genome-wide association studies (GWAS) utilizing a technique referred to as “meta-analysis” has led to the discovery of thousands of variants implicated in hundreds of disease related traits. While most of the analysis of meta-analysis has focused on discovering additional genetic variants associated with the disease trait which were not statistically significant in any of the original studies due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.

Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.


Association studies typically begin with the determination of SNP variants for a set of cases and controls. Genotypes are obtained either from arrays, or more recently by mapping whole-genome sequencing reads to a reference genome for variant discovery. For isolated populations due to low statistical power, meta-analysis opens up the door to the investigation of other interesting phenomena. In particular, it can be observed that the effect sizes of the association often differ among the studies included in meta-analysis. These differences are referred to as heterogeneity. The current method for taking into account these effect size differences is the random effects model which attempts to model the differences when combining the studies. The goal of the random effects model is to increase the power of identifying variants associated with the trait in the presence of heterogeneity. However, the presence of heterogeneity raises several questions which are not addressed by random effects models. Most importantly, what are the sources of the heterogeneity and how can they be discovered utilizing covariate information from the studies? Does heterogeneity suggest the presence of interactions? What is the correct interpretation of the results of a random effects meta-analysis? In this work, I describe a novel framework for meta-analysis which is based in the inference of causal graphs. In this framework, both fixed effects and random effects meta-analysis are just special cases of the framework. The framework allows for the construction of novel meta-analysis algorithms that can incorporate information about the studies leading to higher power and easier interpretability.
Haplotype eQTLs in response to trivalent influenza vaccine. H. Xu \textsuperscript{1,2}, L.M. Franco \textsuperscript{1,4}, R.B. Couch \textsuperscript{4}, Y. Shen \textsuperscript{2}, J.W. Belmont \textsuperscript{1,4}, Y. Guan \textsuperscript{1,2,4}.

1) USDA/ARS Children’s Nutrition Research Center, Baylor College of Medicine, Houston, TX; 2) Department of Biomedical Engineering, Southeast University, Nanjing, Jiangsu, China; 3) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 4) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 5) Department of Systems Biology, Columbia University, New York, NY.

Understanding how gene expressions respond to trivalent influenza vaccine and what genetic factors affect differential response among healthy individuals are of great importance to public health. In a previous study (Franco et al., 2013), a cohort of healthy individuals (225) were recruited, genotyped, and their blood gene-expression were assayed on the day before (designated as day 0) and multiple time points after (at the day 1, 3, and 14 respectively) the trivalent influenza vaccination. Here we reanalyze the data using a new method and report novel findings.

The novel method combines the multi-phenotype analysis (Stephens, 2013) and the haplotype association method (Xu and Guan, 2014), developed particularly for the data set. First, we jointly analyze the four time points simultaneously, treating expressions of a gene over time as correlated phenotypes (Stephens, 2013), which empowers us to detect expression quantitative loci (eQTL) responding to vaccination. Second, in addition to single SNP analysis to detect SNP-eQTLs, we focus on haplotype analysis to detect hap-eQTLs, which is more powerful when there exists allelic heterogeneity (Xu and Guan, 2014). The multi-phenotype analysis produced 425 genes with cis (defined as 1M base pairs up-and-downstream) hap-eQTLs that have Bayes factor > 10\(^6\). These genes are enriched in alternative splicing (p< 10\(^{-4}\)) and splicing variants (p<2\times 10\(^{-5}\)). As expected, these genes are also enriched in immune response (p<2\times 10\(^{-3}\)), antigen presenting and processing (p<2\times 10\(^{-5}\)), and positive regulation of immune system process (p<2\times 10\(^{-3}\), all p-values are Fisher’s exact test obtained from DGV-D). As a baseline for comparison, we obtained SNP-eQTL for each day, and used the max Bayes factor (among four days) to declare significant cis-acting SNP-eQTL (same threshold of 10\(^6\)). There are 60 genes that can only be discovered by the haplo-QTL approach and 77 genes that can only be discovered by the multi-phenotype analysis.

We also discovered a single trans-acting hap-eQTL with Bayes factor of 10\(^{16}\): haplotype variants in an interferon receptor gene affecting expression of an olfactory receptor, whose probe can be uniquely mapped on a different chromosome.
1891S Genetic ancestry associated with obesity and diabetes risks in a Mexican-American population from Houston, Texas. H. Hu, C.D. Huff, Y. Yamamura, X. Wei, S.S. Strom. Epidemiology, UT MD Anderson Cancer Center, Houston, TX.

The Mexican-American population is an admixed ethnic group with varying proportions of Native American, European and African ancestries. African Americans and Native Americans have higher incidence rates of obesity and diabetes compared to Europeans due to a combination of environment and genetic factors. However, the relative contribution of genetic ancestry is not clear. In this study, we focus on a well-defined Mexican-American population in Houston, Texas, which is a socioeconomically homogeneous but genetically diverse population, making it ideal for studying the genetic epidemiology of inherited diseases. We genotyped 96 ancestry informative markers in 4,817 individuals, and used ADMIXTURE to estimate the composition of genetic ancestry in each individual. The mean proportions of Native American, European and African ancestries among all individuals were 62.9±0.3%, 30.7±0.3%, and 6.4±0.1%, respectively. Multivariate logistic regression models were constructed. In females, we found that African ancestry was significantly associated with obesity after controlling for founders. Specifically, a 5% increase in African ancestry corresponded to a 12% increased risk of grade 2 obesity (BMI between 35 and 40; p=0.037) and an 18% increased risk of grade 3 obesity (BMI>40; p=0.006). In contrast, Native American ancestry was associated with increased risk of being overweight (BMI between 25 and 30; p=0.031), while its association with grade 1 obesity was positive but non-significant. In males, we did not observe any significant association between genetic ancestry and obesity (p>0.1). Our preliminary analysis on the risk of diabetes identified a positive and significant correlation between Native American ancestry and diabetes risk in females, after controlling for age, socioeconomic status, physical activity, smoking, alcohol consumption, and BMI (p=0.018). We estimated that 10% increase in Native American ancestry increases the risk of diabetes by 0.7%. We found no association between Native American ancestry and diabetes risk. In conclusion, our study suggests that among the Mexican-American population, African ancestry is a strong risk factor for obesity, and Native American ancestry is a modest risk factor for overweight and diabetes. Interestingly, both associations were only observed in females, indicating that the increased risk could be related to the lifestyle characteristic of females in this population.

1892M Investigating European admixture in GERA East Asians. Y. Banda1, M. Kvale1, T. Hoffmann2,3, S. Hasselx1, D. Ranatunga2, L. Walter2, C. Schaefer2, P. Kwok1,4, N. Risch1,2,3. 1) Institute Human Genetics, University California San Francisco, San Francisco, CA; 2) Kaiser Permanente Northern California Division of Research, Oakland, CA; 3) Department of Epidemiology and Biostatistics, University of California, San Francisco, CA; 4) Cardiovascular Research Institute, University of California, San Francisco, CA.

A number of analyses have been performed to understand East Asian genetic substructure. Most of these studies have utilized samples with 100% East Asian ancestry and not much has been done concerning European admixture in these groups. We describe the genetic structure in a sample of 7,520 subjects of East Asian ancestry from Northern California who are part of the Genetic Epidemiology Research on Aging (GERA) cohort. We observe individuals who are 100% East Asian, but also individuals with mixed East Asian-European ancestry. Subjects of Filipino ethnicity have a more continuous European ancestry proportion distribution compared to other subjects, having, on average, a minimum of 5% European ancestry proportion. For these subjects we analyse expected allele frequencies given the admixture process and also calculate timing of the admixture events that gave rise to the current genetic structure of Filipinos. Analyses of some Y-chromosome loci also give further evidence of European admixture in Filipinos.


One outcome of interspecific hybridization and subsequent effects of evolutionary forces is introgression, which is the integration of genetic material from one species into the genome of an individual in another species. The evolution of several groups of eukaryotic species has involved hybridization, and cases of adaptation through introgression have been already established. In this work, we report on PhyloNet-HMM - a new comparative genomic framework for detecting introgression in genomes. PhyloNet-HMM combines phylogenetic networks with hidden Markov models (HMMs) to simultaneously capture the (potentially reticulate) evolutionary history of the genomes and dependencies within genomes. A novel aspect of our work is that it also accounts for incomplete lineage sorting and dependence across loci. Application of our model to variation data from chromosome 7 in the mouse (Mus musculus domesticus) genome detected a recently reported adaptive introgression event involving the rodent poison resistance gene Vkorc1, in addition to other newly detected introgressed genomic regions. Based on our analysis, it is estimated that about 9% of all sites within chromosome 7 are of introgressive origin (these cover about 13 Mbp of chromosome 7, and over 300 genes). Further, our model detected no introgression in a negative control data set. We also found that our model accurately detected introgression and other evolutionary processes from synthetic data sets simulated under the coalescent model with recombination, isolation, and migration. Our work provides a powerful framework for systematic analysis of introgression while simultaneously accounting for dependence across sites, point mutations, recombination, and ancestral polymorphism.

1894M Rapid radiation of common Eurasian Y chromosome haplogroups occurred significantly later than the out of Africa migration. M. Jâne, International Consortium of the Estonian Centre for Genomics. Estonian Biocentre and Department of Evolutionary Biology, University of Tartu, Tartu, Estonia.

Human genetic diversity outside Africa is low, which is commonly ascribed to a recent out of Africa bottleneck and subsequent rapid colonization of the rest of the world. Previous studies of the male-specific Y chromosome have shown that haplogroups common throughout non-African populations all coalesce to a small number of shared ancestral lineages, the branching order of which is only partly understood. Using 475 high coverage whole Y chromosome sequences, including 317 newly reported here, we selected reliable regions within the Y chromosome based on coverage analysis, mappability and sequence class. Based on these data, we refined the Y chromosome haplogroup tree, applying phylogenetic methods to establish the branching order and temporal dynamics of splits in non-African Y chromosome haplogroups. Compared to the length of the branches that separate African and non-African diversity, the internal branches distinguishing continental and sub-continental differences outside Africa are generally short, consistent with the model of a rapid initial colonization of Eurasia and Oceania. Following the split between African and non-African haplogroups [90 KYA (95% CI: 87-94 KYA)], the differentiation of South and Southeast Asian haplogroups H, S, M, and C did not begin until around 43 KYA, and haplogroups N and R, widely spread among Northeast and Northwest Eurasian populations, started to diversify significantly later [17 KYA (95% CI: 16-19 KYA) and 26 KYA (95% CI: 25-28 KYA), respectively]. Many major phylogenetic groups in different geographic regions seem to originate from the period around 50 KYA.
1895M
High degree of admixture in an urban Brazilian population. M.B. Melo, G. Ananina, M.A. Bezerra, A.S. Araujo, F.P.S. Cruz, M. Simonet, V.L. Gil da Silva Lopes, F.F. Costa. 1) CBMEG, Univ Campinas, Campinas SP, Brazil; 2) Haematology and Hemotherapy Center of Pernambuco/HEMOPE, Recife, PE, Brazil; 3) Department of Medical Genetics, FCM, Univ Campinas, Campinas SP, Brazil; 4) Hematology and Hemotherapy Center/HEMOCE, USP, Osasco, SP, Brazil.

The Brazilian population is the largest Latin American population rather distinct from others by their demographic and cultural history; however, it is still poorly studied genetically. Absence of a detailed genetic profile is an obstacle for many genetic studies. Here we report the results of a genomewide study of a sample of the Brazilian population. Subjects (n=62) were recruited at two Brazilian cities: Campinas (SP) and Recife (PE). This study was approved by the Ethics Committee (FCM-UNICAMP, Campinas, SP, Brazil) in accordance to the Basic Principles of the Declaration of Helsinki. The genotyping protocol was carried out using Affymetrix® Genome-Wide Human SNP 6.0 Array (Affymetrix inc., CA, USA) following manufacturer's recommendations. Genotyping Console v.4.1 was employed to obtain genotype calls. Publically available population data (15 populations) of HapMap and 1000 Genomes Project were used for comparisons. All SNPs were coarsely to the common for all populations list and quality control filters were applied per-individual and per-marker: exclusion of related individuals and individuals with high genotype missing rate (<5%); exclusion thresholds for a SNP - MAF>5%, HWE P-value<1e-4, per SNP call rate <5%; Mendelian errors in any of the population under consideration. Basic manipulations of the data were performed with the aid of PLINK software; population structure and admixture were analyzed applying EIGENSOFT and ADMIXTURE software. All Brazilian samples included in the study were tagged with the population southern from Lima in any of these projects. The ChileGenomico project aims to close this gap for the Chilean population by creating a detailed characterization of its genetic structure. We collected informed consented DNA from over 3000 volunteers born in 208 districts from Arica (northern Chile) to Punta Arenas, i.e. 4000 km of continental Chile. Our first data release includes whole-genome sequencing and genotyping with the Axiom Genome-Wide LAT 1 Array for individuals selected for putative indigenous ancestry (n=45) and the population from the Ñandeño lineage (n=39). Sequencing at the National Center for Genomics and Bioinformatics of Chile gave a 3.6x average, covering 88% of the genome. We identified 4,716,457 SNPs (260,367 novel) and 544,079 INDELs (266,973 novel) using GATK, Precision (0.98), recall (0.74), false discovery rate (0.03), and accuracy (0.83) for our SNPs predictions. Genotype imputation with 66 Mexican and 60 Colombian sequenced at low coverage by 1000G increased accuracy (0.83) for our SNPs predictions. All Brazilian samples passed inclusion control. We started with 840,856 markers. After pruning procedure, we obtained 297,115 wares. PCA analysis was applied to the groups of 42-45 individuals (randomly picked up from the sample) demonstrating that Brazilian samples are gradually distributed between the populations of European and African origins. Admixture analysis identified 7 groups in our dataset revealing a highly admixed pattern for the Brazilian sample. The Brazilian population is highly admixed, challenging for genetic analysis. Studies of extended sample size are desirable for a better characterization of the degree of admixture in the whole population and in regional groups. Financial support: CNPq, Brazil (83672/2011-1, 150398/2013-1); FAPESP, Brazil (2008/57441-0).

1897S
Whole genome association and genetic admixture analysis of EEG phenotypes in a Native American community sample. Q. Peng, N.J. Schork, K.C. Wilhelmsen, C.L. Ehlers. 1) J. Craig Venter Institute, La Jolla, CA; 2) University of North Carolina, Chapel Hill, NC; 3) The Scripps Research Institute, La Jolla, CA.

PURPOSE: Resting EEG features are shown to be genetically influenced. EEG patterns, such as low & high voltage alpha activity, remain highly stable over most of an adult lifespan and have been associated with phenotypes such as addictions. Genetic studies of complex phenotypes such as EEG & addictions can be advantageously conducted in well-defined populations with complex admixture. The purpose of this work is to explore the admixture & risk factors for substance dependence in a NA community. We previously reported that EEG alpha power is highly heritable in this community (h² = 0.62 frontal, 0.67 posterior scalp). We also observed that low voltage alpha (LVA) was correlated with higher fronto-parietal connectivity (FPC). The current study aims to: 1) assess ancestry admixture influence on the EEG LVA & FPC; 2) conduct GWAS on the phenotypes in this NA population (N=612). METHODS. EEG was transformed using fourier analyses and the alpha power determined. Wavelet analysis was used to determine the degree of phase locking between frontal (FZ) & parietal (PZ) electrodes, a measure of connectivity. Blood derived DNA was sequenced using Illumina low-coverage WGS. Ancestry & degree of admixture were estimated with a genomewide panel of SNPs. Multiple linear regression was used to assess the relation between NA ancestry degree and EEG LVA & FPC. A variance component approach was used in the association test to control for admixture and familial relatedness. RESULTS. FPC showed significant correlation with NA ancestry (β=0.05±0.02, p=0.02). LVA was negatively associated with NA ancestry (β=-0.11±0.02, p=0.003). A variance component analysis was negatively associated with FPC with genome-wide significance (β=-0.37, p=4.5e-10). Nearly 20% of the NA samples were heterozygous on the variant, yet it is absent from dbSNP & 1000Genomes, suggesting its uniqueness to the NAs. Analysis of the ancestral background of the genome surrounding the variant showed it resides in a region of lower NA. Of individuals carrying the variant, FPC was uncorrelated with local NA. Of those with the wild-type genotype, FPC was significantly correlated with local NA (β=0.11±0.05, p=0.003) suggesting a polygenic background effect of the EEG phenotype. CONCLUSIONS. Overall higher NA ancestry degree is associated with higher FPC & negatively associated with LVA. A likely NA-specific variant was found to be significantly negatively associated with FPC. (Supported by AA10201, DA030976).
1889M
A Genome Wide Admixed Association Study of a Sleep Disturbance
Phenotype in Adults with Sickle Cell Anemia. C. Liu1, Z. Wang2, L. Dai2, A. Oguhebe1, C. Okung1, K. Vaughan1, D. Darriba1,2, J. Taylor1,2
1) Genomic Medicine Section, Hematology Branch, NHLBI, NIH, Bethesda, MD; 2) Center for Cancer and Blood Diseases, Childrens National Medical Center, Washington, DC.

Population admixture is variable which can lead to spurious associations in genome wide association studies. However, mapping by admixture linkage disequilibrium (MALD) is statistically advantageous for genome wide studies in admixed, rare disease populations. Sickle cell disease (SCD) is a Mendelian disease caused by mutation of beta-hemoglobin. SCD has systemic manifestations including hemolytic anemia, recurrent episodes of severe pain and premature mortality. Our recent work has shown that more than 70% of SCD adults have sleep disturbance defined by the Pittsburgh Sleep Quality Index (PSQI). A significant independent risk factor for sleep disturbance in SCD is more frequent hospitalizations for SCD pain crises. Prior twin and sibling studies in the general population suggest that sleep disturbance is a moderately heritable trait, although no data is available to estimate heritability in SCD. We hypothesized that sleep disturbance in SCD is a complex trait attributable to modifier genes that can be identified by MALD. To test this hypothesis, we identified ancestry informative markers (AIMs) from 3,804,602 SNPs genotyped by HapMap Phase III CEU and YRI populations. These SNPs were combined with a published admixture panel, yielding a new MALD panel of 1795 AIMs. We genotyped these AIMs in 221 HapMap subjects and 489 adults from the Bethesda Sickle Cell Cohort Study using Illumina iSelect arrays. Using our HapMap genotype data, this admixture panel has a mean χ2 (difference in allele frequencies between populations) of 0.707 (SD 0.114) with an average inter-marker distance of 1.692 Mb. We then performed a MALD association study of sleep disturbance, dividing them into cases with PSQI scores >5 or controls with lower PSQI scores. We identified a region on chromosome 13p21 approaching genome wide significance (defined by a locus genome score or LGS of 5 or more) and suggestive of a sleep disturbance association (locus genome score 4.007 at rs17088390). SNPs with a case control statistics >2 span a region of approximately 4 Mb on chromosome 13, including 1 within the DACH1 gene.

1889S
Use of Long-read-sequence Aided Phasing for Inference of Ancestry Assignment in Admixed Populations. F.L. Mendez1, S.S. Shringarpure2, A. Moreno2, E.R. Martin2, M.L. Cuccaro2, C.D. Bustamante1
1) Department of Genetics, Stanford University, Stanford, CA; 2) Center for Genetic Epidemiology and Statistical Genetics, University of Miami, Miami, FL.

Correct phase reconstruction of individual chromosomes is important for numerous genetic analyses, including inferring demographic parameters in admixture processes. Admixture is the result of interbreeding of previously differentiated populations. The chromosomes of admixed individuals are composed of segments that can be traced individually to one of the ancestral populations. The abundance and length distributions of these chromosomal segments provide crucial information on the admixture process; however, the correct inference of their length and ancestry requires phasing data from the admixed individuals. Phase reconstruction can be performed by applying the rules of Mendelian segregation of variants or using statistical methods that rely on population data. Alternatively, molecular phasing (the observation of different polymorphisms in the same chromosomal sequence) provides direct evidence of phase. In this fashion, methods of molecular-phasing, like long-read sequencing, may be used to extend the range of confident phasing. We simulate long-read sequence data based on previously discovered and a genetically admixed population.

1900M
1) John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL; 2) Macdonald Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, FL, USA; 3) Department of Neurology, University of California at San Francisco, San Francisco, CA, USA; 4) Division of Biostatistics, Miller School of Medicine, University of Miami, Miami, FL; 5) Program in Medical and Population Genetics, Broad Institute of Harvard University and MIT, Cambridge, MA, USA; 6) Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatric Epidemiology & Women's Hospital, Boston, MA, USA; 7) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA; 8) Multiple Sclerosis Division, Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA.

Multiple sclerosis (MS) is an inflammatory demyelinating disorder of the central nervous system with variable prevalence across populations. European populations have a higher prevalence and have traditionally been considered more susceptible to MS than either Hispanic or African. This could be due to genetic factors, a higher degree of genetic risk in European populations. Through genetic association studies in individuals of European descent, 110 established MS risk variants in 103 discrete loci outside of the Major Histocompatibility Complex have been identified. Our goal is to characterize the local ancestry at these 110 risk variants, first in a sample of European descent (1691 individuals affected with MS and 5542 controls), and second in a Hispanic sample (200 individuals affected with MS and 1000 population controls). We hypothesize that given the increased prevalence of disease in European populations; there will be an increased percentage of European ancestry in Hispanics affected with MS versus Hispanic population controls at these 110 risk variants. Individuals affected with MS and the controls of European descent have been genotyped using the ImmunoChip genotyping array. The Hispanic population controls have been genotyped using the Affymetrix High Density 6.0 genotyping array. Analyses of a random subset of the 110 risk variants show a difference in risk between Hispanics affected with MS and population controls (p=3.2E-05), indicating that these risk loci are also important in Hispanics. We do not see a significant difference in the risk score between Hispanics and European individuals affected with MS (p=2.0E-01). Support of the hypothesis that the ancestry at these loci is similar amongst cases from both populations. Haplotype phasing based on the reference populations of European, African, and Asian individuals from the 1000 Genomes Project was done using SHAPEIT, with LAMP-LD/LAMP-ANC, we are currently computing local ancestry using a representative subset of possible haplotype pairs and reference populations of Europeans and Africans from the 1000 Genomes Project as well as Native Americans from the Human Genome Diversity Project. Ancestry percentages at each locus will be compared, between individuals affected with MS and controls for samples of both European and Hispanic descent in order to better characterize the local ancestry of these confirmed loci both in the population in which they were discovered and a genetically admixed population.

1901S
Linear Mixed Model-Based Admixture Mapping. L. Brown, T. Thornton
Biostatistics, University of Washington, Seattle, WA.

Genetic studies in recent years have provided valuable insight into novel risk factors contributing to disease. Population admixture results in combined genomes from previously isolated ancestral populations that may have discernible allele frequency differences due to natural selection. We have recently shown that underlie ethnic differences in traits and that show differential risk by ancestry can be identified using admixture mapping. Compared to studies carried out in more ethnically homogenous populations, admixture mapping has potentially greater power to detect certain genetic variants. Linear mixed models have gained traction as a tool for genome wide association studies. Mixed model methods have shown to protect against spurious associations in structured samples, a common pitfall in genetic association studies, by directly accounting for sources of dependence including cryptic relatedness and population stratification. We present a linear mixed model approach for admixture mapping in the presence of population structure and hidden relatedness. We implement this method using local ancestry estimates based on genome-wide SNP data. We apply the method to analyze genetic associations with white blood cell counts and C-reactive protein levels in a genetically diverse cohort of the Women’s Health Initiative study. We demonstrate that our proposed linear mixed model method for admixture mapping provides a substantial improvement over widely used admixture mapping approaches.
1902M

Comparing high-coverage Denisova and Neanderthal whole-genome sequences has revealed significant admixture with all present-day non-African populations. Microblade tool usage from central-India has been reported, yet no genetic-study examined archaic admixture in present-day South-Asians. We report the first evidence of archaic admixture from whole-genome sequence data of 4 present-day Indians. Four individuals, who are at the extremities of a two-dimensional Principal-Component plot summarizing the extant of genomic variance in 237 Indians belonging to 20 linguistically, ethnically diverse populations sampled from different geographic locations in India; were sequenced. Their population identities are Onge, Jamatia, Panniya and Birhore. All individuals showed slight excess of Denisona admixture (D-Statistic 1.6-1.99) when compared with Eurasians, with admixture evidence increasing from Jamatia (1.6) to Onge (1.99). Similar pattern was also observed when compared with Neanderthal. Our findings show evidence of archaic admixture in different present-day populations, not restricted to proximity of archaeological evidence, indicating wide spread admixture.

1903S
Pharmacogenomic patterns for Brazilian and Mexican populations. V. Bonifaz1, A.V. Contreras1, C.J. Struchiner1, R.A. Roella2, T.K. Mazotti3, R. Chammas3, M.J. Gomez-Vazquez2, L. Uribe1, C. Rangel-Escareño1, H.L. McLeod1, B. Randan1, F. Cherende1, F. Romulo1, G. Huerta-Cepas1, C.J. Suarez-Kurtz2, 1 Instituto Nacional de Medicina Genómica, Mexico, Mexico City, Mexico; 2 Programa de Computación Científica, Fundación Oswaldo Cruz, Rio de Janeiro, RJ, Brazil; 3 Laboratório de Oncología Experimental—LIM24, Departamento de Radiología e Oncología, Facultade de Medicina, Universidade de São Paulo, SP, Brazil; 4 DeBartolo Family Personalized Medicine Institute, Tampa, FL, U.S.A; 5 Pharmacogenetics for Every Nation Initiative, Tampa, FL; 6 Department of Anthropology, University of Toronto at Mississauga, Mississauga, ON, Canada; 7 División de Farmacología, Instituto Nacional de Cáncer, Rio de Janeiro, RJ, Brazil.

Latin America is a vast geographic region that is home to more than 600 million people. To a large extent, the current Latin American population reflects a complex history of admixture between people with ancestral roots in the Americas, Europe and Africa. Studies for pharmacogenomics-relates traits are increasingly being performed to identify loci that affect either drug response or susceptibility to adverse drug reactions. However, the effect of a polymorphism can differ in magnitude or be absent depending on the population being assessed. In this work we present the potential impact of admixture on the distribution of genetic markers of pharmacogenetic and pharmacogenomics (PGx) relevance. Differences in admixture history may have an important impact in the distribution of allele and genotype frequencies at population level. Therefore we aim at characterizing the distribution of polymorphisms of PGx relevance of two of the most populous Latin American countries: Brazil and Mexico. The Brazilian sample includes 268 individuals stratified into census categories Branco (White), Pardo (Brown) and Preto (Black). The Mexican sample comprises 45 Native American Zapotecos and 224 self-identified Mestizo individuals located in geographically distant regions. Both samples were genotyped using Affymetrix Drug Metabolizing Enzymes and Transporters (DMET) Plus array. This platform includes 1936 variants, rare and common, in 231 genes involved in drug pharmacokinetics. Our results highlight the complex population history of these two countries. We observed differences in admixture proportions according to geography (Mexico) and census categories (Brazil). Although the overall genetic differentiation of the markers interrogated in the DMET Plus array is low, there are many loci that show high levels of genetic differentiation between the parental populations relevant for contemporary Brazilians and Mexicans. The VKORC7 rs9923231A, a major determinant of warfarin dose in CYPIC guidelines, is an excellent example of the practical PGx implications of admixture history. The frequency of the rs9923231A allele, which associates with high warfarin sensitivity, varies 13-fold (52.2% frequency in Africa and 4.1% in Zapotecos) among proxy parental populations of the Brazilians and Mexicans. Differences in admixture history among Latin American populations may have an important impact in allele frequency distribution and therefore also in haplotypes associated to drug metabolism.

1904M
The genetic ancestry of African, Latino, and European Americans across the United States. K. Bry1, E. Durand2, D. Reich3,4,5, J. Mountain1. 1) 23andMe, Inc., Mountain View, CA; 2) Harvard Medical School, Department of Genetics, Boston, MA, USA; 3) Howard Hughes Medical Institute, Boston, MA, USA; 4) Broad Institute of MIT and Harvard, Cambridge, MA, USA.

Within the past 500 years, North America has been the site of a dynamic mixing of people from populations that were previously separated by oceans and other geographic features. The interactions between Native Americans, European settlers, and Africans brought to the New World via the Transatlantic slave trade shaped both the history of North America and the United States. We studied the genetic ancestry of 5,269 self-reported African Americans, 8,633 Latins, and 1,789 European Americans who are 23andMe customers living in the US and show that the legacy of these interactions is visible in the genetically-inferred ancestry of modern Americans. We shed light on the unique regional differences in genetic ancestry within and across the United States. In addition to well-established variability in individual ancestry proportions, we detect systematic clines in the amount of African and European ancestry in African Americans from different states. We demonstrate that the pervasive mixed ancestry in modern Americans and the relationship between self-reported identity and genetic ancestry reflect regional social and political history. We find that levels of Native American and African ancestry in European Americans, Latinos, and African Americans are highly correlated with the population density of African Americans and Latinos in each state with one notable exception.

We provide evidence that a minor, but measurable, proportion of self-reported European Americans carry African ancestry. Likewise, we demonstrate that a measurable, though not appreciable, proportion of European Americans, both in the US and in Europe carry an admixture signal of African ancestry. The genetic ancestry within each of these combined groups of admixed European Americans varies from a model of complete European ancestry fits a model of admixture on the distribution of genetic markers of pharmacogenetic and pharmacogenomic (PGx) relevance. Differences in admixture history may be differentially restricted to the Iberian Peninsula. The distribution of chromosomal segments of distinct ancestry do show evidence of recent European admixture into the South and Southwestern US, but not in the Salvadoran cohort (NE), while African admixture into admixed populations fits a model of continuous gene flow across the three cohorts. By the genome sequencing we identified between 3.6 M and 4.4 M of autosomal SNPs per each individual, and the high levels of admixture of Brazilians allowed us to identify ~1.4 M of new autosomal SNPs. The EPiGEN Initiative is also performing several GWAS and admixture mapping studies on different complex traits, including longitudinal data. Funding: Brazilian Ministry of Health/FINEP.
Infering patterns of demography and assortative mating in the Thousand Genomes Project admixed populations from the Americas. E. C. Kenny1,2, C. Gignoux3, S. Baharian4, M. Maharaj5, B. Maples6, S. Shingarpure7, A. Auton8, C. D. Bustamante9, S. Gravel10, A. R. Martin2, The 1000 Genomes Consortium. 1) Department of Genetics, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Department of Genetics, School of Medicine, Stanford University, California, CA; 3) Department of Human Genetics, McGill University, Montreal, Canada; 4) Department of Genetics, Albert Einstein College of Medicine, New York, NY; 5) Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY; 6) Center for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY; 7) Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY.

The phase 3 release of the 1000 Genomes project includes genotype and sequence data for 535 individuals from 26 populations around the globe. These include six populations from the Americas with mixed Native American, European, and West African ancestry. We have identified admixture proportions in these six populations, which include African Caribbean’s from Barbados (ACB), African American’s from south-west USA (ASW), Colombians from Medellin (CLM), Peruvians from Lima (PEL), Mexican American from Los Angeles (MXL), and Puerto Ricans from Puerto Rico (PUR). We show presence of Native American, European and African ancestry in all six populations, in particular the LAC, six ASW individuals > 20% Native American ancestry. The European component of these individuals looks most similar to Nordic ancestry, rather than Spanish ancestry often seen in Hispanic/Latino individuals. Among the ACB, PEL, PUR, CLM, and MXL populations, we find an excess of Native American and dicket of European ancestry. When we compare the truth admixture proportions with the estimated admixture proportions, we find a history of non-random mating in these populations. We have also inferred local ancestry tracts (LAT), identifying haplotype specific segments of ancestry across chromosomes. We assessed the accuracy of our tract calls using simulated samples and generated accuracies >0.90, >0.98 and >0.97 in African, European and Native American tracts across all populations. By modeling the distribution of ancestral tract lengths, we inferred the timings of migration in the two populations from the Americas’s that are new to phase 3, ACB and PEL. We estimated the PEL have had more recent admixture with European and African individuals than other Hispanic/Latino groups in the Caribbean and throughout northern South America, consistent with known migration patterns. These analyses have given us an insight into the demographic history and migration patterns among admixed populations in the 1000 Genomes Project. 

Genotype and allele frequencies of RETN -420 C/G polymorphism in three Mexican native populations. A. López-Quintana1, V. Rojas-Cardenas1, A. García Zapíen1, SE. Flores-Martínez1, S. Islas Andrade1, MC. Revilla Monsalve1, J. Sánchez Corona1. 1) Centro de Investigación Biomédica de Occidente (CIBIO), División de Medicina Molecular, Sierra Mojada 800, GDL, Jal; México, 2) Universidad de Guadalajara, Centro Universitario de Ciencias de la Salud (CUCS), Doctorado en Genética Humana, Sierra Mojada 950, GDL, Jal; México; 3) Unidad de Investigación en Enfermedades Metabólicas, Centro Médico Nacional Siglo XXI, IMSS, Belisario Domínguez 1000, Cd. Mex, México. The polymorphism -420 C/G in RETN gene is associated with increased resistin levels and type 2 Diabetes Mellitus. Objective. To determine genotypic and allelic frequencies of RETN -420 C/G polymorphism in individuals from three Mexican native populations: Lacandones (Lacan), Ocósingo, Bethel, Nahá), Yaquis and Tepehuanos. Material and methods. The present study included 419 DNA samples of individuals from three Mexican indigenous populations. The RETN C/G polymorphism was analyzed by PCR-RFLP. After a 331 bp fragment amplification by PCR, the products were digested with the BbsI enzyme restriction to identify the genotypes. The digested products were separated by electrophoresis in 6% polyacrylamide gels and identified by silver staining. Differences between genotype (GF) and allele frequencies (AF) among the three populations were assessed by using a χ2 test with the RXC program and 10000 iterations. There were statistically significant differences among the Lacandones community when assessed by genotype and allele frequencies; Lacan (0.006), Ocósingo (0.007) and Bethel (0.005), respectively. The Yaquis population displayed the genotype frequencies of 0.0004, 0.0007 and 0.0089 for GF and p=0.001, p=0.001 and p=0.001 in case of AF. The Tepehuanos population was statistically different when compared the GF (p=0.0005), and AF (p=0.0006) with Bethel community. The Yaqui population displayed the genotype frequencies of 0.0006, 0.0015 and 0.0089 for GF and p=0.0023 and 0.0072 in case of AF respectively. Conclusion. The genotypic and allelic distribution of the RETN -420 C/G polymorphism is different among the analyzed populations. Furthermore, it is not possible to use these four Lacandones communities and this must be considered in further association studies for the Lacandones community. 

Sub-continental local ancestry inference in U.S. individual. B.K. Maples1, J.K. Byrnes2, J.M. Granka3, K. Noto4, S. Shingarpure1, M.L. Carpeneter1, M.J. Barber4, R.E. Curtis2, N.M. Myers4, C.A. Baill4, K.G. Chapple4, C.D. Bustamante1. 1) Genetics, Stanford University, Stanford, CA; 2) Biomedical Informatics, Stanford University, Stanford, CA; 3) AncestryDNA L.L.C., San Francisco, CA; 4) AncestryDNA L.L.C., Provo, UT.

The phase 3 release of the 1000 Genomes project includes genotype and sequence data for 535 individuals from 26 populations around the globe. This “melting pot” process has led to the majority of current U.S. residents being genetically admixed. Understanding this complex genetic diversity is of great importance for the development of a clinical risk stratification for genetic diseases. Numerous methods have been developed for performing local ancestry inference (LAI) in which the ancestry of each genomic locus is estimated, but the majority of these methods are only accurate at the population level (e.g. admixed populations with ancestry from Africa and Europe). Sub-continental LAI is often more difficult as neighboring populations typically have reduced differentiation. In this study we apply the LAI method RFMix that has been shown to perform well at a sub-continental level (e.g. mixtures of Northern and Southern European ancestry). RFMix is seeded with a reference panel of samples with known origins, and then iteratively learns from a larger collection of test samples. The performance of this method greatly improves with larger reference panel and test sample sizes. Here we use more than 2,000 single-origin reference samples from Ancestry.com and 1000 Genomes, along with over 100,000 research-converted customer samples with admixed origins to train the model to perform inference on individuals with admixed European ancestry. We compare the genome-wide ancestry estimates from RFMix with pedigrees. Using pedigrees as the “ground-truth” we can compare the performance of RFMix with results from the commonly used ancestry estimation method ADMIXTURE run in supervised mode with the same initial single-origin reference panel provided to RFMix. We also use single-origin reference samples to validate our LAI estimates of ancestry patterns to assess the accuracy of RFMix to call individual segments in admixed Europeans. Finally, we apply the highly trained version of RFMix to the National Institute on Aging’s Health and Retirement Study data (HRS) and compare population level geographic summaries of sub-continental ancestry estimates in this data to recent U.S. census data. We find strong evidence of fine-scale population structure with certain localities showing enrichment for particular ancestries (e.g. Irish ancestry in and around Boston and Scandinavian ancestry in the Midwestern states).

Drift and selection contribute to elevated susceptibility for childhood acute lymphoblastic leukemia in individuals with Native American ancestry. A. Lopez-Quintana1, V. Rojas-Cardenas1, A. Garcia Zapien1, SE. Flores-Martinez1, S. Islas Andrade1, MC. Revilla Monsalve1, J. Sanches Corona1. 1) Genetics, Stanford University, Stanford, CA; 2) Universidad de Guadalajara, Centro Universitario de Ciencias de la Salud (CUCS), Doctorado en Genética Humana, Sierra Mojada 950, GDL, Jal; México; 3) Unidad de Investigación en Enfermedades Metabólicas, Centro Médico Nacional Siglo XXI, IMSS, Belisario Domínguez 1000, Cd. Mex, México.

Acute Lymphoblastic Leukemia (ALL) is the most common childhood cancer and one of the leading causes of death in children and adolescents. ALL is characterized by young age at disease onset (particularly in children between 2-5 years of age), which points to a strong genetic basis of disease etiology. Cross-population studies have shown that multiple ancestral backgrounds have identified multiple common variants predisposing children to ALL with relatively large effect sizes. (Trevisó et al. Nat Genet, 2009, Xu et al. JNCI 2013). Interestingly, ALL is more prevalent and severe in populations with high genome-wide Native American (NA) ancestry than those of any other ancestry type (Yang et al. Nat Genet 2011). This suggests that the evolutionary history of NAs has played a role in the high incidence and inferior treatment outcome of ALL in NA populations. Here we analyze germline genotype data from 232,000 NA individuals, many of whom identify as Native American, consisting of children with ALL and their unaffected parents, genotyped with Affymetrix SNP 6.0 and Illumina Exome arrays to better understand the genomic signatures of demographic history and selection on NAs populations, particularly at loci that contribute to genetic predisposition to ALL. Genome-wide association studies (GWAS) in NAs have shown evidence for adaptive selection at rs4982731 (CEBPB) with a minor allele frequency difference in the top 2% between Guatemalans and Asians and an Integrated Haplotype Score (in the top 5.25% of scores indicating selection). We performed a follow-up analysis on four large datasets with a high level (50%) of ALL risk alleles in Native Americans, in contrast, may have resulted from selection for the derived allele). The elevated allele frequencies of other ALL risk alleles in Native Americans, in contrast, may have resulted from selection for the derived allele.)

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Bayesian Computation.

C.D. Quinto

factors other than admixture such as diabetes and socioeconomic status. But, in a fully adjusted model, which included African Admixture, diabetes, hypertension and BMI were not associated with glaucoma. This study shows that, Diabetes and BMI (OR= 1.67) in age-adjusted model. However, this relationship had become not statistically significant after adjusting for socioeconomic status, hypertension, diabetes and BMI (OR= 1.24; 95% CI= 0.91-1.68). In contrast, there was a significant association between diabetes and glaucoma (OR= 1.51; 95% CI= 1.37-1.67) in age-adjusted model. This association persisted after the adjustment for admixture, socioeconomic status, hypertension, and BMI (OR=1.52; 95% CI= 1.37-1.68). In addition, lower socioeconomic status was significantly associated with higher risk of glaucoma (OR=1.07; 95% CI= 1.02-1.12).

But, in a fully adjusted model, which included African Admixture, diabetes, hypertension and BMI, the association between socioeconomic status and glaucoma was marginally significant (p=0.04), while hypertension and BMI were not associated with glaucoma. This study shows that, Diabetes and socioeconomic status, but not the African admixture, were found to be independently associated with glaucoma status. These finding suggest that the high frequency of glaucoma in African Americans may largely due to factors other than admixture such as diabetes and socioeconomic status.

Inference of the demographic history of Japan using Approximate Bayesian Computation. C.D. Quinto1, K.R. Veeramah2, A.E. Woerner3, M.F. Hammer4, 1) Graduate Interdisciplinary Program in Genetics, University of Arizona, Tucson, Arizona, USA; 2) Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, USA; 3) Arizona Research Laboratories, Division of Biotechnology, University of Arizona, Tucson, Arizona, USA.

The genetic exchange between differentiated populations, termed admixture, has increasingly been shown to be an important process in human history. The formation of Hispanic populations in the Americas is one of the best-known examples of this phenomenon. Another important but less well-known example is the origin of modern Japanese. At least two distinct incoming migrations are known to have occurred during the prehistory of Japan. The first took place at least 10,000 years ago and established the Jōmon culture, which was characterized by a semi-sedentary hunter-gatherer way of life and one of the earliest uses of ceramics. Then, around 2,300 years ago, a second migration to the archipelago brought rice agriculture and iron, and established the Yayoi culture. The mixture of the people belonging to these cultures is best known as the modern Japanese population. Although archaeological records provide information about the time of arrival of the Yayoi people to Japan, the dynamics of the admixture process are still unclear. Previous genetic studies, focusing on mitochondrial DNA and the Y chromosome, have supported an admixture model for the origin of the modern Japanese population. While genome-wide data have been used to investigate this question, there are currently no studies that infer the parameters describing the dynamics of the admixture process. Part of the reason for this is that explicit population genetic models do not usually account for the effects of ascertainment bias in the choice of SNPs. To address these issues, we genotype 500,000 SNPs in 282 samples from populations across the Japanese archipelago and East Asia. We then attempt to correct for the admixture bias by using whole genome sequencing data to approximate the discovery sample used to ascertain SNPs. We utilize the SNP genotypes from the different populations to identify ancestry blocks in Japanese samples. The distribution of these blocks provides insights about the time and admixture admixture and the demographic history can be inferred in an Approximate Bayesian Computation analysis to infer other key demographic parameters such as divergence times, migration rates and population sizes.

A Fine-Scale Comparative Analysis of Population Structure, Divergence and Admixture in Han Chinese, Japanese and Korean Populations. S. Xu1, Y. Jiang1, D. Lu2, Y. Chong3, 1) Population Genomics, CAS-MPG Partner Institute for Computational Biology, Shanghai, Shanghai, China; 2) Integrated Research Center for Genome Polymorphism, Department of Microbiology, The Catholic University Medical College, Socho-gu Seoul 137-701, Korea. n East Asia, human origins and dispersals remain poorly understood and debatable. As the major ethnicities of East Asia, Han Chinese, Japanese and Korean people share many similarities in characteristics, language and culture. However, the genetic relationships, divergence times and subsequent gene flow among the three populations have not been well studied or quantitatively estimated. Here, we conducted a genome-wide study using over 900,000 single nucleotide polymorphisms (SNPs) and evaluated the population structure of 182 unrelated Han Chinese, 90 Japanese and 100 Korean individuals, and compared with 663 individuals representing 8 worldwide populations. Our analysis revealed that Han Chinese, Japanese and Korean populations have distinct genetic makeup and can be well distinguished based on the genome wide data, or a panel of ancestry informative markers (AIMs) screened from genome-wide SNPs, indicating they have been isolated for substantially long time. Interestingly, population structure is perfectly corresponding to the geographical distribution of the three populations, indicating geography was an important factor resulted in population differentiation. We identified acline of north/south admixture, which is consistent with either a scenario of isolation by distance (IBD) or that of north/ south migrations or both. We theorized that both IBD effect and migrations could have resulted in such a pattern. On the other hand, our analysis revealed patterns of admixture which occurred after initial splits of populations. We further estimated gene flows among the three populations. We concluded that the genetic structure of the present-day Han Chinese, Japanese and Korean people was shaped jointly by common origin, subsequent dispersal and admixture. Our results advance the understanding of the genetic relationship and population history in East Asia.

Analysis of autosomal and Y-chromosomal DNA Suggests West Asian Population Derivation from Northern Middle Eastern Populations in the post-Glacial Period. P. Zalloua1, F. Utro2, M. Haber1, L. Panda3, E. Matsiko-Smith3, D. Platt4, 1) Genomics Laboratory, Gradute School, Beirut; Lebanon; 2) Harvard School of Public Health, Boston, MA, USA; 3) I.B.M. T. J. Watson Research Center, Yorktown Hgts, NY; 4) University of Otago, Dunedin 9054, New Zealand.

Analysis of Y DNA J and E haplogroups in West Asians (Georgians, Armenians, Turks, Syrians, Lebanese, Jordanians, Saudi Arabsians, Yemenis, Kuwaitis) suggests expansions coming primarily from the north (Turkey, Georgia, Armenia), with an early differentiation between those who headed south along the Tigris-Euphrates, versus those who headed south along the Levantine coast. We sought to resolve whether southern variations represent gene flow and local adaptation. Our results advance the understanding of the genetic relationship and population history in East Asia.

S. Xu1, Y. Jiang1, D. Lu2, Y. Chong3, 1) Population Genomics, CAS-MPG Partner Institute for Computational Biology, Shanghai, Shanghai, China; 2) Integrated Research Center for Genome Polymorphism, Department of Microbiology, The Catholic University Medical College, Socho-gu Seoul 137-701, Korea. n East Asia, human origins and dispersals remain poorly understood and debatable. As the major ethnicities of East Asia, Han Chinese, Japanese and Korean people share many similarities in characteristics, language and culture. However, the genetic relationships, divergence times and subsequent gene flow among the three populations have not been well studied or quantitatively estimated. Here, we conducted a genome-wide study using over 900,000 single nucleotide polymorphisms (SNPs) and evaluated the population structure of 182 unrelated Han Chinese, 90 Japanese and 100 Korean individuals, and compared with 663 individuals representing 8 worldwide populations. Our analysis revealed that Han Chinese, Japanese and Korean populations have distinct genetic makeup and can be well distinguished based on the genome wide data, or a panel of ancestry informative markers (AIMs) screened from genome-wide SNPs, indicating they have been isolated for substantially long time. Interestingly, population structure is perfectly corresponding to the geographical distribution of the three populations, indicating geography was an important factor resulted in population differentiation. We identified acline of north/south admixture, which is consistent with either a scenario of isolation by distance (IBD) or that of north/ south migrations or both. We theorized that both IBD effect and migrations could have resulted in such a pattern. On the other hand, our analysis revealed patterns of admixture which occurred after initial splits of populations. We further estimated gene flows among the three populations. We concluded that the genetic structure of the present-day Han Chinese, Japanese and Korean people was shaped jointly by common origin, subsequent dispersal and admixture. Our results advance the understanding of the genetic relationship and population history in East Asia.
1914M A method to use control data and exploit the structure of genetic ancestry space to enhance case-control studies. C. Bodea1, S. Ripke1,2, B. Neale1,2, M. Daly3, B. Devlin2,3, E. Sanchez-Rodriguez, B.A. Pons-Estel, E. Acevedo, J.M. Cuacho, P. Miranda, L. Catoggio, M.A. Garcia, G. Berbotto, A. Babin, H. Scherbarth, S. Toloza, M. Alarcon-Riquelme, C.D. Bustamante1. 1Department of Genetics, Stanford University, Stanford, CA, USA; 2Center for Human Genomics and Oncological Research (GENYO), University of Granada, Granada, Spain; 3Sanatorio Alarcon-Riquelme, Lima, Peru; 4Hospital Italiano de Buenos Aires, Argentina; 5Hospital Italiano de Córdoba, Córdoba, Argentina; 6Hospital de la Plata, Buenos Aires, Argentina; 7Hospital Evà Peron, Granadero Baigorria, Argentina; 8Hospital Italiano de Córdoba, Córdoba, Argentina; 9Hospital Italiano de Buenos Aires, Argentina; 10Hospital Italiano de Córdoba, Córdoba, Argentina; 11Hospital Italiano de Buenos Aires, Argentina.

The South American continent has experienced multiple migration and admixture events. Here, we examine the genetic history of the Andean region using 551 individuals from Colombia, Ecuador, Peru, Chile, and Argentina. We performed a small-scale association test on the POPRES data (Nelson et al., Am J Hum Genet, 2008) based on simulated signal of varying risk and allele frequency and found that POPRES delivers strongly superior results over a traditional matched control setup, even when the matched controls greatly outnumber cases. We have also replicated the results of a large (5956 cases, 14927 controls) GWAS on Crohn’s disease (Jostins et al., Nature, 2012) via the UNICORN framework by computing the matched allele frequency estimates from the study’s control set. This result highlights that even big studies with ample control collections can benefit from UNICORN.

1916M Fast individual ancestry inference from DNA sequence data leveraging allele frequencies from multiple populations. O. Libiger4,3,2, V. Bansal1,2. 1Scripps Translational Science Institute, La Jolla, CA; 2Department of Pediatrics, University of California San Diego, La Jolla CA; 3MD Revolution, San Diego, CA.

Estimation of individual ancestry from genetic data is useful for the analysis of disease association studies, understanding human population history and interpreting personal genomic variation. We describe a fast method for estimating the relative contribution of known reference populations to an individual’s genetic ancestry. Our method utilizes allele frequencies from the reference populations and individual genotype or sequence data to obtain a maximum likelihood estimate of the global admixture proportions using the BFGS optimization algorithm. It accounts for the uncertainty in genotypes present in sequence data by using genotype likelihoods instead of genotypes. Unlike previous methods, our method does not require individual genotype data from external reference panels and can utilize allele frequencies estimated from the analysis of homogeneous as well as admixed human populations. Simulation studies and application of the method to real data sets demonstrate that our method is 8-10 times faster than ADMIXTURE and has comparable accuracy. Using data from the 1000 Genomes project, we show that our method can estimate genome-wide average ancestry of admixed individuals using exome or low-coverage sequence data. Finally, we demonstrate that our method can be used to estimate admixture proportions using pooled sequence data making it a valuable tool for controlling for population stratification in sequencing based association studies that utilize DNA pooling.

1917S Molecular and cytogenetic analysis of inversions in human and Great Apes. M. Miroballo, A. De Magis, F. Antonacci. Department of Biology, University of Bari, Bari, Italy.

In the last few years the burst of whole genome techniques allowed the detection of all kinds of unbalanced structural variations, while balanced rearrangements, such as inversions, are still hardly discovered. Several studies described inversions as a very frequent interspecific rearrangement and for their high potential power in suppressing recombination they are generally considered as one of the main genomic driving forces among evolutionary mechanisms. The prominent role of inversions coupled with the lack of high-throughput techniques for their detection, force research to take advantage of prediction methods that may imply the possibility to obtain false positives that must undergo experimental validation. In the present work-in-progress we resume a previously published study (Feuk et al., 2005) about the prediction and validation of inversions in human and chimpanzee genomes extending the analysis to the rest of the Great apes. First, we investigated by PCR and FISH a total of 19 predicted inversions in a panel of 16 samples, four per each species. For those regions both direct and inverted orientations in one species, we enlarged the investigations to more individuals in order to evaluate their occurrence and stratification in different geographical areas of origin. 11 out of 19 regions were confirmed as previously published: three inversions occurred in the chimpanzee lineage, six inversions occurred in the human lineage (with one region confirmed as polymorphic) and two were false positives, showing the same orientation in the two species. One of the three previously validated inverted regions in P'TR (7q21.3), was found to be polymorphic among the chimpanzee subspecies: in western chimpanzees the region was found to be inverted with a 67% allele frequency, while the eastern chimpanzees uniquely showed the direct orientation. We also characterized seven predicted inversions found to be false positives, mainly due to the presence of inverted duplications at the breakpoints of the predicted rearrangements. Further analysis will include at least three steps: (i) selection and validation of additional predicted inversion regions to increase the reliability of the predictive method; (ii) sequence analysis of the inversion breakpoints to investigate the molecular mechanisms of inversions; and (iii) study of genes mapping in the inverted regions to investigate their role in the recent evolutionary history of Great apes.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1918M
Identification of pleiotropic association signals in multiple autoimmune diseases at 2q24. J.E. Molineros, C. Sun, S.K. Nath. Oklahoma Medical Research Foundation, Oklahoma City, OK.

Overlapping etiological factors among the autoimmune diseases (ADs) have been suspected owing to shared clinical and immunological characteristics, and familial aggregation of multiple ADs. As a follow up to our previous study of Systemic Lupus Erythematosus (SLE) association on 2q24, where we identified three functional independent variants in IFIH1, we performed association analysis on Six additional autoimmune diseases (type 1 diabetes N=823, celiac disease N=1716, vitiligo N=1089, psoriasis N=919, rheumatoid arthritis, N = 1999 and systemic sclerosis N=835). One common control dataset (N=8833) was used to facilitate comparisons of this study. Given the different SNP densities in each dataset, missing SNPs were imputed using MACH. We focused on the genomic region including genes GCG, GCA, FAP, IFIH1 and KCN7. We performed conditional/logistic regressions to identify which SNPs were statistically independent. We were able to replicate association (P<0.05) of our three reported IFIH1 variants (rs1990760, rs10930046 and rs10232380) in all ADs, although not all three in all of them. Surprisingly, we identified stronger association signals originating from potassium channel gene KCN7, that after conditional analyses explained much of the association in the region for celiac disease (rs867824 P=3.0x10^-7), psoriasis (rs2389739 P=6.6x10^-10), rheumatoid arthritis (rs12623464 P=9.31x10^-17) and type 1 diabetes (rs1352071, P=4.8x10^-3), whereas the strongest signal for systemic sclerosis originated at GCA (rs17783344 P=2.8x10^-10), and at FAP in the case of vitiligo (rs13422767 P=4.9x10^-6). Bioinformatic analysis of these independent variants, identified eQTLs as well as transcriptional regulatory elements. Thus, together with IFIH1, another KCN7 and GCA, KCN7-GCA points towards the involvement of 2q24 in susceptibility for multiples ADs.

1919S
1) Human Genetic Research cum Counselling, Jammu, J & K, Jammu, India; 2) University of Nottingham, United Kingdom.

Human genomic diversity is the result of differential accumulation of genetic variations in individuals and populations throughout the evolution. The identification of such distinctive characteristics in the DNA represents the basis of human identification, genetic diversity and population genetics. The state of Jammu and Kashmir, India harbors heterogeneous population groups inhabiting the different geographical regions. A little work is being carried out regarding the study of genetic diversity of the people of the state. In present study, genomic diversity study in six of the prominent population groups (Brahmins, Rajputs, Bhagats, Chamar, Gujars and Jatt Sikhs) was carried out using ten autosomal DNA markers belonging to seven Alu insertion/deletion polymorphisms namely Alu ACE, Alu APO, Alu PV-92, Alu PLAT, Alu FXIIIB, Alu D1, Alu CD4; LPL Pvull, ESR Pvull and ESR Xbal polymorphism. Blood samples were collected randomly from 800 unrelated healthy individuals after prior consent. DNA was extracted and amplified by PCR using target specific oligonucleotide primers and finally subjected to agarose gel electrophoresis. Further, for LPL Pvull, ESR Xbal and ESR Pvull, the PCR product was subjected to restriction digestion using Pvull and Xbal restriction enzymes. Allele frequencies were used to calculate average heterozygosity. All the markers except Alu CD4 were found to be highly polymorphic with high heterozygosity values in almost all the population groups of the state. It was observed that most of the genomic diversity was attributed to individuals within the population. The study is a preliminary work on the population groups of the J&K state which may help in future work on the genetic heterogeneity in other population groups of the state which may give a genetic insight and genetic basis underlying the different genetic diseases prevailing in the state.

1920M
Acetylation of RNA Polymerase II Evolved in the Early History of Animals. C. Simonit1, K.S. Pollard2, S. Schroeder4, D. Her3, B.G. Bruneau2, M. Ott1, J.A. Capra1,2 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Biological Sciences, Vanderbilt University, Nashville, TN; 3) Gladstone Institutes, University of California, San Francisco, CA.

The evolution and diversification of animals from their unicellular ancestors required the development of novel functional capabilities that enabled the differentiation and coordination of different cell types. We argue that a recently discovered posttranslational modification of RNA polymerase II—acetylation of its RBPI subunit—arose in the early history of animals and provided a substrate for developing gene regulatory mechanisms, such as polymerase pausing, essential to animal diversification. RBPI contains a modifiable C-terminal domain (CTD) that consists of variations of a heptad repeat sequence (Y_{336}S_P_{3}T_{334}P_{3}S_{3}). Modeling the evolutionary dynamics of CTD repeat count and sequence content across diverse eukaryotes revealed an expansion of the CTD in the ancestors of Metazoa. The new CTD repeats introduced the potential for RBPI acetylation (acRBPI) due to the appearance of repeats with lysine at position seven. This was followed by an increase in the number of lysine-containing repeats in Deuterostomia. Mouse genes enriched for acRBPI occupancy at their promoters and genes with significant expression changes when acRBPI is disrupted are enriched for functions essential to the development and diversification of animals, such as growth factor response, cellular adhesion, and vascular development. Genes sensitive to acRBPI show significant enrichment for origins near the appearance of lysine-containing repeats, and p300, the acetyltransferase enzyme known to acetylate the CTD, also appeared on the same autosomal loci in a common ancestor of birds and mammals. Mouse and human analyses show that RBPI CTD acetylation specifically regulates processes essential to multicellularity and suggest that this regulatory mechanism was involved in the expansion of animals.

1921S

Human Profilagrin gene (FLG) is produced in the outer layers of the epidermis and encodes important components of the skin. Mutations in this gene are associated with a variety of skin disorders, such as, atopic dermatitis (atopy). These mutations have a high prevalence in European and Asian populations. When this mutations are analysed, a higher frequency of non-synonymous variation is found, which is indicative of positive selection and may suggest a benefit for the organism. This benefit is described in one hypothesis called "natural vaccination". This hypothesis suggests a heterozygote advantage in which a higher number of antigens enter through skin - during childhood- and thus serve to strengthen the immune system, as zygote advantage in which a higher number of antigens enter through skin - during childhood- and thus serve to strengthen the immune system, as an adult. High prevalence of mutations in different populations, higher non-synonymous variations and "natural vaccination" hypothesis suggest that filaggrin is evolving under positive selection. The aim of my study is to determine the structure of FLG in Macaque, Orangutan, Gorilla and Chimpanzee, compare it with the one in human and test their sequences for evidence of positive selection. To look for this, phylogenetic analysis; maximum-likelihood tests were used and filaggrin monomers were analysed for functional independent variants in IFIH1, suggesting that this regulatory mechanism was involved in the expansion of animals.
1922M
Population genomics analysis in whole genome sequencing of 152 rhesus macaques. F. Yu1, C. Xue1, M. Raveendran2, G.L. Fauci1, S. White1, R.A. Harris1, M. Dahdouh1, W. Salemo4, Z. John- son2, E. Vallender3, R. Wiseman1, H.M. Kubisch1, L. Cox4, S. Kanhas- wamy2, D.G. Smith5, B. Ferguson2, J. Horvath2, D. Muzny1, R.A. Gibbs1, J. Rogers1. 1) Human Genome Sequence Ctr, Baylor College Med, Houston, TX; 2) Yerkes National Primate Research Center, Atlanta, GA; 3) New England NPRC, Southborough, MA; 4) Wisconsin NPRC, Madison, WI; 5) Tulane NPRC, Covington, LA; 6) Tulane NPRC, Covington, LA; 7) Tulane NPRC, Covington, LA; 8) Oregon NPRC, Oregon Health University, Portland, OR; 9) Tulane NPRC, Covington, LA.

Rhesus macaques (Macaca mulatta) are themost widely studied nonhu- man primate species in biomedicine. Rhesus share many fundamental bi- ological and physiological processes with humans that make them an ideal model system for vaccine research and for studying molecular mechanisms of human diseases. However the patterns of the genetic variation and the evolutionary and population genomics processes that shaped rhesus genomes have not been well studied. We applied next generation whole genome sequencing for 152 unrelated individuals (144 Indian-origin, 8 China-origin) using four deep (30x) or low coverage (6x) strategies. Our analysis using SNPTools identified 51.6 million SNPs. And rigorous quality control procedure was applied, the average transition transversion ratio is ~2.2. We characterized evolutionary and population genomics from five different aspects: (1) characterization of population genomics parameters; (2) detailed demographic modeling; (3) identification of loci affected by Darwiniain selection; (4) characterization of the patterns of linkage disequilibri- um and recombination hotspots in comparison to human and chimpanzees; and (5) functional annotation of SNPs. Analysis revealed that the rhesus genome is substantially higher than values for humans. Analyses also reveal dramatic demographic changes over the past 10 million years. Functional annotation in coding sequences found >250,000 missense variants and >4600 stop-codon-gained muta- tions. In addition, about 110,000 SNPs mapped to conserved ENCODE transcription factor binding motifs. We used position weight matrices from the JASPAR database to assess these SNPs and found >25,000 candidate variants that may significantly affect TF binding, and thus gene expression. We mapped rhesus env-derived genes that were identified as from across 29 mammals, and found reduced SNP density and MAF, consistent with negative selection in those regions. We also applied anumer of site frequency spectrum tests and found significant new evidence for both posi- tive and negative selection in both coding and noncoding regions. Analyses of LD and local recombination rates are in progress.

1922N
Posters: Evolutionary and Population Genetics
445
How population growth affects linkage disequilibrium. A. Rogers. Anthropology, University of Utah, Salt Lake City, UT.

The LD curve relates the linkage disequilibrium (LD) between pairs of nucleotide sites to the distance that separates them along the chromosome. It is used to map disease genes and to search for adaptive evolution. But it also responds to the history of population size. The present research describes new theoretical results about the effect of population history. When a population expands in size, the LD curve grows steeper, and this effect is especially pronounced following a bottleneck in population size. When a population shrinks, the LD curve rises but remains relatively flat. As LD converges toward a new equilibrium, its time path may not be monotonic. Following an episode of growth, for example, it declines to a low value before rising toward the new equilibrium. These changes happen at different rates for different LD statistics. They are especially slow for estimates of $\alpha_2^n$, which therefore allow inferences about ancient population history. For the human population of Europe, these results suggest a history of popula- tion growth.

1923S
RNA-seq analysis of endogenous retroviral elements in bovine concep- tuses during the period of placenta formation. S. Nakagawa1,2,4, R. Koga1,2, T. Gojobori1, K. Imakawa4,1. 1) Department of Molecular Life Science, Tokai University School of Medicine, Isehara, Kanagawa, Japan; 2) Department of Cell Biology, Institute for Viral Research, Kyoto University, Kyoto, Japan; 3) Center for Genome Information Biology, National Institute of Genetics, Mishima, Shizuoka, Japan; 4) Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan.

In evolution of mammals, some of essential genes for placental develop- ment are known to be of retroviral origin, as syncytin-1 derived from an envelope (env) gene of an endogenous retrovirus (ERV) aids in the cell fusion of placenta in humans. Although the placenta serves the same func- tion in all placental mammals, env-derived genes responsible for trophoblast cell fusion and maternal immune tolerance differ among species and remain largely unidentified in the bovine species. To examine env-derived genes playing a role in the bovine placental development comprehensively, we determined the transcriptomic profiles of bovine conceptuses during three critical windows of implantation periods using a high-throughput sequencer.

The sequence reads were mapped into the bovine genome and the ERV candidates were annotated using RetroTector© (7624 and 1542 for ERV- derived and env-derived genes, respectively). The mapped reads showed that about 18 percent (284 genes) of env-derived genes in the genome were expressed during placenta formation, and about four percent (63 genes) were detected for all days examined. We verified three env-derived genes that are expressed in trophoblast cells by polymerase chain reaction. Out of these three, the sequence of env-derived gene with the longest open reading frame (named BERV-P env) was found to show high expression levels in trophoblast cell lines, and to be similar to those of syncytin-Car1 genes found in dogs and cats, despite their disparate origins. These results suggest that placenta depends on various retrovirus-derived genes that could have replaced endogenous predecessors during evolution.

1924S
Population structure and linkage disequilibrium in modern Indian populations. Y. Zhang, J. Jung, B. Grant. National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Rockville, MD.

The linkage disequilibrium (LD) in human population is found to be stronger with increasing geographic distance from Africa, which reflects the Africa origin of human history. Recently admixed populations (such as African Americans and Hispanic Americans) are more likely to harbor a larger num- ber of genetic variants, relative to their inferred ancestral populations. How- ever, the pattern of linkage disequilibrium in these admixed populations are not well studied. Here, we conduct an analysis of linkage disequilibrium at 659,184 single nucleotide polymorphisms (SNPs) in 924 unrelated samples from 11 Hanmap3 populations and 24 samples from Kartliana population (Native American in Brazil from Human Genome Diversity Project). African Americans (ASW) derive their genomic ancestry from African and European with an average of 77.3% African and 20.0% European ancestry. Hispanic Americans (MXL) lie on a cline of an average of 45.5% European ancestry, 42.9% Native American ancestry, 4.9% East Asian ancestry and 4.4% Afri- can ancestry. The mean of SNP based haplotype heterozygosity across the whole genome in these two admixed populations is greater than that of their major inferred ancestral populations. We further use $r^2$ between all possible SNP pairs in various distance classes as a measure of LD and also focus on the proportion of SNP pairs with $r^2$ greater than 0.8. Both of these two admixed populations show intermediate LD (as measured in $r^2$ and the proportion of SNP pairs with $r^2$>0.8), compared with their two major inferred ancestral populations. The extent of LD ($r^2$) in African Americans (ASW) is more closer to that in African population (YRI) in the short distance classes, while the values of LD in African Americans (ASW) is more likely to be similar to the European Americans (CEU) with the increased distance classes. The findings on the structure of LD in admixed populations are helpul to better understand the evolution of human population and the design of genetic association studies in admixed populations.
1926M
Statistical genetic considerations for expansion of panel of DNA markers for forensic applications: Lessons learned from the panel of 29 autosomal STR loci. M.R. Notan, R. Nakrabo. Department of Biostatistics and Medical Genetics, University of North Texas Health Science Center, Fort Worth, TX., USA.

With increasing success of using DNA markers for forensic applications, there are several recent attempts to expand the panel of autosomal short tandem repeat (STR) markers to solve more complex forensic problems and to achieve worldwide uniformity of DNA typing to aid in forensic investigations that cross international boundaries. This research deals with the relevant statistical genetic considerations that are needed to validate genotype databases of such expanded panels of markers. Recently, the US National Institute of Standards (NIST) has published genotype data for 29 autosomal STR loci typed in 1,036 unrelated individuals belonging to four populations (African-Americans - n = 342, US Caucasians - n = 361, US Hispanics - n = 236, and Asians - n = 97). Since these loci include all markers of the CODIS 13, Identifier®15, Globalfiler®21, and PowerPlex®Fusion 22 panels, we conducted a comparative analyses of this expanded panel against the above 4 more restrictive panel of markers. Our analyses suggest that even with the increase of the number of loci, deviations from Hardy-Weinberg equilibrium and linkage equilibrium for this expanded panel did not exceed the number of deviations expected by chance alone (in each population). This is so even in the presence of 7 pairs of syntenic loci in this expanded panel (F13B and D15S1656 on 1q, TPOX and D25S441 on 2p, D5S818 and CFS1PO on 5q, SE33 and D6S1043 on 6q, vWA and D12S381 on 12p, FESFPS and Penta E on 15q, D21S11 and Penta D on 21q). This expanded panel improves the statistical power of human identification as well as kinship determination. We therefore conclude that this panel has been validated for forensic use. In total, the number of loci also more explicitly confirms some of the population genetic features of STR polymorphism, not clearly seen in the four more restrictive panels stated above (e.g., negative association of F2ST with heterozygosity at individual loci). However, while the average F2ST values (across the four subpopulations in the database) for the four restrictive panels are very similar (1.66 to 1.74%), addition of 7 more loci in this expanded panel increased the F2ST to 2.15%. This is largely due to the inclusion of the F13B locus, which is known to have one of the lowest frequencies (42%) in the Asian population. Further expansion of panels of forensic markers should attempt to control further increase of F2ST for improving their power of forensic utility.

1927S
Multidrug-resistant pulmonary tuberculosis in Mexican population. Evidence of association of HLA class II and TNF-308 G/A polymorphism. B. Silva1, C. Saenz2, L.A. Bracho-Vela3, M.A. Bermúdez de León1, S. Sait-Fernández2, J. Granados2, CIBIN-IMSS, Monterrey, Nuevo León, Mexico, 1) CIBIN-IMSS, Monterrey, Nuevo León; 2) Fac de Medicina UANL, Nuevo León; 3) INCMN-SZ, México.

Tuberculosis (TB) remains a global emergency. It is estimated that one third of the world’s population is infected with Mycobacterium tuberculosis (Mtb), and only 10% of those infected will develop clinical tuberculosis (TB), which indicates the existence of host factors regulating disease expression. We investigated the association of human leukocyte antigens (HLA) class II and tumor necrosis factor (TNF) gene with susceptibility to pulmonary TB in Mexicans with special emphasis on their association with drug resistance. HLA-DRB1 and DQB1 gene polymorphism were analyzed in a 100 Mexican patients with Multidrug-resistant tuberculosis (MDR-TB denotes bacillary resistance to at least isoniazid and rifampicin) and 150 ethnically matched healthy controls using PCR-SSO (Luminex). Polymorphisms in the promoter region of the tumor necrosis factor (TNF) gene at positions -308 and -238 were studied using the TaqMan® allelic discrimination assays for both.

The frequencies of HLA-DRB1*14:06 were significantly higher in patients with MDR-TB as compared to healthy controls (P<0.05). Higher frequency of TNF-308 rs1800629 G/A genotype was observed in MDR-TB cases. The results suggest that HLA-DRB1*14:06 in combination with TNF-308 rs1800629 G/A genotype influence to MDR-TB. Further studies are needed to confirm our findings using larger number of patients with MDR-TB.

1928M
A Renewal Theory Approach to IBD Sharing. S. Carmi1, P. Wilton2, J. Wakeley3, I. Pe’er1. 1) Department of Computer Science, Columbia University, New York, NY; 2) Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA.

Sharing of long, identical-by-descent (IBD) genomic segments is common in many human populations. Dense genotyping of large cohorts, together with sophisticated detection tools, have transformed IBD sharing into an important and popular tool in population genetics, with numerous applications ranging from demographic inference to selection detection, pedigree reconstruction, disease mapping, imputation, and phasing. Nevertheless, many key questions in the theory of IBD sharing are still open. Here, we introduce and analyze a novel theoretical framework for the IBD process, based on renewal theory. Using our renewal approach, we were able to derive, in some cases for the first time, several key quantities, including the distribution of the number and total length of shared segments. Specifically, we consider the IBD process for a pair of chromosomes under two Markovian approximations of the coalescent with recombination, SMC and SMC’. We then propose, and justify by simulations, the renewal process, in which lengths of successive shared segments are independent. This research deals with the relevance of IBD sharing to forensic applications: Lessons learned from the panel of 29 autosomal STR loci typed in 1,036 unrelated individuals belonging to four populations (African-Americans - n = 342, US Caucasians - n = 361, US Hispanics - n = 236, and Asians - n = 97). Since these loci include all markers of the CODIS 13, Identifier®15, Globalfiler®21, and PowerPlex®Fusion 22 panels, we conducted a comparative analyses of this expanded panel against the above 4 more restrictive panel of markers. Our analyses suggest that even with the increase of the number of loci, deviations from Hardy-Weinberg equilibrium and linkage equilibrium for this expanded panel did not exceed the number of deviations expected by chance alone (in each population). This is so even in the presence of 7 pairs of syntenic loci in this expanded panel (F13B and D15S1656 on 1q, TPOX and D25S441 on 2p, D5S818 and CFS1PO on 5q, SE33 and D6S1043 on 6q, vWA and D12S381 on 12p, FESFPS and Penta E on 15q, D21S11 and Penta D on 21q). This expanded panel improves the statistical power of human identification as well as kinship determination. We therefore conclude that this panel has been validated for forensic use. In total, the number of loci also more explicitly confirms some of the population genetic features of STR polymorphism, not clearly seen in the four more restrictive panels stated above (e.g., negative association of F2ST with heterozygosity at individual loci). However, while the average F2ST values (across the four subpopulations in the database) for the four restrictive panels are very similar (1.66 to 1.74%), addition of 7 more loci in this expanded panel increased the F2ST to 2.15%. This is largely due to the inclusion of the F13B locus, which is known to have one of the lowest frequencies (42%) in the Asian population. Further expansion of panels of forensic markers should attempt to control further increase of F2ST for improving their power of forensic utility.

1929S
Using linkage disequilibrium to refine estimates of accelerating growth in human populations. M. Reppeli1, J. Carlson1, S. Zöllner2, The BRIDGES Consortium, 1) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 2) Department of Psychiatry, University of Michigan, Ann Arbor, MI.

Correctly modeling the effective size of a population is critical to making accurate inferences about mutation and migration rates, and the strength of selective pressures. In humans, several large sequencing studies have given us novel insight into a genome characterized by an abundance of extremely rare genetic variation, consistent with a history of recent massive population growth. These large sequencing studies offer us unprecedented resolution for distinguishing between models of recent growth. To improve on conventional inference methods we propose a novel likelihood based approach that incorporates pairwise r2, a measure of linkage disequilibrium, in addition to the site frequency spectrum. We observe that over short genetic distances, pairwise r2 is a function of the variance in ancestral tree branch lengths, and therefore contains information about ancestral population sizes lacking from the site frequency spectrum, which is a function only of the mean total ancestral branch lengths. Using simulations we show that increasing sample size increases the accuracy of inferences about recent demographic and magnifies the improvement our method yields versus conventional approaches. Lastly, we apply our method to regions defined as neutral in whole genome sequence data from ~4,000 European ancestry individuals sequenced as part of the BRIDGES consortium. This cohort is ideal for our purposes: providing both a large sample and non-coding genetic regions free from evidence of ongoing selection, a mixture unavailable from exome only or functional sequencing projects. We use a Monte Carlo method to estimate the likelihood of the observed data under a range of realistic growth models, including those incorporating continental, accelerating, faster than exponential growth. With our data we are able to simultaneously make inferences about the mutation rate, µ, and the rate of accelerating growth experienced by the European population from which our sample is drawn.
Non-visual Opsin Evolution and Implications for Human Health. A.B. Popadin1, S.E. Antonarakis1,2, V. Pritchard3,4, X. Lan5, 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Human Genetics, University of Chicago, Chicago, IL; 3) Department of Biology, Stanford University, Stanford, CA; 4) Howard Hughes Medical Institute, Stanford University, Stanford, CA.

Gene duplication is regarded as an important driving force for the evolution of new biological functions. This study seeks to understand the mechanisms driving gene duplication events, when they occurred, and how they contributed to the evolution of tissue-specific functions. We constructed a database of duplicated genes pairs in the human genome based on the pairwise sequence similarity of protein coding genes. Using the synonymous substitution rate between a pair of duplicated genes and the conservation of these genes to their orthologs in other species, we estimated the time of the duplication event of the gene pairs. We found that pairs of genes that were duplicated earlier generally have lower expression similarity among different tissues, lower Dn/Ds and increased gene length compared to genes that were duplicated more recently. The majority of gene duplications appear to be due to segmental duplications, and a relatively small fraction appears to be due to retrotransposition. Duplicate genes tend to have more tissue-specific expression than singleton genes (i.e., genes with no identified duplicates in the genome). We found that duplicated genes are significantly enriched for involvement in neurological functions compared to singleton genes; however, interestingly, the majority of these duplication events are relatively ancient, preceding the diversification of mammals. Similarly, the majority of gene duplicates that are primarily expressed in brain and cerebellum tissue arose before the divergence of mammals. Finally, we describe patterns relating to classic models of the evolution of gene duplicates, including loss-of-function, neofunctionalization and subfunctionalization.

Can phylogenomic analysis of Hemopexin repeat-containing proteins provide insights into the evolution of adaptive immunity? L. Likins, A. Smith, J. Wyckoff. Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri Kansas City (UMKC), Kansas City, MO.

Hemopexin (HPX), a glycoprotein in the blood plasma of many vertebrate organisms, has been shown to have the highest known binding affinity for heme. Many studies have shown that the primary function of the HPX protein product, Hemopexin is to sequester unbound heme released into the plasma from the breakdown of hemoglobin. The crystal structure of Hemopexin has been solved and the molecule consists of two structurally similar N-terminus and C-terminus beta-propeller domains joined by a flexible, unstructured linker peptide. This linker region is known to be part of the heme binding site. The gene that codes for Hemopexin (HPX) has been shown to have numerous sequences identified as recognizable hemopexin-like (HX) amino acid repeats in the primary structure of the protein. Structural mapping has revealed some interesting relationships between the functional motifs and the secondary structure of the HPX. Understanding the evolution of HPX proper, as well as its historical relationship to the several other classes of proteins that contain HX repeats, including: the myriad Matrix Metalloproteinases (MMP’s), Proteoglycan-4 (PRG4) and Vitronecin (VTN). All these proteins are associated with the Extracellular matrix of multicellular animals, and research has implicated most of them in immune system dynamics. Earlier studies had deduced that the HPX protein was a complex of sequences that had evolved from ancestral molecular precursors, and as such, is the most derived of the proteins containing the hemopexin domain. However, more recent phylogenetic analyses of both primary amino acid sequences, gene nucleotide sequences, and tertiary structure of these proteins has determined that hemopexin is most likely the evolutionary ancestral molecule. We conclude that the other proteins that have identifiable HX domain structure have adapted the 4-blade beta-propeller which has then been modified for variable functionality. Importantly, our studies show for the first time that the beta-propeller has not arisen from amino acid identity over long periods of evolutionary time; instead, the propeller structure has remained unchanged even as the components of the propellers have labile over long periods of evolutionary history. Our initial results show that these new insights reveal key details of the early evolution of the adaptive immune system in the Vertebrata and have clear implications for Human health.
1934M
POTE: an example of gene family evolution. F. Anaclerio, G. Gian
nuzzi, M. Ventura. 1) Department of biology, University of Bari, Bari, Italy; 2) Center for Integrative Genomics (CIG), University of Lausanne, Lausanne, Switzerland.

Eukaryotic genomes are characterized by the presence of gene families, set of genes derived from duplication events of a common ancestor. We described the genetic organization and evolution history of POTE, a specific gene family, POTE, whose members are expressed both in normal tissues and in many types of cancers (testis, prostate, ovary and breast). We studied the evolution of this gene family using both in silico and molecular cytogenetic approaches. Six species, including gorilla, chimpanzee and human. POTE gene family evolution has been characterized by both intra and inter-chromosomal duplication events leading to the increase of gene copies from one in marmoset to 14 (located on seven chromosomes) in human. In the analyzed species, we identified four different POTE gene isoforms: isoform 1, the unique encompassing exons 12-14, showed to be the ancestral form, isoform 2 appeared in the ancestor of Catarrhini and isoforms 3 and 4 appeared after the divergence of Old World Monkeys and Great apes. Isoform 3 and 4 are described as the most expressed in cancers and they are characterized by both the presence of a Long Inverted Repeat (LIR) within the last intron and by a processed actin retrogene in the last exon. The presence of a premature stop codon within the 15th exon of isoform 3 and the duplication of exons 8-9 in isoform 4 allowed us to distinguish these two isoforms. We proposed a model for the evolution of all the four isoforms from the first POTE gene on the ancestral chromosome 8 to the most recent human specific POTEF located on chromosome 2 and characterized by the presence of two additive exons harbored in the 5' UTR and responsible for a positive regulated expression. At 20 POTE genes map at the ancestral centromeric region that lost its function due to the fusion of the two ancestral acrocentric chromosomes that originated human chromosome 2. Analyzing the increase in the number of POTE genes and their localization on different chromosomes, we showed that a burst duplication of this gene family occurred since New World Monkeys split from the common ancestor of Catarrhini, moreover, exons loss and gain and the distribution of these events along the gene let us to suppose not only the acquisition of new function for this gene family in most recent species but also the potentiality to a further evolution.

1935S
Evolutionary Triangulation: informing Genetic Studies with Evolutionary Evidence. M. Huang, B. Graham, L. Maglia, S. Williams. 1) Department of Genetics, Dartmouth College, Hanover, NH; 2) IQBS, Dartmouth College, Hanover, NH; 3) Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

Genetic studies of disease have been successful in identifying many variants associated with clinical outcomes. However, most have performed straightforward analyses without incorporating other factors that may affect risk. One such factor is evolution, as human populations have evolved under a variety of environments that have shaped disease risk. Using the unique evolutionary trajectories in diverse populations may improve our ability to extract important genetic factors in disease. Given that a large portion of common variants differ among populations in allele frequency, as do disease prevalence, we hypothesized that patterns of disease and population genetic structure may together inform association studies. Specifically, the distribution of allelic variants that associate with certain diseases will mirror the distribution of the diseases. Moreover, variants should be detectable and less than 10 thresholds used. The results indicate that ET can filter associated variants based on evolutionary comparisons among populations, and that it can be effectively filter association results.

1936M

Previous analyzing the maternal heritage of Taiwanese through mitochondrial DNA (mtDNA) unanimously agree for an ancestral origin Taiwan Indigenous people coming from mainland Southeast Asia (MSEA) Southeast China and Indo-China. Contemporary Non Taiwan Aborigines (often referred as Taiwan Han) have a more diversified range of origin with many finding their roots in Northern Asia and others in Southern Asia. In both groups, Taiwan Aborigines and Non-Aborigines some mtDNA lineages can trace their origin back to 15,000 YBP. Using association between mtDNA and systemic functional measurements (ATP, mitochondrial membrane potential, 80HdG, lipid peroxidation, systolic BP, diastolic BP, glucose, total cholesterol, TG, LDL, and HDL) we have shown that several measurements were significantly associated with mtDNA variation. Since many of these measurements are used in disease diagnosis, we asked whether mtDNA sequence variation could be used as a tool to foresee disease outcomes. We used DNA samples from Taitung area, where one third of the population is Taiwan Aborigines. We also included samples from the Taipei area where most people are Taiwan Han and have a very different mtDNA polymorphism. Efforts were directed toward determining significant associations between mtDNA SNP and disease determining functional measurements. Several mtDNA haplogroup/SNPs found potentially good indicators of disease development were further investigated in relation to ethnicity in view to pre-determine the likely health status of a group, with greater attention to disadvantaged groups in the Taiwanese population.

1937S

1) Instituto de Genética Humana. CUCS. UdG; 2) Laboratorio de Inmunobiología. CUBCA. UdG; 3) Cirugía de Colon y Recto. Hospital Civil “Juan I. Menchaca”; 4) Hospital Civil “Fray Antonio Alcalde”; 5) Departamento de Clínicas. CUAltos. UdG.

Background: The human RUNX3 gene is mapped to chromosome 1p36.1, contains six exons and its overall size is approximately 67 kb. The RUNX3 protein is required for neurogenesis of dorsal root ganglia and in intestinal epithelial participate with WNT and TGF-β pathways. The inactivation of RUNX3 gene has been associated with colorectal cancer. In this study, we analyzed the variants rs6672420 (exon 1, e3 A>T), rs11249206 (1tron 1, T>G), rs760805 (1tron 3, T>A) and rs2236982 (1tron 4, G>A) in colorectal cancer patients from Mexico. Material y Methods: Genomic DNA samples were obtained from peripheral blood of 176 Mexican patients with CRC at diagnosis and 195 individuals as control group. The haplotypes were determined by calculating the odds ratio (OR). The haplotypes and linkage disequilibrium (LD) were established in Arlequin v3.5 software. Results: The RUNX3 rs2236982 AA genotype and A allele showed association with CRC (OR = 0.95% CI = 0.21 - 0.73 P < 0.01; OR = 0.65% CI = 0.49 - 0.87 P < 0.01, respectively) while rs6672420, rs11249206 and rs760805 polymorphisms displayed not significant results. The SNPs rs760805 and rs2236982 showed LD with r2 = 0.70 for controls and r2=0.44 for patients. The haplotype analysis revealed the most frequent was TG, constructed by combining the wild alleles. Moreover, the analysis showed CRC risk associated to TA haplotype OR = 2.52 95% CI = 1.47 - 4.30 P<0.01). Conclusion: The AA genotype and A allele of rs2236982 polymorphism have a decreased risk and TA haplotype have a risk to CRC development in Mexican patients. 1.Tsunematsu T, Kudó Y, Iizuka S, Ogawa I, Fujita T, Kurihara H, Abiko Y, Takata T: RUNX3 has an oncogenic role in head and neck cancer. PLoS One 4:e5892, 2009. 2.Kim WJ, Kim EJ, Jeong P, Quan C, Kim J, Li QL, Yang YO, Ito Y, Bae SC: RUNX3 inactivation by point mutations and aberrant DNA methylation in bladder tumors. Cancer Res: 65:9347-54, 2005. 3.Zhang Z, Wang S, Wang M, Tong N, Fu G, Zhang Z: Genetic variants in RUNX3 and risk of bladder cancer: a haplotype-based analysis. Carcinogene-
1938M

Deep sequencing of the human MHC region reveals widespread and ancient structural variation. A.Q. Fu1, B. Howier2, E.Y. Kim3, M. Stephens4, S. Wolinsky4. 1) Department of Human Genetics, University of Chicago, Chicago, IL; 2) Adaptive Biotechnologies Corporation, Seattle, WA; 3) Division of Infectious Diseases, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; 4) Department of Statistics, University of Chicago, Chicago, IL.

The human major histocompatibility complex (MHC) region of 3.6 Mb harbors over 200 genes, most of which have important roles in innate and adaptive immunity. Several genetic variants in this region have been associated with risk for autoimmune and disease progression in HIV infection and AIDS. The DNA sequence of the MHC is highly heterogeneous across individuals, and much of its population variation remains to be characterized. Here, we studied the genetic diversity of the extended MHC region in 64 HIV patients who carry no CCR5-delta32 allele but at least one copy of HLA-B*57:01 or B*57:03. We performed sequence capture and targeted enrichment of the MHC and its flanking regions for a total length of 4.7 Mb followed by ultra-high-throughput DNA sequencing, generating long reads of ~400 bp at 55x coverage on average. Alignment of the sequencing reads to the human reference genome (hg19) led to the identification of ~50,000 SNPs, about 10 of which are in perfect linkage disequilibrium with HLA-B*57:01/03. De novo assembly of the reads revealed sequence segments that are more similar to the genomes of apes (chimp, gorilla, or orangutan) than to the human reference genome, instead than to the 8 available human MHC reference haplotypes. These ape-like segments map to several loci in the MHC region of the ape genomes, and their presence correlates with the occurrence of several evolutionarily related HLA-A alleles. These ape-like segments of ~1000 bp were observed in about 20 of the 64 individuals in our study, and we hypothesize that they represent ancient structural polymorphisms. We further confirmed the presence of these long structural variants in several 1000G individuals by aligning sequencing reads in the 1000 Genomes project to the consensus ape-like sequence segments.

1940M


Background: Folate is used by many biological processes to donate methyl groups to other compounds. Its active form is L-methylfolate which is produced through a reaction catalysed by Methylene tetrahydrofolate Reductase (MTHFR). Mutations in MTHFR have been linked to everything from heart disease to cancer, but not, until recently, addiction. Objective: The focus of this study is to understand the association between MTHFR (MTHFR C677T, Rs1801133) and two popularly accepted tests in determining risk of opioid abuse/Misuse: Screener and Opioid Assessment for Patients with Pain (SOAPP®) and opioid risk tool (ORT). ORT (Low Risk 0 - 3, Moderate Risk 4 - 7, High Risk > 8) and SOAPP® (High risk SOAPP®-R score = 22 or greater, moderate risk SOAPP®-R score = 10 to 21, low risk SOAPP®-R score < 9) Subjects: Subjects for this study were made up of two independent groups of patients. The SOAPP®-R test was taken by 2228 chronic pain patients in a clinic in Arizona. Another 3517 chronic pain patients across 55 clinical research sites in the US took the ORT test. Methods: The study evaluated data from both groups of subjects. The subjects were genotyped with Taqman single nucleotide polymorphisms (SNP) assays using the proprietary Proove Narcotic Risk Genetics Profile Test (Proove Biosciences Inc, Irvine, California). The subjects were stratified into low, moderate, and high risk groups, based on the results from both questionnaires (i.e. ORT & SOAPP ®-R). Results: A cross tab analysis using IBM SPSS found an association between MTHFR (MTHFR C677T, Rs1801133) and risk of opioid abuse from both ORT and SOAPP® scores (p=0.028 and p= 0.042 respectively). A multinomial logistic regression showed that MTHFR homozygous mutation is more likely to be associated with the low risk group derived from SOAPP®-R compared to high risk group (P= 0.020, OR=4.273). In addition, it was found that Homozygous mutations in MTHFR is more associated with subjects with low ORT scores (P=0.028, OR =4.273) when compared to those with high risk stratified by ORT scores. Conclusion: This study suggests that an association exist between MTHFR and these two popular tests used to determine risk of opioid abuse/misuse. Findings in this study and further studies in this direction could improve understanding about the role of MTHFR in prescription opioid medication abuse or misuse.

1939S


Across mammals, decreased generation times (i.e., mean ages of reproduction) are known to correlate with increased neutral substitution rates. This observation is believed to reflect increased mutation rates due to higher numbers of cell divisions per unit time. Perhaps reflecting the same phenomenon, whole genome sequencing studies of human pedigrees, which reflect the mutation rate at present, yield a yearly estimate that is about two-fold lower than the one obtained from phylogenetic methods, which average over millions of years. One hypothesis is that increases in the generation time in the human lineage have led to a slow down of the mutation rate per unit time. Motivated by these considerations, we introduce an approach to estimate the generation time based on the insight that distinct types of mutations arise through different mechanisms and hence have different dependences on the generation time. Notably, transitions at CpG sites are thought to arise through spontaneous deamination, implying that their mutation rate should depend primarily on absolute time and should be relatively insensitive to the generation time. In contrast, replication-driven mutations at non-CpG sites occur in higher numbers following male puberty, and yearly mutation rates should depend on generation times. We characterize these different time dependences from human pedigree data. By relating them to the observed average pairwise diversity values at CpG and non-CpG sites, we then obtain joint estimates of the mean coalescence time in years and the generation time over evolutionary time scales. With this ingenious approach, the age of reproduction has changed recently, we estimate the generation time from doubleton mutations, which reflect a more recent time depth than do average pairwise differences. Our results are consistent with the ethnographic literature and suggest a slight increase in the generation time towards the present. Beyond this application, our method provides a novel population genetics based estimator of the generation time, applicable over different time depths of human evolution.

1941S

Y Chromosome STR Mutation Rates: the Factor of 3 Conundrum. D.E. Platt. IBM T. J. Watson Research Center, Yorktown Hts., NY, USA.

Y chromosome STR mutations serve as a prime clock for phylogeographic studies. However, there is a discrepancy between mutation rate estimates of roughly a factor of 3 between pedigree derived data vs. population based approaches. In 1995, Zhivotovsky et al constructed a Wright-Fisher based multi-step STR mutation model, extending Moran’s ladder model of 1975. In 2004, Zhivotovsky, et al published an STR mutation rate using population data from a range of indigenous peoples that diverged from the Y chromosome apes between 1.2 and 2.3 million years ago. By derived results presented by Di Giacomo et al in 2013. In 2005, Zhivotovsky, et al published a paper indicating that diversity fluctuations could produce a diversity bottleneck, yielding an apparently low mutation rate. Interestingly, their test population dynamics model was Poisson rather than Wright-Fisher. This left the question open of whether the Wright-Fisher model is inadequate for modeling diversity fluctuations in real populations. We present an argument showing that the Zhivotovsky Wright-Fisher dynamics model and the Poisson-based model may both be viewed as limiting cases of a more general model with coinciding population sizes at each generation. We show that the generation-by-generation Zhivotovsky formulation is preserved in both of these limits. Therefore, we conclude that Wright-Fisher dynamics is adequate for predicting the level of fluctuations expected for a Poisson model. Alternatively, Xu, Peng, and Fang proposed a model for STR mutations that allowed for a length dependent rate for decreasing single step mutations which could perhaps account for limited diversity, but assumed a Markov chain equilibrium state. We extend the Xu model first to consider the non-eradicator Markov case, and second, incorporated this into a Wright-Fisher based model. We show that the Markov model is robustly convergent, with equilibration sharing the same law of information as predicted for the Jukes-Cantor model at equilibration, and dominating Wright-Fisher convergence. We identified regimes of back mutation length dependence based on both phylogeographical observations of human genetic diversity. We explored how BAT-WING (formulated on a single step model) parameter estimation responses to data generated from the extended Xu model. We see a reduction in predictions of mutation rates across a range of divergence times, while still seeing correlations to adequate depth to capture modern human variation.
had effects on the genetic composition of these animals, but did not reduce Apparently the severe bottleneck in the founding of the Mauritius population exhibit significant levels of SNV with diversity higher than found in humans. while MCM have reduced diversity in MHC haplotypes, they nonetheless using Polyphen and DAVID are in progress. These results indicate that (diabetes). Analyses of the predicted impact of amino acid substitutions higher than in human (4.2 million per person). We identified SNVs throughout a 0.23% sequence divergence between MCM and IRM. The average number are fixed differences between these 20 MCM and the rhemac2, suggesting 22.59 million SNVs polymorphic among the 20 MCM. 7.16 million variants mapped using BWA to the Indian rhesus macaque (IRM) reference genome at a deep coverage using Illumina Hi-seq platform. The Illumina reads were variation can influence results. However, little is known about variation (MHC) alleles and haplotypes. This makes these MCM more attractive for there by humans ~500 years ago. These Mauritian cynomolgus macaques species extends through Vietnam, Cambodia, the Philippines and Indonesia. used in pharmacology and drug development. The natural habitat of this Center, Madison, WI; 3) University of Wisconsin, Madison, WI.

---

4194M Forensic Phenotyping in Brazilian population: SLC24A5 and ASIP as phenotypic predictors genes of skin, eye and hair color. C. Friedman, F.A. Lima, F.T. Goncalves, Dept of Legal Medicine, Ethics and Occupational He, University of São Paulo, São Paulo, São Paulo, Brazil.

Pigmentation is a very variable and complex trait in humans and it is determined by the interaction of environmental factors, age, disease, drugs, hormones and exposure to ultraviolet radiation and genetic factors, including pigmentation genes. Many of these genes and their variants have been associated with phenotypic diversity of skin, eyes and hair color in homogeneous populations. SLC24A5, TYR, MC1R, SL C45A2, ASIP, OCA2 and KIT genes are noteworthy because they are related to pigmentation process. Prediction of phenotypes by using genetic information has benefited forensic area in many countries because it has made possible to infer physical characteristics from biological samples and, thus, lead criminal investigations. The aim of this study was evaluate polymorphisms in TYR, ASIP, SLC24A5 and SLC45A2 genes in a sample of 350 individuals of admixed population from Brazil, intending to use the data in forensic genetics casework in several situations. Volunteers answered a questionnaire where they self-reported their skin, eye and hair colors, sun sensitivity and lifestyle. No significant results were observed except for SLC24A5 and ASIP. The polymorphic homozygous allele of rs1426654 and rs6058017 in SLC24A5 (OR 32.86 < p<0.0001) and ASIP (OR 8.68 < p<0.007) respectively, showed strongest association with fairer skin. Besides, the polymorphic homozygous allele in SLC24A5 exhibited relation to light eye color - green (OR 9.82 - < p<0.0001), blond hair (OR 50.14 < p<0.0001) and also to increased sensitivity to sun exposure (OR 7.86 - < p<0.0002). Our data suggests that polymorphic allele (A) in the SLC24A5 and ASIP genes is correlated with characteristics of low pigmentation, while allele (G) is related to lighter traits. Our findings corroborate previously published data on studies in European and African populations. These associations between pigmentation genes and skin, eyes and hair color shows that it is possible to use molecular information of an individual to access its phenotypic traits and use the obtained data in attempt to help forensic investigations. Additional analyzes are ongoing as part of a project that evaluates 600 samples to check possible associations of phenotypic pigmentation in the Brazilian population with the following genes: FTO, TYR, ASIP, OCA2, SLC24A5, SLC45A2. Financial Support: FAPESP (2013/0243-6), LIM 40/ HCFMUSP and Department of Legal Medicine, Ethics and Occupational Health - FMUSP.

---

4194S Interpretation of the high allele frequency of GJB2 c.109 G>A variant in Chinese population: a pathogenic mutation or coincidental polymorphism? Y. Lu1,2, J. Cheng2, H. Yuan2. 1) Otolaryngology Dept., No. 150 Central Hospital of PLA, Luoyang, Henan, China; 2) Otolaryngology Dept., Chinese PLA General Hospital, Beijing, China.

Hearing loss is a common sensory disorder in humans. Genetic factor plays more and more important role in the etiology of hearing loss than the environmental factors, which are controlled gradually and effectively with the development of the economy in the past two decades. The high frequency Nonsyndromic Hearing Loss is a hereditary condition with heterogeneity of genetic defects of hearing loss. GJB2 mutations cause more than 17% nonsyndromic hearing loss in Chinese population. The pathogenic role of a common c.109 G>A variant in GJB2 gene keeps uncertain in hearing loss cases. In this study, we recruited 37 subjects with homozygous c.109 G>A variant in GJB2, and 17 subjects with both heterozygous c.109 G>A and a heterozygous mutation in GJB2. All the subjects were excluded hotspot mutations in SLC26A4 gene and miRNA 125b-3p, which were also common causative genes in Chinese deaf population. The hearing status of the subjects included normal(7 cases), mild to moderate (both prelingual and postlingual) (21 cases), severe to profound(26 cases). DNA samples from the parents were available. Fourteen SNPs in GJB2 were selected based on the HapMap LD data set (r2>0.9, MAF>0.1) in DFN1 loci. Haplotypes of the subjects were derived from genotyping by SNaPshot technique. The results inferred that c.109 G>A in GJB2 gene in Chinese population originated from a founder with a common L-related rather than a recurrent hotspot. Recent functional studies have provided evidence that the defective p.Ala37threonine variant is likely to be the cause of the c.109 G>A variant. More than 6% allele frequency of c.109 G>A were detected both in hearing loss subjects and normal hearing controls in Chinese population. Although the penetrance of mild to moderate hearing loss is low, the deafness is serious and is a significant public health problem. Although the possibility to detect the c.109 G>A variant in Chinese population is higher than in other population, the severity of hearing loss caused by GJB2 c.109 G>A subject is still unpredictable in clinical practice. It is possible that other defects, such as CNV in the upstream sequence of GJB2, causes functional defect in correlation with the c.109 G>A variant. Therefore, the c.109 G>A variant of GJB2 should be further studied in the next generation sequencing. In conclusion, the pathogenic role of GJB2 c.109 G>A needs further interpretation under next generation sequencing and functional study.

---

1943S Whole genome sequencing of twenty Mauritian cynomolgus macaques (Macaca fascicularis). M. Raveendran1, R.A. Harris3, R. Deinios1, A.J. Ericsson1, G. Fawcett4, F. Yu5, C. Xue1, R. Wiseman2, D. Munz1, R.A. Gibbs1, D. O’Connor1, J. Rogers1. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Wisconsin National Primate Center, Madison, WI; 3) University of Wisconsin, Madison, WI.

Cynomolgus macaques (Macaca fascicularis), also known as long-tailed or crab eating monkeys, are used in several aspects of biomedical research including studies of AIDS, diabetes and alcohol abuse. This species is widely used in pharmacology and drug development. The natural habitat of this species extends through Vietnam, Cambodia, the Philippines and Indonesia. An isolated population of cynomolgus macaques inhabits the island of Mauritius, and is descended from a small number of founder animals transported there by humans ~500 years ago. These Mauritian cynomolgus macaques (MCM) carry a very limited number of major histocompatibility complex (MHC) alleles and haplotypes. This makes these MCM more attractive for studies of infectious disease and other research questions where MHC variation can influence results. However, little is known about variation across the genome. In this study, we compared the average number of SNVs (SVNH) through whole genome sequencing of twenty MCM samples at a deep coverage using Illumina Hi-seq platform. The Illumina reads were mapped using BWA to the Indian rhesus macaque (Macaca mulatta) reference genome (rhemac2) and variant calls made using Samtools. In total we identified 29,754,615 genome positions showing differences from the reference, and 22.59 million SNVs polymorphic among the 20 MCM. 7.16 million variants are fixed differences between these 20 MCM and the rhemac2, suggesting a 0.23% sequence divergence between MCM and IRM. The average number of polymorphic variants found in each animal is 9.96 million which is much higher than in human (4.2 million per person). We identified SNVs throughout the genome, including non-synonymous coding (51,132), synonymous coding (58,022), stop altered (1,619), splice altered (1,849), 3’ UTR variant (82,307) and 5’ UTR variant (21,429). Non-synonymous variants are found in genes associated with human diseases, such as GJB2 (Charcot-Marie-Tooth), GAB2 (Gartenstein’s disease) (CACHN1) (Neuroblasta) and CAPN10 (diabetes). Analysis of the predicted impact of amino acid substitutions using Polyphen and DAVID are in progress. These results indicate that while MCM have reduced diversity in MHC haplotypes, they nonetheless exhibit significant levels of SNV with diversity higher than found in humans. Apparently the severe bottleneck in the founding of the Mauritius population has no effects on the genetic composition of these animals, but did not reduce overall SNV levels dramatically.

---

1942M Extending the Ewens Sampling Formula to structured populations: Recursive computation of exact probabilities of allele frequency spectra. A. M. Uyenoyama, S. Kumagai, B.D. Redelings. Dept Biol, Duke Univ, Durham, NC.

To accommodate the enormous amount of information available in a sample of entire genomes, many evolutionary analyses have relied upon summary statistics or data reduction methods that marginalize over full genealogical trees. An approach that appears to have received relatively little attention involves the Ewens Sampling Formula (ESF), which provides a closed-form expression for the probability of the allele frequency spectrum (AFS) observed in a sample of arbitrary size under the infinite-alleles model of mutation. We extend the ESF to accommodate population structure, here accommodating migration between two demes. We develop a recursion which relates a sample to its immediate ancestor sample (vertical in time). We then characterize properties of the next-sampled gene, given the AFS already observed (horizontal in time). While the AFS probabilities are fully determined by the vertical recursion, their derivation is greatly simplified by the horizontal argument. A rescaling of time or mutation rate that would permit the use in structured populations of the ESF for unstructured population does not exist, even in the case in which the sample is derived entirely from a single deme in a multimode population. For that case, we propose an approximate scaling, which interpolates between no migration and the high migration limit. This approximation provides a qualitative description of the AFS probabilities; higher migration rates increase the effective mutation rate. Under other sampling conditions and for applications in which the exact likelihood is required, our method provides the AFS probabilities through a double recursion that accommodates successively greater numbers of observed alleles and successively greater sample sizes.
1946M Target sequencing analysis of Parkinson's disease genes in a healthy Amerindian Population from Puno-Peru. M. Cornejo-Olivas1, 2, I.F. Mata1, 2, M.O. Dorsher1, 3, M. Inca-Martino1, 2, A. Medina1, 2, A.C. Sheth1, 2, P. Kumar1, 2, K. Espinoza-huertas1, 2, D. Veliz-Otani1, 2, D. Incacutipa1, 2, V. Marca1, 2, O. Ortega1, 2, D. Mazzeotti1, 2, C. Zabellini1, 2, T. O'Connork2, 2, Latin American Research Consortium on the Genetics of Parkinson's Disease (LARGE PD). 1) Neurogenetics Research Center, Instituto Nacional de Ciencias Neurologicas, Lima, Peru; 2) Northern Pacific Global Health Research Training Consortium, Bethesda, MD; 3) Veteran Affairs Puget Sound Health Care System, USA; 4) University of Washington, USA; 5) Institute for Genome Sciences, University of Maryland School of Medicine; 6) Program in Personalized and Genomic Medicine, University of Maryland School of Medicine; 7) Universidad Nacional del Altiplano, Puno, Peru; 8) School of Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru.

Background: Both causal and susceptibility genes for Parkinson's disease (PD) known to date were discovered in either European or Asian populations. Studies on the genetics of PD in Latin-American populations are very much needed yet all such studies face a major obstacle in that there is no systematic information available on common genetic variability in healthy Amerindians to facilitate distinguishing pathogenic mutations from benign variants.

Objective: To characterize genetic variability and estimate the frequency of genetic variants in 15 known causal PD genes in a Peruvian Quechua native population.

Methods: Our study was performed using DNA samples and linked demographic data from 50 unrelated Amerindian subjects recruited in Puno, Peru through the Latin American Research Consortium on the Genetics of Parkinson's Disease. Sequencing analysis, including intron/exon boundaries was performed using a panel based on modified molecular inversion probes (MIPs) that included the 15 PD genes and 29 ancestry informative markers (AIMs). The resulting variants were categorized into different groups for comparison with other populations/datasets. We compared the frequencies of the single nucleotide variants (SNVs) to those in the 1000 Genomes Project (1KG) and Exome sequencing project databases (ESP).

Results: Analysis of the 29 AIMs indicated that the average proportion of Amerindian ancestry in our sample of Peruvian Quechas was 90%. Within the 15 PD genes, we detected 400 SNVs. Of these variants 302 were novel. The vast majority of these unique variants were rare, but 18 had a minor allele frequency ≥5%. We found 124 novel non-synonymous variants and 9 novel nonsense mutations.

Conclusions: Within a sample of healthy Americans from Peru, we found a surprisingly large number of novel, unreported substitutions in the coding region of known PD genes. Annotating these variants in publicly available databases will aid future studies on the genetics of PD in Latin American populations. Furthermore, our results suggest that generating exomes or whole genomes in additional Amerindian controls (to supplement publicly available data) might be of benefit to genetic studies of other diseases conducted in populations with substantial Amerindian admixture.

1947S Prevalence and sources of genetic variation in human mitochondria. E. Glassberg1, A. Harpak1, D. Calderon1, D. Cusanovich1, M. Caliskan2, V. Gilad3, C. Ober2, J. Pritchard1. 1) Biology, Stanford University, Stanford, CA; 2) Biomedical Informatics, Stanford University, Stanford, CA; 3) Human Genetics, University of Chicago, Chicago, IL.

Heteroplasm refers to the presence of multiple distinct organellar genomes within a single individual. In humans, heteroplasmic mutations in the mitochondrial genome are associated with a variety of bioenergetic defects and maternally inherited diseases. Here, we set out to quantify the prevalence of human mitochondrial heteroplasm at the individual and population levels. We use the observed patterns of variation to infer the contributions of stochastic genetic drift and natural selection to mitochondrial evolutionary dynamics within and between human hosts. To identify heteroplasmic sites from next-generation sequencing data, we examine the log-likelihood ratio of simple statistical models of homoplasm and heteroplasm at each position of the mitochondrial genome. We apply this method to 2535 individuals from the 1000 Genomes Project and to RNAseq data from 432 individuals of varying degrees of relatedness from a founder population of Hutterites in South Dakota. In both cases, we observe widespread heteroplasm. In the Hutterites, we find that heteroplasmic sites are significantly more likely to be shared between mother-child pairs than between unrelated individuals or father-child pairs. This pattern is more pronounced among sites at which the heteroplasmic allele is present at greater than five percent frequency in the mother. However, while some heteroplasmic sites are inherited maternally, the vast majority arise de novo within an individual. We then develop a novel inference algorithm to estimate the intensity of genetic drift experienced by the mitochondrial population during transmission from mother to child and during expansion from the germline to somatic cells. Separating the contributions of germline drift from those of somatic drift will allow us to clearly determine the population genetic setting in which inherited and de novo heteroplasmic variants are lost. Further, to understand the effects of selection on heteroplasmic sites, we compare observed patterns of heteroplasm at sites that are likely to be neutral to those at sites that are likely to be influenced by selection, especially focusing on known disease-associated sites. These analyses form a basis for disentangling the contributions of stochastic genetic drift and natural selection to mitochondrial genetic variation.

1948M Maternal Age Effect and Severe Germline Bottleneck in the Inheritance of Human Mitochondrial DNA. M. Su1, 2, B. Rebollar-Jaramillo3, 3, N. Stoler1, 2, J.A. McElhoe1, 3, B. Dickens1, 4, D. Blankenberg5, 7, T. Korneliussen1, 6, F. Chiaromonte6, 7, R. Nielsen5, 7, M.M. Holland1, 3, I.M. Paul1, 2, A. Nekrutenko2, 2, A. Nekrutenko2, 2, K.D. Makova1. 1) Department of Biology, Penn State University, USA; 2) Department of Biochemistry and Molecular Biology, Penn State University, USA; 3) Forensic Science Program, Penn State University, USA; 4) School of Science and Technology, Nottingham Trent University, UK; 5) The Department of Integrative Biology, the University of California at Berkeley, USA; 6) Department of Statistics, Penn State University, USA; 7) Department of Pediatrics, College of Medicine, Penn State University, USA.

The manifestation of mtDNA diseases depends on the frequency of heteroplasm (the presence of several alleles in an individual), yet its transmission across generations cannot be readily predicted due to the lack of data on the size of mtDNA bottleneck during oogenesis. For deleterious heteroplasmies, a severe bottleneck may abruptly transform a benign (low) frequency in a mother into a disease-causing (high) frequency in her child. Here we present a high-resolution study of heteroplasm transmission conducted on blood and buccal mtDNA of 39 healthy mother-child pairs of European ancestry (a total of 156 samples, each sequenced at ~20,000x/site). On average, each individual carried one heteroplasm, and one in eight individuals carried a disease-causing heteroplasm, with minor allele frequency ≤1%. We observed frequent drastic heteroplasm frequency shifts between generations and estimated the size of the bottleneck at only ~29-35 mtDNA molecules. Strikingly, we found a positive association between the number of heteroplasmies in a child and maternal age at fertilization, likely attributable to oocyte aging. Accounting for heteroplasmies, we estimate mtDNA germline mutation rate to be 1.3×10-8 mutations/site/year - lower than in previous pedigree studies but in agreement with phylogenetic studies, thus solving a long-standing controversy and informing the use of mtDNA in dating evolutionary events. This study takes advantage of droplet digital PCR (ddPCR) to validate heteroplasmies and confirm a de novo mutation. These results have profound implications for predicting the transmission of disease-causing mtDNA variants and illuminate mitochondrial genome evolutionary dynamics.
1949S

Cilioretinal artery: is it a variant angiogenesis under the effect of PAI-1 5G allele? I. Akalin1, A. Ardagil Akacakaya2, S. Basaran Yilmaz2, Y. Dag2, M. Guzin Altekin3, E. Kurum1, E. Koyun1, S. An Yaylali1, H. Bayramlar1
1) Medical Genetics, Istanbul Medeniyet University, Faculty of Medicine, Istanbul, Turkey; 2) Ophthalmology, Istanbul Medeniyet University, Goztepe Training and Research Hospital, Istanbul, Turkey; 3) Ophthalmology, Istanbul Medeniyet University, Faculty of Medicine, Istanbul, Turkey; 4) Biostatistics, Istanbul Medeniyet University, Faculty of Medicine, Istanbul, Turkey.

Purpose: Cilioretinal arteries (CRA) are small accessory arteries contribute to macular blood supply. The reason why some people have CRA while the others don't has not been elucidated yet, except its higher concordance in monzygotic twins. Here, we hypothesized whether genetic tendency to thrombosis due to well-known gene polymorphisms may induce CRA as a variant angiogenesis in embryonic life. Methods: We assessed plasminogen activator inhibitor-1 (PAI-1) 4G/5G, methylenetetrahydrofolate reductase (MTHFR), factor V leiden and prothrombin gene polymorphisms on 130 patients [82 females, 48 males; Median age: 57 (18-84) with CRA and 100 (64 females, 36 males; Median age: 55 (19-90)] without CRA. Genotyping was performed by melting curve analyses using a Rotor-Gene Q Real-Time PCR system (QIAGEN, Hilden, Germany) and PAI-1 4G/5G, MTHFR (C677T and A1298C), Factor V Leiden, Prothrombin G20210A mutation analyses kits (NLM, Settala, Italy). Multiple Logistic Regression model used to predict the dependent variable (occurrence of cilioretinal artery) adding a set of independent variables (gene polymorphisms) to test the association significance. Results: We found PAI-1 4G/5G; MTHFR (C677T and A1298C) have significant effects on the probability of occurrence of CRA at significance level 0.05. The odds ratio (OR) corresponding to PAI-1 4G/5G was 1.904 (95% CI: 1.050-3.499, p=0.026) which implies that keeping all other variables constant and having at least one 5G allele would increase the odds of having the cilioretinal artery by 98.4%. Additionally, we observed that while keeping all other variables constant, having at least one MTHFR C677T or A1298C allele would decrease the odds of having the cilioretinal artery by approximately 38% (OR=0.618, 95% CI: 0.394-0.961, p=0.035) or 44% (OR=0.558, 95% CI: 0.354-0.871, p=0.011), respectively. No significant differences were found between occurrence of CRA and Factor V Leiden or prothrombin. Conclusion: This is the first study to suggest the existence of strong association between the presence of CRA and genetic factors, up to the literature. Here we suggest that, not only the lack of genetic predisposition to thrombosis by MTHFR gene polymorphisms, but also the presence of anti-thrombotic polymorphisms determine the pressure on endothelial cells as a result of decreasing the blood viscosity.

1950M

Tumor Necrosis Factor-alpha Gene Polymorphism in Turkish Patients with Psoriasis. H. Akar1, Y. Yildray2, E. Koç3, E. Caliskan3, A. Akar3, Y. Tunca3
1) Medical Genetics, Gulhane Military Medical Faculty, Ankara, Turkey; 2) Dermatology, Gulhane Military Medical Faculty, Ankara, Turkey.

Introduction and objectives: Psoriasis is a chronic dermatosis that includes genetic and environmental factors in the etiology. Although complex cytokine network of psoriasis has not been well enlightened yet, tumor necrosis factor-alpha (TNF-α) is considered as main mediator in the pathogenesis of psoriasis. In addition to this, an association between single nucleotide polymorphisms at positions -238, -308 and -857 in the promoter region of TNF-α and psoriasis were reported previous studies in different populations. In this study, we aimed to investigate potential association of TNF-α promoter gene polymorphisms with psoriasis compared with healthy controls in Turkish population. Materials and methods: We investigated 40 patients with psoriasis vulgaris and 40 healthy controls. SNPs at positions -208, -308, -857 in the promoter region of TNF-α were analysed by competitive allele specific-polymerase chain reaction method in both groups. SNPs frequencies were compared between the two groups and significance was evaluated by chi-squared test and two-tailed Fisher’s exact test. Results: At the end of the study, we found that the TNF-α -308 AA and -857 CC homozygotes were significantly increased in patients with psoriasis in comparison with healthy controls (82.5% vs. 10%, p<0.01, OR=15.060; 47.5% vs. 7.5%, p<0.01, OR=11.593). TNF-α -308 AG and -857 CC heterozygotes were significantly decreased in patients with psoriasis in comparison with healthy controls (32.5% vs. 75%, p<0.01, OR=0.160; 15% vs. 47.5%, p<0.01, OR=0.195). No significant difference was observed in SNP at -238 position.

Conclusion: Our results suggested that TNF-α -308 AA and -857 CC homozygotes were associated with psoriasis vulgaris in Turkish population and these polymorphisms might be used as biological markers for psoriasis risk prediction. In contrast to previous studies, no significant differences were observed at -238 position. Further studies with larger populations are needed to confirm associations between -308 AA and -857 CC homozygotes and risk of psoriasis development in Turkish population.

1951S

Whole genome sequencing of a gibbon parent-offspring quartet to examine mutation rate variation in apes. D.M. Bobo1, O. Gokcumen2, L. Carbonzi3, M.F. Hammer4, J.D. Wall5, K.R. Veeramah6
1) Ecology & Evolution, Stony Brook University, Stony Brook, NY; 2) Department of Biological Sciences, University at Buffalo, Buffalo, NY; 3) Department of Behavioral Neuroscience, Oregon Health and Science University, Portland, OR; 4) Arizona Research Laboratories Division of Biotechnology, University of Arizona, Tucson, AZ; 5) Institute for Human Genetics, University of California San Francisco, San Francisco, CA.

Understanding the evolution of the human germ line mutation rate (μ) is vital for numerous aspects of medical, population and evolutionary genetics. Traditionally, μ is estimated using a phylogenetic approach in which genetic divergence between two species is calculated. This divergence is then calibrated using some external estimate of when the species split and an assumed generation time to give the phylogenetic mutation rate per generation. More recently, direct observation of de novo mutations from whole genome sequencing of parent-child trios in humans have resulted in pedigree estimates of μ per generation that are approximately half of the phylogenetic rate. This discrepancy may be explained by the ‘hominoid slowdown’ where μ is thought to have decreased per unit time during the transition from small ancestral primates to larger apes with a longer generation times and lower mass-specific metabolic rates. As gibbons lie intermediate of old world monkeys and great apes with regard to phylogeny, body size and generation time, they are a potentially important extant primate family for testing the robustness of the hominoid slowdown hypothesis. Therefore we generated high coverage short read second-generation sequencing data for a family quartet from Nomascus gabriellae (the yellow-cheeked gibbon) in order to calibrate mutation rates. The male parent was sequenced using Illumina’s 1TB technology and yielded ~67x mean mapped read depth. The Illumina HiSeq 2500 was used for sequencing the parents and female sibling, with genome coverage ranging from 25-35x. A custom bioinformatics pipeline was implemented incorporating published likelihood-based methods to weight candidate gene line de novo variants and determine the probable parent of origin. Extensive validation of de novo candidates across a range of confidence scores was performed. The mutation rate was then compared to those from previously published data for developing a machine learning algorithm to estimate μ for the two individuals in this quartet for comparison to the recent pedigree estimates in humans. We were also able to identify de novo copy number variation using comparisons of normal and affected gibbon genomes. Further phylogenetic analysis was performed using coalescent models to infer inheritance states across the genome and obtain fine-scale resolution of paternal and maternal recombination events, demonstrating that gibbons show a female recombination bias similar to that observed in humans and mice.
1952M
Significant association of Pro12Thr polymorphism in the fatty acid amide hydrolase (FAAH) gene with body mass index in Oceanic populations. I. Nakai, N. Nishida, T. Furusawa, R. Kimura, T. Yamache, K. Natsuura, M. Nakazawa, Y. Ataka, T. Ishida, T. Inakoa, Y. Matsuura, R. Ohtsuka, J. Ohashi, 1) University of Tsukuba, Tsukuba, Ibaraki, Japan; 2) Research Center for Hepatitis and Immunology, International Medical College, (IMI) Kanonoi Hospital, Ichikawa, Japan; 3) Graduate School of Asian and African Area Studies, Kyoto University, Kyoto, Japan; 4) Transdisciplinary Research Organization for Subtropics and Island Studies, University of the Ryukus, Nakaagami, Okinawa, Japan; 5) Department of Health Sciences, Kansai University, Osaka, Hyogo, Japan; 6) Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo, Tokyo, Japan; 7) Department of International Health, Kobe University Graduate School of Health Sciences, Kobe, Hyogo, Japan; 8) School of Policy Studies, PUCPR, Curitiba, Brazil; 9) Department of Genetics, Universidade Federal do Paraná, Curitiba, Brazil; 10) Department of Human Ecology, Faculty of Agriculture, Saga University, Saga, Japan; 11) Faculty of Health and Nutrition, Bunyoko University, Chigasaki, Kanagawa, Japan; 12) Japan Wildlife Research Center, Taito, Tokyo, Japan.

Oleoylthalamide (OEA), an appetite-suppressing mediator, is degraded by an integral membrane enzyme encoded by the fatty acid amide hydrolase (FAAH) gene. A previous study demonstrated that the activity of FAAH enzyme is influenced by a non-synonymous SNP (rs324420; Pro129Thr) in the FAAH gene. The FAAH activity has been suggested to be higher in subjects homozygous for rs324420-C (129Pro) than in those for rs324420-A (129Thr) due to a post-translational mechanism. The possible association of rs324420 with overweight/obesity has been assessed in various populations to date. However, the association studies have yielded contradictory results. In the present study, the associations of rs324420 with anthropometric measurements (i.e., height, weight, and body mass index) was examined in 694 adult subjects living in Solomon Islands and Tonga. A multiple regression analysis adjusted for age, sex, and population revealed that a copy of rs324420-C allele significantly increased weight by 2.3 kg (P-value = 0.0201), whereas rs324420-C was not associated with height. The rs324420-C allele was also significantly associated with increase in BMI (P-value = 0.00627 and slope = 0.91 kg/m2). Our results suggest that the efficient degradation of OEA due to 129Pro-FAAH protein may lead to excessive food intake and significant weight gain in Oceanic populations.

1953S
Brazilian population data on 26 non-CODIS STR loci used foraternity and kinship analysis. V.S. Sotomaior, T.F.C. Pereira, N. Gaburo Júnior, J.C.M. Magalhães, M. Malaghini, 1) Graduate Program in Health Sciences, PUCPR, Curitiba, Brazil; 2) Diagnósticos da América S.A., Brazil; 3) Department of Genetics, Universidade Federal do Paraná, Curitiba, Brazil. Microsatellites or short tandem repeats (STRs) are molecular markers that are widely used in population genetics and forensic analyses because of their high degree of polymorphism and heterozygosity. However, only a few population studies have been carried out in Brazil, particularly using markers that are not included in the US criminal justice DNA database systems - Combined DNA Index System (CODIS). In this study, we investigated 26 autosomal non-CODIS STRs (D1GATA113, D1S1627, D1S1677, D2S441, D2S1776, D3S3053, D4S4529, D4S2364, D4S2482, D5S2500, D6S474, D6S1017, D8S1115, D9S1122, D9S2157, D10S1248, D10S1435, D11S4463, D12S4949, D12S1017, D14S1434, D17S974, D17S1301, D18S853, D20S482, D20S1082, and D22S1045) in a sample of the Brazilian population. This sample was composed of 1,068 unrelated individuals who had undergone genetic kinship analyses between 2007 and 2013. The individuals (n = 1068) were from 5 Brazilian regions: south: 717; southeast: 163; central west: 36; northeast: 129; and north: 23. For some analyses and whenever data was available, the individual's self-declared ethnicity was taken into consideration, according to IBGE's criteria (white: 614, black: 43, brown: 174, and yellow: 19). All the loci analyzed showed great diversity, the number of alleles varied from 5 (D4S2364) to 14 (D10S1435) and the mean of alleles number in all loci was 8.6. The majority of alleles exhibited a frequency that was higher than 1%, being allele 9 at locus D4S2364 the most frequent allele (q = 0.572). No significant deviations from the Hardy-Weinberg equilibrium were observed in any locus or sample; similar findings were obtained for F<sub>ST</sub> values. The heterozygosity observed for each locus varied from 0.559 (D4S2364) to 0.851 (D9S2157). All 26 markers exhibited information capacity of use in kinship tests, with PIC > 0.5. The PDM was 81%, the PCE was 99.999999%, and the ITP was 99.999999%. In order to investigate population structure, the F<sub>ST</sub> values for each pair of samples were calculated according to region and ethnicity. Although differences were observed for some loci, none of the obtained values reached significance, indicating limited genetic differentiation among subsamples. Therefore, it has been demonstrated that the panel of markers investigated in this study are an important tool in the analysis of complex cases during kinship investigation, allowing the use of Brazilian population allele frequencies to infer the evidential value of kinship tests in Brazil.

1954M

The effects of inbreeding on human health depend critically on the number and severity of the recessive deleterious mutations carried by an individual. In humans, estimates of the burden of recessive mutations per individual are based either on comparisons between consanguineous and non-consanguineous couples, an approach that confounds socioeconomic and genetic effects, or on carrier screening for disease-causing mutations, which suffers from other biases, notably the highly incomplete catalogue of disease-causing mutations. To circumvent these limitations, we sought to estimate a lower bound of the burden by focusing on recessive lethal disorders in a founder population with almost complete Mendelian disease ascertainment and a known pedigree. By considering all autosomal recessive lethal diseases recognized in the population and simulating allele transmissions along the pedigree, we estimated that each haploid human genome carries on average approximately one autosomal recessive allele that leads to severe disorders at or after birth in homozygous condition. When compared with previous estimates, our result suggests that recessive mutations that are lethal constitute a substantial fraction of the total burden of recessive deleterious mutations in humans.

1955S
Testing the infinite sites assumption using the 1000 genomes dataset. S. Huang, D. Yuan, State Key Laboratory of Medical Genetics, Central South University, Changsha, Hunan, China.

The infinite sites model of the neutral theory is a fundamental assumption underlying nearly all population genetic and phylogenetic studies today but has yet to be properly tested. We here examined if from two novel perspectives using the 1000 genomes dataset. First, we studied the genetic diversity patterns of different human populations using a variety of different types of SNPs, such as noncoding, stop codon, nonsyn, syn, etc. Patterns shown by noncoding SNPs are expected to be similar to those shown by known functional stop codon SNPs, if most SNPs are not neutral. In contrast, neutral SNPs should show a most different pattern from stop codon SNPs. Second, it has long been well known that most genetic variations are shared among different human groups, which has been interpreted to mean few genetic differences among the ethnic groups (Lewontin, 1972). But the possibility of saturation or independent mutations to account for this phenomenon has yet to be examined and excluded. Saturation would mean limited rather than sufficient sites and fast evolving DNAs are expected to reach saturation faster. We compared the number of shared SNPs in DNAs of different evolutionary rates among different human populations to see if shared SNPs are more common in fast evolving DNAs relative to slow ones. Results from these tests and their implications for phylogenetic studies will be presented.
Inference of mutation rates using hidden relatedness, P.F. Palamara,1,2,3, P. Wilkinson,1,4 M. Fromer3,8,9, G. Kiryv,1,4, S. McCarthy,3,8,9, P. Sklar2,3, M. O’Donovan1,2,4, M. Wilmer1,2,4, J. Ryder1,2,4, 1) Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; 2) Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA; 3) Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Boston, MA, USA; 4) Department of Organismic and Evolutionary Biology, Harvard University, Boston, MA, USA; 5) Division of Psychiatric Genomics, in the Department of Psychiatry, and Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 6) Center for Psychiatry Research, Broad Institute of Harvard and MIT, Boston, MA, USA; 7) Medical Research Council Centre for Neuropsychiatric Genomics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK, 8) Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA; 9) Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 10) Analytic and Translational Genomics Unit, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; 11) Center for Computational Biology & Bioinformatics, Columbia University Medical Center, New York, NY, USA; 12) Department of Computer Science, Columbia University, New York, NY.

Reliably estimating the mutation rate in modern humans has several important implications for our understanding of demographic history (Scally and Durbin, Nature Reviews Genetics 2012). Recent estimates of the mutation rate obtained using de novo mutations in next-generation sequencing of families, however, were found to disagree with phylogenetic mutation rates derived from fossil evidence, motivating the development of new analytical methods. We describe an approach for the inference of mutation rates based on sharing of identical-by-descent (IBD) segments in sequencing data across purportedly unrelated individuals from a population. Using coalescent theory, we derive theoretical results for the distribution of mutation events found on IBD segments longer than a specified centimorgan threshold, for arbitrary demographic settings, under the SMC and SMC′ models. Leveraging the relationship between the length and the age of shared IBD haplotypes, we devise a method to estimate both genotype error rates and mutation rates. The proposed approach estimates both genotype error rates and mutation rates, and increases in statistical power compared to family-based analysis of de-novo mutations. This gain in power occurs despite the fact that the fraction of genome shared through long (e.g. >1cm) IBD segments across purportedly unrelated individuals is small.

Furthermore, analysis of de-novo mutations in trio-based studies is limited to genomic regions transmitted through known pedigree relationships, while when accurately phased data is available, mutation events can be analyzed on IBD segments across the quadratically larger set of all pairs of unrelated individuals. We validate the proposed methodology using synthetic datasets for a variety of demographic scenarios, and analyze mutation rates in 1246 trios that contained unrelated individuals from a recent exome sequencing study (Fromer et al., Nature 2014) of schizophrenia patients.

The evolution of genes underlying autism and schizophrenia and its impact on disease susceptibility, M.-J. Favé,1, J. Hussin,2, P. Awadalla1,2, 1) CHU Sainte-Justine, Department of Pediatrics, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 3) CARTaGENE, 3333 Queen Mary Road, Office 493, Montreal, Quebec.

It has been hypothesized that a genotypic and phenotypic continuum exists between autism and schizophrenia, which is supported by both clinical data showing a gradient in diagnostic severity and patterns in the impairment of language, communication, and social behaviour. Recent work proposes that the human susceptibility to develop these disorders may originate as a collateral consequence of the selection for cognitive traits. Over the last decade, hundreds of both rare and common variants associated with the development of these disorders have been identified: several common variants have been found by genome-wide association studies, and rare de novo mutations have been captured by whole-exome sequencing studies of sporadic cases, for which no family history exists. We can therefore test whether genes harbouring common and de novo variants have evolved under different selective regimes, and whether local genomic features influence their variation and evolution. Using a recombination map from a French Canadian population recruited by the CARTaGENE project, we found that de novo variants associated with these disorders are enriched in regions of low recombination rate, whereas no such biases are found for common variants. Low recombining genes harbouring de novo variants are predominantly involved in fundamental cellular processes, whereas the few genes (49, or 2.7%) found in high recombination regions are enriched in membrane-proteins, axonal guidance and neuronal differentiation. In order to detect if genes harbouring common and de novo variants associated with autism and schizophrenia evolved under different selective regimes, we measured signals of selection using dN/dS ratios on the whole gene, and using sliding windows within genes. We performed these tests along the mammal and the primate lineage to estimate the timing of any selective events. Although most genes show patterns of variation consistent with purifying selection, some genes in which many recent de novo mutations have been detected, such as ABCA13, LRP1 and MYH9, show strong signals of positive selection in at least one domain of the evolutionary tree. These genes encode proteins that interact with 12 pathogen species. Human-pathogen interactions are a constantly evolving interface between host and disease. We created a human pathogen interaction database and used the integrated haplotype score (IHS) to detect recent positive selection in genes that interact with proteins from 24 different infectious organisms. The Human Genome Diversity Panel was used to identify specific populations harboring pathogen-interacting genes that have undergone positive selection. We found that genes that interact with 12 pathogen species have undergone recent positive selection. The 12 pathogens are human immunodeficiency virus 1, Yersinia pestis, human respiratory syncytial virus, human herpesvirus 8, Bacillus anthracis, measles virus, dengue virus, murid herpesvirus 1, Francisella tularensis, Zaire ebolavirus, hepatitis C virus, and influenza A virus. Which human-pathogen interactions that interact with Yersinia pestis show significant signs of selection. Our results show strong positive selection in genes that interact with human immunodeficiency virus 1, providing further support for the hypothesis that ancient humans were repeatedly exposed to lentivirus pandemics. These results indicate that infectious disease created distinct genetic footprints within affected populations.
1961S Evaluating the impact of recent human demographic history on the frequency spectra using numerical solution of time-inhomogeneous diffusion equation. E. Koch1, J. Novembre2. 1) Department of Ecology and Evolution University of Chicago, Chicago, IL; 2) Department of Human Genetics University of Chicago, Chicago, IL.

Differences in recent demographic history appear to be an important driver of observed levels of genetic diversity among human populations. Recent attention has particularly centered on how populations that went through the out-of-Africa bottleneck have lower heterozygosity and polymorphic sites that are proportionally more likely to be neutral or predicted to be damaging. These results have suggested differences in the frequency spectrum of deleterious variation are also caused by varying population demographic histories. To investigate these phenomena in more detail, we perform numerical solutions to time-inhomogeneous diffusion equations for the allele frequency spectrum under the Poisson Random Field Model. This allows us to efficiently examine how the frequency spectra has evolved in survival in severe hypoxia, demonstrating their functional relevance in hypoxia adaptation. Using whole genome genotyping data from three high-altitude human populations, we investigate the frequency and time distribution of selective effects. We also are able to easily stratify variation through time under a large number of possible human demographies and allows us to efficiently examine how the frequency spectra has evolved under the Poisson Random Field Model. This perform numerical solutions to time-inhomogeneous diffusion equations for demographic histories. To investigate these phenomena in more detail, we have also exome sequenced 83 Khomani San individuals to high coverage, generating one of the largest indigenous African exome datasets sequenced outside of the 1000 Genomes Project. In this study, Khomani individuals have 11.5% admixture with Europeans and 10.9% admixture with Bantu speakers on average. European ancestry significantly lightens skin and explains 13.3% of the variance in pigmentation, and Bantu ancestry significantly darkens skin and explains 16.1% of the variance in pigmentation on average. We estimate that pigmentation is highly heritable (h2 = 0.887 ± 0.188 standard error) and find that most of the heritability can be explained by 50 known pigmentation genes (0.527 ± 0.310 or 64.1% on average). After controlling for admixture with European and Bantu-speaking populations, a linear mixed model GWAS approach does not identify variants significantly associated with pigmentation. However, pigmentation genes are among the most globally differentiated between the Khomani San and European or Bantu individuals, and aggregating differentiation with association data improves power to detect variants influencing selected traits. We identify highly differentiated variants through the Khomani exome data and find potential signatures of positive selection in pigmentation candidate genes in populations of East Asian ancestry. Based on the list of genes that show putative signatures of selection in East Asia, we prioritized a number of polymorphisms based on allele frequency information (e.g. differences in frequency between East Asian and non-East Asian populations) 2/ potential functional effects (e.g. Polyphen, SIFT and CADD scores) and 3/ conservation (e.g. GERPO++ scores). The panel of SNPs selected includes 3 markers in the OCA2 gene (rs375722964, rs7522053 and rs4659610), one marker in the MLPH gene (rs2292881), 2 markers in the OPRM1 gene (rs1799971 and rs9171661), one marker in the EGFR gene (rs2227983), 4 markers in the BNC2 gene (rs9406647, rs3739784, rs10756778 and rs10962591), one marker in the TH gene (rs1090220), 2 markers in the OCA2 gene (rs1800414, rs74655330 and rs7497270), one marker in the TRPM1 gene (rs3809578) and 2 markers in the MC1R gene (rs33932559 and rs885479). We evaluated the association of these polymorphisms with skin pigmentation measured quantitatively using a DSM II colorimeter in a sample comprising 452 individuals of East Asian ancestry. Two previously described nonsynonymous polymorphisms within the OCA2 gene, rs1800414 (His615Arg) and rs74655330 (Ala481Thr) were strongly associated with melanin levels in this sample. Under an additive model, the common rs1800414 G allele, coding for Throneine, is associated with higher levels of melanin in the skin. These associations were consistent with rs1800414 having a stronger effect on melanin levels than rs1800414 (decrease of 1.3 melanin units). Other significant associations with skin pigmentation were observed for any of the other variants.

1962M The Genetic Architecture of Skin Pigmentation in the Southern African Khomani San. A.R. Martin1, J.M. Granka2, C.R. Gignoux3, C. Lint4, C. Uren4, M. Möller4, C.J. Wereley4, J.M. Kidd5, M.W. Feldman5, E.G. Hoare4, C.D. Bustamante4, B.M. Henn5. 1) Genetics Department, Stanford University, Stanford, CA; 2) Department of Biological Sciences, Stanford University, Stanford, CA, 94305; 3) Department of Ecolog and Evolution, SUNY Stony Brook, NY 11794; 4) Division of Molecular Biology and Human Genetics, Stellenbosch University, Tygerberg, South Africa; 5) Department of Human Genetics, University of Michigan, Ann Arbor MI.

Skin pigmentation is one of the most recognizable and diverse phenotypes in humans across the globe, but its highly genetic basis has been primarily studied in northern European, Asian, and African American populations. The Eurasian pigmentation alleles are among the most differentiated variants in the genome, suggesting strong selection for light skin pigmentation. Light skin pigmentation is also observed in the far southern latitudes of Africa among KhoeSan hunter-gatherers of the Kalahari Desert. The KhoeSan hunter-gatherers are among the oldest human populations, believed to have diverged from other populations 100,000 years ago, and maintain extraordinary levels of genetic diversity. It is unknown whether light skin pigmentation represents convergent evolution or the ancestral human phenotype. We have collected ethnographic information, pigmentation phenotypes, and genotype data from 136 individuals in the Khomani San from the Kalahari. To understand the genetic basis for light skin pigmentation, we have also exome sequenced 83 Khomani San individuals to high coverage, generating one of the largest indigenous African exome datasets sequenced outside of the 1000 Genomes Project. In this study, Khomani individuals have 11.5% admixture with Europeans and 10.9% admixture with Bantu speakers on average. European ancestry significantly lightens skin and explains 13.3% of the variance in pigmentation, and Bantu ancestry significantly darkens skin and explains 16.1% of the variance in pigmentation on average. We estimate that pigmentation is highly heritable (h2 = 0.887 ± 0.188 standard error) and find that most of the heritability can be explained by 50 known pigmentation genes (0.527 ± 0.310 or 64.1% on average). After controlling for admixture with European and Bantu-speaking populations, a linear mixed model GWAS approach does not identify variants significantly associated with pigmentation. However, pigmentation genes are among the most globally differentiated between the Khomani San and European or Bantu individuals, and aggregating differentiation with association data improves power to detect variants influencing selected traits. We identify highly differentiated variants in the Khomani exome data and find potential signatures of positive selection in pigmentation candidate genes in populations of multiple canonical pigmentation genes, including OCA2 and MITF. Our results highlight the strength of diverse population studies to explain phenotypic variation impacted by human evolutionary history.
1964M

Positive Selection on Loci Associated with Drug and Alcohol Dependence. B. Sadler,1,2, G. Haller1,3, A. Goate1,3,5. 1) Department of Psychiatry, Washington University, St. Louis, MO, USA; 2) Department of Genetics, Washington University, St. Louis, MO, USA; 3) Department of Neurology, Washington University, St. Louis, MO, USA.

Much of the evolution of human behavior remains a mystery, including how certain disadvantageous behaviors are so prevalent. Nicotine addiction is one such phenotype. Several loci have been implicated in nicotine related phenotypes including the nicotinic receptor gene clusters (CHRNs) on chromosomes 8 and 15, and the nicotine metabolizing gene CYP2A6 on chromosome 19. Here we use 1000 Genomes sequence data from 3 populations (Africans, Asians and Europeans) to examine whether natural selection has occurred at these loci. We used Tajima’s D and the integrated haplotype score (iHS) to test for evidence of natural selection. Our results provide evidence for strong selection in the nicotinic receptor gene cluster on chromosome 8, previously found to be significantly associated with both nicotine and cocaine dependence as well as evidence of weaker, but still detectable, selection acting on the region containing the CHRNA5 nicotinic receptor gene on chromosome 15, that is genome wide significant for risk for nicotine dependence. To examine the possibility that this selection is related to memory and learning, we utilized genetic data from the Collaborative Studies on the Genetics of Alcoholism (COGA) to test variants within these regions with three tests of memory and learning, the Wechsler Adult Intelligence Scale (WAIS) Block Design, WAIS Digit Symbol and WAIS Information tests. Of the 17 SNPs genotyped in COGA in this region, we find one significantly associated with WAIS digit symbol test results. This test captures aspects of reaction time and memory, suggesting that a phenotype relating to mem-
ory and learning may have been the driving force behind selection at these loci.

1965S

Neanderthal Origin of the Haplotypes Carrying the Functional Variant Val92Met in the MC1R in Modern Human Populations. Q. Ding1, Y. Hu1, S. Xu1, S. Wang1, H. Li1,2, R. Zhang1, S. Yan1, J. Wang1,1, Jin1,2,1. 1) State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China; 2) CAS-MPG Partner Institute for Computational Biology, Shanghai Institute for Biological Sciences (SIBS), Chinese Academy of Sciences (CAS), Shanghai, China.

Skin color is one of the most visible and important phenotypes of modern humans. Melanocyte-stimulating hormone and its receptor played an important role in regulating skin color. Here we present evidence of Neanderthal introgression encompassing the melanocyte-stimulating hormone receptor gene MC1R. The haplotypes from Neanderthal introgression diverged with the Altai Neanderthal 103.3 KYA, which postdates the anatomically modern human - Neanderthal divergence. We further discovered that all of the putative Neanderthal introgressive haplotypes carry the Val92Met variant, a loss-of-function variant in MC1R that is associated with multiple dermatological traits including skin color and photoaging. Frequency of this Neander-
thal introgression is low in Europeans (~5%), moderate in continental East Asians (~30%), and high in Taiwanese aborigines (60-70%). Since the putative Neanderthal introgressive haplotypes carry a loss-of-function variant that could alter the function of MC1R and is associated with multiple dermatological traits related to skin color, we speculate that this Neanderthal introgression, together with the previously reported Neanderthal introgression at HYAL2, may have played an important role in the local adaptation of modern Eurasians to sunlight intensity.

1966M

Altitude adaptation in Tibet caused by introgression of Denisovan-like DNA. E. Huerta-Sanchez1,2,3, X. Jin2,4, A. Asan2,3,5, Z. Bianga1, B. Peter4, H. Huschke1, Y. Li2,5,6,7, M. He1,2,5,6,7, P. Nif1, B. Wang1,2,4, Y. Liu1, H. Huassang1, Z. Luosang1, Z.X. Ping1,2, K. Li1,6, G. Gao1,7, Y. Yin1, W. Wang1, X. Zhang2,3,12,13, X. Xu1, H. Yang1,4,12,13, Y. Li1, J. Wang1,2,4, J. Wang1,2,4. 1) University of California at Berkeley, Berkeley, CA; 2) BGI-Shenzhen, Shenzhen, China; 3) University of California at Merced, Merced, CA; 4) School of Bioscience and Bioengineering, South China University of Technology, Guangzhou, China; 5) Binhai genomics institute, BGI-Tianjin, Tianjin 300308, China; 6) Tianjin Translational Genomics, BGI-Shenzhen, Shenzhen, China; 7) The People’s Hospital of Lhasa, Lhasa, 850000, China; 8) Bioinformatics and Computational Biology Program, Iowa State University; 9) Department of Biological Sciences, Middle East Technical University, Ankara, Turkey; 10) The No.2 people’s hospital of Tibet Autonomous Region, Taipa, Macau 999078, China; 11) Department of Medicine, University of Hong Kong, Hong Kong; 12) Department of Statistics, University of Califor-
ia, Berkeley, CA; 13) Department of Biology, University of Copenhagen, Ole MaaløesVej 5, 2200 Copenhagen, Denmark; 14) Macau University of Science and Technology, AvenueiWa long, Taipa, Macau 999078, China; 15) Department of Medicine, University of California, Berkeley, CA; 16) Department of Biology, University of Copenhagen, Ole MaaløesVej 5, 2200 Copenhagen, Denmark.

As modern humans migrated out of Africa, they encountered many different environmental conditions including temperature extremes, new pathogens, and high altitude. These diverse environments have likely acted as agents of natural selection and led to local adaptations. One of the most illustrious examples in humans is the adaptation of Tibetans to the hypoxic environment of the high-altitude Tibetan plateau. A hypoxia pathway gene, EPAS1, was previously identified as having the most extreme signature of positive selec-
tion in Tibetans and was shown to be associated with differences in hemi-
globin concentration at high altitude. Re-sequencing the region around EPAS1 in 40 Tibetan and 40 Han individuals, we find that this gene has a highly unusual haplotype structure that can only be convincingly explained by introgression of DNA from Denisovans into humans. Scanning a larger set of worldwide populations, we find that the selected haplotype is only found in Denisovans and in Tibetans, and at very low frequency among Han Chinese. Furthermore, the length of the haplotype, and the fact that it is not found in any other populations, makes it unlikely that the Tibetan/Denisovan haplotype sharing was caused by incomplete ancestral lineage sorting rather than introgression. Our findings illustrate that admixture with other hominin species has provided genetic variation that helped humans adapt to new environments.

1967S

Whole genome sequencing to uncover adaptation to high altitude in the Andes. M. Muzzio1,2, K. Slivinski1,2, M.C Yee3, T. Cooke2, C.Busta-
mante1, G. Bailliel4, C.M.Braun1, E.E.Kenny1,2,4,6,7,8,9,1) Consejo Nacional de Investigaciones Cientificas y Tecnologicas, La Plata, Buenos Aires, Argentina; 2) Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Argentina; 3) The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, NY; 4) Delaney Lab, Carnegie Institution of Washington, Department of Plant Biology, CA; 5) Stanford University School of Medicine, CA; 6) Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, NY; 7) Center for Statistical Genetics, Icahn School of Medicine at Mount Sinai, NY; 8) The Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, NY.

There is interest in human adaptation to a diversity of environments, including finding the genetic basis to phenotypic adaptations to pressures such as hypoxia. We have preliminary illumina Exome Array data on a set of 43 individuals from high altitude villages in the Andes from the Humahuaca area, Argentina (~2500 meters above sea level) and 11 individuals from a neighboring lowland population, Tartagal, Argentina (less than 500 meters above sea level), all with over 90% Native American ancestry estimated using the Admixture software. Currently, we are sequencing full genomes of 10 individuals from each of these populations, in search for new population-specific variants. We will use the population branch statistics (PBS) to identify highly differentiated genomic regions between the highlanders (Andean) and lowlanders (Chaqueos). We will discuss the results of our scan in light of related work on the adaptation of Tibetans, Ethiopians, and other Andean populations to hypoxia.
1968M

IFNL3/IFNL4 region shows evidence for recent positive selection specific to Asian populations. G.L. Wojick, C.D. Bustamante. Department of Genetics, Stanford School of Medicine, Stanford, CA.

Hepatitis C virus (HCV) is a global health burden, chronically infecting 130-150 million people and causing 350,000-500,000 deaths per year from HCV-related liver disease. Twenty-five years after the discovery of HCV, there is still no vaccine and treatment remains ineffectual for the majority of individuals. Heterogeneity in clinical outcomes such as spontaneous clearance of the virus, as well as sustained virologic response (SVR) after treatment, has been observed between individuals of different genetic ancestry. Previous genetic studies have pinpointed single nucleotide polymorphisms (SNP) in the interferon-3 and 4 (IFNL3/IFNL4) region (rs12979860) as being strongly associated with clinical outcome. While the derived and favorable allele of rs12979860 (C) is present globally, its frequency is greatly differentiated by continent with the lowest in African populations (34-49%).

Natural selection studies have impacted genetic research and our understanding of human adaptations, including malaria resistance, skin pigmentation and Statistics, Utah State University, Logan, Utah; 5) Department of Mathematics and Statistics, Utah State University, Logan, Utah.

Natural selection studies have impacted genetic research and our understanding of human adaptations, including malaria resistance, skin pigmentation, and the highest in Asian populations (89-96%). To determine if these differences are due to selective pressures, data from the phase 3 release of the 1000 Genomes Project (TGP) was analyzed for population-specific signatures of selection. Derived allele frequency (DAF), Fst, nucleotide diversity (π), and haplotype structure were examined and compared in populations from Europe, Africa, Asia, and the Americas. A 5 kilobase (kb) region around IFNL3/IFNL4 showed decreased nucleotide diversity, high DAF, and increased haplotype homozygosity in Asian populations. This pattern is not found in Native American populations, suggesting recent positive selection specific to Asian. Historical selective pressures from HCV, or likely a related ancestral virus, may have driven the favorable rs12979860 allele to near fixation. However, Asia currently has disproportionately high HCV-related disease burden, as much as 70% of the world's HCV infections are found there. Using SNPs with strong evidence for association with PC2 and strong evidence for selection from the X-HEH test as a starting point, we performed a genome-wide scan for adaptive mutations without foreknown, adaptive phenotypic traits. This genome-wide scan was performed in all human populations at the same time, allowing for the simultaneous assessment of the global genetic diversity of both binding and DNaSe sites, including those of the MTRR, SEPP1, HLA, RAP-GEF5, AUT52, TNKS, HBG1, and L2HGDH genes. Enhancer enrichment analysis of the top 1,000 SNPs revealed a 10.8 fold significant enrichment of these loci in human populations, but not in other species. In summary, we identified several functional variants that differentiate the Gullah from the Sierra Leoneans, suggesting that recent selection may be operating at these loci. Identification of functional regions that might be under selection in the Gullah has the potential to elucidate disease risks in AA.

1970M


Many common diseases are more prevalent in specific ethnic groups. Given the growing number of disease-associated loci in regions that show evidence of selection, identification of alleles under selection may provide insight into disease susceptibility. Relative to other African-Americans (AA), the Gullah are a population that has been isolated from other Africans and the highest in Asian populations (89-96%). To determine if these differences are due to selective pressures, data from the phase 3 release of the 1000 Genomes Project (TGP) was analyzed for population-specific signatures of selection. Derived allele frequency (DAF), Fst, nucleotide diversity (π), and haplotype structure were examined and compared in populations from Europe, Africa, Asia, and the Americas. A 5 kilobase (kb) region around IFNL3/IFNL4 showed decreased nucleotide diversity, high DAF, and increased haplotype homozygosity in Asian populations. This pattern is not found in Native American populations, suggesting recent positive selection specific to Asian. Historical selective pressures from HCV, or likely a related ancestral virus, may have driven the favorable rs12979860 allele to near fixation. However, Asia currently has disproportionately high HCV-related disease burden, as much as 70% of the world's HCV infections are found there. Using SNPs with strong evidence for association with PC2 and strong evidence for selection from the X-HEH test as a starting point, we performed a genome-wide scan for adaptive mutations without foreknown, adaptive phenotypic traits. This genome-wide scan was performed in all human populations at the same time, allowing for the simultaneous assessment of the global genetic diversity of both binding and DNaSe sites, including those of the MTRR, SEPP1, HLA, RAP-GEF5, AUT52, TNKS, HBG1, and L2HGDH genes. Enhancer enrichment analysis of the top 1,000 SNPs revealed a 10.8 fold significant enrichment of these loci in human populations, but not in other species. In summary, we identified several functional variants that differentiate the Gullah from the Sierra Leoneans, suggesting that recent selection may be operating at these loci. Identification of functional regions that might be under selection in the Gullah has the potential to elucidate disease risks in AA.

1971S


Zinc transporters play very important roles in all eukaryotes by maintaining the rational zinc concentration in the cells. The genetic diversities of those transmembrane proteins are expected to reflect the biologic functions of both zinc transporters themselves and the other zinc-related enzymes. Although there are many works about the huge diversity of Zn content in soils or crops around globe-scales, the diversity of Zinc transporter genes (ZTGs) in across human populations have been not well studied. Here, we investigated the global genetic diversity of 24 human ZTGs, including 10 SLC30A and 14 SLC39A family genes basing on full sequencing data of 1000 Genomes project. Intriguingly, we found some of ZTGs are very evolutionarily convergent across all human populations, such as SLC30A6 with significant low GA-FST (0.015), while some other ZTGs exhibited the extremely high differentiations among populations, such as SLC30A9/GA-FST, 0.284, SLC30A3 (GA-FST, 0.154), etc. Moreover, ZTGs harbored higher differentiated SNPs than random genes, suggesting that large genetic variations among populations are kept in ZTGs. Moreover, we found that SLC30A9 and SLC30A3 were more likely to be selected both in East Asians and Africans but the selective pressures were different in the continental groups. In SLC30A9, the frequencies of two different alleles of a non-synonymous SNP, rs10476266, are almost fixed in East Asians and East Asians, respectively, I.e. 96.4% A in CHB, 92% G in YRI. Therefore, there are two different functional haplotypes exhibited dominated abundance in Africans and East Asians, respectively. Furthermore, according to previous studies of global distributions Zn contents in soils, we also found that the high zinc occupies frequency of SLC30A9 among different human populations. We speculate that population was forced to utilize advantageous functional haplotype to adapt to the local zinc state or diets and the genetic differentiation of ZTGs could contribute to population health. Shifting environments or zinc transport gene in different living environment so that the balance of zinc concentration in serum or cell can be kept properly. Our investigations should facilitate to further functional studies of ZTGs and medical studies on worldwide and regional nutrient problem and zinc-related diseases.
**1972M**


Background: The innate immunity genes are the first line of host defence, being developed before the separation of invertebrates and vertebrates. There is considerable interest to decipher the effect of natural selection on these genes across diverse human populations, and also across invertebrates and vertebrates. To account for the change in pathogen load and diversity, Among them the most important are the cell surface Toll-like Receptors (TLRs) that engage in direct host-pathogen interactions in humans and other mammals. They are the homologs of Toll genes identified in Drosophila. Furthermore, genes involved in the N-Glycosylation pathway that shapes the cell-surface glycome moeately regulates the innate immune response through host-pathogen interactions and are hence potential candidates for natural selection. Objectives: To study the differential natural selection patterns on the Toll-Like Receptors and the N-Glycosylation genes across human populations and characterizing the TLH homologous Toll genes in multiple Drosophila genomes. Methods: We have generated DNA sequence data on the important cell surface TLR genes in pre and post agricultural populations (411 individuals) and analysed them in conjunction to the data available in the 1000 Genomes Project (n=1092). The TLR homologs in Drosophila Toll genes are identified and genome-level data are downloaded from the Drosophila Reference Gene Panel for further analysis. For studying the N-Glycosylation genes across Indian populations, we have generated sequence data of the same TLR genes from 300 individuals sequenced in the 1000 Genomes Project (n=1092). The TLR homologs in Drosophila Toll genes, SNVs with MAF<0.05 are predominant. The results are quite similar to that observed in humans. The N-Glycosylation genes in humans show significantly lower heterozygosity values compared to neutrality. Population differentiation indices (FST) are lower for all the genes across multiple populations. Inference: The results obtained suggest possible role of purifying selection operating on key innate immunity genes across arthropod vertebrates, thus focusing on their immense importance in the survival of the host.

**1973S**

Evolutionary history of pigmentation candidate gene diversity in a Melanesian population. H. Norton, M. Herman. Department of Anthropology, University of Cincinnati, Cincinnati, OH.

Pigmentation of the skin, hair, and eyes are complex polygenic traits determined by multiple loci. Human skin pigmentation is a trait that is believed to have evolved under strong natural selection in response to varying levels of ultra-violet radiation (UVR) intensity. Lighter skin color has evolved multiple times in human evolutionary history, but it is unclear if the darker skin color observed in many high UVR populations is also the result of evolutionary convergence (suggesting that population-specific mutations may have been favored by positive selection) or if instead ancestral variants associated with darker skin color have been maintained in high-UVR populations via purifying selection. To begin to address this question we compare DNA sequence variation from multiple pigmentation candidate genes in a Melanesian population to variation observed in European, East Asian, and African populations sequenced in the 1000 Genomes Project. Summaries of the site frequency spectrum, including Tajima’s D (TD), for three genes, ASIP, OCA2, and TYR, are presented. Inference: Of all the variants targeted by positive selection in the Melanesian population (ASIP TD = 0.037, OCA2 TD = -0.85, TYR1 TD = 0.55). With the exception of a single novel haplotype in the OCA2 locus observed at a frequency of ~10%, there is little evidence that Melanesians exhibit any high frequency population-specific haplotypes at these loci, suggesting that if an independent adaptation to high UVR conditions occurred in Melanesians then other pigmentation loci are responsible. However, there is also little evidence that Melanesians are similar to Africans at these loci, which one might expect if Melanesians share ancestry with Africans. Other papers have shown significant enrichment in high ranking FST signals. In summary the proposed approach based on three genomes per population required to further reduce the false positive rate. These promising results, given the limitations imposed by the small sample sizes, make our method suitable to be applied on newly sequenced populations (expected to be released on Mid June 2014, during the SMBE conference).

**1974M**


Interdepartmental Program in Bioinformatics, University of California, Los Angeles, CA; 2) Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA; 3) Department of Human Genetics, University of Chicago, IL.

The strength of purifying selection is a central factor underlying levels of genetic diversity in a population and is important to characterize to understand the expected genetic architecture of disease traits. Recent sequencing studies with large sample sizes have revealed a much higher proportion of non-synonymous compared to synonymous mutations. These results are consistent with neutral expectations only if the same selective coefficient is applied to both types of mutations. This finding suggests that natural selection is acting against such variants to keep them at low frequencies in the population. To estimate the strength of purifying selection, we have developed a method that uses the lengths of pairwise haplotype identity among rare-variant-carrying haplotypes. Unlike previous approaches, our method conditions on the present-day frequency of the allele and is based on the intuition that alleles under purifying selection are on average younger than neutral alleles and, therefore, should have higher average levels of haplotype identity among variant carriers. To obtain the probability distribution on the lengths of pairwise haplotype identity, one needs to perform two integrations: one over all possible allele frequency trajectories and another one over all pairwise coalescent times given a certain allele frequency trajectory. The integration over the space of possible allele frequency trajectories is done using a fast importance-sampling algorithm while the integration over the coalescent times is done using an analytical solution. Using the probability of the lengths of the haplotypes under different selective coefficients, we can calculate the likelihood of observing a single variant set of haplotypes. We use simulations to test how accurately the method estimates the selective coefficient under different demographic scenarios, such as a constant population size and a realistic model of European population growth. Variants with the same selective coefficient are harder to differentiate from neutral variants in scenarios of recent population growth. These methods will be applied to a set of 202 drug target genes sequenced in 14,002 individuals (Nelson et al, 2012, Science) to identify which genes are most likely to harbor damaging variants that may predispose to disease.

**1975S**

Highlighting strongly differentiated regions using three high coverage genomes each from a set of worldwide human populations. L. Pagani, T. Kivisild.


1) Department of Anthropology, University of Cambridge, Cambridge, Cambridge, United Kingdom; 2) The Wellcome Trust Sanger Institute, CB10 1SA, Hinxton, UK; 3) Molecular Anthropology Lab, Department of Biological Geological and Environmental Sciences, University of Bologna, Italy.

Following the steady reduction in sequencing costs, several international projects will shortly make available sets of 2-4 high coverage genomes each from hundreds of worldwide human populations. While these resources can be used to refine the demographic histories of the studied populations, little can be done to detect signatures of differentiation, possibly driven by natural selection, on these populations. The selection scan methods available to date indeed focus on various genomic components (SNPs, Haplotypes, LD blocks) rather than on genome frequencies rather than on the full sequence information. Here we show how the top 1% of genic regions analysed using only three genomes each from two populations (CEU and YRI) contains as many as 25% of the top 5% FST candidates obtained using 160 low coverage individuals from each of the same populations. If the results from each population are combined in three pairs, and FST based on average pairwise differences is calculated between populations. The average FST is computed on a sliding window of 10000 or 50000 bp across all the pop1-pop2 sets of genomes per gene. The top 1% windows showing the highest differentiation were selected and inspected for their gene content. Of the 1785 genes identified by the FST scan based on the 160 low coverage individuals (taken as the gold standard), 98 were found among the 439 genes included in the top 1% 50000bp windows of the YRI-CEU pairs. This shows a 24-fold enrichment with a p-value of 1.2 × 10^-10. The empirical ranking nature of the gold standard did not allow a formal assessment of the false positive rate of our newly developed method. However, the overlap between the top genes retrieved using the 10000 and 50000bp window approach showing a significant enrichment in high ranking FST signals. In summary the proposed approach based on three genomes per population is capable of retrieving at least 25% of the genes under putative natural selection found from traditional methods. Ongoing power assessment on these genome sets will provide these genome sets per population required to further reduce the false positive rate. These promising results, given the limitations imposed by the small sample sizes, make our method suitable to be applied on newly sequenced populations (expected to be released on Mid June 2014, during the SMBE conference).
1976M
Identification of functional variants under positive selection in the Gullah African American population of South Carolina. P.S. Ramos1, S. Sajutin2, J. Divers3, Y. Huang4, U. Nayak3, W.M. Chen5, K.J. Hunt6, D.L. Kamen7, G.S. Gilkeson4, J.K. Fernandes8, J. Spruill8, W.T. Garvey8, M.M. Sale3, C.D. Langefeld1. 1) Department of Medicine, Medical University of South Carolina, Charleston, SC; 2) Department of Biostatistical Sciences and Center for Public Health Genomics, Wake Forest Medicine, Winston-Salem, NC; 3) Department of Medicine and Center for Public Health Genomics, University of Virginia, Charlottesville, VA; 4) Department of Nutrition Sciences and Birmingham VA Medical Center, University of Alabama, Birmingham, AL.

Relative to other African-Americans (AA), the Gullah population has lower European admixture and higher ancestral homogeneity from the Sierra Leone (SL) area in Far-West Africa. The shorter genetic distance between the Gullah and SL suggests that these genetic signals, such as regions under recent selection, may be more easily detected in the Gullah than in other AA populations. Since both protein-coding and regulatory variation have important roles in recent human adaptation, the goal of this study was to integrate evidence for natural selection with functional annotation for the identification of biologically relevant signals. We computed the cross-population extended haplotype homozygosity test (XP-EHH) to identify alleles with higher than expected frequency relative to their haplotype length in the Gullah (n=277) relative to SL (n=400), to HapMap Phase II YRI (n=203), and CEU (n=165). In total, 679,513 SNPs with MAF > 5% met standard GWAS QC criteria. Variants that met suggestive significance (XP-EHH > 4, P < 0.04) were annotated and prioritized based on the potential impact of amino acid changes and regulatory functions using RegulomeDB and HaploReg. Nearly the same number of loci showed significant evidence for selection between the Gullah and YRI (0.15% of all SNPs), and Gullah and SL (0.14%), although only 106 SNPs in 12 regions showed evidence for selection in both comparisons. Fewer loci showed evidence for selection between Gullah and CEU (1%). Enhancer enrichment analyses revealed a significant enrichment of strongest enhancers in H1 human embryonic stem cells. Several regions harbor missense SNPs, including those showing evidence of selection between the Gullah and the YRI (CENPO, FSHR, C7orf67, GRK5, EKMT1, USP31), Gullah and SL (PRR11L4), and Gullah and CEU (ADPRHL2, LCT). Other regions harbor multiple SNPs with high regulatory scores based on the simultaneous presence of eQTLs, transcription factor binding and DNAse sites, including those showing evidence of selection between Gullah and CEU (ADPRHL2, LCT, PRR11L4, USP31), and Gullah and SL (e.g., CCR2, ADcy2, TNKS), and Gullah and CEU (e.g., NARS2). These results reveal several novel regions with evidence for selection and concomitant high functional potential in the Gullah AA population. Given that many common diseases are more prevalent in specific ethnic groups, identification of functional regions under selection in the Gullah has the potential to elucidate disease risks in AA.

1978M
The Clade Fitness Proxy haplotype score: delineating the progression of soft selective sweeps at the haplotype level. R. Ronen1, G. Tesler2, S. Zako3, N.A. Rosenberg4, V. Balan2. 1) Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA; 2) Department of Mathematics, University of California, San Diego, La Jolla, CA; 3) Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA; 4) Department of Biology, Stanford University, Stanford, CA.

The dynamics of selective sweeps have long been of interest in population and evolutionary genetics. Of particular interest is the problem of selection on standing variation, when the beneficial allele already segregates in the population at a non-negligible frequency at the time that selection begins. In this case, multiple haplotypes all carrying the beneficial allele will drift relative to one another, while outcompeting other haplotypes. Moreover, the beneficial allele may fix without any of the carrying haplotypes reaching fixation. Consequently, we observe a soft selective sweep, or a weakening of the classical signature of a selective sweep. To identify regions undergoing soft sweeps, we develop a new coalescent-based statistic dubbed the Clade Fitness Proxy (CFP) score, which scores individual haplotypes based on the frequencies of alleles they carry. We develop the statistical properties of the CFP score in the framework of a neutral coalescent model. We use theoretical calculations and simulated data to demonstrate its utility both as a general test of selection and as a proxy for the relative fitness of individual haplotypes. Under a soft sweep, we observe a bimodal distribution of CFP scores, where ‘fit’ haplotypes have high CFP score and ‘unfit’ haplotypes have low CFP score. This bimodality can be used in a general test of selection. Perhaps more interestingly, it suggests that CFP scores can serve as a proxy for the relative fitness of individual haplotypes. Thus, the CFP score may be useful for elucidating the state of a selective sweep within a population sample (i.e., at the level of individual haplotypes). Once a sweep has completed and the beneficial mutation has reached fixation (all haplotypes have ‘fit’), the in-sample distribution of CFP scores is no longer bimodal. While at this stage the relative fitness of haplotypes is not of interest, as it is by definition equal, the distribution of CFP scores is still highly informative about the occurrence of a sweep in the population’s history for many genera-

1977S
A model for gene expression level evolution to identify expression conservation, divergence, and diversity. R. Rohlf, R. Nielsen. Integrative Biology, University of California, Berkeley, Berkeley, CA.

As RNA-Seq becomes more available, we see more extensive comparative expression datasets, particularly datasets with multiple individuals sampled per species. These new data facilitate transcriptome-scale analyses of expression variance between and within species. We have developed a phylogenetic model of expression level evolution which can be used to specifically investigate the ratio of within to between species expression variance. We have applied this method to a phylogeny of 15 mammals, mostly primates. We have identified genes with high expression divergence between species as putative targets for expression level adaptation. Conversely, we have identified genes with high expression conservation across species as putative targets for selection and conservation between species, which likely have plastic responses to environmental inputs. Additionally, we identified genes on the human and catarrhine lineages which show a shift in expression level, putatively due to lineage-specific expression level adaptation. This analysis shows the flexibility and utility of our model in facilitating comparative expression analyses to illuminate the biological meaning of expression levels in humans and other organisms.

1979S
Searching for soft selective sweeps in worldwide human populations. Z.A. Szpiech1, R.H. Hernandez2,3,4, S. Zakov1, A.A. Rosenberg2, V. Balan2. 1) Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA; 2) Institute for Human Genetics, University of California, San Francisco, CA; 3) Institute for Quantitative Biosciences (QB3), University of California, San Francisco, CA; 4) Institute for Human Genetics, University of California, San Francisco, CA.

There is ample debate about the strength and mode of natural selection that has occurred in recent human evolution. This is particularly so for classical hard sweeps, during which an adaptive allele quickly drags a single haplotype to high frequency. An alternative model of adaptation involves soft sweeps, whereby multiple haplotypes are brought to high frequency (i.e. when a previously segregating neutral or slightly deleterious allele becomes adaptive in a new environment). Existing haplotype-based tests—such as the integrated haplotype score (iHS) that scans for positive selection by tracking the decay of haplotype homozygosity—work under the assumption that a positively selected region will be dominated by a single haplotype. However, iHS is expected to lose power under a soft sweep. Here we develop a statistic, inspired by iHS and recent work in Drosophila population genetics, designed to detect recent soft sweeps by tracking the decay of homozygosity of multiple haplotypes away from a core locus. We evaluate our statistic with rigorous simulations under multiple realistic models of human demography. We find that it has high power to detect both hard and soft sweeps and has improved power compared to iHS. In particular, for a fixed selection coefficient, our simulations suggest that we have greatest power to detect soft sweeps in African populations, which have been under-

Posters: Evolutionary and Population Genetics
1980M
Natural selection at the melanocortin-3 receptor gene loci. I. Yoshiuchi. Dept Diabetes Mellitus and Medicine, Yushiichi Medical Diabetes Institute, Kamakura, Kanagawa, Japan.

Obesity is significantly associated with type 2 diabetes mellitus, metabolic syndrome, hypertension, stroke, and cardiovascular diseases. The worldwide prevalence of obesity is increasing steadily. Obesity is highly heritable disease that causes serious health problems. During the traditional cycles of feast and famine, natural selection of obesity-related genes would be significant because these genes control body weight and fat levels. Human adaptation to environmental changes in food supply, lifestyle, and geography may have influenced the selection of genes associated with the metabolism of glucose, lipids, carbohydrates, and energy. The melanocortin-3 receptor (MC3R) gene is one of obesity-associated genes, and MC3R mutations have been shown to be associated with obesity. MC3R-deficient mice showed increased fat mass. Here, We aimed to uncover evidence of selection at the MC3R gene loci. We performed a three-step method to detect selection at the MC3R gene loci using the HapMap population data. We used Wright’s F statistics as a measure of population differentiation, the long-range haplotype test to test extended haplotypes, and the integrated haplotype score (iHS) test to detect selection at the MC3R gene loci. We observed natural selection at the MC3R gene loci by the integrated haplotype score test in the African population. This finding provides evidence of natural selection at the MC3R gene loci. Further discoveries are warranted on the adaptive evolution of obesity-associated genes.

1981S
Identification of putative high altitude adaptation determinants in Tibetan whole genome sequences. J. Downie, H. Hu, T. Simonson, D. Wilkson, G. Glusman, J. Prchal, T. Simonson, 2) Interdepartmental Program in Bioinformatics, University of California Los Angeles, Los Angeles, CA; 3) Department of Anthropology, Penn State University, University Park, PA; 4) Department of Obstetrics and Gynecology, University of Colorado, Aurora, CO; 5) Department of Exercise Science, Syracuse University, Syracuse, NY.

Despite its extreme high-altitude environment, the Tibetan Plateau has been inhabited by humans for thousands of years. Tibetans possess heritable adaptations that allow them to better survive in hypoxic conditions. A number of genes that could account for Tibetans’ ability to live at high altitude have been found, including EGFL1 and EPAS1, using SNP arrays and low-coverage exome-sequencing. However, there has been little success thus far in finding specific variants responsible for the adaptive traits seen in Tibetans. To address this, we performed Complete Genomics whole-genome sequencing of 17 Tibetans from the United Kingdom (12) and Utah (5). The Tibetan ancestry of these genomes was confirmed by comparing them with genomes of other Tibetan and Eurasian samples using EIGENSOFT. Variants in Tibetans frequently found in Tibetans but rare in Mesoamericans were identified. We then performed an integrated haplotype score (iHS) and cross-population extended haplotype homozygosity (XP-EHH) test, using 1000 Genomes CHB+CHS as a comparison, to identify putative candidate regions. We calculated the number of iHS values >12.01 and the max XP-EHH value for 200kb non-overlapping windows. We compiled a candidate gene list based on their function and the max XP-EHH value for 200kb non-overlap-

1982M
Positive selection in smallpox associated genes among Mesoamericans. O.A. Garcia1, K. Arslanian2, D. Whorff1, M. Shriver1, L.G. Moore1, T. Brutsaert2, A.W. Bingham1. 1) Department of Anthropology, University of Michigan, Ann Arbor, MI; 2) Department of Anthropology, Yale University, New Haven, CT; 3) Department of Anthropology, Penn State University, University Park, PA; 4) Department of Obstetrics and Gynecology, University of Colorado, Aurora, CO; 5) Department of Exercise Science, Syracuse University, Syracuse, NY.

During the colonization of Mesoamerica, one of the major causes of death was the introduction of novel infectious diseases. Among the most lethal infectious diseases was smallpox. Therefore, studying signatures of natural selection in genes related to smallpox infection and immune response not only provides a window to our evolutionary past but is also a particularly attractive strategy to identify host factors for modern infectious disease. To characterize host risk factors within Mesoamerican populations, we interrogated 906,600 SNPs assayed using the Affymetrix 6.0 genotyping array for signatures of natural selection in 231 immune response genes. Our populations included: Mesoamerican: 25 Maya and 14 Nahua, Mixtec, and Tripanec speakers from Mexico, Andean: 25 Aymara from Bolivia, and 24 Quechua from Peru. Additionally, we used available data from 60 Europeans of northern European ancestry and 90 East Asians from China and Japan. We applied three statistical tests to identify signatures of natural selection: locus specific branch length (LSBL), the natural log of the ratio of heterozygosities (inhH), and Tajima’s D. Furthermore, we analyzed partial and hard sweeps with two haplotype tests: integrated haplotype score (iHS) and cross-population extended haplotype homozygosity (XP-EHH). We determined statistical significance based on an empirical distribution. Among our strongest test results for positive selection were CD74, ZAP-70, and IKZF1 that were significant in all the statistical tests at the 5% and 1% level for Mesoamericans between East Asians and European Americans comparisons. Furthermore, they were statistically significant in comparison to the Andean populations. CD74 is major histocompatibility complex class II (MHC II) invariant chain. Several studies have shown CD74’s protein to function as a receptor for cytokine MIF, a critical immune response factor. ZAP-70 is an integral part of the T-cell signaling pathway thereby regulating adaptive immune response. Several human-mouse chimeras show a higher frequency of Neanderthal alleles in East Asians relative to Europeans. Several hypotheses to explain this difference have been proposed. One hypothesis posits that there was a single admixture event in the population ancestral to modern Europeans and East Asians and that many of the Neanderthal alleles were weakly deleterious in modern humans. Because East Asians have historically had smaller population sizes than Europeans, purifying selection may have been less effective at removing the Neanderthal alleles from East Asian populations, leading to the observed high proportion of Neanderthal ancestry in East Asians. Here we test this hypothesis using forward-in-time population genetic simulations. These simulations include plausible models of European and East Asian population history which have been estimated from data as well as models of the fitness effects of Neanderthal alleles in humans that include different dominance scenarios and a distribution of selection coefficients. Starting with the same initial amount of Neanderthal ancestry in both populations, we find that the differences in population size between European and East Asians combined with purifying selection cannot lead to the observed increase in the amount of Neanderthal ancestry in East Asian populations. Furthermore, when starting with the same initial amount of Neanderthal ancestry in both populations, realistic population size changes alone are insufficient to decrease or increase the Neanderthal ancestry in one population relative to the other. The observed data must be explained by some other process, such as additional waves of Neanderthal admixture into East Asian populations.

1983S
Selection and reduced population size cannot explain higher amounts of Neanderthal ancestry in East Asian than European human populations. J. Kim1, K. Lohmueller1,2,3, 1) Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA; 2) Interdepartmental Program in Bioinformatics, University of California Los Angeles, Los Angeles, CA.

Understanding the Neanderthal ancestry of modern humans may provide crucial insights into the evolution of different human populations. It is believed that Neanderthals admixed with European and Asian populations to a much greater degree than with African populations. Additionally, recent studies show a higher frequency of Neanderthal alleles in East Asians relative to Europeans. Several hypotheses to explain this difference have been proposed. One hypothesis posits that there was a single admixture event in the population ancestral to modern Europeans and East Asians and that many of the Neanderthal alleles were weakly deleterious in modern humans. Because East Asians have historically had smaller population sizes than Europeans, purifying selection may have been less effective at removing the Neanderthal alleles from East Asian populations, leading to the observed high proportion of Neanderthal ancestry in East Asians. Here we test this hypothesis using forward-in-time population genetic simulations. These simulations include plausible models of European and East Asian population history which have been estimated from data as well as models of the fitness effects of Neanderthal alleles in humans that include different dominance scenarios and a distribution of selection coefficients. Starting with the same amount of Neanderthal ancestry in both populations, we find that the differences in population size between European and East Asians combined with purifying selection cannot lead to the observed increase in the amount of Neanderthal ancestry in East Asian populations. Furthermore, when starting with the same initial amount of Neanderthal ancestry in both populations, realistic population size changes alone are insufficient to decrease or increase the Neanderthal ancestry in one population relative to the other. This suggests that further data must be collected, such as additional waves of Neanderthal admixture into East Asian populations.

Copyright © The American Society of Human Genetics. All rights reserved.
1984M
Identifying incomplete selective sweep using Sequential Markov Coalescent. K. Liao, W. Hsieri, C. Tang. 1) Department of Computer Science, National Tsing-Hua University, Hsinchu City, Hsinchu, Taiwan; 2) Institute of Statistics, National Tsing-Hua University, Hsinchu City, Hsinchu, Taiwan.

Positive selection leaves signatures in genomes of the population and identifying such signatures provides insight into the genetic basis of how humans are different from other species. Among these signatures, long-haplotype types are useful for identifying incomplete selective sweep, but it is usually difficult to compute the likelihood of statistics based on long-haplotype under the neutrality assumption. Hence, the statistical significance must be evaluated by empirical distribution derived from a large number of simulations. Sequentially Markov Coalescent (SMC) is a promising framework of haplotype-based analysis and it has the potential to be extended to various population genetics models. We implemented a long-haplotype based test for incomplete selection under SMC framework. The likelihood of a sample is computed with Product of Approximate Conditionals (PAC) and the statistical significance can be evaluated by the likelihood-ratio test. In addition, the strength of selection is estimated through maximizing PAC-likelihood function. The procedures will be evaluated with simulated data and compared with other methods based long-haplotype signatures.

1985S
Asian diversity project: a survey of population structure and local adaptations in Asian populations. X. Liu, D. Lu, W. Yang, J. Li, T.H. Ong, C. Liao, F. Dong, L. Hong, K.-C. Lin, Y.Y. Teo. 1) Saw Swee Hock School of Public Health, National University of Singapore, Singapore; 2) NUS Graduate School, National University of Singapore, Singapore; 3) Max Planck Independent Research Group on Population Genetics, Department of Human Genetics and Max Planck Institute for Evolutionary Biology Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences., Shanghai, China; 4) Oxford University Clinical Research Unit, Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam; 5) Division of Bioinformatics and Data Management for Research, Mahidol University, Bangkok, Thailand; 6) Genome Institute, National Center for Genetic Engineering and Biotechnology, Pathumtani, Thailand; 7) Institute of Medical Molecular Biotechnology (IMMB), Faculty of Medicine, Universiti Teknologi MARA (UiTM) Malaysia. Sg Buloh, Selangor, Malaysia; 8) Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan; 9) Department of Statistics and Applied Probability, National University of Singapore, Singapore.

As the largest continent on Earth, Asia hosts more than 60% of the human populations in the world. Great genetic diversity exists in the Asian populations. The HUGO Pan-Asian SNP consortium provided a valuable genetic resource of Asian populations and performed a thorough survey of genetic diversity and population history of Asian populations. However, the sparse coverage of SNPs made the analysis of natural adaption difficult to perform. In this study, we collected dense genotyping data from 46 populations across Asia. More than 4093 individuals from East Asia, Central Asia, Southeast Asia and South Asia were genotyped on various genotyping platforms. Principal components analysis (PCA) and admixture analysis were performed to elucidate the population structure in ADP populations. It was revealed that geographic played an important role in the population structure of Asian populations; and the ADP populations were further grouped into East Asian, Central Asian, Southeast Asian and South Asian subgroups. We performed a genome wide scan of positive selection signals in the ADP populations using iHS, XP, and nSLD. A total of 36 candidate selection regions were detected across the 46 ADP populations. A PCA analysis on the selection signals were performed to investigate the degree of sharing of the selection signals in the 46 populations. It was found that clustering of populations by selection signals resembles the clustering inferred from population structure analysis. East and Southeast Asian groups share the largest number of selection signals; and the South Asian group possesses distinct selection signals from the rest of the Asian populations. For selection signals shared by multiple populations, we studied the origin of the selection, i.e., either the selection originated from a single mutation in the common ancestor followed by subsequent gene flow, or it was the result of convergent evolution, where the selection emerged separately from multiple mutation events. The origins of positive selection signals were investigated by using the iHS, XP, and nSLD to calculate the population similarity index identified 36 selection regions under convergent evolution, and most of them involve aboriginal populations from Southeast Asia.

1986M
The pleiotropic effects of EDARV370A in an admixed Uyghur population. Q. Peng, J. Li, J. Tian, Y. Yang, Y. Guan, L. Zhang, Y. Jiao, P. Sabeti, L. Jin, S. Wang. 1) CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China; 2) MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai, China; 3) CMC Institute of Health Systems Biology, Taizhou, Jiangsu Province, China; 4) Department of Biochemistry, Preclinical Medicine College, Xinjiang Medical University, Urumqi, Xinjiang, China; 5) The Broad Institute of Harvard and MIT, Cambridge, USA; 6) Center for Systems Biology, Department of Organismic and Evolutionary Biology.

An adaptive variant of the human Ectodysplasin receptor, EDARV370A, showed one of the strongest signals of recent positive selection from genome-wide scans. In transgenic mice and in humans, it is found that EDARV370A affects ectodermal related phenotypes, including hair thickness and color, teeth development, and formation of nails and sweat glands. The pleiotropic effects of EDARV370A in an admixed Uyghur population were explored by using a population-based association study. Our study provides a more complete picture for the adaptive evolution of EDARV370A in human history.
1988M

Gene-wide survey of positive selection signals in African Americans since admixture. A. Wang, Y. Choo, X. Wang, T. Tayo, K. Broecker, C. Hanis, S. Kardia, S. Redline, R. Cooper, H. Tang, X. Zhu. 1) Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH; 2) Department of Genetics, Stanford University, Stanford, CA; 3) Departments of Preventive Medicine, Biomedical Informatics, and Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY; 4) Department of Public Health Science, Loyola University Medical Center, Maywood, IL; 5) Human and Molecular Genetics Center, Medical College of Wisconsin, Milwaukee, WI; 6) Department of Epidemiology, Human Genetics and Environmental Sciences, University of Texas Health Science Center at Houston, Houston, TX; 7) Department of Epidemiology, University of Michigan, Ann Arbor, MI; 8) Department of Medicine, Harvard Medical School, Boston, MA, USA.

In an admixed population such as African Americans, over or deficient ancestry in a local genomic region may suggest natural selection. We scanned three large African American cohorts of 20,153 individuals but failed to identify any genome-wide significant over or deficient signals. We showed that the failure to identify any significant selection signals can be attributed to the estimated variance of the test, which consists of two components: variance due to sampling error and variance due to genetic random drift. The proportion of variance due to genetic random drift increases when sample size increases. Thus, a test based on examining local ancestry excess is not efficient and its power will not increase when increasing sample size. We also showed that the high correlations of local ancestries between different cohorts are due to the historical recombination and genetic random drift. Assuming African-Americans have been admixed for 8 to 12 generations, we estimated the effective population size as between 32,000 to 48,000.

1989S

Differential purifying and positive selection across genes stratified by X chromosome inactivation status. A. Slavney, F. Geos, A. Clark, A. Keinan. 1) Genetics, Genomics & Development, Cornell University, Ithaca, NY; 2) Biological Statistics & Computational Biology, Cornell University, Ithaca, NY.

In eutherian mammals, dosage compensation between XX females and XY males occurs through ChrX inactivation (XCI). XCI randomly silences transcription from one of the two X chromosomes in each female cell, but at least 25% of human ChrX loci escape XCI to varying degrees, creating three broad gene groups: genes that (i) are consistently inactivated, (ii) escape XCI in some individuals, and (iii) consistently escape XCI. Evolutionary mechanisms underlying XCI status are poorly understood, but the predominant model posits that XCI evolved in response to ChrY degeneration. In this model, XCI escapes encode products that are dosage-sensitive and/or female-biased, and thus highly conserved. We therefore tested the hypothesis that mutations in group ii) and iii) genes are more likely to be deleterious than those in group i) genes.

To evaluate the effect of mutations in each XCI group, we first used European American single nucleotide variant (SNV) data from the NHLBI Exome Sequencing Project to compare synonymous (S) and non-synonymous (NS) SNV frequency distributions and estimated the individual burden of private mutations (IBPM) for all groups. These analyses showed a greater NS:S excess at low SNV frequencies and greater IBPM estimates in groups ii) and iii) than in i), suggesting stronger conservation in XCI escapers. We then incorporated divergences from chimpanzee to estimate statistics derived from the McDonald-Kreitman test, including the fractions of strongly (d) and weakly deleterious (b) sites, for each group. We found significantly higher d and lower b values in groups ii) and iii) than in group i), which support stronger conservation in escapers. However, we were surprised to find that group ii) showed evidence of experiencing stronger purifying selection than group iii) as per the IBPM, d and b values.

In conclusion, we report that polymorphism and divergence data support a model of XCI escape driven by ancient and ongoing evolutionary conservation, but that heterogeneous escapers show stronger purifying selection than consistent escapers. These results suggest a potentially important role for XCI escapers in phenotypes that affect fitness, but also demonstrate a need for further investigation into the true extent of XCI profile heterogeneity among tissues and individuals.

1990M


The sequencing of multiple Neandertal specimens has opened up the possibility of estimating Neandertal phenotypes based on the effect of “Neandertal alleles” present at low frequency in modern human populations. To capture a portion of the functional effects of Neandertal variants, we consider whole genome sequence and mRNA expression profiles obtained using an Illumina Sentrix Human Whole Genome BeadChip microarray from lymphocytes drawn from 921 participants in the San Antonio Family Study (SAFS). 56,294 single nucleotide variants (SNVs) were identified as monomorphic in the three Vindja Neandertal genomes, but polymorphic in the SAFS. Of these SNVs, 2,950 are located upstream of genes represented by high-quality, heritable probes on the microarray in a putatively cis-regulatory position. 81 SNV-probe pairs were found to be significantly associated (p < 1E-5) in a variance components-based association analysis performed in SOLAR. Additionally, the effect of genome-wide Neandertal ancestry on transcript expression levels was considered. The full set of SNVs were used to generate a covariance matrix derived from LDAK which calculates kinship coefficients while correcting for linkage disequilibrium. A variance component model was fit for each of the 2,869 probes and the variance explained by the Neandertal ancestry covariance matrix was determined. Neandertal ancestry had significant (p < 0.05) explanatory power for 61 probes. To contextualize these results, all significant probes were considered in a gene-set enrichment analysis of biological domains defined from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Enrichment was determined by an empirical p-value derived from comparing the test set of significantly associated mRNA transcriptions to 1000 randomly selected sets of equal numbers of transcripts represented on the array. These results indicate an enrichment of related to the growth, development, and innervation of the central nervous system and immune-related pathways. Thus, Neandertal ancestry has been shown to have a significant impact on the expression of genes involved in development and physiology across many tissues and individuals.

1991S

Khoisan hunter-gatherers have been the largest population throughout most of modern human demographic history. HL. Kim, A. Ratan, GH. Perry, A. Montenegro, W. Miller, SC. Schuster. 1) Singapore Centre on Environmental Life Sciences Engineering, Nanyang Technological University, Singapore; 2) Center for Comparative Genomics and Bioinformatics, Pennsylvania State University, PA, USA; 3) Department of Anthropology, Pennsylvania State University, PA, USA; 4) Department of Geography, Ohio State University, OH, USA; 5) Campus do Litoral Paulista, Unesp - Univ Estadual Paulista, Brazil.

We sequenced the complete genome sequences of five Khoisan hunter-gatherers from the Kalahari Desert and one Bantu-speaking agriculturalist individual also from southern Africa, with a high accuracy. Compared the 420K SNP genotyping dataset from 490 worldwide individuals, admixture analyses showed that three of our Khoisan genomes from the Ju’hoansi group (northern Khoisan) each contain 90-95% of non-Khoisan populations, allowing us to assess the early demographic history of the human species. Population genomic analyses for our complete genome sequences along with those from eight non-Khoisan humans were performed to infer their effective population sizes and demonstrated that the Ju’hoansi population have maintained their large effective population size and been the people most isolated from all the other human populations, since the earliest population split between the Khoisan and other populations ~100-150 thousand years ago (kya). In contrast, all other human populations, including the ancestral Bantu-speaking agriculturalists (currently the largest population within Africa in terms of census size), have experienced severe bottlenecks and lost more than half of their genetic diversity from ~120 to 30 kya. According to paleoclimatic records and models, west-central Africa became drier, while southern Africa experienced increased precipitation, ~80-100 kya. We hypothesize that these climate differences might be related to the divergent ancestral population history within African human populations.
1992M 
Insights from low-coverage whole Y chromosome sequencing of 1,244 individuals. Y. Xue for The 1000 Genomes Project 1000Y Group. Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

The 1000 Genomes Project Phase 3 has sequenced 1,244 males belonging to 26 populations from Africa, South and East Asia, Europe and the Americas. In addition to these low coverage (~4-6x) sequences, the project also includes over two hundred males sequenced to high coverage by Complete Genomics and several sets of SNP genotype data used for validation. The group generated a union set of 80,895 Y-SNP, 2,830 Y-MNP and 6,076 short indel calls by combining results from seven different callers. Approximately 6,000 Y-STRs and ~1,000 large structural variants were also called, using two callers for each class. Validation suggests that the Y-SNP, Y-STR and large deletion and duplication calls have very high quality, but that the Y-MNPs and indels do not. Using 59,666 stable high confidence Y-SNPs, we have constructed a phylogenetic tree, to which the more complex classes of variant can be added. The tree recapitulates and extends the established phylogeography. It confirms a very rapid Paleolithic expansion (in number) of Y lineages post-dating the movement out of Africa, and Neolithic or later expansions of independent Y lineages in Africa, Europe, East Asia and South Asia. We observe different patterns in different continental regions, suggesting that this male expansion was extremely rapid in Europe, rapid in Africa, and less rapid in South and East Asia. These data thus provide powerful new insights into male evolutionary history and promise further insights into Y-chromosomal mutation and selection processes.

1993S 
Exploring the Y-Chromosome Variation of Modern Panamanians. A. Achilli1, V. Battaglialia, G. Grunghi, U.A. Perez2, L. Lanconii, M. Tribaldos2, A. Colletti, I. Cardinalli, E. Rizzi, A. Raveane, M.R. Capodiferro, S.R. Woodward3, J.M. Pascale4, R. Cooke5, N. Myres3, A. Torroni5, J. Motta5, O. Semino5. 1) Dept. of Chemistry, Biology and Biotechnology, University of Perugia, Perugia, Italy; 2) Dept. of Biology and Biotechnology, University of Pavia, Pavia, Italy; 3) Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah, USA; 4) Gorgias Memorial Institute for Health Studies, Panama City, Panama; 5) Institute of Biomedical Technology, National Research Council (ITB-CNR), Segrate, Milano, Italy; 6) Ancestry, Provo, Utah, USA; 7) Smithsonian Tropical Research Institute, Panama City, Panama.

The Isthmus of Panama - the narrow neck of land connecting the northern and southern American landmasses - was a forced corridor for the Paleolithic Indian expansion that originated from Beringia ~15-17,000 years ago. Archaeological findings suggest that some descendants of the earliest migrants remained on the isthmus, while accounts from early European explorers witnessed the presence of two main indigenous groups (the Cueva and the Coolie) in pre-Columbian times - populations that have since disappeared due to disease, warfare, and enslavement following the Spanish conquest. Today's indigenous groups total about 5.3% of the Panamanian population, and are mainly represented by the Ngöbe, Buglé, Kuna, Emberá, and Wounaan tribes, which traditionally appear to have settled in Panama from surrounding regions after the autochthonous natives were decimated. However, there is no evidence that the ancestral indigenous gene pool was completely replaced. If this was the case, the populations of modern Panama should have retained at least a fraction of the native pre-Columbian gene pool, possibly at a variable extent, given the differential degree of geographic and genetic isolation of the different Panamanian communities during the past five centuries. A recent study of the mtDNA history of the modern Panamanian population (Perego et al., 2012), based on a sample of 1565 individuals with Native American maternal ancestry, concludes that (1) the first settlement of Panama occurred quite rapidly after the initial colonization of the American continent, (2) based on complete sequence analyses, the founder ages of the most common lineages point to an ancient expansion supporting the antiquity of the Pacific coastal route, 3) the mitochondrial gene pool exemplifies the link between pre-Columbian and modern Panamanian populations (in fact, 83% of modern Panamanians clusters into native pre-Columbian gene pools). It appears that the Spanish conquistadores and additional European demographic influences did not contribute significantly to today's genetic composition of Panama, at least with regard to the maternal side. In this study, we have now tested the same scenario from the paternal side by employing the analysis of the Y-chromosome variation in modern Panamanians.

1994M 
Population specific patterns of novel haplotype groups at the PAH locus. G. Povysli1, S. Wieser2, S. Hochreiter3, J. Zschocke2. 1) Institute of Bioinformatics, Johannes Kepler University Linz, Linz, Austria; 2) Division of Human Genetics, Medical University Innsbruck, Innsbruck, Austria.

The phenylalanine hydroxylase (PAH) gene is of particular interest for population genetic studies because the distribution patterns of well-defined phenylketonuria (PKU [MIM 261600]) mutations can be linked with distinct SNP haplotypes for the assessment of ancient migration. Through family segregation analysis and molecular haplotyping with long-range PCR in PKU patients (232 PAH mutant alleles) and controls (157 PAH normal alleles) from various European countries we identified five major haplotypes in the distal 15 kb region of the PAH gene. Haplotypes differ by 3-16 specific SNP each and have been quite stable over the last millennia. The 29 common European PKU mutations can be linked to specific haplotypes with little evidence of recombination in the PAH gene. The results were compared to available sequencing data of Africans, East Asians, Europeans, and Admixed Americans from the 1000 Genomes Project. Additional data from chimpanzee, orangutan, and macaque, as well as high coverage sequences of Neandertals and Denisova, were used in conjunction with these data to establish a possible evolutionary tree of haplotype emergence. There are five major distal PAH haplotypes that can be found in all continental populations, but at different relative frequencies. For Europeans and Asians they make up more than 95% of all PAH alleles. While Europeans have comparatively high frequencies for all of them, the most common haplotype in Asians amounts to almost 78% of PAH alleles. In contrast, Africans have many very rare haplotypes that can only be found in Africans, or Africans and Admixed Americans. The ancestral haplotype that matches the sequences of chimpanzee, orangutan, and macaque, has only been found in Africans and one Admixed American individual. The haplotype that matches the Neandertal and Denisova sequence can be found in Africans, Admixed Americans and one Asian individual. Interestingly, additional variations on the PAH gene are more similar to the respective haplotypes of the present day individuals with this haplotype. The combination of disease mutations and common gene variants with molecular haplotyping and available genetic data from different countries allows a unique insight in the genetic history of human populations.

1995S 

As the longest stretch of non-recombining DNA in the human genome, the Y chromosome provides unique insight into the demographic and evolutionary history of modern humans. However, the lack of single nucleotide polymorphisms makes it difficult to discover new mutations, and the small number of SNPs and STRs makes it difficult to study the demographic history of populations. In order to gain a greater understanding of human Y-chromosomal diversity.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
1996M

The Kalash isolate from Pakistan. Q. Ayub1, L. Pagani1,2, M. Mezzavilla1,3, C. Tyler-Smith1, 1) The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom; 2) Division of Biological Anthropology, University of Cambridge, Cambridge, United Kingdom; 3) Institute for Maternal and Child Health — IRCSS “BurloGarofolo” — Trieste, University of Trieste, Trieste, Italy.

The Kalash represent an enigmatic isolated population that has been living for centuries in the Hindu Kush mountain ranges of present-day Pakistan. Previous uni-parental (Y and mitochondrial) DNA markers provided no support for their claimed Greek descent following invasion of this region by Alexander III of Macedon, and analysis of autosomal loci provide evidence of a strong genetic bottleneck. To understand their origins and demography further, we genotyped 23 unrelated Kalash samples on the Illumina HumanOmni2.5 BeadChip and sequenced a male individual at high coverage on an Illumina Hi-Seq 2000. Comparisons with neighboring populations confirmed results based on genotyping 650,000 common single-nucleotide polymorphisms in the Kalash samples from the Centre Etude Polymorphism Humain (CEPH) Human Genome Diversity Project (HGDP) Cell Line Panel. However, we observed no evidence for admixture as suggested recently by Hellenthal et al. The mean time of divergence between Kalash and other populations currently residing in this region, that also speak Indo-European languages, was estimated to be 11.8 (10.6 -12.6) KYA. Since the split the Kalash have experienced little, or no, gene flow from their geographic neighbors and have maintained a low long-term effective population size (2,247-2,780). They could represent some of the earliest migrants into the Indian sub-continent.

1997S

Identifiability and efficient inference of population size histories and locus-specific mutation rates from large-sample genomic variation data. A. Bhaskar1,2, Y.X. Wang2, Y.S. Song1,2,3
1) Simons Institute for the Theory of Computing, University of California, Berkeley, Berkeley, CA; 2) Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA; 3) Department of Statistics, University of California, Berkeley, Berkeley, CA; 4) Department of Integrative Biology, University of California, Berkeley, Berkeley, CA.

Several recent large-sample human genetics studies have found a massive excess of rare variants compared to predictions of previously inferred demographic models of human history. A widely cited explanation is that such polymorphism patterns are indicative of explosive and accelerating population growth in recent human history. Using the site frequency spectrum (SFS), a summary of genetic variation in a set of sequences that counts polymorphisms as a function of the mutant allele frequency, we develop an efficient method for inferring recent population demography that can scale to samples involving tens or hundreds of thousands of individuals. Using analytic results for the expected SFS under the coalescent and by leveraging the technique of automatic differentiation, we develop a very efficient algorithm to infer piecewise-exponential models of the effective historical population size from the distribution of sample allele frequencies. Our method is orders of magnitude faster than previous demographic inference methods based on the frequency spectrum and can also accurately estimate locus-specific mutation rates. We show that our method can accurately infer multiple recent epochs of rapid exponential growth, a signal that we detect in the 1000 Genomes Project data. Across populations, we find that the allele-frequency spectrum is well-approximated by piecewise-exponential models, and piecewise-generalized-exponential models, which are often assumed in population genomic inferences, provide explicit values for the sample sizes that are sufficient for identifying the demographic model from the SFS.

1998M

Does genetic hypermutability contribute to the current prevalence of connexin 26 deafness? D.C. Braun1, E.A. Cr PDF1, B.K. Herold1, K.S. Amor1, M. Tekin2,3, A. Pandya1, 1) Department of Science, Technology, and Mathematics, Gallaudet University, Washington, DC; 2) John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 3) Department of Pediatrics, Ankara University School of Medicine, Ankara, Turkey. 4) Department of Human Genetics, Virginia Commonwealth University, Richmond, VA.

Approximately 35% of congenital deafness in North America results from recessive mutations in the GJB2 (connexin 26) gene, making it the most prevalent cause for hereditary deafness [MIM 230300] in this population. There are more than 100 known GJB2 mutations, placing it in the top 1% of mutational diversities reported by the Human Gene Mutation Database. Another remarkable characteristic of GJB2 is that some mutations have ethnic associations: 35delG in Caucasians, 235delC in Asian and 167delT in Ashkenazi Jews. These associations would normally suggest single origins in founders. Some researchers have proposed that GJB2 is hypermutable. For example, the 35delG mutation occurs within a homopolymorphic run of six guanines and could have arisen repeatedly due to strand slippage. GJB2 also possesses sequence motifs associated with hypermutability in humans or other organisms. The hypermutability hypothesis is important because it could explain the prevalence of GJB2 deafness. To test the hypothesis of generalized hypermutability in GJB2, we analyzed the historical mutation rate of GJB2 by directly sequencing an 8 kilobase region of genomic DNA, containing GJB2, from participant samples from two different repositories: Virginia Commonwealth University, which holds the largest repository of genetic deafness in North America, and Ankara University in Turkey. First, we were directly sequencing that region and found that the Turkish samples carried a single haplogroup A or B. We conclude that the 8 kb locus containing GJB2 is not generally hypermutable. However, our discovery of two 8 kb haplogroups for 35delG confirms that it has multiple origins and therefore, strand-slippage mutation, were not within the area of the Turkish samples. However, the prevalence of 35delG mutations. Our research is significant because it directly addresses the causes for the prevalence of GJB2 deafness, which is not yet understood.

1999S

Identity by descent segments within and across worldwide populations from sequence data. S.R. Browning1, B.L. Browning1,2, 1) Department of Biostatistics, University of Washington, Seattle, WA; 2) Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA.

Segments of identity by descent (IBD) shared by individuals within and across populations provide information on key aspects of demographic history, such as effective population sizes and migration rates. Sequence data present opportunities and challenges for IBD analysis. Sequence data are more informative than SNP array data, improving power to accurately detect smaller IBD segments and hence obtain higher levels of information about demographic history. On the other hand, low-coverage sequence data have high rates of error, whereas SNP array data are usually extremely accurate.

We recently developed two IBD segment detection methods: Refined IBD and IBDseq. Refined IBD is a haplotype-frequency-based method designed for low-coverage sequence data, while IBDseq is an allele-frequency-based method designed for high-coverage sequence data. Both methods were developed in the context of samples from a homogeneous population. When using frequency-based methods in a heterogeneous setting we expect increased rates of false-positive IBD with subpopulations.

We use 1000 Genomes Project data and simulated data to investigate the performance of the IBDseq and Refined IBD methods when analyzing sequence data from world-wide populations. We find that the allele-frequency-based IBDseq method suffers from increased rates of false positive detections, IBD segments due to population heterogeneity, whereas the haplotype-frequency-based Refined IBD approach is much less affected. We develop a strategy using multiple runs of Refined IBD and a process of filling small gaps between adjacent detected segments in order to recover near-complete IBD segments which have high power to detect short segments. Our approach enables powerful IBD detection in the 1000 Genomes project data.
2000M
The Population Genomic Landscape of Human Genetic Structure, Admixture History and Local Adaptation in Peninsular Malaysia. L. Deng, B. Hoff, D. Lu, R. Fu, M. Phipps, S. Lu, A. Nur-Shatwati, W. Hatin, E. Ismail, S. Mohktar, L. Jin, B. Zilfalil, C. Marshall, S. Scherer, F. Al-Mulla, X. Xu. 1) Max Planck Independent Research Group on Population Genomics, Chinese Academy of Sciences and Max Planck Society (CAS-MPG) Partner Institute for Computational Biology (PICB), Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai; 2) Institute of Medical Molecular Biotechnology, Faculty of Medicine, Umeå University, Umeå; 3) Sungai Buloh Campus, Jalan Hospital, Universiti Kebangsaan Malaysia, Selangor 46150, Malaysia; 4) Ministry of Education (MOE) Key Laboratory of Contemporary Anthropology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China; 5) Department of Pediatrics, School of Medical Sciences, Universiti Sains Malaysia, Kelantan 16150, Malaysia; 6) Human Genome Center, School of Medical Sciences, Universiti Sains Malaysia, Kelantan 16150, Malaysia; 7) School of Biosciences & Biotechnology, Faculty of Science & Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia; 8) The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada; 9) McLaughlin Centre and Department of Molecular Genetics, University of Toronto, Toronto, Canada; 10) Department of Pathology, Faculty of Medicine, Kuwait University, Safat 13110, Kuwait.

Peninsular Malaysia is a strategic region which might have played an important role in the initial peopling and subsequent human migrations in Asia. However, the genetic diversity and history of human populations—especially indigenous populations—inhabiting this area remain poorly understood. Here, we conducted a genome-wide study using about 900,000 single nucleotide polymorphisms (SNPs) in four major Malaysian ethnic groups (MEGs; Malay, Proto-Malay, Senoi and Negrito), and made comparisons with 17 worldwide populations. Our data revealed that Peninsular Malaysia has greater genetic diversity corresponding to its role as a contact zone of both early and recent human migrations in Asia. However, each single Orang Asli (indigenous) group was less diverse with a smaller effective population size than the four MEGs. The geographical isolation of some duration for these four groups. All four MEGs were genetically more similar to Asian populations than to other continental groups, and the divergence time between MEGs and East Asian populations (120,000–60,000 years ago) is much shorter than that between East Asians and Europeans. Thus, Malaysian Orang Asli groups, despite their significantly different features, may share a common origin with the other Asian groups. Nevertheless, we identified traces of recent gene flow from non-Asians to MEGs. Finally, natural selection signatures were detected in a batch of genes associated with immune response, human height, skin pigmentation, hair and facial morphology and blood pressure in MEGs. Notable examples include SYN3 which is associated with human height in all Orang Asli groups, a height-related gene (PNNPT1) and two blood pressure-related genes (CDH13 and PAX5) in Negritos. We conclude that a long isolation period, subsequent gene flow and local adaptations have jointly shaped the genetic architectures of MEGs, and this study provides insights into the peopling and human migration history in Southeast Asia.

2001S
Fine-Scale Genetic Structure in the Open Population of Western France. C. Dina, F. Simonet, P. Olivier, F. Gros, S. Lecointe, S. Kury, S. Bézieux, H. Le Marec, J. Schott, M. Karakachoff, R. Redon. 1) institut du Thorax, Nantes, France; 2) CNRS, UMR 6291, Nantes, France; 3) Université de Nantes, Nantes, France.; 4) CHU Nantes, l'institut du thorax, Service de Cardiologie, Nantes, France.; 5) CHU Nantes, Service de Genetique Medicale, 44093 Nantes CEDEX 1, France; 6) Institute of Clinical Physiology, National Research Council, Pisa, Italy.

The ‘Common Variant - Common Disease’ hypothesis was only partly verified through effective discovery of statistical associations. This empirical observation led the research community to reconsider the involvement of rare genetic variation in predisposition to common disease. Rare alleles of recent origin are likely to cluster geographically in communities with limited reproduction rates, such as the rural church populations before the 20th century. These communities are likely to be particularly susceptible to diseases such as cancer. The use of very high density SNP datasets over the whole genome can reveal the fine-scale genetic structure of a population. We performed a systematic examination of the whole genome of a large number of SNPs to determine if fine-scale structure is present. Our approach is based on the Principal Component Analysis (PCA). This method provides insights into demographic history and population structure. We observed a fine-scale genetic structure in the open population of Western France which is highlighted by PC3. It seems that there is a close relationship between some SNP clusters and geographical differences. These differences can be related to the migration history of France. Interestingly, we demonstrated that the genetic structure is present at the local level with PC4. This result suggests that rare variants inducing disease susceptibility can benefit from a strategy focusing on small geographic units. Moreover, we will apply spatially explicit methods like spatial-PCA in order to identify geographic subpopulations of interest.

2002M
Shared identity by Descent segments within current Italian population reveals new details about recent population history. G. Fioritto, D. Di Gaetano, F. Rosa, S. Quattrero, B. Pardini, A. Piazza, G. Malutello. 1) Human Genetics Foundation, Turin, TORINO, Italy; 2) Department of Medical Sciences, University of Turin, Turin, Italy.

The inference of identity by Descent (IBD) shared segments were recently enabled by high-resolution genomic data from large cohorts and novel algorithms for IBD detection. This approach permits to examine in detail the genetic structure of a population as well as to get information about recent demographic events such as bottlenecks and migrations. This study aims to characterize the genetic variability within the Italian population. We present analytical results on the relationship between IBD sharing across 301 unrelated Italian individuals genotyped for about 2.5 million Single Nucleotide Polymorphisms (SNPs). Each sample has well-defined geographical origins (four grandparents coming from the same geographical region). Due to the well-known common ancestral origin of the Italian population we focused our attention on long-range and relatively recent shared IBD segments. By using Principal Component Analysis (PCA) and ancestry estimation, we ascertain Sardinia as the genetic outlier within Italy. Moreover a certain degree of differentiation is still detectable within Astoa Valley population. For each of the 11 subpopulation, we find a significant highest number of shared IBD segments within vs. between population, suggesting isolation by distance. Samples sharing the highest number of internal IBD blocks are Sardinians as expected, followed by those living in Astoa Valley, Tuscany and Sicily. We also evaluate the relationship between shared IBD segments and geographical distance. Contrary to what is expected, the decay of IBD with distance is not steeper for longer (recent) blocks. Such result suggests a constant exchange due to several migratory waves within Italy and/or to the considerable high number of populations that have lived in Italy. Finally we demonstrate that regions of increased IBD sharing are enriched for structural variation and loci implicated in natural selection and we highlighted the relationship between shared IBD haplotypes and demographic events occurred both in Sardinia and in the Italian peninsula. In conclusion, our results suggest that the study of shared IBD segments between populations is a useful method to detect novel details about relatively recent population history.
2003S
Reconstruction of ancestral human haplotypes using genetic and genealogical data. J.M. Granka1, R.E. Curtis2, K. Noto1, Y. Wang1, J.K. Byrne1, M.J. Barber1, N.M. Myres3, C.A. Ball1, K.G. Chakhchouk1. AncestryDNA, San Francisco, CA; 2) AncestryDNA, Provo, UT.

The genomes of individuals who lived long ago may persist in modern populations in the form of genomic segments broken down by recombination and inherited by their descendants. We developed a novel computational method to reconstruct the chromosomal haplotypes of human ancestors given genetic data from a sufficient number of their present-day descendants. After genealogical information is used to identify descendants of an ancestor, haplotypes of their progeny are simulated using a single nucleotide polymorphism (SNP) data to find regions of the genome that are identical-by-descent (IBD) among them. We develop a novel stitching algorithm to reconstruct up to four chromosomal haplotypes of an ancestor and their partner given the descendant IBD segments and haplotypes. The method aims to remove spurious IBD segments caused by false inference, inaccurate genealogical information, or multiple common ancestors. Short regions of the genome with missing data can also be imputed given flanking reconstructed haplotypes. Lastly, given descendants of other individuals related to the ancestral couple, we show that it is sometimes possible to tease apart the personal identity of each of the reconstructed chromosomal haplotypes (i.e., which are the ancestor’s, and which are their partner’s). Through simulation, we calculated accuracy of the genotypes that can theoretically be reconstructed given the number of generations back to the ancestor and the number of actual and sampled descendants. Given sufficient data, we can reliably reconstruct the haplotypes of in silico ancestors with high precision and recall; performance is sensitive to genealogical tree quality and number of IBD segments inferred. We also used inferred IBD segments to capture the diversity Linz, Linz, Austria.

Identity by descent between humans, Denisovans, and Neandertals. S. Hochreiter, G. Pysoliv. Institute of Bioinformatics, Johannes Kepler University, Linz, Linz, Austria.

We analyze the sharing of very short identity by descent (IBD) segments between humans, Neandertals, and Denisovans to gain new insights into their demographic history. Short IBD segments convey information about events far back in time because the shorter IBD segments are, the older they are assumed to be. The identification of short IBD segments becomes possible through next generation sequencing (NGS), which offers high variant density and reports variants of all frequencies. Only recently HapFABIA has been proposed as the first method for detecting very short IBD segments in NGS data. HapFABIA utilizes rare variants to identify IBD segments with a low false discovery rate. We applied HapFABIA to the 1000 Genomes Project whole genome sequencing data to identify IBD segments which are shared within and between populations. Some IBD segments are shared with Neandertals or Denisovans, which extends our understanding of Neandertal and Denisovan ancestry and suggests the feasibility of reconstructing the genomes of human ancestors and has immediate applications in population genetics, medical genetics, and genealogy research.

2005S

Reconstructing past demography from neutral genetic data is essential as it both improves our knowledge about human history and provides an accurate model against which selective hypotheses can be tested. Approximate Bayesian Computation (ABC) has proven to be useful for inferring demography from microsatellite or SNP data. This approach consists in simulating genetic data under a large range of complex demographic scenarios and realistic biological processes. Simulations are then compared to observed data using a set of informative summary statistics. Whole-genome data are expected to be extremely rich in information about past demography but, because simulations were, until recently, computationally too costly, ABC methods have not been thoroughly tested on such very long sequences. Dense polymorphism data contain extra information that is not available from unlinked site polymorphisms, and will, therefore, hopefully improve the reconstruction of demographic history. They allow computing specific statistics, such as the decay of linkage disequilibrium with distance, the distribution of length of haplotypes shared between two or more individuals, or the “allele frequency identity-by-state” as described by Theunert et al. (2012). The power given by some of these statistics to infer demographic parameters has been investigated. However, studies were done independently on single classes of statistics, and not always under the approximate Bayesian framework.

Here, we examine how combining these “dense data statistics”, with “classical statistics” (e.g. pairwise differences, heterozygosity) in an ABC framework improves the inference of demographic history. Furthermore, we describe how sequence errors that are usually more frequent in full sequence SNPs data impact the summary statistics and the ABC inference. To diminish these effects we propose to either (i) filter data DNA sequence and prune summary statistics that are highly sensitive to errors, or (ii) model errors within our ABC and incorporate them into simulations. We benchmark these different approaches for simple demographic scenarios, and focus more specifically on population expansion events that happened in recent human history.


2006M
Genetic Structure of North-Indian Punjabi Population Based on Autosomal Microsatellite Loci. M. Kaur, B. Badaruddoza. Dept. of Human Genetics, Guru Nanak Dev University, Amritsar, India.

The population of Punjab, India, possesses an exclusive genetic profile, primarily due to the many major historical events in this region which caused an extensive range of genetic diversity. Hence, the present study is an attempt to find out the genetic similarity and phylogenetic position of north-west Punjab population with respect to past history of admixture of foreign populations, especially, Central Asianoid Populations. In this study, six microsatellite markers: TH01, TPOX, CSF1PO, VWA, D7S820 and FGA have been analyzed among 516 samples from five endogamous population groups, Jat Sikhs, Mazhabi Sikhs, Brahmans, Ramdassias and Muslims of north-west border districts of Punjab. The number of alleles ranged from 8 to 12 at six STR loci. The exact test probabilities for HWE suggested some significant departures in certain loci and population groups. In general, the average observed heterozygosity was lower than expected heterozygosity in six STR markers among the five population groups. The average sub-ethnic genomic differentiation (Fst) among five population groups of northwest Punjab was 0.0335. The CSF1PO showed highest sub-ethnic differentiation (0.0649), whereas, the lowest Fst has been observed for FGA locus (0.013). The polymorphic gene flow might be attributable to different populations relative to the total genomic diversity (Gst) varied between 6.1% for CSF1PO locus and 0.9% for D7S820. When all the loci were jointly considered, 3.0% of the total genomic diversity was attributable to the five population groups. The average pairwise gene flow observed in FGA (Nm=18.77%), which was followed by TH01 (15.92%). The loci showed significant positive departures in CSF1PO (3.6%). However, in general with all loci the gene flow was observed to be 7.2% among these studied population groups. To understand the extent of sub-structuring among five north-west Punjab population. Structure analysis was also performed with different values of K. The log probability of data values and the membership proportion of each group showed clear sub-structure among the population groups. Overall, five Punjabi speaking population groups of northwest Punjab are regionally well differentiated and are characterized by their unique ethnogenetic affinities based on their origin, settlement and their shared ethnico-historical background.

The demographic inference of human population is important not only for archeological and evolutionary studies but also for guiding sampling design for medical genetic study. Recent increase in the amount of human genetic data combined with the population genetic method contributed to the understandings of the demographic history of modern human. Both theoretical and experimental studies have shown that some populations have experienced the rapid population growth resulting in more low-frequency alleles than expected under the neutral and constant population size model. Thus, large samples, which include substantial amount of low-frequency alleles, are necessary for the demographic inference of rapidly growing population. The coalescence-based inference methods are commonly used for demographic inference with such large samples. With the aid of progress in computational power, a forward-time simulation under Wright-Fisher diffusion model is also applicable for observed allele frequency spectrum. We developed a method for demographic inference based on a diffusion approach, in which variable population size is estimated by nonparametric regression. By applying this method to the complete sequences of mitochondrion DNA from Japanese and Basque populations, we observed the recent population growth in both populations but with different growth rates and different time periods of growth.


With the development of Next Generation Sequencing Technology (NGS), the field of hominin paleogenetics has transformed significantly from studying specific DNA markers to revealing whole genome information. However, ancient DNA of interest is usually highly fragmented so an NGS library protocol optimized to capture short DNA fragments (40bp to 200bp) was developed. The improved workflow includes the use of column-based DNA purification and concentration and automated gel-based size-selection. This workflow permitted production of “shotgun” genomic libraries from very limited input DNA (6ng to 30ng). Methods that permit the use of such low input, degraded DNA enable the partitioning of exceedingly rare samples into multiple analytical workflows. For example, to establish highest confidence SNP calls from ancient genomes it is best to sequence the sample on multiple orthogonal platforms. To pilot this approach, DNA extracted from 4 human tooth specimens from Bulgaria (ancient Thracians) that date to the Bronze and Iron Ages (1500-400 yr BCE) were sequenced on a semiconductor-based platform. Per individual sample, 259 million to 312 million sequence reads were produced. Deep sequencing (467 Mb) on one of the samples (P192-1) yielded detection of ~400,000 SNPs. Using principal component analysis that included more than 1,300 modern Europeans, this large number of SNPs indicated clustering of P192-1 (from an ancient farming community) closely with modern Sardinians; this result resolves ambiguous inferences from previous studies. Similar frequency in the ancestral lineage of this individual that was recently reported based on ~10-fold fewer SNP calls and supports the hypothesis that the ancient Thracians and Sardinians share a heritage that dates to the initial spread in Europe on ~10-fold fewer SNP calls and supports the hypothesis that the ancient Thracians and Sardinians share a heritage that dates to the initial spread in Europe (~100,000 yrs). The results supported many mainstream viewpoints on the demographic histories of human populations, and at the same time also produced several interesting observations worth further and more careful investigations.

Exome sequencing of 3,000 individuals reveals differences in recent demographic history between East Asian and European populations. K.E. Lohmueller, M. He, Y. Li, B. Kim, L. Sun, X. Zhang, X. Jin, K. Kristiansen, T. Hansen, J. Wang, O. Pedersen, E. Sanchez, R. Nielsen. 1) Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA; 2) Bioinformatics and Computational Biology Program, Iowa State University, Ames, IA; 3) BGI-Shenzhen, Shenzhen, China; 4) Department of Dermatology, First Affiliated Hospital, Anhui Medical University, Hefei, China; 5) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 6) Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; 7) The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; 8) Department of Health Sciences, Aarhus University, Aarhus, Denmark; 9) Institute of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark; 10) Integrative Biology, University of California, Berkeley, Berkeley, CA.

Studies in the last 200 years have shown that the Han Chinese (from the Han population) have been growing at an average rate of 1% per year. However, if we compare the Han Chinese individuals sampled from the Anhui province of China to the Danish individuals, we observed that the Han Chinese population is expanding at a rate of 1.5% per year, whereas the Danish population is declining at a rate of 0.5% per year. This difference can be explained by greater variance in reproductive success in the Han Chinese population than in the Danish population. Alternatively, this result could be due to greater migration into the Han Chinese population than the Danish population. While it is appreciated that effective population size of human populations are smaller than the census sizes, here we demonstrate that the magnitude of this difference varies across populations, even after accounting for population size changes. Finally, we examine patterns of deleterious variants in the Han Chinese and Danish populations. We find that the proportion of private variants that are nonsynonymous is higher in the Han Chinese sample (67.6%) than in the Danish sample (64.6%; P<10^-6), consistent with recent population growth increasing the input of weakly deleterious mutations into the population that have sufficient time to remove. Our study provides the first analysis of recent population history and exploration of neutral and deleterious rare variants in an East Asian population.
2011S
Analysis of Genetic Diversity Representation of the 1000 Genomes in Worldwide Human Populations. D. Lu, S. Xu. Partner Institute for Computational Biology, Shanghai, Shanghai, China.

The 1000 Genomes Project (1KG) aims to provide a deep characterization of human genetic diversity, by design was expected to provide a comprehensive resource on human genetic variation. With an effort of sequencing 2,500 individuals, 1KG is expected to cover the majority of the human genetic diversities worldwide. However, it would be interesting to evaluate to what extent the 1KG data represent the genetic diversity of human populations in each region, which will give insight into the power of 1KG and also give guidance to regional efforts for further sequencing project and study design. In this study, using analysis of population structure based on genome-wide single nucleotide polymorphisms (SNPs) data, we examined and evaluated the coverage of genetic diversity of 1KG samples with the available genome-wide data from 3,831 individuals representing 140 worldwide population samples. We demonstrated that the 1KG does not have sufficient coverage of human genetic diversity in Asia, especially in Southeast Asia. We thus suggest a better coverage of Southeast Asian populations be considered in 1KG or a regional effort to be initialized to provide a more comprehensive characterization of the human genetic diversity in Asia, which is important for both evolutionary and medical studies in the future.

2012M

One of the core features of any genetic variant, beyond its potential phenotypic effects or its frequency, is its geographic distribution. The geographic distribution of a genetic variant can shed light on where the variant first arose, in what populations it survived and spread within, and in turn help us learn about historical patterns of migration and natural selection. Collectively, the geographic distribution of genetic variants can help to explain how populations have been related through time (e.g. levels of gene flow and divergence). For variants with large effects, it can also help us understand the geographic distribution of spatially-varying phenotypes. For these reasons, visualization of geographic maps for genetic variants is common practice in genetic studies. Here we develop a series of reusable interactive visualizations for illuminating the geographic distribution of genetic variants. We specifically address several non-trivial challenges of this type of visualization; in particular, how to represent non-uniform levels of uncertainty in allele frequencies due to variable sample sizes; how to represent results from data with >10,000 individuals in which allele frequencies can vary over 4 orders of magnitude; how to display data for regions of the globe with dense sampling of populations; and how to quickly access frequency data from large samples. To meet these challenges, we implement a flexible REST API for allowing for easy access to allele frequency and sample size data from large scale public genomic datasets. Built upon this API we develop a web-based browser, entitled the Geography of Genetic Variants (GGV) browser for visualizing the geographic distribution of genetic variants. The GGV browser rapidly provides maps of derived allele frequencies in populations distributed across the globe. The GGV browser builds upon past tools such as the HGDP Selection browser by allowing for more interactive features, new representations of rare variation, as well as incorporating uncertainty in allele frequency estimation. As ancillaries, we also develop a research visualization toolkit that includes a method for displaying high Fst outlier SNPs from the joint site frequency spectrum and an interactive version of commonly used PCA figures. We hope the GGV browser will be a valuable research and education tool for exploring population genetics data.

2013S
A Novel Likelihood Ratio Test Framework for Sex-Biased Demography and its Application to Human and Dog Genomic Data. S. Musharoff1, S. Shringarpure2, T. Cooke1, A. Adams1, C. Bustamante1. 1) Genetics, Stanford Univ Sch Medicine, Stanford, CA; 2) School of Biological Science, Washington State University, Pullman WA; 3) Ecology and Evolutionary Biology, Brown University, Providence, RI.

Sex-bias is linked to an unequal number of breeding males and females in a population. This can be caused by variance in reproductive success, demographic events involving unequal numbers of males and females, and/or different selection at sex-linked genomic loci. A commonly used estimator of the proportion of females is based on the test statistic Q where Q is the ratio of current neutral genetic diversity estimated from the X chromosomes to that estimated from the autosomes. This is problematic if the population changed in size because X chromosomal diversity recovers from size changes at a different rate than autosomal diversity due to unequal effective population sizes, this estimator of the proportion of females based on current diversity will be biased. To this end we present a novel likelihood ratio test framework for sex-bias in a single population based on the Poisson random field model. We use the program dadi to estimate demographic parameters jointly from autosomal and X chromosomal data and test first for a persistent sex-bias and then for a sex-biased demographic event. When applied to simulated data, our test has more power to detect sex-bias from unlinked or partially linked sites than the commonly used test statistic Q for a range of demographic scenarios. Encouragingly, our test is well powered for events relevant to human history including recent rapid expansion whereas the test statistic Q is not. We recover the true proportion of females with our framework whereas the estimator based on Q underestimate the magnitude of sex-bias. We test for sex-bias using GBS data and genetically linked markers to genome-wide data from three Thousand Genomes populations and find evidence for female sex-bias in Europeans, Asians, and Africans after modeling recent demographic events including a bottleneck and recent growth. Our analysis of breed dogs to breed dogs we find evidence of female bias in the breed dogs showing a more extreme bias than village dogs. These findings argue for the importance of modeling demography when assessing sex-bias in populations and highlight the role of sex-bias in the history of humans and of dogs.

2014M
Inferring the effective founder size of a spatially expanding population. B.M. Peter, M. Slatkin. University of California, Berkeley, Berkeley, CA.

The gradual loss of diversity associated with range expansions is a well known pattern that can be explained with a serial founder model. Using a branching process model, we show that this loss in diversity can be mainly attributed to sex-bias in the history of humans and of dogs. We show that many classical properties of the effective population size, such as the two-sex formula and the harmonic mean for time-varying models extend naturally to expanding populations. We demonstrate that the predictions from the branching process model fit very well with Wright-Fisher forward simulations and backwards simulations under a modified Kingman coalescent, and further show that estimates of the effective founder size are robust to possibly confounding factors such as migration between subpopulations.
2015S
Finding the oasis of humanity in Neanderthal deserts. B. Vernet, JM. Akey. Department of Genome Sciences, University of Washington, Seattle, WA.

As anatomically modern humans dispersed out of Africa, they encountered Neanderthals in Eurasia and low levels of hybridization occurred such that approximately 2% of each non-African’s genome is inherited from Neanderthal ancestors. Recently, we developed an approach to identify surviving Neanderthal lineages in contemporary individuals, and recovered over 600 Mb of the Neanderthal genome present in modern non-African populations [1]. The map of surviving Neanderthal sequences shows marked heterogeneity across the genome, and we identified many “deserts of Neanderthal sequence” that are almost entirely devoid of Neanderthal sequence. These genomic regions are of particular interest because they delimit sequences that may confer uniquely human characteristics. For example, the largest Neanderthal desert is a 15Mb region on Chromosome 7, centered around the FOXP2 gene, which has previously been implicated in speech and language. Here, we present a detailed characterization of Neanderthal deserts by analyzing surviving archaic sequences in an expanded sample of geographically diverse individuals. We have developed a formal statistical test to identify genomic regions significantly depleted of Neanderthal lineages, and performed extensive simulations to infer the strength of purifying selection acting on these Neanderthal deserts. Additionally, we have utilized extensive bioinformatics analyses superimposing heterogenous functional genomics data to identify candidate causal variants. These analyses provide significant new insights into regions of the human genome that harbor sequences that have played a critical role in the evolution of anatomically modern humans, and suggest that regulatory sequences responsible for muscle, bone, and brain development were key differences between humans and Neanderthals. [1] Vernet and Akey, Science, 2014.

2016M
Coalescent times of 63 males estimated using the complete Y-chromosomes. E. WONG1, S. Limkonska2, A. Valouev1. 1) Department of Preventive Medicine, Keck School of Medicine of University of Southern California, Los Angeles, CA; 2) Department of Human Molecular Genetics, Russian Academy of Sciences, Moscow, Russia.

The Y-chromosome contains the longest non-recombining region in the genome and is useful for inferring human population history. Previous studies on the Y-chromosome mostly rely on rapidly mutating microsatellites or selected single nucleotide polymorphisms ascertained from small panels of individuals. These approaches likely underestimate the genetic diversity of the Y-chromosomes. Complete sequencing of the Y-chromosome to high coverage would enable us to more accurately estimate the coalescent times of populations across the world and reconstruct past events in human evolution. We sequenced the complete Y-chromosomes of 12 males sampled from Asia and Europe. We analyzed these chromosomes together with publicly available high-coverage complete Y-chromosomes of 51 males from across the world with known ethnicities. In our analysis, we excluded regions of the Y-chromosome that are subject to recombination and regions that are ill-suited for short-read sequencing due to low complexity and homology to the X-chromosome. By focusing our analysis on 10.45 million sites, we identified over 15,000 single nucleotide variants (SNVs) in 63 males from 38 populations across the world and constructed a maximum likelihood tree. We performed haplotype analysis for these Y-chromosomes, based on the SNVs present in the International Society of Genetic Genealogy (ISOGG) database. We estimated the Y-chromosome coalescent time to be around 129 thousand years, which is in line with previous estimates. By estimating the coalescent time of these Y-chromosomes, we are able to understand more about the history of these populations and the migrations of their ancestors.

2017S
Detecting and dissecting the fine-scale genetic population structure of Spain. C. Bycroft1, A. Carracedo2, C. Fernandez-Rozadilla1,2, C. Ruiz-Ponte1, I. Quintela-Garcia2, S. Myers1, P. Donnelly1,2,3,7,7T1)Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxford, UK; 2) Galician Public Fundation of Genomic Medicine (FPGMX)-Grupo de Medicina Xenómica-Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERer)-IDIS, Santiago de Compostela, Spain.

The Iberian Peninsula has been the subject of diverse population movements, both from Europe and North Africa, over millennia. The patterns of genetic variation within Spain reflect these movements, together with local isolation of populations due to cultural, linguistic, and geographical factors. Here we describe the largest genome-wide study of population structure across Spain, involving over 500,000 markers in ~600 individuals, analysed with powerful, recently developed statistical tools which exploit patterns of linkage disequilibrium. These analyses reveal striking patterns of genetic population structure at very fine geographic scales and shed light on the demographic history of populations within Spain. The genetic differences we identify distinguish individuals from many parts of Spain, including Galicia, the Basque Country, Catalonia, and the Balearic Islands. These differences correlate with both physical barriers to migration, and different languages spoken within Spain. Interestingly, we detect some genetic clusters that span from the north to the south of Spain, while groups differ more strongly east-west, suggesting predominantly north-south movements of people within Spain. Differing amounts of ancestry shared with individuals from Basque-speaking regions in populations from different parts of Spain also suggests a historic population movement out of the Basque region, largely in a southerly direction. Finally, we use a mixture model-based test to interpret genetic differentiation in genomic data. [1] Detecting and dissecting the fine-scale genetic population structure of Spain. C. Bycroft1, A. Carracedo2, C. Fernandez-Rozadilla1,2, C. Ruiz-Ponte1, I. Quintela-Garcia2, S. Myers1, P. Donnelly1,2,3,7,7T1)Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxford, UK.

2018M
Using constraints on FST to interpret genetic differentiation in genomic data. M.D. Edge1, M. Jakobsson2, N.A. Rosenberg1. 1) Department of Biology, Stanford University, Stanford, CA; 2) Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

The availability of genomic polymorphism data from large, worldwide samples has made it possible to study population structure at loci with rare alleles and low genetic diversity. Computation of FST is often a first step in population-genomic analyses of data from multiple populations. FST is a theoretically rich measurement of genetic differentiation, interpretable as an index of excess homozygosity due to population structure, as an excess coalescence time for alleles drawn from different subpopulations, and as the proportion of variance in allelic indicator variables attributable to population structure. At the same time, FST has sometimes behaved in surprising ways in genomic studies of human populations. For example, subpopulations within Africa often have low estimated FST, despite having considerable population structure as detected via individual-based clustering. Rare variants can generate low FST—despite being more geographically localized than common variants. We suggest that one way to understand these surprising phenomena is to consider mathematical bounds on FST as a function of other population-genetic statistics. We have previously determined strict upper bounds on FST as a function of homozygosity and as a function of the frequency of the most frequent allele for loci with arbitrarily many alleles. Here, we extend these bounds to accommodate loci with specified numbers of alleles, showing that FST is highly constrained both for low and high levels of diversity and can only reach values near the theoretical maximum of 1 for loci with intermediate diversity levels. These bounds provide an explanation for the surprising behaviors of FST; for example, the low values for African subpopulations can be explained in terms of the particularly high genetic diversity in Africa. As genomic studies continue to examine rarer and more localized alleles - which generate loci with low genetic diversity -- our bounds can facilitate the sensible interpretation of FST results.
2019S

Population structure of Amerindian 19 populations in South America. S. Nishikawa1, R. Saito1, N. Konno1, G. Tamiya1, N. Fuse1, Y. Nakamura2, J. Yasuda1, J. Danjoh1. 1) Department of Integrative Genomics, Tohoku Medical Megabank Organization, Tohoku University, Sendai, Miyagi, Japan; 2) Cell Engineering Division, BioResource Center, RIKEN Tsukuba Institute, Tsukuba, Ibaraki, Japan.

Introduction

Ancestors of Amerindians migrated from Eurasia to North American continent via Bering Strait, and then reached South American continent over 10 thousand years ago. The study of HTLV-1 in South American populations indicates that they can be divided into two major ethnic groups. However, detailed population structure and genetic diversity in Amerindians are still unclear. To address these questions, we obtained genome-wide SNP data from the Sonoda-Tajima Cell Collection, that is a vast collection of B lymphoblastoid cell lines established from 550 native Americans belong to 29 tribes in South America (Danjoh et al., 2011), and analyzed population structure within and among tribes, and then analyzed the genetic structure of Amerindian populations.

Methods

The standard SNP and samples quality control (QC) procedures including the estimation of “cryptic relatedness” were applied to the genome-wide SNP data on Genome Studio software and PLINK software, and removed closely related samples. After this step, 10 tribes retained insufficient number of samples, therefore we did not process them for further analysis. Remaining 19 tribes were used for the following analysis. The between-population Fst estimation, neighbor-joining (NJ) tree construction, and PCA were carried out with Arlequin, MEGA, and EIGENSTRAT software, respectively. The number of clusters (K) in Amerindian populations was estimated with ADMixTURE.

Results and Conclusion

NJ-tree showed that Samuna and Chipaya are distant from other populations. Moreover, PCA and cluster analysis showed that Samuna formed a distinct cluster along PC1 and PC2, on the other hand, Chipaya was clustered with the other populations. The result of cluster analysis showed that our 19 populations of Amerindian was constituted of 7 clusters. This is the first report for the detailed population structure of these Amerindians including Samuna.

2020M


Presence of polymorphisms in the CYP450 enzymes genes is one of the principal determinants for individual variability in drug response. Particularly, CYP2C9 enzyme metabolizes almost 20% of the existing drugs. CYP2C9 gene exhibits more than 57 variants, of which CYP2C9*2 (rs1799853) and CYP2C9*3 (rs1057910) have been associated with a reduced catalytic activity compared with wild type (CYP2C9*1). Variation in frequencies has been described for these two variants in Mexico, mainly between mestizos and native groups. This variation might be the cause, at least in part, of the disparities in drug disposition between ethnic groups. The present study is aimed to analyzing the CYP2C9*2 and CYP2C9*3 variants in two Native populations of Mexico to increase the information about the polymorphisms of pharmacogenomic relevance in the Mexican population. Methods. DNA samples from two Native Mexican populations (Lacandonians n=223 and Yaquis n=103) were genotyped for two variants in the CYP2C9 gene (CYP2C9*2 and CYP2C9*3). The variants were identified using PCRRFLPs. Genotype and allele frequencies and the Hardy-Weinberg equilibrium were performed using a package of statistical analysis for genetic marker data, ARLEQUIN v3.5.1.2. Results. In Lacandonian population, both heterozygous genotypes to CYP2C9*2 and CYP2C9*3 variants were found in a frequency of 0.4%. In Yaquis population, the heterozygous genotype to CYP2C9*2 variant was found with a frequency of 3.9%, while heterozygous genotype to CYP2C9*3 variant with a frequency of 14.6%. The compound heterozygous genotype CYP2C9*2/3 was observed only in Yaquis population. The CYP2C9*2 or CYP2C9*3 variants in homozygote state were no detected in none population. No significant deviations from HWE expectations were observed in any of the two studied populations. Conclusions. Variation in allele and genotype frequencies is evident between northern and southern Mexican native population, reflecting the differences among Mexican subpopulations. The presence of these variants could represent a risk factor of drug resistance or adverse reactions to those drugs metabolized by CYP2C9 enzyme in individuals of both populations.

2021S

Full ancient mitochondrial genomes using the Ion Proton platform: Analyzing the genetic diversity of the Neo-Eskimo. J. Tackney1, A.M. Jensen2, S. Watkins3, J. Brenner-Coltrain3, D.H. O’Rourke3. 1) Department of Anthropology, University of Utah, Salt Lake City, UT; 2) UIC Science LLC, Barrow, AK; 3) Department of Human Genetics, University of Utah, Salt Lake City, UT.

A cultural shift occurred in the North American Arctic archaeological record beginning around 1000 AD. This culture, known as the Neo-Eskimo Thule, quickly occupied the region from Alaska to Greenland, and their descendants are modern Inupiat/Inuit. Nuvuk is a long-term Thule village at Pt. Barrow, AK. Its cemetery has yielded archaeological material with radiocarbon dates between 980-1300 AD, suggesting that human remains from the site should span from the classic Thule period to modern Inupiat Eskimo.

Mitochondrial haplogroup D is the second most common haplogroup in northern Asia and it is present at high frequency in central and eastern Asia, as well as in the Americas. One subclade of D4 seems to have originated in southern Siberia and spread northward into northern Asia and the Americans post-Last Glacial Maximum. This subclade is present in the Inuit/Inupiat populations of North America at low frequency (~5%), is shared with some Siberian Chukotkans, and is absent in more southern Amerindian populations. To help further characterize the sequence variation present in this haplotype, we assessed the mitochondrial genomes of two samples previously typed for HVS1 SNPs, and assigned to D3/D4b1a2a1a. The bone samples were directly dated with median intercept calibrated dates of 1269 and 1313 AD.

We optimized the Ion Torrent sequencing library creation and the Torrent Suite bioinformatics pipeline for ancient DNA analysis by limiting reaction clean-up steps, determining PCR cycle numbers using qPCR, and expanding the mapping algorithms performed on the unaligned reads. Using mtDNA hybridization probe capture, we generated complete mtDNA for each sample on the Proton sequencing platform. The full genomes allowed us to place the samples with high phylogenetic confidence in the mtDNA tree. Sequences suggest population continuity with modern Inupiat/Inuit and the D3/D4b1a2a1a haplotypes confirm that the Thule belong in the Beringian, and not Asian or Pan-American, gene pools.

This research was supported by NSF grants OPP-0732846 AND OPP-0637246 to DHO’R, OPP-0820790 to JBC, and ARC-0726253 to AMJ.
2022M

Large scale whole-genome sequencing of the Icelandic population isolate. D.F. Gudbjartsson1,2, H. Helgason1,2, S.A. Gudjonsson3, F. Zink4, A. Oddsson4, A. Gylfason4, G. Magnusson4, B. Halldorsson1,3, S. Besenbacher2, A. Kong1, G. Masson1, U. Thorsteinsdottir1,2, A. Helgason1,2, P. Sulem1, K. Stefansson1,2. 1) Dept. Statistics, deCODE Genetics/Amgen, Inc, 101 Reykjavik, Iceland; 2) School of Engineering and Natural Sciences, University of Iceland, 101 Reykjavik, Iceland; 3) Institute of Biomedical and Neural Engineering, Reykjavik University, 101 Reykjavik, Iceland; 4) Bioinformatics Research Centre, Aarhus University, C.F. Møllers Alle, 8000 Aarhus C, Denmark; 5) Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland; 6) Department of Anthropology, University of Iceland, 101 Reykjavik, Iceland.

Here we describe the insights gained from sequencing the whole genomes of 2,636 Icelanders to a median depth of 20X. Twenty million single nucleotide polymorphisms (SNPs) and 1.5 million insertions/deletions (indels) passed stringent quality control. We report the number of sequence variants by minor allele frequency (MAF) and functional annotation. Almost all SNPs with derived allele frequency (DAF) over 2% in Iceland have been recorded in international databases of sequence variants, while for SNPs with a DAF of 0.1%, half of those in protein coding regions and 20% of those in non-protein coding regions have been recorded outside Iceland. The corresponding fractions of Icelandic indels that have been recorded elsewhere are substantially lower. Sequence variants in genes involved in the function of the nucleus have the lowest density and the highest fraction of rare variants (DAF < 0.5%, FRV), while variants in genes involved in sensory perception have the highest density and the lowest FRV. In particular, variants in olfactory receptors have a higher density and are more common than the genome average. Loss of functional (LOF) and missense variants have a higher density and are more common in the first and last exons of genes than they are in the middle exons. Missense and LOF variants in genes that have been associated with disease in the Online Mendelian Inheritance in Man, cataloging a dominant mode of inheritance, are less dense and rarer than variants in genes that have not been associated with disease. Variants with a chromatin state annotation have a density and FRV that fall between intergenic and untranslated regions. Sequence conservation between mammal species, as measured by the GERP score, predicts variant density and frequency. However, after accounting for the GERP score, sequence annotation has a substantial effect on variant frequency but not on variant density. Consistent with a founder effect and the small size of the Icelandic population, the oldest 56% most top gained SNPs in the frequency range between 0.1% and 1% in Iceland than in the European American part of the Exome Sequence Project data. Icelanders are more likely to be homozygotes than Hardy-Weinberg equilibrium would predict by a factor of 3 for variants with a MAF of 0.1%. We have described the sequence diversity and structure of the Icelandic population isolate and evaluated several types of sequence annotations based on the strength of selection signature.

2023S

Identifying Clusters of Rare Skeletal Genetic Disorders in Brazil - preliminary data. D. Cavalcanti, C. Moreno, Skeletal Dysplasia Group. Medical Genetics Department, FCM, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil.

The identification of clusters of rare genetic diseases has significant medical and scientific implications. The establishment of an investigation group of skeletal genetic disorders [osteochondrodysplasias (OCD) and skeletal malformations or dysostosis] at our institution has contributed to the creation of a register of these conditions in which familiar and demographic data in addition to the clinical ones have been systematically collected. This presentation aims to show the preliminary results related to the identification of clusters of OCD and dysostoses in Brazil. The data base was revised with special attention to the clinical and demographic data. The more common registered cases as FGFR3 group, osteogenesis imperfecta as well as campomelic dysplasia, type IV osteomelia, and incontinentia pigmentation were all excluded of the present analysis. Thus, among 359 cases 52 different diagnoses were established. The autosomal recessive (AR) inheritance pattern was observed in 62% (32/52) of the cases and parental consanguinity was present in 28% (30/107). The molecular investigation found a pathogenic mutation confirming the diagnosis in 25 cases of 21 families. Among 11 cases for whom parental consanguinity was referred as negative a homozygous mutation was found in seven patients and for five ones both paternal and maternal ancestors came from small and identical/closed cities, suggesting a common origin with an unknown and distant consanguinity. The following were considered evidence of a cluster: the presence of a rare mutation in different families associated or not with families from the same region, and cluster evidence already known from the literature. So far, five new clusters were found, that are distributed in different states: Blomstrand dysplasia (Alagoas), Pychodysostosis (Ceará), Short-rib polydactyly type 3 (Maranhão and Pernambuco), Opsismodysplasia (Pernambuco), Spondyloepicondylodysplasia (Minas Gerais). All these conditions are associated with parental consanguinity. We also found cases with one of Greig syndrome (Bahia), and a case of Riechier-Costa-Pereira syndrome (São Paulo), reinforcing clusters already known from the literature. In conclusion, we found five new clusters of AR skeletal conditions in Brazil, for whom the associated inbreeding suggests founder effect. [Supported by grants from CNPq # 402008/2010-3, 590148/2011-7].

2024M

The CARTaGENE Genomics Project: How successful bottlenecks have shaped the present French Canadian founder population. H. Gauvin1,2, JP. Goulet1, JC. Grenier1, M. Capredon1, T. de Maillard1, V. Bruot1, E. Gbeha1, E. Hip-Ki1, A. Hodgkinson1, P. Awadalla1,2, 1) Sainte-Justine Hospital Research Center, Université de Montréal, Montréal, QC, Canada; 2) Department of Social and Preventive Medicine, Université de Montréal, Montreal, QC, Canada; 3) Department of Pediatrics, Université de Montréal, Montreal, QC, Canada.

French Canadians (FC) are a population of Nouvelle-France's settlers who established themselves before the British conquest. The population has undergone recent rapid growth, with limited genetic exchange with other populations, and has colonized the whole province in successive waves leading to regional founder effects contributing to population structure. It is known that levels of kinship and disease prevalence can differ region to region within Quebec, making the population a particularly interesting case-study for understanding the origins of rare genetic diseases; in this study we attempt to disentangle and describe the processes driving these patterns. Using participants from the CARTaGENE project, the provincial biobank, we generated whole genome genotyping data for ~1000 individuals sampled in three regions (Montreal, Quebec City and Saguenay) with the OMNI2.5 chip, and then examined the fine structure of the Quebec population and searched for signals of genetic drift and selection using a haplotype-based approach. Focussing on individuals with FC descent, which was inferred from principal component analysis and information provided by participants, we performed a high-resolution analysis of the genetic structure of the population using the Chromopainter and FineSTRUCTURE clustering algorithms. FineSTRUCTURE analysis showed differences between two geographically close regional populations, Montreal and Quebec City, and provided greater resolution of the migration patterns within the province. Furthermore, using HaploFS, we identified genomic signatures of positive selection and genomic regions that have potentially drifted to particular high frequency among FC. Among these regions we located one, at a 50% frequency among FC from Saguenay, encompassing the RYR1 gene found to be implicated in malignant hyperthermia. This is a pharmacogenetic disorder for which cases were previously identified in families of French-Canadian heritage. Further support is provided by the fact that patients in these cases were previously identified in families of French-Canadian heritage. Collectively, these results provide a refinement of previous analyses of the structure of the Quebec population and suggest that the unique demographic history of FC may have affected specific regions of the genome that are relevant to the population's health.
2025S
Dominant mutation in GJB2 causes hearing loss in northeast mexican family. P. Loza-Becerra1,2, R. Rivera-Vega3, E. Guilezner-Contreras2, P. Berruecos-Villalobos3, S.A. Cuevas-Covarrubias1,3, UMSNH Morelia, Michoacán, Mexico; 2) Secretaria de Salud; 3) UNAM, Mexico DF; 4) Centro de Rehabilitacion Gpe. H. de De las Fuentes DIF-Torreon Coah.
To know mutations in GJB2 and GJB6 that causes hearing loss in mexican population, we make genotyping and audiometric study in n=95 families (trios and more) from the Mexican Republic to the north, north-east, north-west and west with at least one case non-syndromic sensorineural hearing loss (NSSNHL). It was 70% wild type (mutations in no-GJB2 nor GJB6 or mADN), eighteen different mutations and this: c.366G (p.G12R). Most of the cases are sporadic or recessives, and only one north-eastern family has the mutation c.366G (p.G12R) in nine members, three generations, both sexes affected and transmission male-to-male. All affected has neuro-sensory hearing loss. Now corroborated no anatomical anomalies. We remarks that its importance is clinical and epidemiological because the hearing loss by GJB2 is классically recessive and it is northern that mutan.

2026M
A large number of studies on nonsyndromic hearing impairment (NSHI) have been performed using samples collected from the Indian subcontinent, Middle East and Europe. There has been limited study of NSHI in African-Americans (AA) and sub-Saharan Africans. We evaluated the frequency of previously reported “pathogenic” variants in NSHI genes using data from the NHLBI-Exome Sequencing Project (ESP) which is a population-based study of AA (N=2203) and European-Americans [EA (N=4300)]. Of 201 observed variant sites those which are pathogenic according to ClinVar and/or the Deleasness Variation Database, we reclassified 121 (60.2%) variant sites as non-pathogenic based on literature, high allele frequencies in ESP and bioinformatics tools. In ESP, 80 pathogenic variant sites were observed in 5 autosomal dominant (AD) and 14 autosomal recessive (AR) NSHI genes, of which 24 variant sites in 7 genes cause syndromic hearing impairment (HI), e.g., MYO7A (MIM 276903) variants cause both NSHI and Usher syndrome. Of these variant sites, 49 (91 alleles) were found only in EA, 18 (23 alleles) only in AA, and 13 (194 alleles) in both EA and AA. GJB2 (MIM 121011) c.35delG (p.G12Vfs) was the AR NSHI variant site with the highest allele frequency in EA [N=89; 1.09% (95%CI:0.88%, 1.34%)] but with a much lower allele frequency in AA [N=4; 0.094% (95%CI:0.03%, 0.2%)]. Additionally, OTOF (MIM 603681) c.2348delG (p.G783fsA) was identified in 4 AA alleles [0.085% (95%CI:0.03%, 0.2%)] but is very rare in EA [N=1; 0.01% (95%CI:0.0003%, 0.07%)]. For EA, pathogenic variant sites were identified in 11 AR and 5 AD NSHI genes including variant sites in 7 genes which cause syndromic HI, while for AA pathogenic variant sites were identified in 13 AR and 2 AD NSHI genes, of which 6 genes have variant sites which cause syndromic HI. MYO7A c.3764delA (p.K1255Rfs) which was reported to cause Usher syndrome, is homoyoxous in 1 AA and 3 EA individuals. The use of EVS in order to exclude nonpathogenic variants must be done cautiously due to NSHI-causal variants within EVS. Knowledge of the frequency of NSHI variants in large population samples from different ethnic backgrounds is not only important to evaluate clinical significance but also aids in evaluating pathogenicity. Although population-based samples of AA aids in evaluating population-specific frequency of pathogenic variants, AA and sub-Saharan Africans with HI need to be studied to better understand the genes and variants underlying disease etiology.

2027S
Human population growth and purifying selection have increased the burden of autosomal and X-linked private mutations. F. Gao, A. Keinan. Department of Biostatistics and Computational Biology, Cornell University, Ithaca, NY.
Several recent projects have sequenced the genome or whole-genome of a large number of individuals. A common feature discovered in these data sets is the significant elevation of private variants, which are the segregating sites that have their minor allele in only one or few chromosomes across the entire sequenced sample. This phenomenon can be explained by the effect of recent human population growth and purifying selection. Here, we considered the amount of rare variants in the prism of the burden of private mutations. We defined the burden of private mutations as the proportion out of all heterozygous sites in an individual that are unique compared to the rest of the sample. This quantity also answers the following question: Considering a large number of sequenced individuals, how many new variants will be ascertained with each additional individual sequenced? We studied the burden of private mutations by calculating the prediction from different demographic models and comparing with empirical estimates based on the Neutral Regions (NR) dataset that we recently published and that from the NHLBI Exome Sequencing Project (ESP). By analyzing autosomal SNVs, we observed a significant excess in the proportion of private mutations in the empirical data compared with models of demographic history without a recent epoch of population growth. Incorporating recent growth into the model provides a much improved fit to empirical observations. This phenomenon becomes more marked for larger fit sample sizes. For example, after sequencing 10,000 individuals from the same population, still about 1 in 400 heterozygous sites (~6,000 variants) at the 10,011th individual are predicted to be novel, 18-times more than predicted in the absence of recent population growth. The burden of private mutations is additionally increased by purifying selection, generally with intergenic SNVs exhibiting the lowest burden and splice sites the highest. Finally, we contrasted the burden of private mutations between the autosomes and the X chromosome, across different functional annotations and across matched allele sample sizes ranging from n = 1 to n = 3027. The results of this comparison point to a combination of sex-biased demographic history and purifying selection in effect during all epochs of human expansion and contraction X SNVs in particular. Our findings suggest that careful consideration is needed in the design and analysis of sequencing-based association studies.

2028M
Whole Genome Sequence of Japanese from Tohoku Medical Megabank Prospective Cohort Study. M. Nagasaki1,2, F. Katsuoka1, N. Nariai1, K. Kojima1, D. Danjo1, Y. Kawai1, S. Saito1, X. Pan1, Y. Yokozawa1, R. Saito1, Y. Saito1, T. Mimori1, Y. Yamaguchi-Kabata1, K. Tsuda1, Y. Kuroki1, K. Kikuchi1, J. Yasuda1, M. Yamamoto1,2, Tohoku Medical Megabank Organization Tohoku University Project. 1) Department of Integrative Genomics, Tohoku Medical Megabank Organization, Sendai, 980-8575, Japan; 2) Department of Medical Biochemistry, Tohoku University Graduate School of Medicine, Sendai, 980-8575, Japan; 3) Department of Applied Information Sciences, Graduate School of Information Sciences, Tohoku University, 6-05 Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi 980-8579, Japan.
Tohoku University Tohoku Medical Megabank Organization (ToMMo) located at the north-eastern Japan is now developing a biobank (the final goal is 150,000 volunteers with millions samples) that combines medical and genome information during the process of rebuilding the community medical system and supporting health to foster the reconstruction from the Great East Japan Earthquake on Mar/11/2011. One of the missions of ToMMo is to reveal a fine genetic architecture of Japanese population to tackle the further GWAS analysis by combining the knowledge, which is daily accumulated in the ToMMo prospective genome cohort project, e.g. questionnaire data, physiological data, medical treatment records and other omics data from serum, plasma and immortalized lymphocytes. The first goal is to sequence 1K samples to cover MAF > 0.5% variants including short indels and large structural variants in Japanese for constructing a Japanese 1K reference panel in our experimental design, to minimize the bias caused by the differences of ethnics, protocol and bioinformatics analy- sis, we performed whole genome sequencing of 1K samples with 30X high coverage using the HiSeq 2500 rapid run mode and analyzed by the same bioinformatics pipeline. We have catalogued reliable SNPs with MAF > 0.5%. In the rest of the sample, we detected variants with MAF > 0.05% in the individual were pre-
2029S
Analyzing the impact of consanguinity and admixture on the Middle East Variome. E.M. Scott1, E. Spencer1, B. Copeland1, M. Abdelalef2, S. Gabrielli1, J. Gleeson1. 1) Neurogenetics Laboratory, Howard Hughes Medical Institute, Department of Neurosciences, University of California, San Diego, La Jolla, CA 92039, USA; 2) The Broad Institute of MIT and Harvard, Cambridge, MA 02141, USA.

Consanguineous matings are frequent among the Middle Eastern and North African populations. This affords a unique opportunity to study the regional effects of consanguinity and admixture on the genome. Several large-scale human migrations into the Middle East of recent and more ancient origins have been documented, and the Middle East is a hub of human migration, population admixture, and contribution of genetic diversity. Although details are disputed, several large-scale human migrations into an out of Africa, along with many more recent invading populations could have contributed to a complex heterogeneous Middle Eastern genome. Additionally, cultural practices predisposing consanguineous marriage have resulted in some of the highest known inbreeding coefficients, which theoretically should depress burdens of deleterious variants. Despite the size and uniqueness of these populations, so far they have been severely underrepresented and understudied. Here we investigate the genetic landscape of the Middle Eastern genome by repurposing a cohort of 2000 exomes samples resulting from three separate recruitment efforts focusing collection in the Gulf region, North Africa and Central Asia, collectively termed Middle East. Analysis of this cohort provides evidence for unique admixture components, and confirmatory evidence for inter-mixing of Sub-Saharan African and European populations. High levels of consanguinity resulted in a markedly higher rate of inbreeding coefficients compared with 1000G controls along with longer runs of homozygosity. Finally, we investigated potential impact these trends on the load of deleterious variant, demonstrating surprisingly a higher than expected burden of deleterious variants compared with European ancestry controls.

2030M
Rare variant stratification in a German/Austrian sample set. E. Tich1,2, B. Schormair1,2, B. Pütz3, D. Czamara4, M. Müller-Nurasyid5,6, P. Lichtenblau7, C. Trenkwalder2, W. Paulus8, B. Hög9,10,11, K. Berger9,1, I. Fietz9, W. Oertel1,2,3, T. Meltzer12,13, C. Gieger6, B. Müller-Myhsok14, J. Winkelmann1,2,14,15,16. 1) Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 2) Institut für Humangenetik, Klinikum rechts der Isar, Technische Universität München, Munich, Germany; 3) Max Planck Institute of Psychiatry, Munich, Germany; 4) Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-University, Munich, Germany; 5) Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 6) Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-University, Munich, Germany; 7) Paracelsus Elena Klinik, Kassel, Germany; 8) Department of Neurosurgery, University Medical Center, Georg August Universität Göttingen, Göttingen, Germany; 9) Neurologische Klinik, Medizinische Universität Innsbruck, Innsbruck, Austria; 10) Institut für Epidemiologie und Biostatistik, Westfälische Wilhelms Universität Münster, Münster, Germany; 11) Zentrum für Schlafmedizin, Charité Universitätsschulmedizin, Berlin, Germany; 12) Neurologische Klinik, Philipps Universität Marburg, Marburg, Germany; 13) Munich Cluster for Systems Neurology (SyNergy), Munich, Germany; 14) Center for Environmental Health, Neuherberg, Germany; 15) Ludwig-Maximilians-University, Munich, Germany; 16) Department of Neurology and Neurosciences, Stanford Center for Sleep Medicine and Sciences, Stanford University, Palo Alto, California, USA.

Genome-wide association studies of complex traits with common genetic variants could only explain a part of a trait’s heritability. The missing heritability is expected to be partially explained by rare genetic variation. But this variation was shown to already stratify populations across relatively short geographical distances within Europe. A German/Austrian sample set comprising 3,654 patients with the same common complex disease from geographically distinct recruitment areas as well as 2,911 convenience controls (KORA) were genotyped for mainly rare variants using Illumina®'s Infinium HumanExome BeadChip. After quality control, all polymorphic markers were grouped into bins of 300k SNPs each, covering 60% of the genome, based on all median conservation scores (PhyloP, GERP++, and SiphY annotations provided by CHARGE), and were pruned by LD (r2 > 0.2). The bins were identified by identifying SNPs between individuals and the subsequently grouping SNPs into bins (as described by our group in an earlier publication). Differentiation of samples within these clusters is negatively correlated with MAF and positively with degree of conservation. Furthermore, IBS calculations were used for multidimensional scaling (MDS) to visualize the groups. Further, we identified 65 additional genomic annotations in predicting ASE, including DNase hypersensitivity, transcription factor binding, allele frequency and evolutionary conservation. We use publicly available RNA sequencing data in 373 lymphoblastoid cell lines from the GEUVADIS consortium to perform over 3.3M tests of ASE at roughly 10K genes. This data provides sufficient power to evaluate the association of diverse genomic sequence annotations at candidate regulatory loci with allelic imbalance of nearby genes. Specifically, individuals heterozygous at non-coding loci are pooled to estimate the odds of allelic imbalance in the context of each annotation of interest. Our results demonstrate a positive association between ASE and the level of heterozygosity, with the strongest modulation by the position of heterozygous variants with respect to the TSS. Additionally, we find that the likelihood of ASE is associated with the severity of nearby heterozygous mutations; in particular heterozygous SNPs have a weaker effect than indels, consistent with a general observation in cell lines. Furthermore, we find that roughly 65 additional genomic annotations in predicting ASE, including DNase hypersensitivity, transcription factor binding, allele frequency and evolutionary conservation. Finally, we extend this work to develop a statistical model that accounts for confounding from individual and environmental sequence information. Using this framework, we characterize the effects of rare variation and contrast the genomic context of rare versus common regulatory variants.

2031S
Population structure in the UK: Rare variant analysis using whole genome sequencing in 3,621 samples in the UK10K cohorts project. K. Walter1, S. Metzler2, E. Ziegler1, Y. Memari1, J. Min2, J. Huang2, S. Coca3, S. Schieflf1, I. Mathieson1, D. Lawson2, N. Soranzo3, UK10K Consortium Cohorts Group. 1) Human Genetics, Wellcome Trust Sanger Institute, Hinxton, United Kingdom; 2) Twin Research & Genetic Epidemiology, University College London, United Kingdom; 3) MRC CATE Centre, University of Bristol, United Kingdom; 4) Institute for Maternal and Child Health-IRCSS “Burlo Garofolo”-Trieste, University of Trieste, Italy; 5) Harvard Medical School, Boston MA 02115, United States; 6) Heilbronn Institute, School of Biological Sciences, University of Bristol, United Kingdom.

Population structure is a well-characterized potential confounder of association studies based on common variants, but the structural pattern for rare variants and their influence on association studies is less understood. The cohorts arm of the UK10K project undertook whole-genome sequencing at low-read depth (median ~7x) in nearly 4,000 individuals from two large population samples in the UK (TwinsUK, N=1,754 and the Avon Longitudinal Study of Parents and Children (ALSPAC), N=1,867) in a comprehensive exploration of associations between rare and common genetic variants and a set of 61 bi-medically important quantitative phenotypes. The two study cohorts have marked differences in demographic profile with ALSPAC participants originating from a geographically restricted area (Bristol) in the South West of the UK, while the TwinsUK participants were born in all parts of the UK. After stringent QC steps, the data set comprises 42 million SNPs, 3.5 million INDELS and about 18,000 large deletions across 3,621 study participants. Here we describe the extent to which geographic stratification exists at rare variants by focusing on 31 shared ‘core’ phenotypes in 1,139 twins from each available place sample. We further modelled correlation of genotypic and phenotypic data at these geographical locations, and compared them to simulated datasets. Finally, we applied Mantel tests to analyze the significance of genotypic and phenotypic relationships given the distance metrics. Overall, these analyses suggested that there is a moderate genetic structuring of very rare alleles (MAF=0.1-0.3%), however this structure is not associated with phenotypic variation and is unlikely to pose a serious concern for association studies of complex quantitative phenotypes and rare variation in the UK.
**2033S**

Targeted transcriptomics to compare the susceptibility of human naive and pre-immune volunteers to an infection challenge with viable *Plasmodium vivax* sporozoites. ML. Rojas-Pena1, A. Vallejo4, S. Herrera2, M. Arévalo-Herrera2, M. Gibson1. 1) Biology, Georgia Institute of Technology, Atlanta, GA; 2) Faculty of Health, Universidad del Valle, Cali, Colombia; 3) Malaria Vaccine and Drug Development Center, Cali, Colombia; 4) Cauca-seco Scientific Research Center, Cali, Colombia.

Malaria is one of the most serious infectious diseases around the world, with an estimated 300 to 500 million cases annually, most of which occur in tropical areas. Infection rates vary across regions of Colombia due to different geographical conditions. Individuals living in high-prevalence malaria regions acquire a natural immunity that gradually eases the symptoms of the disease, and is likely to impact the efficacy of vaccines for malaria. Additionally, little is known about the immunological response to early infection and how the immune system may be boosted during vaccination. This study hypothesizes that the delayed onset and decrease of clinical and parasitological manifestations of malaria observed in pre-immune volunteers from Buenaventura (high prevalence of malaria) compared to naive individuals from Cali (low prevalence of malaria) is associated with altered peripheral blood gene expression. Here we describe the use of nanofluidic Fluidigm quantitative RT-PCR arrays targeting a set of 96 transcripts that are broadly informative of the major axes of variation, to explore the difference in gene expression between Cali and Buenaventura individuals experimentally exposed to viable *Plasmodium vivax* sporozoites and monitored at six time-points (Pre-challenge, day 5, day 7, day 9, week 3 and month 4). The largest sources of variation are individuals and time. Site has a relatively minor effect, although it does separate the two clusters in the overall profiles of expression among the samples. Our analyses show that week 3, at the peak of parasitemia, is the most perturbed of the different time points, while among the axes, axis 3 related to B-Cell activation is most perturbed. RNA of the samples will be used to assess transcript abundance genome-wide and to perform gene expression profiling by RNA-seq of the samples to identify differentially expressed genes. Gene expression profiling of lymphocytes can thus be used to identify the type and duration of the immune signals that are biomarkers for vaccine immunogenicity, and establish how pre-immune exposure modifies their activation.

**2034M**

Conflations of short IBD blocks can bias inferred length of IBD. C.W.K. Chiang1, P. Ralph2, J. Novembre3. 1) Ecology and Evolutionary Biology, UCLA, Los Angeles, CA; 2) Computational Biology and Bioinformatics, USC, Los Angeles, CA; 3) Human Genetics, University of Chicago, Chicago, IL.

Identity-by-descent (IBD) is a fundamental concept in genetics with many genetic applications. A common working definition of IBD blocks is that they are contiguous segments of the genome inherited from a recent shared common ancestor without intervening recombination. Long IBD blocks (>1cM) can be efficiently detected by a number of programs using high-density SNP array data of a population sample. However, all programs detect IBD based on contiguous segments of identity-by-state (IBS). As such, detected IBD blocks could often be due to the conflation of smaller IBD blocks inherited from different common ancestors. Here, we show through theory and simulation that the conflation of small IBD blocks can occur with appreciable frequencies and can lead to errors in estimating the length distribution of IBD blocks, thereby affecting downstream inferences. Specifically, we used coalescent simulations where we know the precise genealogy of the sample and found that, under a realistic demographic model of human history, >35% of the detected IBD segments of 1cM or longer are composed of at least two subsegments, oftentimes each of appreciable length. This effect is universally observed across different IBD detection programs and demographic histories, and was more pronounced for shorter segments (1 to 2cM long) compared to longer segments (>2cM long). Furthermore, we characterized the rate of subsegment conflations as a function of minimum subsegment lengths and gap sizes, and obtained a conflation rate of 0.67 per 1000 pairs of individuals due to segments as small as 0.2cM. Based on this rate, ~28% of all IBD segments between 1 to 2 cM would be due to conflation. Finally, we propose a novel estimator of the de novo mutation rate using IBD blocks detected in population samples. We have found via simulations that such an approach will only work if it explicitly models the conflation of short IBD segments (e.g. 15-fold higher estimates are obtained if conflation is ignored). Thus, the conflation effect should be carefully considered as methods to detect shorter IBD blocks using sequencing data are being developed.
A systematic analysis of genetic forms of dilated cardiomyopathy reveals numerous ubiquitously expressed and muscle-specific genes. F.W. Asselbergs1, G. Kummeling1, A. Sammani1, M.P.M. Linschoten1, A.F. Baas2, J.J. van der Smagt1, P.A. Doodeman1, D. Dooijes2, M. Moky1, M. Harakalova1, 1) Department of Cardiology, University Medical Center Utrecht, Utrecht, Netherlands; 2) Department of Medical Genetics, University Medical Center Utrecht, Netherlands; 3) Division of Pediatrics, Wilhelmina Children’s Hospital, University Medical Center Utrecht, Netherlands.

Background: Dilated CardioMyopathy (DCM) is a genetic disorder with several tens of genes known to underline its etiology. Despite considerable progress being made in developing next generation sequencing-based diagnostic panels using a subset of the most prevalent 40 DCM genes, the genetic cause remains unsolved in the majority of patients. This suggests that to increase the yield of diagnostic testing all relevant DCM-implicated genes need to be catalogued and carefully considered. Methods and Results: We have conducted a systematic literature search on PubMed, Embase, and OMIM, to find genes that have been implicated in syndromic and non-syndromic DCM and PeriPartum CardioMyopathy (PPCM). Besides strong evidence for mitochondrial inheritance, our search yielded a total of 113 nuclear protein-coding genes. In addition to 42 genes sufficiently reviewed elsewhere we provide a comprehensive annotation of the level of genetic evidence for the set of remaining 71 genes. Next, we investigated the tissue specificity of collected genes using public RNA sequencing data. We show that genes primarily expressed in heart more likely result in DCM with possible myopathies while genes expressed ubiquitously cause DCM with extramuscular manifestations. Conclusions: This comprehensive analysis of genetic evidence of DCM revealed a much higher number of genes than routinely screened at the moment. However, these findings have to be carefully considered and validated in larger cohorts. Our results suggest that targeted sequencing of all known protein-coding genes and the whole Mt-DNA together with consideration of the tissue specificity of mutated genes, will improve genetic testing and facilitate further genotype-phenotype studies in DCM.

The -844 GA PAI-1 polymorphism is associated with Acute Coronary Syndrome in Mexican population. I.J. García-González1,2, Y. Valle1, E. Sandoval-Pinto1,3, E. Valdes-Alvarado1,3, F.J. Muñoz-Valle1, HE. Flores-Salinas1, LF. Figuera-Villanueva1, NO. Dávalos-Rodríguez2, JR. Padilla-Gutiérrez1. 1) Instituto de Investigación en Ciencias Biomédicas, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, México; 2) Doctorado en Genética Humana, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, México; 3) Doctorado en Ciencias Biomédicas, Centro Universitario de Ciencias de la Salud, UdeG. Sierra Mojada 950, 44350, Guadalajara, Jalisco, Mexico; 4) Doctorado en Genética Humana, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, México; 5) Centro de Investigación Biomédica de Occidente, 44350, Guadalajara, Jalisco, México; 6) Instituto de Investigación en Genética Humana, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, México.

Introduction: Acute coronary syndrome (ACS) has an important impact in public health with high morbidity and mortality. Pro-thrombotic and pro-inflammatory states are involved in the pathogenesis of the disease. Plasminogen activator inhibitor-1 (PAI-1) is the major inhibitor of the fibrinolysis and also is part of immune response. The -844 GA gene polymorphism is related to increased PAI-1 protein levels. Aim: We evaluate the association of -844 GA PAI-1 polymorphism with ACS. Methods: A total of 646 individuals were recruited from Western Mexico: 350 unrelated healthy subjects and 296 patients with diagnosis of ACS. PCR-RFLP and PAGE with silver stain were used to identify the polymorphism. Results: The most important risk factor in our population was hypertension, followed by smoking. The genetic distribution showed an association of the A allele (OR=1.27, p=0.04) and AA genotype (OR=1.86, p=0.02) with ACS. The recessive model displayed similar results (OR=1.76, p=0.02). As additional finding, we observed significant differences in the genetic distribution of ACS dyslipidemic patients (p=0.04). Conclusion: The A allele and AA genotype of -844 polymorphism of PAI-1 gene are risk factors for ACS. The AA genotype might be associated with the development of dyslipidemia in ACS patients.
2038M
Associations of Endothelial System Genes with Blood Pressure Changes and Hypertension Incidence: the GenSalt Study. F. Liu1, J. Ho2, D. Gu1, D. Rao3, J. Huang1, J. Nixoo1, B. Luckett-Chen1, C. Li1, X. Yang1, J. Li1, T. Rice1, L. Shimmin1, T. Kelly1. 1) Department of Evidence Based Medicine, Fijiwai Hospital, National Center of Cardiovascular Diseases, PUMC & CAMS, Beijing, China; 2) Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine; 1440 Canal Street, Suite 2000, New Orleans, LA 70112, USA; 3) Division of Biostatistics, Washington University School of Medicine, St. Louis, MO; Campus Box 8067, 660 South Euclid Ave. St. Louis, MO 63110-1093, USA; 4) Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine; 1440 Canal Street, Suite 2000, New Orleans, LA 70112, USA; 5) Division of Biostatistics, Washington University School of Medicine, St. Louis, MO; Campus Box 8067, 660 South Euclid Ave. St. Louis, MO 63110-1093, USA; 6) Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine; 1440 Canal Street, Suite 2000, New Orleans, LA 70112, USA; 7) Dept of Human Genetics and Disease Diversity, Tokyo Medical and Dental University, Tokyo, Japan.

High blood pressure (BP) is a complex trait, influenced by both environmental and genetic determinants. Although established as a heritable trait, the genomic mechanisms underlying BP regulation remain largely unknown. The objective of the current study was to examine the associations of 206 common variants in 15 endothelial system genes with BP changes and hypertension incidence among 1,775 Han Chinese participants of the family-based Genetic Epidemiology Network of Salt Sensitivity (GenSalt) follow-up study. Nine BP measurements were obtained at baseline and during each of two follow-up observations using a random-zero sphygmomanometer. The associations of 206 SNPs in 15 endothelial system genes with BP changes and hypertension incidence were assessed using mixed models to account for the correlations of repeated measures among individuals and within families. A genotype by time interaction term was used to model differences in longitudinal BP change according to genotypes over time. Gene-based analyses were conducted using the truncated product method. The Bonferroni method was used to adjust for multiple testing in all analyses. Among those free from hypertension at baseline, 513 (32.1%) GenSalt participants developed hypertension during the average 7.2 years of follow-up. In single-marker analyses, each copy of the minor alleles of SELE markers rs4565704, rs4927212 and rs5368 was associated with increased risk of hypertension with relative risks (95% confidence intervals) of 1.42 (1.12, 1.79) and 1.77 (1.16, 2.69) for hypertension and 1.66 (1.04-2.4) for hypertension respectively. SELE marker rs3917436 predicted longitudinal DBP change (P=8.27x10^(-5)). Results of gene-based analyses showed the SELE gene was significantly associated with SBP change and DBP change and hypertension incidence (alpha=3.1x10^(-6)). Moreover, the DDAH1, SELP, and COL18A1 genes were associated with hypertension incidence (P=2.00x10^(-6)). Furthermore, the DDAH1, SELP, and COL18A1 genes were associated with SBP change (P=2.00x10^(-6)), DBP change and hypertension incidence (P=2.00x10^(-6)). In conclusion, the current study provides strong evidence of a role of endothelial system genes in BP progression and hypertension incidence. Future studies will be required to identify the causal variants underlying the observed associations. 

2040M
A replication study for fifteen coronary artery disease susceptible loci in a Japanese population. K. Ozaki1, Y. Sakata2, S. Suna3, Y. Onouchi4, T. Tsuzugawa5, M. Kuroda3, T. Konmuro5, T. Tanaka1,2. 1) Lab for Cardiovascular Diseases, RIKEN,Ctr Integrative Medical Sci, Yokohama, Japan; 2) Dept of Cardiovascular Medicine, Osaka University Graduate School of Medicine, Suita, Japan; 3) Dept of Public Health, Chiba University Graduate School of Medicine, Chiba, Japan; 4) Lab for Medical Sci Mathematics, RIKEN, Ctr for Integrative Medical Sci, Yokohama, Japan; 5) Lab for Genotyping Development, RIKEN Ctr for Integrative Medical Sci, Yokohama, Japan; 6) Dept of Cardiovascular Med, Grad School of Med, The University of Tokyo, Tokyo, Japan; 7) Dept of Human Genetics and Disease Diversity, Tokyo Medical and Dental University, Tokyo, Japan.

Recent large-scale association analysis for coronary artery disease (CAD) identified fifteen new loci with genome wide statistical significance. Replication for the genetic association in other ethnic population is one of important issues to investigate a clinical benefit in the future. We conducted here the reproducibility for the association of the 15 single nucleotide polymorphism (SNP) loci (IL6R, rs48456825, APOB, rs715153, ZEB2,AC074093.1, rs2277923, VAMP5, VAMP8-GGCX, rs1291119.1, ABCG5-ABCG8, rs4554713, GUCY1A3, rs7692387, EDNRA, rs66942241, SLC22A4, rs17689550, KCNN5, rs10947789, PLG, rs4252120, HDAC9, rs2023938, LPL, rs264, TRIB1; rs2954029, FLT1; rs3194928, FURIN(rs17514846) and CAD with 7,990 cases and 6,582 controls in a Japanese population. We found a convincing association with statistical significance for rs3919428 in intron 6 of FLT1, encoding fms-related tyrosine kinase 1, for CAD susceptibility (P = 5.98 x 10^(-6)). We also found the replication for the association of two SNPs, rs6842241 in 5-flanking region of EDNRA, encoding endothelin receptor type A, and rs17514846 in intron1 of FURIN, encoding a calcium-dependent serine endopeptidase, with CAD (P = 0.000207 and P = 0.00208, respectively). Our validation results revealed that the three loci were genetic risk factor for CAD in the Japanese population.

2041S
Relations between PAI-1 plasma levels and adipsic tissue expression of PAI-1 and hsa-mir-421 microRNA. DA. Tregouet1, M. Civelek2, L. Markku3, MC. Alesa4, AJ. Lussi5, PE. Morange2, 1) UMR 5116, INSERM, Paris, France; 2) Departments of Microbiology, Medicine, and Human Genetics, UCLA School of Medicine, Los Angeles, USA; 3) Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland; 4) Insier U1062, Aix-Marseille University, Marseille, France.

Several works have demonstrated that increased plasminogen activator inhibitor (PAI)-1 levels, the main inhibitor of the fibrinolytic system, could predict the risk of metabolic syndrome (MetS) and associate with its separate components, such as visceral obesity, increased blood glucose levels and dyslipidaemia. Biological mechanisms responsible for these associations are still debated and several arguments are in favor of a key role of adipose tissue in these phenomena. As part of the METSIM study, we measured plasma PAI-1 activity in a sample of 850 healthy males for which the subcutaneous adipose tissue expression of the SERPINE1 gene coding for PAI-1 was also available. We demonstrated that adipose tissue SERPINE1 expression was correlated to plasma PAI-1 activity (p = 0.23, p < 0.01, beta = 0.02) and adipose tissue PAI-1 activity (p = 0.01, beta = 0.01). The influence of SERPINE1 expression on MetS components was mainly independent on plasma PAI-1 activity levels as their correlations were hardly modified by the adjustment on plasma PAI-1 activity. Conversely, all correlations were strongly reduced (p = 0.14, p < 0.01, beta = 0.10) after adjusting for BMI. In a subsample of 194 METSIM individuals, we were also able to show that both plasma PAI-1 activity and adipose tissue SERPINE1 mRNA expression positively correlated with adipose tissue expression of hsa-mir-421 microRNA (p = 0.22, p = 0.0018, p = 0.23; p = 0.04, respectively), a microRNA whose plasma levels have been recently shown to correlate with plasma PAI-1 activity. Interestingly, the association between adipose tissue levels of hsa-mir-421 and plasma PAI-1 activity was still significant after adjusting for BMI (p = 0.015) and SERPINE1 mRNA levels (p = 0.013). This study confirms in a large epidemiological cohort the strong association between PAI-1 plasma levels and adipose tissue serpin 1 expression, and provides novel elements suggesting, for the first time, that this expression may be influenced by adipose tissue expression of the hsa-mir-421 microRNA.

2039S
Novel NKX2.5 mutation associated with congenital heart disease in South Indian patients. S. Mattapally1, S. Nizamuddin2, KS. Murthy2, K. Thangaraj2, SK. Banerjee1. 1) Division of Medicinal Chemistry and Pharmacology, CSIR-Indian Institute of Chemical Technology, Uppal Road, Hyderabad 500007, India; 2) Innova Children’s Heart Hospital, Tarnaka, Hyderabad 500017, India; 3) CSIR-Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India. 

Background: NKX2.5 is one of the transcription factors, it is a cardiac specific homeobox gene and act as early marker genes for heart field development. Heterozygous NKX 2.5 germline mutations were reported to cause congenital heart disease (CHD), hence the aim of this study is to find the association of NKX2.5 mutations with CHD among the south Indian CHD patients. Method: NKX 2.5 gene was sequenced in 100 CHD patients (ASD, VSD, TOF and SV) and 200 controls. Functional significance of the observed NKX2.5 mutations were analyzed using in silico software such as: Polyphen, SIFT, PMut, plink, etc. Results: Our analysis with NKX2.5 gene revealed a total 7 mutations, out of which 3 were in intronic region, 3 mutations were in coding region and 1 mutation was in 3’ UTR. Of the above mutations, one was found to be associated with Tetralogy of Fallot (TOF) and two (rs2277923 and a novel mutation) were strongly associated with VSD. Interestingly, one novel missence mutation (p-value = 0.008744, Asp16Asn) was most significant findings of this study. Our ‘in silico’ analysis also provides evidence that some of the mutations reported above are pathogenic. Conclusion: The present study found that NKX2.5 genetic variations are associated with TOF and VSD in South Indian patients.
2042M Functional fine mapping of the genes involved in plasma lipid metabolism in the LD-block of NCAN/CILP2/PBX4 region. S. Boonvisut, A. Boonchaya, S. Jitsumote, M. Sato, H. Kadowaki, Center for Molecular Medicine, Jichi Medical University, Tochigi, Japan.

Dyslipidemia is a major risk factor of cardiovascular diseases. The estimated heritability ranges from 0.20-0.60. NCAN/CILP2/PBX4 region is the initially identified locus by GWAS that substantially influences plasma lipid concentrations. In addition, the genetic analysis of non-alcoholic fatty liver disease (NAFLD) reported that risk allele increasing LDL and TG levels in this region reduced the risk of NAFLD. These studies suggested that liver is one of the main organs where the responsible gene acts to regulate plasma lipid metabolism. However, NCAN/CILP2/PBX4 region is in a tight linkage-dis-equilibrium block spanning over 300kb, in which 11 genes are encoded and the responsible gene has not yet been identified. This study aims to identify the molecule involved in lipid metabolism from this region. Four genes expressed in liver, Trm6s12, Suga1, Gata2a, and Yfemn3, were chosen from the 11 genes referring to ENCODE database. The high-throughput RNA (shRNA) templates against them were synthesized and inserted in adenovi- ruses vectors. The recombinant adenoviruses were administrated into mice and the plasma lipid levels were followed up. Knocking-down of Suga1 showed increased serum cholesterol level in both feeding and fasting states (p = 0.04). The mRNA of Suga1 decreased its fasting state (p = 0.04). While, transcriptome Analysis of these mice revealed that knocking down of Gata2a and Suga1 affected on several lipid metabolism pathways, such as fatty acid metabolism, steroid synthesis and bile acid biosynthesis, which was not noteworthy that shSuga1 enhanced the expression of Apob experimentally, evaluated to study the effect of SUGP1 on the expression of APOB in another main organ, intestinal epithelium, Caco-2 cells with enforced expression and knockdown of SUGP1 were established using lentiviral gene transduction. The established cells were used to study the expression of APOB. In order to study the effect of SUGP1 on the expression of APOB, secreted APOB-48 and -100 protein levels and the ratios of them were increased by the knock-down of SUGP1. The APOB mRNA-editing for the conversion from APOB-100 to -48 in differentiated Caco-2 cells was increased by the knock-down and decreased by the enforced expression of SUGP1. SUGP1 (SURF and G patch domain containing 1) is estimated to be involved in mRNA processing. These data identified SUGP1 as the most powerful candidate gene regulating serum lipid levels in NCAN/CILP2/PBX4 region.

2043S Mutations in genes NKX2.5, GATA4 and TBX5, associated to congenital heart disease with septal defect in pediatric patients from the Guadalajara Civil Hospital Fray Antonio Alcalde, Guadalajara, Mexico; 2) Laboratorio de Diagnóstico Molecular, Tabasco, Mexico; 3) Pediatric Cardiology, Hospital Civil de Guadalajara, Guadalajara, Mexico.

Introduction. Congenital heart disease is the leading cause of birth defects, affecting approximately 1% of all live births. It has been associated with single mutations in transcription factors that regulate heart development. The genes frequently reported are GATA4, NKX2.5 and TBX5. These interact during cardiogenesis. Objectives. Investigate the mutations frequently reported in literature of the genes, GATA4, NKX2.5 and TBX5 in children with congenital heart disease involving septal defects in HCFAA. Material and Methods. We selected 30 children with congenital heart disease (CHD) with septal defect, a genetic history was established and by electrocardiography CHD type in each patient was determined. DNA was extracted and sequenced for mutations in GATA4, NKX2.5 and TBX5. Results. In total 84 samples of all participants, 30 of these samples were from patients diagnosed with CHD with septal defect and 5 of their parents. We found seven patients with mutations in genes previously selected for this study, the mutations seen in patients were also seen in all occasions in one or both parents. We found 3 female parents with mutations, and 5 male. The sex ratio was 3:3. Of patients with CHD with septal defect found association between two or more of the studied mutations. In patients with mutations we found that one or both parents had the same mutations. The frequency of mutations in index cases (n = 30) was 13.33%. (4) for mutations in TBX5 - 16.6%; (5) for mutation in GATA4 and 10%; (3) for mutation in NKX2.5 (c.73C > T). Finding that approximately 60%; of patients with this mutations appeared in double or triple, mainly related to the GATA4 gene. Conclusions. We found a frequency of these mutations in the index cases (n = 30) in TBX5 - 16.6%, for the mutation in GATA4 and 10%; mutations in NKX2.5 (c.73C > T). No patient had mutation c.533C > T in gene NKX2.5. From 30 families we found that 23%; of them had single or combined mutations of the genes studied. The type of inheritance resembles autosomal dominant. In carrier parents no abnormalities were reported, however they inherited the mutation.

2044M Standard Schnauzer dogs with dilated cardiomyopathy have a 22 bp deletion and frame shift in RBM20. D. Gilliam1, M. Harmond2, G.S. Johnson1, T. Mhatre-Mutangadura3, L. Hansen1, J.F. Taylor2, R.D. Schnabel1, S. Leach1. 1) Veterinary Pathobiology, University of Missouri, Columbia, MO; 2) Department of Veterinary Medicine, University of Missouri, Columbia, MO; 3) Animal Sciences, University of Missouri, Columbia, MO.

To identify the molecular genetic cause for dilated cardiomyopathy (DCM) in Standard Schnauzers, we generated a 30-fold average coverage whole genome sequence (WGS) with DNA from a male Standard Schnauzer that died at 13.5 months of age with DCM and congestive heart failure. This WGS contained 69 homozygous sequence variants that were predicted to alter the primary structure of the gene product and absent from the WGSs of 101 canids not exhibiting DCM. A homozygous 22 bp deletion and frame shift in exon 11 of RBM20 was the sequence variant considered most likely to be causal because mutations in the human and rodent orthologs cause DCM. All 9 of the confirmed DCM cases for which DNA was available tested homozygous for the deletion allele (del/del). A prospective screen for the mutation identified 63 Standard Schnauzers that were homozygous for the wild type allele, 225 Standard Schnauzers that were heterozygous, and 13 additional Standard Schnauzers that were homozygous for the deletion allele. Eight of these 13 del/del Standard Schnauzers have subsequently been diagnosed with occult DCM via echocardiography. The other 5 were not available for evaluation and 3 of them were < 1 year old, before clinical signs typically appear. Echocardiographic evaluation of 17 Standard Schnauzers that were homozygous for the deletion allele, including a 15 year old Standard Schnauzer, has revealed no echocardiographic evidence of myocardial dysfunction. Thus, RBM20-associated DCM appears to be rare. This is the first large, NCAN/CILP2/PBX4 region associated DCM is reported to be a dominant trait. In addition, there appears to be a gender difference in survival time among Standard Schnauzers. Male Standard Schnauzers with the homozygous RBM20 deletion have survived to a mean age of 43.4 months (range 19 months to 55 months). To our knowledge, a similar gender difference has not been reported for human RBM20-associated DCM patients.

2045S Disruption of the SEMA3D gene in a patient with congenital heart defects. C. Le Caignec1, I. Sanchez-Castro2, O. Pichon3, D. Poulian1, A. Briand1, V. Gounay4, A. David1, 1) Medical genetics department, CHU Nantes, Nantes, France; 2) INSERM, UMR1087, l’institut du thorax, Nantes, France; 3) Université de Nantes, Nantes, France; 4) Pediatric cardiology department, CHU Nantes, Nantes, France.

Congenital heart defects are associated with abnormal migration of cardiac neural crest cells during heart development, and the migration of cardiac neural crest cells into the outflow tract during cardio- genesis. SEMA3D null mice present with anomalous pulmonary venous connec- tions and atrial septal defects but the role of SEMA3D in humans remains unclear. The results suggest that a truncated SEMA3D may have hampered the migration of cardiac neural crest cells into heart development, and consequently contributed to CHD in our patient.
2046M

P. Nallari. Dept. of Genetics, Osmania University, Hyderabad, Andhra Pradesh, India.

Congenital Long QT syndrome (cLQTS) is an inherited arrhythmogenic disease characterized on ECG by a prolonged QTc interval. The ECG manifestation reflects an abnormally prolonged ventricular action potential, which can be the substrate for life-threatening arrhythmias that lead to syncope and sudden cardiac death. Heterogeneity of clinical manifestations corroborated with genetic etiology is well established. Screening of ion channel encoding genes was carried out by PCR-SSCP analysis followed by commercial sequencing in 46 LQTS probands and 69 available family members. We report novel variations in four LQTS probands and their family members in KCNQ1 gene. In-silico analysis predicts that these variations may lead to abnormal IKs channel and altered ion fluxes. The variation is also observed in the family members of the probands which helps in predictive testing and prognosis. The study is aimed at highlighting the importance of family screening in clinical management of LQTS with emphasis on personalized medicine.

2047S


Introduction: The acute coronary syndrome (ACS) is a complex disease where genetic and environmental factors are involved. TNF-α is a candidate gene for ACS progression due to its contribution in the inflammatory process and endothelial function. The rs1799964 polymorphism in the TNF-α gene consists of a T>C change at the -1031 promoter region. The C allele has been associated with changes in gene expression and plasmatic levels thus associated with a decrease in the risk of cardiovascular disease. Objective: To associate the TNF-α rs1799964 polymorphism with ACS and to measure the serum levels of TNF-α.
Methods: We recruited 251 patients with ACS classified according to American College of Cardiology and 184 healthy subjects (HS) age-matched from Western Mexico. The study was made in accordance with the Declaration of Helsinki. All individuals accepted to participate and an informed written consent was obtained. The rs1799964 polymorphism was identified by PCR-RFLP. Fragments were separated in polyacrylamide gel electrophoresis. The TNF-α was measured using an enzyme-linked immunosorbent assay. The genotype and allele differences were estimated by Fisher's exact test. The association measure was evaluated by OR and 95% of confidence intervals. The Mann-Whitney U test was applied in order to compare the TNF-α serum levels. The significance level was p<0.05. Results: The genotype and allele frequencies of the rs1799964 polymorphism showed statistically differences between groups (OR= 0.317, p= 0.01; OR= 0.688, p= 0.03, respectively). These results suggest that the C allele carriers have 1.44 less susceptibility to ACS. Also, the goodness of fit was performed to identify the most likely heritage model. In this context, significant differences were found in the recessive genetic model (OR: 0.333, p= 0.02). The TNF-α levels were significantly higher in ACS patients compared to HS (38.05 vs 23.92 ng/mL, p= 0.0001). The C/C genotype carriers showed lower levels of TNF-α when compared to heterozygous ACS patients; this finding could not be replicated in HS (36.95 vs 33.52 ng/mL, p=0.03). Conclusion: The TNF-α gene polymorphism rs1799964 (-1031T>C) is a susceptibility genetic marker for Acute Coronary Syndrome in Western Mexico population. In addition, serum levels of TNF-α may be a biological marker of ACS.

2048M


Background: Hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM) are genetically and phenotypically heterogeneous. NEXN, which encodes a component of the cardiac Z disc, has been previously reported to contain mutations in patients with HCM and DCM. The age of onset and disease severity associated with NEXN mutation has shown some inter-individual variability.

Methods: Whole exome DNA sequence analysis was performed on a small nuclear family with maternal left ventricular noncompaction (LVNC) cardiomyopathy and in-utero “dilated” cardiomyopathy with severe hydrops fetalis in two affected children. All 3 affected individuals also had a large ASD. Post mortem analysis on one of the children revealed LVNC cardiomyopathy. Interestingly the other affected child's cardiac studies normalised ex-utero.

Results: Standard clinical molecular analyses including array CGH, candidate nuclear gene analysis and selected mitochondrial genome mutation analysis had failed to identify a pathogenic variant. Exome capture was undertaken using the NimbleGen SeqCap EZ v2.0 library and sequenced using an Illumina HiSeq2000. The data were masked and analysed for 54 genes previously implicated in causing HCM or DCM or LVNC. Results: An infarct single amino acid deletion (p.Gly650del) in the last exon of the NEXN gene was identified in all three affected individuals. The p.Gly650del variant has been previously reported in adult patients with dilated cardiomyopathy; animal studies supported pathogenicity of this variant, however segregation studies were not reported. The extreme phenotypic diversity seen in our family and the normalisation of the cardiac studies in the proband remain unexplained. Further clinical evaluation of the maternal family is being planned.

Conclusion: Using an exome sequencing strategy we have successfully identified a likely pathogenic variant within the NEXN gene and provided evidence that LVNC is associated with variants in this gene. Further analysis is underway to examine the extreme phenotypic variability seen with the p.Gly650del variant and to determine whether this result can be used for prenatal diagnosis in this family.

2049S

Association of rs4340 ACE polymorphism with acute coronary syndrome in Mexican population. A. Valdez Haro1,2, Y. Valle1, J.I. García-González1,2, E. Sandoval-Pinto1,3, E. Valdés-Alvarado1,3, JF. Muñoz-Valle5, JR. Padilla-Gutiérrez3, 1) Instituto de Investigación en Ciencias Biomédicas, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, Mexico; 2) Doctorado en Genética Humana, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, 44350, Guadalajara, Jalisco, Mexico; 3) Doctorado en Biomedicinas, Centro Universitario de Ciencias de la Salud, UdeG. Sierra Mojada 950, 44350, Guadalajara, Jal., Mexico.

Introduction: Acute Coronary Syndrome (ACS) is an important public health problem and a main cause of death, in which the classic risk factors can barely explain half of the cases. A proposed explanation is that the sum of unfavorable gene polymorphisms and a proper environment may lead to the disease. Angiotensin converting enzyme (ACE) increase angiotensin II (AI1) plasmatic levels, which is a powerful vasoconstrictor that leads to endothelial and myocardial oxidative stress and atherosclerosis. The polymorphism (I/D) consists in an Insertion/deletion (I/D), in which the D allele has been related to increase ACE levels, therefore to risk for ACS. Aim: To determine the association of rs4340 polymorphism of ACE gene in ACS. Methods: 266 ACS patients and 147 healthy subjects (HS) matched by age were recruited from western Mexico. The ACS selection based on the American College of Cardiology criteria. An informed written consent was obtained from all patients. The rs4340 polymorphism was identified by PCR and PAGE with silver stain. Results: The genetic frequencies in healthy subjects were on Hardy-Weinberg equilibrium expectations (p=0.13). Allelic and genotype distributions were compared. Significant differences between groups were found. The allele D was 1.47 more frequent in ACS patients than HS subjects (OR= 1.47, p= 0.008). The genotype D/D was also more commonly found in the SCA group (OR = 1.86, p = 0.014). The recessive model supported this result. Conclusion: The D/D and DD genotype is a marker for susceptibility to ACS in Western Mexican population.
HLA-DRB1*01 allele associates with Acute Coronary Syndrome (ACS) in Finnish population. E. Vlachopoulos1, M. Marchesi1, J. Nokelainen1, M.S. Nieminen2, J. Sinisalo2, M-L. Lokki1, 1) Transplantation Laboratory, Haartman Institute, Haartmaninkatu 3, University of Helsinki, Helsinki, Finland; 2) HUCH Heart and Lung Center, Division of Cardiology, Helsinki University Central Hospital, Finland. Background: Acute Coronary Syndrome (ACS) is one of the leading causes of death in the world, still being a diagnostic and management challenge. Inflammation has a central role in the pathophysiology of ACS. However, the ultimate reasons for the inflammation remain unrevealed. The system that regulates inflammation and immunity lies in big part in the Human Leukocyte Antigens (HLA) system on chromosome 6p21.31. Our previous smaller candidate gene study suggested that HLA-DRB1*01 allele of MHC class II is associated with ACS. The aim of this study is to confirm this association in a larger material. Methods: We studied 2090 ACS patients and 1580 geographically-matched controls (Corogene study) having information on various risk factors and survival data. All subjects were genotyped for HLA-DRB1*01 allele to determine allele copy numbers with the genomic real-time quantitative polymerase chain reaction. The HLA-DRB1*01 was tested for association with ACS. Results: were replicated in two different replication materials from Finland. Results: HLA-DRB1*01 was associated with ACS (frequencies cases vs. controls 38.9% vs. 32.4%; odds ratio 1.33; P = 7.7×10-5). Interaction analysis showed that total cholesterol, low-density lipoprotein cholesterol, and triglyceride levels interacted with HLA-DRB1*01 (odds ratio 1.23 [95% CI 1.01-1.49]; P = 0.04, odds ratio 1.20 [95% CI 1.06-1.36]; P = 0.004, odds ratio 1.18 [95% CI 1.02-1.37]; P = 0.03, respectively). The survival curves stratified for HLA-DRB1*01 copies did not show any statistical difference. Conclusion: We showed that HLA-DRB1*01 is associated with ACS in Finnish materials. This gene may play a role in promoting ACS by affecting antigen presentation. Peptide fragments of oxidized low-density lipoprotein cholesterol may connect the cholesterol pathway to inflammation and immunology in atherosclerosis progression.

2053M
GWAS-identified loci for coronary heart disease are associated with intima-media thickness and plaque presence at the carotid artery bulb. M. den Hoed1,2, R. Strawbridge1,2, P. Almqvist3, S. Gustafsson1, T. Axelsson4, G. Engström4, U. de Faire4, B. Hedblad5, S.E. Humphries6, C.M. Lindgren7, A.P. Morris8,9, G. Östling3, A-C. Syvänen3,4, E. Tremoli10, A. Hamsten2, E. Ingelsson2, Ö. Melander2, L. Lind11, 1) Medical Sciences, Mol. Epidemiol, SciLifeLab, Uppsala University, Uppsala, Sweden; 2) Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital Solna, Stockholm, Sweden; 3) Department of Clinical Sciences, Diabetes and Endocrinology, Lund University and Lund University Diabetes Centre, Malmö, Sweden; 4) Department of Medical Sciences, SNP & SEQ Technology Platform, Uppsala University, Uppsala, Sweden; 5) Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; 6) Centre for Cardiovascular Genetics, University College London, London, UK; 7) Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; 8) Department of Biostatistics, University of Liverpool, Liverpool, UK; 9) Department of Medical Sciences, Molecular Medicine, Uppsala University, Uppsala, Sweden; 10) Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano & Centro Studi di Genetica delle Malattie Rare, IRCCS, Milan, Italy; 11) Department of Medical Sciences, Cardiovascular Epidemiology and EpiHealth, Uppsala University, Akademiska sjukhuset, Uppsala, Sweden.

Large-scale genome-wide association studies (GWAS) have so far identified 45 loci that are robustly associated with coronary heart disease (CHD) in data from adult men and women of European descent. We examined whether these loci are also associated with measures of atherosclerosis in vivo. Methods: Four large UK biobanks (IMPROVE, MDC-CC, ULSAM and PIVUS) provided data from up to 9,582 individuals of European ancestry. Forty-five SNPs representing the CHD-associated loci were genotyped in middle-aged to elderly individuals of European descent from four independent population-based studies (IMPROVE, MDC-CC, ULSAM and PIVUS). Intima-media thickness (IMT) was measured by external B-mode ultrasonography at the far wall of the bulb and common carotid artery. Plaque presence was defined as a maximal IMT of the bulb >1.5 mm. We meta-analysed single-SNP associations across the four studies, and combined them in a genetic predisposition score (GRS) using a score with equally weighted contributions from the 45 loci. Results: Higher values of the GRS were associated with a 0.24% increase in IMT at the far wall of the bulb (p=4.0×10-7), and with a 2.8% increased odds of plaque presence (p=7.4×10-6), independently of traditional risk factors. The genetic predisposition score was not associated with IMT of the common carotid artery (p=0.47). Our results suggest that the association between the 45 previously identified loci and CHD at least partly acts through atherosclerosis.
2053S
Transthetic replication of the gene-gene interaction of the prostaglandin E2 system in determining blood pressure reactivity. X. Kong1,2, Q. Zhao1, TN. Knelly1, C. Li1, D. Gai1, J. He1. 1) Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, US; 2) Department of Endocrinology, China-Japan Friendship Hospital, Beijing, China.

Increased blood pressure (BP) reactivity to cold stress has been suggested as a risk factor for hypertension and cardiovascular disease. Prostaglandin E2 (PGE2), the most abundant prostanooid in humans, plays a critical role in BP regulation. This study aimed to systematically investigate individual and interaction effects of genetic variants of the PGE2 system on BP reactivity to cold stress in both Chinese and Whites. The cold pressor test (CPT) was performed among 1,881 Han Chinese individuals from the Genetic Epidemiology Network of Salt Sensitivity (GenSalt) study. BP reactivity variables were calculated as the changes between the highest BP levels during the CPT and those prior to the CPT. A total of 397 single nucleotide polymorphisms (SNPs) covering 10 genes from the PGE2 system, including synthases (PLA2G4A, PTGS1, PTGS2, PTGES, PTGES3) and receptors (PTGER1, PTGER2, PTGER3, PTGER4) of the PGE2, were genotyped for the analysis. In addition to single marker analyses, gene-based and gene-gene interaction analyses were conducted using the truncated P value method. Results were further replicated in 1,448 participants of European ancestry from the Coronary Artery Risk Development in Young Adults (CARDIA) study, in which a similar CPT was conducted. SNPs from the PTGER3 and PTGER4 genes were significantly associated with BP reactivity variables of diastolic BP (DBP) or mean arterial pressure (MAP) in both Han Chinese and Whites. In gene-based analyses, PTGER4 was associated with DBP and MAP responses in both Han Chinese (both P < 1.0×10−5) and Whites (P = 0.029 and 0.038). In addition, gene-gene interactions in PTGER2 were associated with BP reactivity variables in Han Chinese (both P < 1.0×10−5). For example, PLA2G4A and PTGES2 genes showed the strongest interaction for systolic BP and MAP responses in both Han Chinese (both P < 1.0×10−5) and Whites (both P < 1.0×10−5). Transthetic replication of genetic association of genetic variation in the PGE2 system in affecting BP reactivity to stress.

2054M
Role of common sarcomeric gene polymorphisms in genetic susceptibility to left ventricular dysfunction. B. Mittal1, S. Kumar1, A. Mishra1, A. Srivastava1, N. Garg2, S. Agarwal2, S. Pandey2. 1) Genetics, SGPGIMS, Lucknow, U.P, India; 2) Cardiology, SGPGIMS, Lucknow, UP, India; 3) CVTS, SGPGIMS, Lucknow, U.P, India.

Background: Mutations in sarcomeric genes are common genetic cause of cardiomyopathies. An intronic 25-bp deletion in MYBPC3 at 3’ region is associated with dilated (DCM) and hypertrophic (HCM) cardiomyopathies in Southeast Asia. However, the frequency of sarcomeric gene polymorphisms and associated clinical presentation has not been established with left ventricular dysfunction (LVD). Therefore, the aim of the present study was to explore the association of MYBPC3 25 bp deletion, TTN 18 bp I/D, TNNT2 5bp I/D, and Myospryn K296N polymorphisms with LVD. Methods: The study included 988 consecutive patients with angiographically confirmed CAD and 300 healthy controls. Among 988 CAD patients, 253 with reduced left ventricle ejection fraction (LVEF<45%) were categorized as LVD. MYBPC3 25 bp deletion, TTN 18 bp I/D and TNNT2 5bp I/D polymorphisms were determined by direct polymerase chain reaction (PCR) method while Myospryn K296N polymorphism by TaqMan assay. Results: Our results showed that MYBPC3 25 bp deletion polymorphism was significantly associated with elevated risk of LVD (healthy controls vs LVD: OR=3.85, p-value<0.001; and Non-LVD vs LVD: OR=1.65, p-value=0.035) while TTN 18 bp I/D, TNNT2 5bp I/D, and Myospryn K296N polymorphisms did not show any significant association with LVD. The results also showed that MYBPC3 25 bp deletion polymorphism was significantly associated with other parameters, such as LV remodeling, i.e. LV dimensions (LV end diastolic dimension, LVESD; p-value=0.037, and LV end systolic dimension, LVEDS; p-value=0.032). Conclusion: Our data suggests that MYBPC3 25 bp deletion may play significant role in conferring LVD risk in Southeast Asian populations. Financial support from DBT, Government of India.

2055S
Identification of TEMEM241 as the underlying gene in the chromosome 18q11.2 triglyceride region in Mexicans. A. Rodriguez1, L. Gonzalez2, Y. Shagaghi1, E. Nikkola1, T. Tusie-Luna3,4, C.A Aguilar-Salinas2, P. Raju-Kanta1,2. 1) Human Genetics, UCLA, Los Angeles, CA; 2) Molecular Biology Institute at UCLA, Los Angeles, USA; 3) Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubiran, Mexico City, Mexico; 4) Instituto de Investigaciones Biomedicas (IIBN, UNAM, Mexico City, Mexico.

High serum triglyceride (TG) levels are a critical risk factor for cardiovascular disease (CVD). We recently identified a locus on chromosome 18q11.2 to be associated with high serum TGs in Mexicans in a genome-wide association study (GWAS). However, the biological mechanism(s) underlying this signal has not been elucidated yet. Characterization of GWAS hits is an important area of research in the field of functional genomics. The 18q11.2 locus in particular could have implications in the treatment and diagnosis of CVD in the rapidly growing Hispanic populations. To characterize the biological mechanism, we performed a cis-eQTL analysis utilizing 856 publicly available human adipose, skin, and lymphocyte RNA microarrays of the MuThER resource, and discovered that the lead SNP rs9949617 is a genome-wide significant cis-eQTL, regulating the expression of one of the 5 regional genes on 18q11.2, the transmembrane protein 241 (TMEM241) gene. However, lead GWAS SNPs are often not the underlying causal variant. Therefore, to identify the causal variant, we performed a regional linkage disequilibrium (LD) analysis using PLINK and found a total of 9 single nucleotide polymorphisms (SNPs) in high LD (R2>0.8) with rs9949617, suggesting that anyone of these SNPs can be the underlying susceptibility variant. To prioritize the variants for functional analysis, we filtered the variants through systematic data mining including ENCODE ChIP-seq data and leveraging the GWAS-identified SNPs (rs17259126 and rs17259127) to narrow down the candidate gene. TMEM241 was prioritized through systematic data mining including ENCODE ChIP-seq data and leveraging the GWAS-identified SNPs (rs17259126 and rs17259127) to narrow down the candidate gene. TMEM241 was identified as a potential candidate gene underlying the TG phenotypes. To prove our hypothesis, we cloned a 500-bp sequence surrounding the SNP upstream of a minimal promoter in the pgSL4.23 vector (Promega) and transfected the firefly/luciferase reporters via viro-transfection into human hepatocellular carcinoma cells (HepG2). Luciferase assays at 48 hrs post transfection confirmed that the A allele in rs17259126 has an increased reporter expression (p<0.05) compared to the reference G allele in 3 biological replicates. These results suggest that TMEM241 is associated with high serum TG levels and the TG GWAS signal on chromosome 18q11.2 in Amerindian origin populations.

2056M
Contribution of genetic variation of ATP-binding cassette transporter ABCA1 to the regulation of plasma lipid/lipoprotein levels in US Non-Hispanic Whites. F.Y. Demirici1, V. Nieminski1, X. Wang1, M.M. Barrmada1, J.E. Hokanson2, R.F. Hamman1, M.I. Kamboh1. 1) Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA; 2) Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO.

Abnormalities in plasma lipid/lipoprotein levels have been linked to the risk of coronary heart disease (CHD). ABCA1 (ATP-binding cassette, subfamily A (ABC1), member 1) encodes a member of the ATP-binding cassette transporters superfamily that functions as a cholesterol efflux pump and ABCA1 mutations are known to cause hereditary high-density lipoprotein (HDL) deficiency. In this study, we sequenced ABCA1 (exons, most introns, and ~1 kb of each of 5’ and 3’ flanking regions) in 95 US Non-Hispanic Whites (NHWs) with extreme HDL cholesterol levels using the Sanger method and identified 404 variants (402 bi-allelic and 2 tri-allelic). A total of 237 bi-allelic variants, including those identified by sequencing and tagSNPs with MAF≥5% and selected uncommon/rare variants plus additional tagSNPs from the HapMap project, were genotyped in our entire NHW sample comprising 623 individuals. The resulting quality control (QC)-passed genotyping data (including 183 variants; 117 common and 66 uncommon/rare) were then evaluated for associations with 5 major lipid/lipoprotein traits (HDL-C, low-density lipoprotein cholesterol (LDL-C), triglycerides (TG), apolipoprotein (apo) A1, and apoB levels). In single-site analysis, 35 common variants showed significant association (p<0.05) with at least one lipid/lipoprotein trait. Although significant associations were observed with all but at least one lipid/lipoprotein trait (for at least 2 lipids/lipoproteins), the strongest and highest number of associations (21 of 35 significant SNPs) were detected with TG levels. The rare variant and haplotype analyses are currently underway. Our preliminary findings suggest a role for common ABCA1 variants in modulation of plasma lipoproteins (especially the apoB levels), which is in line with the recent studies implicating ABCA1 in the metabolism of various major plasma lipoproteins as a contributor to the metabolic link between them.
2057S
A high yield of variants with a putative role as modifiers in patients with hypertrophic cardiomyopathy, S. Bard1, F. Girolami1, M. Benelli1, B. Tomberli2, E. Conti1, G. Marsaglia1, C. Pesucchini1, G. Castelli1, A. Fornaro2, F. Cechini2, I. Olivotti2, F. Torricelli1. 1) Diagnostic Genetic Unit, Careggi University Hospital, Florence, Italy; 2) Reference Centre for Myocardial Diseases, Careggi University Hospital, Florence, Italy; 3) University of Florence, Department of Clinical and Experimental Medicine, Florence, Italy.

Next Generation Sequencing enables simultaneous screening of multiple genes for multiple patients in a single run. We designed a panel of 111 genes known to be associated to CMs to study 94 unrelated patients (80 with Hypertrophic Cardiomyopathy, HCM; 18 with Dilatative Cardiomyopathy, DCM and 6 with Arrhythmogenic Cardiomyopathy, AC). Targeted resequencing was performed on Illumina platform (98.13% of the regions with a depth of coverage of 20X or more, mean coverage on target of 530X). A mean of 1016 variants were found for each patient. Rare (frequency <0.05), non-synonymous, loss-of-function and splice-site variants were defined as candidates. Pathogenic or likely-pathogenic variants were all confirmed by Sanger and cosegregation was tested when possible. Excluding titin missense variants, we identified 48 variants (27 novel) in sarcomeric or associated genes in 48/70 HCM patients (68%), with 14% of complex genotype. MYH7, MYBPC3 and TNNI3 resulted the high-yield genes; 19 additional candidate variants (13 novel) in desmosomal and ion-channel genes in 14 patients (20%) were identified in this group. We identified 10 candidate variants (7 novel) in 7/18 DCM patients (39%) and 5 candidate variants in 3/6 AC patients (50%). A targeted protocol allowed the identification of likely pathogenic variants in a large proportion of patients with CMs, irrespective of phenotype. The unexpected finding of rare non-synonymous variants in desmosomal and ion-channel genes among HCM patients raises important issues regarding their role as previously unappreciated modifiers of the disease, potentially relevant to risk prediction and counseling.

2058M
Genetic dissection of a novel X-linked congenital heart syndrome. C. Preuss1, S. Yang2, M. Samuels3, P. Awadalla3, P. Chetelat2, H. Björkås1, S. Mohamed2, P. Eriksson2, A. Andelfinger1. 1) Department of Pediatrics, Centre Hospitalier Universitaire Sainte Justine, Montreal, Quebec, Canada; 2) Department of Cardiology, Nanjing Children’s Hospital, Nanjing Medical University, Nanjing, China; 3) Department of Pediatrics, Faculty of Medicine, Sainte-Justine Research Center, University of Montreal, Montreal, QC, Canada; 4) Cardiology Service, Centre Mére-Enfants, Centre Hospitalier Universitaire de Quebec, Université de Laval, Québec City, Quebec, Canada; 5) Atherosclerosis Research Unit, Center for Molecular Medicine, Department of Medicine, Karolinska Institutet, Stockholm, Sweden; 6) Department of Cardio and Thoracic Vascular Surgery, University Clinic of Schleswig-Holstein, Luebeck, Germany.

The major burden of congenital heart disease (CHD) is caused by left ventricular outflow tract obstructions (LVOTO). This heterogeneous group of cardiac malformations constitutes to an increased risk of valve replacements and aortic aneurysms. Although a strong male predominance has been described for the highly heritable trait, the genetic basis for these cardiac malformations remains poorly understood. Here, we describe a comprehensive genetic study in two French Canadian pedigrees with 15 affected male individuals with septal defects, aortic valve lesions and bicuspid aortic valve. Linkage analysis and genetic fine mapping using the Illumina Omni 5.0 platform revealed a significant interval on chromosome Xq28 (LOD score = 3.29) harboring a 260kb haplotype co-segregating with disease in both families. The disease associated haplotype is rare (MAF < 5%) among the aortic media of dilated aortas compared to all expressed transcripts in affected male individuals with septal defects, aortic valve lesions and bicuspid aortic valve. Linkage analysis and genetic fine mapping using the Illumina Omni 5.0 platform revealed a significant interval on chromosome Xq28 (LOD score = 3.29) harboring a 260kb haplotype co-segregating with disease in both families. The disease associated haplotype is rare (MAF < 5%) among 960 genotyped French Canadian controls. Whole-exome sequencing of six patients and re-sequencing of the disease haplotype did not reveal rare (MAF < 2%) coding or splice-site mutations. Validation of all identified rare and novel mutations in the disease haplotype using a Sanger re-sequencing panel confirmed only one mutation co-segregating with disease in an X-linked manner. This highly conserved intronic mutation (GERP score > 4) in FLNA is absent among public (dbSNP138, 1000 Genomes) and in-house sequencing data sets. In silico analysis predicts the hemizygous mutation to create a high-affinity MyoD binding site. Expression data from the aortic valve revealed that FLNA shows the highest expression patterns in the aortic media of dilated aortas compared to all expressed transcripts in the disease associated haplotype. Reconstruction of ascending genealogies supports the notion of a founder effect for this novel X-linked syndrome, dating back to a common founder couple in 1788.

2059S
Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. A. Ganna1, S. Salihić2, J. Sundström1, C.D. Broeckling3, A.K. Hedman4, P.K.E. Magnusson1, N.L. Pedersen1, A. Larsson1, A. Siegbahn1, M. Zilmer5,7, J. Prenti1,8, J. Amlöv9, L. Lind2, T. Fall1, E. Ingelsson1,2,10. 1) Medical epidemiology and biostatistics, Karolinska institutet, Stockholm, Sweden; 2) Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden; 3) Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, Colorado, USA; 4) Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 5) Department of Biochemistry, The Centre for Excellence for Translational Medicine, University of Tartu, Tartu, Estonia; 6) Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA; 9) School of Health and Social Studies, Dalarna University, Falun, Sweden; 10) Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, United Kingdom.

Background: Analyses of circulating metabolites in large prospective epidemiological studies could lead to improved prediction and better biological understanding of coronary heart disease (CHD).

Methods and Results: We performed a mass spectrometry-based non-targeted metabolomics study for association with incident CHD events in 1,028 individuals (131 events; 10 y. median follow-up) with validation in 1,670 individuals (231 events; 4 y. median follow-up). Four metabolites were robustly replicated and independent of established risk factors (lysophosphatidylcholine 18:1 (hazard ratio [HR] per standard deviation [SD] increase = 0.77, P-value<0.001), lysophosphatidylcholine 18:2 (HR=0.82, P-value=0.002), monoglyceride 18:2 (MG 18:2; HR=1.16, P-value=0.032) and sphingomyelin 28:1 (HR=0.86, P-value=0.010)). Together they contributed to moderate improvements in discrimination and re-classification in addition to traditional risk factors (C-statistic: 0.76 vs. 0.75; NRI: 9.2%). MG 18:2 was associated with CHD independently of triglycerides. Lysophosphatidylcholines were negatively associated with BMI, glucose, C-reactive protein and triglycerides, whereas a positive association was found with HDL,-weekly physical activity and with less evidence of subclinical cardiovascular disease in 970 PIVUS participants; a reverse pattern was observed for MG 18:2. MG 18:2 showed an enrichment (P-value=0.002) of significant associations with CHD-associated SNPs (P-value=1.2x10^-4 for association with rs964184 in the ZNF259/ABCA1 region and a weak enrichment of P-value=0.05 per SD increment in MG 18:2. P-value=0.05) on CHD, as suggested by Mendelian randomization analysis.

Conclusions: We identified four lipid-related metabolites with evidence for clinical utility, as well as a causal role in CHD development.

2060M
Genetic and Metabolic Causes of Neonatal Cardiomyopathy. C. Prada1,4, I. Villamizar-Schiller4, J. Castro4, L. Pabon4, A. Duran4. 1) Centro de Medicina Genómica y Metabolismo, Cardiovascular Foundation of Colombia, Bogotá, Colombia; 2) Dept Pediatrics Genetics, Cincinnati Children's Hosp Med Ctr, Cincinnati, OH; 3) Division de Cardiologia Pediatrica, Cardiovascular Foundation of Colombia, Floridablanca, Colombia.

Background: Neonatal cardiomyopathy is an important cause of pediatric deaths. This is a heterogeneous disease with a strong genetic component. The objective of this study was to determine the prevalence of genetic and metabolic etiologies in patients with neonatal cardiomyopathy.

Study Design: A retrospective analysis of data from clinical records of 40 neonates with cardiomyopathy seen at the Cardiovascular Foundation of Colombia between 2011 and 2014 was performed. Results: In a total of 29 neonates (72.5%) a genetic or metabolic etiology was identified. The most common group of causes was metabolic (13/29, 45%) followed by syndromic (10/29, 34%), and maternal diabetes (6/29, 21%). The single most common diagnosis were RASopathies (7/25, 28%) followed by maternal diabetes, mitochondrial and fatty acid oxidation defects. Mortality was higher in the group of metabolic disorders (7/13, 54%). Hypertrophic cardiomyopathy was identified in 2/29 (6.9%). The 30-day survival rate was 90% of the patients with cardiomyopathy (24/40) survived beyond their first year of life. Extra-cardiac manifestations were more prevalent in the syndromic group. Family history of a sibling with unexplained death was seen in the metabolic group. Conclusions: Neonatal cardiomyopathy has a strong genetic component in the majority of patients. An underlying metabolic or syndrome cause was identified in 72.5% of neonates. Identification of etiology is important for management and family counseling and recurrence.
2061S Targeted Oligonucleotide-Selective Sequencing for Genetic Diagnostics of Pulmonary Arterial Hypertension. E.H. Seppälä1, S. Vattulainen1, J. Ahokas2, J. Talilla1, M. Gentile1, M. Sankklo1, T. Laitinen1, J.W. Koskenvuo1,2, T.-P. Alastalo1,2, S. Myllykangas1,2. 1Blueprint Genetics, Helsinki, Finland; 2Childrens Hospital Helsinki, University of Helsinki, Helsinki, Finland; 3Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; 4Department of Internal Medicine, Tampere University Hospital, Tampere, Finland; 5Division of Medicine, Department of Pulmonary Diseases and Allergology, Turku University Hospital and University of Turku, Turku, Finland; 6Institute of Biomedicine, University of Helsinki, Helsinki, Finland.

Pulmonary arterial hypertension (PAH) is a severe and progressive disease. The genetic basis of idiopathic PAH is well recognized but rarely utilized in a diagnostic setting. Hundreds of mutations in seven genes have been reported to associate with idiopathic and familial forms of PAH. Next-generation sequencing for comprehensive genetic testing improves diagnostics, prognostics and treatment optimization of the index patient and allows an effective screening of asymptomatic family members. We utilized novel next-generation sequencing and bioinformatics approaches for genetic diagnostics of PAH. We applied Oligonucleotide-Selective Sequencing (OS-Seq) and custom data analysis and interpretation pipelines to identify pathogenic base substitutions, insertions and deletions in seven genes associated with PAH (BMPR2, BMPR1B, ACVRL1, ENG, SMAD9, CAV1 and exon 2 of KCNK3). Targeted sequencing covered all coding exons, exon-intron boundaries and known intronic mutations (12,638 bases). In OS-Seq, oligonucleo-tide-functionalized illumina flow cells are used for both capture and sequencing of DNA. We automated target DNA capture and sequencing using the Illumina NextSeq500 system and the Oxford Nanopore MinION platform. We analyzed 21 Finnish PAH patients. Sequencing depth was at least 30-fold. We identified 18 different mutations causing stop, frameshift and missense changes and had minor allele frequency of <0.001 reported in the 1000 genomes database. Mutations or potential variants were not identified from 6 other analyzed genes. Our results demonstrate that a comprehensive next-generation sequencing approach is an effective tool for genetic diagnostics of PAH and 29% of the Finnish PAH patients carried mutations in BMPR2.

2062M Use of a Gene Expression Score in a Primary Care Setting to Evaluate African American Patients Presenting with Symptoms Suggestive of Obstructive Coronary Artery Disease. L. Wilson1, M. Brown2, B. Rhees2, D. Smith2, B. Rhodes3, 1CardioDX, Inc., Redwood City, CA; 2Providence - Dayton Primary Care, Dayton, OH; 3Novant Heart and Vascular Institute, Huntersville, NC.

Purpose: Approximately 3 million adults without diabetes present annually to primary care clinicians with symptoms suggestive of obstructive coronary artery disease (CAD). After a detailed examination, physicians still rely heavily on advanced imaging diagnostic tests, to determine the etiology of symptoms. Given the heterogeneity in clinical manifestations of CAD amongst different ethnic populations in the US, the use of a gene expression score (GES) to identify CAD in African Americans could provide a useful tool to assess the spectrum of primary care pts, thereby appropriately avoiding unnecessary referrals and advanced diagnostic tests. We hypothesized that use of the GES would show clinical utility for a clinician evaluating African American pts presenting with symptoms suggestive of CAD:

Methods: A previously validated gene expression diagnostic test (Corns® CAD, CardioDX, Inc.) has a 96% NPV in ruling out obstructive CAD among symptomatic pts with no previous history of diabetes or myocardial infarction. GES results are predefined as low (GES ≤15) or elevated (GES >15), with low scores having a low likelihood of obstructive CAD. Previous evaluation showed no significant difference in test performance between non-white and white populations. This was a single primary care practice study with a large African American pt population. Approximately 15% of pts from 2011-2013 were evaluated. Results: Of 1062 pts included 852 African American pts who were received a GES, with 325 (56%) female pts and 201 (35%) being >65yrs. Approximately 90% (518/582) of pts presented with typical or atypical symptoms suggestive of obstructive CAD. Mean GES was 19 (range, 1-40), and 245 (23%) pts (18%) had low scores. In this population, 32% of pts with low scores were referred to cardiology and/or further diagnostic testing, whereas 248/337 (74%) of elevated GES pts were referred to cardiology and/or further diagnostic testing (p-value<0.0001). Conclusions: The personalized GES should be utilized by primary care clinicians to evaluate African American pts presenting with symptoms suggestive of CAD. Given the lack of comprehensive approaches for extended genetic diagnosis of FH which can increase the yield of FH diagnosis and improve phenotype-genotype correlation studies in pedigrees with phenocopy or non-penetrant FH mutations.

2063S Extended genetic diagnosis of Familial Hypercholesterolemia using next-generation sequencing. M.M. Motazacker1, B. Sjouke2, O.R.F. Mook3, M.A. Haagmans1, G.K. Hovingh1, J.C. Defesche1, A.R. Mensen-kamp1, M.M.A.M. Mannens1. 1Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands; 2Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands; 3Department of Experimental Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands; 4Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands.

Familial Hypercholesterolemia (FH) is a major risk factor for coronary artery disease. FH is caused by mutations in the low-density lipoprotein receptor (LDLR), apolipoprotein B (APOB) and proprotein convertase subtilisin/kexin 9 (PCSK9). Routine genetic diagnosis of FH is often limited to sequencing LDLR followed by partial sequencing of APOB and PCSK9 in cases with no LDLR mutations. This is mainly due to the large size of APOB and rarity of PCSK9 mutations which makes Sanger sequencing inefficient. Using DNA from 20 patients with 31 known mutations (including single-nucleotide coding and promoter variants, insertions, deletions, indels and large copy-number variants), we employed an Ion AmpliSeq™ enrichment of coding sequence (and 25 flanking intronic regions) of all three FH genes followed by sequencing using Ion-PGM™ sequencing platform (Life technologies™). The sequence data were then analyzed using SeqNext software (JSI medical systems). We could successfully identify all previously detected mutations. Interestingly, we also identified additional 8 rare variants including PCSK9, c.6679X and 7 APOB variants (e.g. p.Arg523Trp, p.Asp1113His, p.Lys3076Met, etc.), some of which already reported in the literature to be functional while others predicted to be functionally neutral. We also identified a non-synonymous variant in the histidine-rich C-terminal domain of the APOB gene may help in explaining unexpected phenotypes (which are usually justified by incomplete penetrance of the phenotype) seen frequently in the families with FH. Our study suggests a fast, cost-effective and accurate approach for extended genetic diagnosis of FH which can increase the yield of FH diagnosis and improve phenotype-genotype correlation studies in pedigrees with phenocopy or non-penetrant FH mutations.

2064M Evidence for the novel variant, c.1937 C>T (p.Ser464Phe) in the membrane binding domain of the ANK2 gene contributing to Long QT syndrome in a First Nations community of Northern British Columbia. L.T Arbour1, A.K.J. Boyce2, J. Christensens2, S. MacIntosh3, S. Lauzon3, S. Tungt4, C. Karr4, L.A. Swaynes5, 1Department of Medical Genetics, University of British Columbia, Victoria, BC; 2Division of Medical Sciences, University of Victoria, Victoria, BC; 3Island Health and the BC Inherited Arrhythmia Program, Victoria BC; 4Department of Medicine, University of British Columbia, Vancouver, BC.

Long QT syndrome (LQTS) is a rare inherited cardiac condition named for a prolonged QT interval on ECG (corrected for rate) conferring susceptibility for life-threatening ventricular arrhythmias. An estimated 75 percent of hereditary LQTS is caused by mutations on channels or functionally related proteins. In Northern British Columbia, through participation methods, a First Nations community was previously identified to have a disproportionately high rate of LQTS, accounting for largely by a mutation in the ANK2 gene. This is the first description of a novel missense founder mutation (p.V205M) in the ANK2 gene (NM_001148.4) resulting in a serine to phenalanine substitution at amino acid 205, which is predicted to have a significantly reduced affinity for the membrane binding domain of the ANK2 gene contributing to Long QT syndrome. Given the heterogeneous clinical manifestations of LQTS, this evidence supports pathogenicity of the identified mutation.
2065S

Novel variants in VINCULIN and TROPOMYOSIN1 combinatorially predispose patients to dilated cardiomyopathy. D.C. Deacon1,2,3,5, A.M. Manso1,2,3, B.C. Nelson1,2, R.S. Ross1,2,4, E.D. Adler1, N.C. Chi1,2,3. 1) School of Medicine, University of California San Diego, La Jolla, CA; 2) Biomedical Sciences Graduate Program, UCSD; 3) Medical Scientist Training Program, UCSD; 4) Veterans Affairs San Diego Healthcare System; 5) California Institute of Regenerative Medicine Fellow.

Dilated cardiomyopathy (DCM) may affect as many as 1 in 250 individuals and is the leading cause of heart failure necessitating heart transplant. We have identified a large family presenting with multiple cases of DCM across four generations. The proband for this study, a 14-year-old male patient, presented with severe heart failure and required heart transplantation. His father was also diagnosed with DCM while his sister, who had a normal heart by echocardiography, died suddenly and, on autopsy, showed mild cardiac hypertrophy. We have identified sequence variants in the genes encoding the costameric protein VINCULIN (VCL) and the sarcomeric regulatory protein TROPOMYSIN1 (TPM1) in this family. Targeted sequencing at these loci in 31 family members showed that the combination of VCL and TPM1 variants cosegregated with all family members diagnosed with cardiomyopathy. Given the genetic cosegregation of these novel heterozygous variants in VCL and TPM1, we hypothesize that the functional interaction between costameric and sarcomeric structures is critical to the proper regulation of cardiomyocyte force generation and transmission. In order to study the effects of these variants on cardiac structure and function, we have utilized CRISPR-mediated genomic editing to introduce these patient-specific variants into a mouse model. Both of these genes are over 99% conserved between human and mouse at the protein level and mutations in both have been previously associated with cardiomyopathies, though the molecular mechanisms governing these pathophysiology have not been fully described. We will determine the consequences of these variants on cardiac contractility and sarcomeric and cytoskeletal organization in single and double variant heterozygous mice to elucidate the mechanisms underlying the observed combinatorial disease inheritance pattern. A further understanding of the genetic and molecular interactions between costameric and sarcomeric proteins in proper cardiac function could pave the way for new therapies to treat heart failure.

2066M

A novel mutation in the RYR2 gene (c.527 G>T, p.R176L) identified in a 4 generation family presents with a catecholaminergic polymorphic ventricular tachycardia (CPVT) phenotype with variable penetrance. S. Lauson1,2, B. Sinclair1,2, A.A. Collier4,5, P. Curtis4,5, F. Van Petegem6, B.A. Fernandez4,5, A.E. Williams7, S. Connors2, C.G. Templeton7, K. Hodgkinson5,6, R. Leather8, S. Sanatan9,10, L. Arbour8,9. 1) Medical Genetics, Island Health, Victoria, BC, Canada; 2) BC Inherited Arrhythmia Program, Victoria/Vancouver, BC; 3) Pediatric Cardiology, Island Health, Victoria, BC; 4) Provincial Medical Genetics Program, Eastern Health, St. John's, NL; 5) Discipline of Genetics and Medicine, Memorial University of Newfoundland, St. John’s, NL; 6) Department of Biochemistry, University of British Columbia, Vancouver, BC; 7) Division of Cardiology, Discipline of Medicine, Faculty of Medicine, Memorial University of Newfoundland, St. John’s, NL; 8) Division of Cardiology, Island Health, Western Cardiology, Victoria, BC; 9) Clinical Epidemiology, Discipline of Medicine, Memorial University, St. John’s, NL; 10) Department of Pediatrics, Children & Women’s Hospital, University of British Columbia, Vancouver, BC; 11) Department of Medical Genetics, University of British Columbia, Vancouver BC.

Catecholaminergic Polymorphic Ventricular Tachycardia (CPVT) is a rare condition with population incidence estimated to be 1:10,000 and confers substantive risk for sudden cardiac death in the young. Four known genes cause the condition (RYR2, CASQ2, TRDN, and CALM1) but the identification of large extended families are rare. Stress and exercise-induced releases of catecholamines can lead to aberrant calcium ion flux, and the hallmark feature of bidirectional ventricular tachycardia (VT). Other inherited arrhythmia conditions can mimic the presentation, including Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC) and Long QT syndrome (LQTS). This family also has a known pathogenic LQTS mutation, TMEM43 (p.S358L). Subsequent comprehensive testing for ARVC included the RYR2 gene, revealing a novel RYR2 gene variant (c.527 G>T, p.R176L). Confirmation of bidirectional VT in her son suggested the CPVT phenotype. He had no evidence of ARVC. Of interest, this family also has a known pathogenic LQTS mutation (R518Q in KCNQ1), and two siblings are heterozygous for both. The R176L variant was not seen in 600 alleles from controls, and results in a non-conservative amino acid substitution of a polar Arginine with a non-polar Leucine. Lee and colleagues have noted this region of conservation (515–526) in the RyR2 Cterminus. These interactions stabilize the RyR2 closed state, and any weakening through R176L would result in facilitated channel opening, leading to premature or prolonged release of calcium ions. Currently 16 individuals in 4 generations have been identified with the mutation. Cardiac arrest during surgery (age 29), was documented in a variant carrier (cousin to the proband), and bidirectional VT has been confirmed in another during exercise testing supporting clinically that the phenotype segregates with the mutation alone. However, the carrier father of the proband (age 76) remains apparently unaffected. ECGs, echocardiograms, Holter monitors and stress tests are being carried out on family members. This family provides a unique opportunity to explore the variable penetrance and expression of an RYR2 mutation causing CPVT.
2067S Harnessing genomic data to identify drug targets for reduction of LDL cholesterol and CAD risk that do not impact upon glycaemic status. V. Tragante do O1, F. W. Asselsberg1, M. T. V. Holmes2, D. Divie Hart & Longen, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, Netherlands; 2) Division of Transplant Surgery, Perelman School of Medicine, University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104, USA.

Introduction LDL cholesterol (LDL-C) reduction is recognized as an efficacious therapeutic means to reduce risk of coronary artery disease (CAD). However, the most widely prescribed LDL-reducing drugs (statins) are associated with an increased risk of type 2 diabetes (T2D). We used data from genome-wide association studies for CAD and glycemic traits to identify potential drug targets for LDL-C reduction that have no impact on glycemic status. Methods and Results We obtained summary-level results from a GWAS on LDL cholesterol from Global Lipids Genetics Consortium (GLGC), CAD/MI from CARDIoGRAMplusC4D, T2D from DIAGRAM consortium and glucose, insulin and related traits from MAGIC consortium, and standardized all effects to unit of increase in LDL-C. A p-value filter of P<5E-08 for GLGC results was employed, and the SNPs surpassing this threshold were further filtered with P<0.05 for CARDIoGRAMplusC4D and DIAGRAM. Directions of effect were the same for GLGC and CARDIoGRAM for 81 out of 84 SNPs passing the significance threshold (binomial P=9.3E-21), whereas all 17 SNPs passing significance thresholds for DIAGRAM and MAGIC had the same direction of effect (binomial P=7.63E-08). We then performed a multiple-trait meta-analysis of all MAGIC subcomponents (referred to as “metabolic burden”) and searched for SNPs that are significant for LDL and CAD/MI, but did not associate with metabolic burden (P>0.05). Eighteen loci met these criteria, and were further investigated for their potential druggability. Six of those present interaction with one or more commercially available drugs, including LDLR, HNF1A, OASL, PLG, PTPN11 and SLC22A3. Conclusions These findings provide novel information for prioritizing therapeutic targets for reduction of LDL-C and CAD risk that should be free from adverse consequences on glycaemic status. Keywords Coronary artery disease, type 2 diabetes, LDL cholesterol, drug target.

2068M From Identification of Differing TIE2 Mutations with Distinct Cellular Characteristics in Four Types of Venous Anomalies towards a Murine Model and a Therapeutic Pilot Study. M. VIKKULA1, N. LIMAYE1, J. SOBELT1, M. UEBELHOER2, M. NETYKNIK2, E. BOSCOLO2, L. EKLUIND3, J. BISCHOFF1, L.M. BOON1,4, 1) Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Oulu Center for Cell-Matrix Research, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, Finland; 3) Vascular Biology Program and Department of Surgery, Boston Children’s Hospital, Harvard Medical School, MA, USA; 4) Center for Vascular Anomalies, Division of Plastic Surgery, Cliniques universitaires Saint Luc, Brussels, Belgium.

Venous anomalies are composed of ectatic veins with irregular smooth muscle covering. They are commonly cutaneous. They usually occur as a single lesion without family history (sporadic Venous Malformation, VM). Some sporadic patients have multifocal lesions (Multifocal Sporadic Venous Malformation, MSVM). In the sporadic Blue Rubber Bleb Nevus syndrome (BRBN), patients also have multifocal lesions; pathogenic are rubby/palmpiont lesions and those located in the Gl tract. In rare cases, venous malformations are multifocal because of autosomal dominant inheritance (Mucocutaneous Venous Malformation, VMC). VMs progressively expand causing functional and cosmetic alteration. Furthermore, diffuse, telangiectatic or nodular lesions that are indistinguishable from telangiectasias or nodules. The capacity to form lesions clearly resides in mutant endothelial cells, which, when injected into immunodeficient mice generate lesions mimicking human VM. Interestingly, an mTOR inhibitor is able to delay lesion development. Finally, we conducted a pilot study comprising five patients with VMs refractory to standard-of-care, an mTOR inhibitor diminished pain, intravascular coagulopathy and improved quality of life.

2069S A novel pathway involved in the susceptibility of non-alcoholic fatty liver diseases. S. Makishima, S. Boonvisut, K. Watanabe, K. Nakayama, S. Iwamoto, Division of Human Genetics, Center for Molecular and Medical University, Shimotsuke, Tochigi, Japan.

Background. Mammalian target of rapamycin 1 (TRIB1) is a locus that has convincing impact on cardiovascular diseases and levels of plasma triglyceride (TG) and LDL-cholesterol across several ethnic groups. The genetic associations have been shown in both genome-wide SNP association studies (GWAS) and the replication studies. In addition, we recently showed a deep association of TRIB1 SNP with non-alcoholic fatty liver diseases (NAFLD) in Japanese (P = 9.3E-7). The risk allele of NAFLD decreased transactivation activity in reporter gene assay. Furthermore, knockdown of TRIB1 expression in mouse liver increased plasma and hepatic lipid levels, whereas the over expression decreased them. Although the enhanced lipogenesis in mouse liver was estimated to result from the reduced decay of carbohydrate response element binding protein (ChREBP), molecular pathways for the hepatic lipid accumulation were still uncovered. Methods. The novel molecular targets of TRIB1 were explored using Yeast two-hybrid system. The positive clones were further screened functionally using shRNA template via adenoviral gene transduction system into mice liver. Plasma and hepatic lipid levels of the mice were measured to identify the genes involved in the hepatic lipid accumulation. Results and discussion. Nine cDNA clones were repeatedly identified through the Yeast two-hybrid screening, four of them simultaneously showed molecular interaction with TRIB1 protein in mammalian cells. Knocking-down one of the four genes, SAP18, reduced plasma TG levels and increased hepatic lipid accumulation, suggesting impaired lipid secretion. Transcriptional analysis of the mice liver demonstrated that SAP18 revealed significant reduction of the expression of microsomal triglyceride transfer protein (MTTP), one of the causal genes of apolipoproteinemia. SAP18 is a subunit of Sin3A-HDAC complex. ChIP studies using haio-chip system and anti-Sin3A showed enrichment of MTTP promoter sequence. These results showed a possible molecular mechanism of NAFLD associated with TRIB1 expression levels.

2070M Association of the eNOS -786T>C gene polymorphism and coronary artery disease in Iranian population. S. Mehtash1,2, M. Safarpour1, A. Esmaeili Khateri1, A. Ebrahimif2, 1) Biology Dept, Guilan University, Rasht, Guilan, Iran; 2) Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran; 3) Sasan Hospital, Tehran, Iran.

Introduction: Coronary artery disease (CAD) is the most common type of cardiovascular disease and one of the major causes of mortality worldwide. CAD is a complex, multi-system disease that is caused by plaque formation along the coronary arteries, which restricts the heart blood supply. This process is called atherosclerosis. CAD is a multifactorial disease which is influenced by both genetic and environmental factors. One of the most important genes involved in cardiac disease is nitric oxide synthase 3 (NOS3) gene which its product synthesizes nitric oxide from L-Arginine. The endothelium plays an important role in maintaining vascular tone and blood pressure and that is largely mediated by nitric oxide (NO). Since reduced NO synthesis has been involved in the development of coronary atherosclerosis, polymorphisms of the NOS3 gene can be associated with increased susceptibility to CAD. So, The aims of the present study was to evaluate the possible association between the endothelial nitric oxide synthase (eNOS) gene polymorphism and occurrence of Coronary artery disease. Methods: In this case-control study, one variation with more clinical significance in cardiac diseases (T786C) in Japanese (P = 9.3E-7). The risk allele of NAFLD decreased transactivation activity in reporter gene assay. Furthermore, knockdown of TRIB1 expression in mouse liver increased plasma and hepatic lipid levels, whereas the over expression decreased them. Although the enhanced lipogenesis in mouse liver was estimated to result from the reduced decay of carbohydrate response element binding protein (ChREBP), molecular pathways for the hepatic lipid accumulation were still uncovered. Methods. The novel molecular targets of TRIB1 were explored using Yeast two-hybrid system. The positive clones were further screened functionally using shRNA template via adenoviral gene transduction system into mice liver. Plasma and hepatic lipid levels of the mice were measured to identify the genes involved in the hepatic lipid accumulation. Results and discussion. Nine cDNA clones were repeatedly identified through the Yeast two-hybrid screening, four of them simultaneously showed molecular interaction with TRIB1 protein in mammalian cells. Knocking-down one of the four genes, SAP18, reduced plasma TG levels and increased hepatic lipid accumulation, suggesting impaired lipid secretion. Transcriptional analysis of the mice liver demonstrated that SAP18 revealed significant reduction of the expression of microsomal triglyceride transfer protein (MTTP), one of the causal genes of apolipoproteinemia. SAP18 is a subunit of Sin3A-HDAC complex. ChIP studies using haio-chip system and anti-Sin3A showed enrichment of MTTP promoter sequence. These results showed a possible molecular mechanism of NAFLD associated with TRIB1 expression levels.

2070M Association of the eNOS -786T>C gene polymorphism and coronary artery disease in Iranian population. S. Mehtash1,2, M. Safarpour1, A. Esmaeili Khateri1, A. Ebrahimif2, 1) Biology Dept, Guilan University, Rasht, Guilan, Iran; 2) Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran; 3) Sasan Hospital, Tehran, Iran.

Introduction: Coronary artery disease (CAD) is the most common type of cardiovascular disease and one of the major causes of mortality worldwide. CAD is a complex, multi-system disease that is caused by plaque formation along the coronary arteries, which restricts the heart blood supply. This process is called atherosclerosis. CAD is a multifactorial disease which is influenced by both genetic and environmental factors. One of the most important genes involved in cardiac disease is nitric oxide synthase 3 (NOS3) gene which its product synthesizes nitric oxide from L-Arginine. The endothelium plays an important role in maintaining vascular tone and blood pressure and that is largely mediated by nitric oxide (NO). Since reduced NO synthesis has been involved in the development of coronary atherosclerosis, polymorphisms of the NOS3 gene can be associated with increased susceptibility to CAD. So, The aims of the present study was to evaluate the possible association between the endothelial nitric oxide synthase (eNOS) gene polymorphism and occurrence of Coronary artery disease. Methods: In this case-control study, one variation with more clinical significance in cardiac diseases (T786C) in Japanese (P = 9.3E-7). The risk allele of NAFLD decreased transactivation activity in reporter gene assay. Furthermore, knockdown of TRIB1 expression in mouse liver increased plasma and hepatic lipid levels, whereas the over expression decreased them. Although the enhanced lipogenesis in mouse liver was estimated to result from the reduced decay of carbohydrate response element binding protein (ChREBP), molecular pathways for the hepatic lipid accumulation were still uncovered. Methods. The novel molecular targets of TRIB1 were explored using Yeast two-hybrid system. The positive clones were further screened functionally using shRNA template via adenoviral gene transduction system into mice liver. Plasma and hepatic lipid levels of the mice were measured to identify the genes involved in the hepatic lipid accumulation. Results and discussion. Nine cDNA clones were repeatedly identified through the Yeast two-hybrid screening, four of them simultaneously showed molecular interaction with TRIB1 protein in mammalian cells. Knocking-down one of the four genes, SAP18, reduced plasma TG levels and increased hepatic lipid accumulation, suggesting impaired lipid secretion. Transcriptional analysis of the mice liver demonstrated that SAP18 revealed significant reduction of the expression of microsomal triglyceride transfer protein (MTTP), one of the causal genes of apolipoproteinemia. SAP18 is a subunit of Sin3A-HDAC complex. ChIP studies using haio-chip system and anti-Sin3A showed enrichment of MTTP promoter sequence. These results showed a possible molecular mechanism of NAFLD associated with TRIB1 expression levels.
SNP with BP and suggests that vanin-1 misfolding and degradation are the study provides strong biological evidence for the association of the identified variants, we show that N131S vanin-1 was degraded significantly faster reticulum-associated degradation (ERAD) as the underlying mechanism for retains apparent protective effects in individuals of lowering systolic BP. patient plasma samples; we observed that the N131S mutation led to significant-ly faster of African Americans (P=0.01). This association was further validated using St Louis, MO.

Genetics, Stanford University School of Medicine, Stanford, CA; 8) Department of Biophysics, School of Medicine, Case Western Reserve University, Cleveland, OH; 4) Department of Public Health Sciences, Loyola University Chicago, Maywood, IL; 5) Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC; 7) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 8) Department of Genetics, Geisel School of Medicine, Dartmouth College, Hanover, NH; 9) Division of Biostatistics, Washington University School of Medicine, St Louis, MO.

High blood pressure (BP) is the most common cardiovascular risk factor worldwide and a major contributor to heart disease and stroke. We previously discovered that patients with mutations in the Dyrk1B gene, encoding vanin-1, a glycosylphosphatidylinositol (GPI)-anchored membrane protein, have a reduced risk of arterial hypertension. To further investigate this association, we conducted genetic linkage and association studies in a large cohort of families with FHt. In this study, we replicated the association with rs2272996 and BP traits with a total sample size of nearly 30,000 individuals from the Continental Origins and Genetic Epidemiology Network (COGENET) of African Americans (P=0.01). This association was further validated using patient plasma samples; we observed that the N131S mutation led to significantly lower plasma vanin-1 protein levels, and thus the N131S mutation retains apparent protective effects in individuals of lowering systolic BP. We hypothesized that the N131S mutation reduces endothelial nitric oxide synthase (eNOS) activity, leading to decreased NO production, which can result in impaired vascular function and increased risk of cardiovascular disease. To test this hypothesis, we performed in vitro and in vivo experiments using primary human endothelial cells and a murine model of endothelial dysfunction. Our findings support the role of the N131S mutation in regulating eNOS activity and suggest potential therapeutic targets for the management of endothelial dysfunction and associated diseases. In conclusion, the N131S mutation in Dyrk1B is associated with lower plasma vanin-1 levels, reduced endothelial dysfunction, and a decreased risk of arterial hypertension.
2077S Regulatory Polymorphisms in DBH Affect Peripheral Gene Expression and Sympathetic Phenotypes. E.S. Baner1, D. Weinschenker2, A. Verma3, S. Pendergrass2, L.A. Lange4, M.D. Ritchie5, J.G. Wilson6, H. Kuvin7, G. Tromp8, D.J. Carey2, G.S. Gerhard2, M.H. Brilliant4, S.J. Hembri9, J.F. Cubells1, W. Sadee1. 1) Pharmacology, Center for Pharmacogenomics, The Ohio State University Wexner Medical Center, Columbus, OH; 2) Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA; 3) Center for Systems Genomics, Pennsylvania State University, University Park, PA 16802, USA; 4) Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA; 5) University of Mississippi Medical Center, Jackson, MS 30921, USA; 6) The Sighed and Janet Weis Center for Research, Geisinger Health System, Danville, PA 17322, USA; 7) Institute for Personalized Medicine, The Pennsylvania State University College of Medicine, Hershey, PA 17033, USA; 8) Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, WI 54449, USA.

Rationale: Dopamine β-hydroxylase (DBH) catalyzes the conversion of dopamine to norepinephrine in the CNS and peripherally. DBH variants have been associated with large changes in circulating DBH and norepinephrine and are implicated in multiple disorders, yet causal relationships and tissue-specific effects remain unresolved. Objective: To characterize regulatory variants in DBH mRNA, effect on expression in human tissues, and role in modulating sympathetic tone and disease risk. Methods and Results: Analysis of DBH mRNA in human tissues confirmed high expression in the brain (locus coeruleus, LC) and adrenal gland, but also unexpectedly in sympathetically innervated organs (liver—lung—heart). Allele-specific expression assays of mRNA in LC and adrenals yielded small allelic differences (±0.01 fold change) and did not differ from wild type (±11 fold) revealed regulatory effects. The minor alleles of two variants, promoter region rs1611115 and exon 2 rs1108580, were associated with significantly reduced DBH mRNA expression in liver and lung, but not brain and adrenals. In mice, Dbh mRNA levels in the liver correlated with cardiovascular risk phenotypes. Using a PhelWAS (phenome-wide association study) analysis, the minor alleles of rs1611115 and rs1108580 were associated with sympathotenic phenotypes including angina pectoris. Testing combined effects of rs1611115 and rs1108580 indicated robust protection against myocardial infarction in two separate clinical cohorts, which was replicated in a third cohort demonstrating increasing protective effect with an increasing number of minor alleles. Conclusions: These results demonstrate profound effects of DBH mRNA expression in sympathetically innervated organs, modulating clinical phenotypes responsive to peripheral sympathetic tone. This work was supported in part by NIH grant U01092655.
Notch1 haploinsufficiency increases risk of congenital heart defects in the setting of maternal diabetes by an epigenetic mechanism. M. Basu1, K. Bocc2, V. Garg1. 1) Center for Cardiovascular and Pulmonary Research, The Research Institute at Nationwide Children’s Hospital, 700 Children’s Drive, Columbus, OH 43205, USA; 2) Departments of Molecular Genetics & Pediatrics, The Ohio State University, Columbus, OH, USA.

**Rationale:** Congenital heart disease (CHD) is the leading noninfectious cause of infant morbidity and mortality. Epidemiologic studies have demonstrated the importance of genetic and environmental factors in the multifactorial etiology of CHD. Pre-gestational maternal diabetes is one of the non-genetic risk factors that predispose individuals to CHD. Hyperglycemia is associated with endothelial cell dysfunction and we recently demonstrated a genetic interaction between endothelial nitric oxide synthase and Notch1, which encodes a transmembrane receptor, is important for heart development.

**Objectives:** We tested the hypothesis using streptozotocin-induced mouse model of diabetes in mice heterozygous for a null allele for Notch1. To elucidate the molecular mechanism underlying this gene-environment interaction, studies were associated using cardiomyblast (H9C2) and endocardial-derived cell lines, chick embryos and mouse models. Gene expression was measured using qRT-PCR and Western blotting.

**Results:** H9C2 and endocardial-derived cells exposed to hyperglycemia showed that hyperglycemia was associated with a gradual decrease in Notch1, while a rise in a downstream target Hey2, which encodes a transcription factor. Similarly, in vivo, Notch1 expression increased with hyperglycemia. Further link between hyperglycemia and Notch signaling was demonstrated by reduced CBF-luciferase activity in H9C2 cells transfected with a constitutively active Notch1 intracellular domain that were exposed to hyperglycemia. Studies demonstrating the relationship between Jarid2 and Notch1 locus with hyperglycemia by ChIP-qPCR are in process.

**Conclusions:** Our data demonstrate that maternal hyperglycemia disrupts cardiac development by deregulating Notch1 signaling pathway. Our preliminary findings suggest this gene-environment interaction is mediated by an epigenetic mechanism involving Jarid2.

---

A transcriptomic study reveals KLF15 as a circadian metabolic switch in the heart. L. Zhang, M. Jain. Dept Med/Human Gen, Case Western Univ SOM, Cleveland, OH.

Physiological parameters of the cardiovascular (CV) system and pathological states such as heart attacks show a circadian rhythm. We have previously identified that in multiple CV diseases, including atherosclerosis and heart failure, KLF15, a zinc-finger transcription factor, expression is reduced. Interestingly, its expression also oscillates in a circadian fashion under the direct regulation of core clock. To study the role of KLF15 in the circadian regulation of CV health and disease, we performed a transcriptomic study of the heart in wild type and cardiomyocyte specific Klf15 null mice (cKO).

We have identified 1335 genes oscillating in the wild type mice heart. These genes fall into two distinctive groups, one peaks at rest to active phase, another peaks at active to rest phase. Interestingly, in addition to the core clock genes, genes peaking at the beginning phase of activity are strongly associated with metabolic processes including amino acid and fatty acid catabolism, ABC transporters, as well as “de-tox” pathways, such as drug and xenobiotic metabolism. This is consistent with the increased energy demand and food intake. The genes expressed highly at the resting phase are associated with cell migration, cell shape and cell cycle regulation. This suggests the resting phase is vulnerable to both hypertrophy and hyperplasia, and is a critical phase for cardiac repair and remodeling. This novel discovery has important implication for chronotherapy in management of patient with CV disease. There are 1003 oscillating genes showed autonomous KLF15 dependency, specifically, cKO mice showed a loss of all metabolic gene induction upon active phase, including lysine, tryptophan, branch chain amino acids and fatty acid metabolism. The oscillation of core clock genes, "de-tox" genes and ABC transporters are preserved. This observation suggests KLF15 is the switch for metabolic control in the heart for the active phase and loss of function of KLF15 in the diseased state likely results in an energy deficiency state in the heart. We have also identified 473 genes, which gained oscillation in the cKO mice without an increase of other cell types including inflammatory cells. De novo motif prediction has identified enrichment of EGR2, NR5A2, MYC and several others. This is the first report of gaining circadian oscillation in a genetic mutant mice.

We are currently confirming and studying the molecular mechanism of this phenotype.

---


Varicose veins are a common chronic condition in which veins become enlarged and twisted due to deficient functioning of one-way valves that normally return blood to the heart. They usually appear in the legs, and are more common in women than men. While they are most commonly considered a cosmetic problem, in some cases they can lead to, or signal, more serious circulatory problems. We carried out a GWAS of self-reported venous varicose veins in the 23andMe participant cohort, including approximately 20,000 cases and 65,000 controls of European ancestry, imputed against 1000 Genomes reference haplotypes. We identified several novel associations for varicose veins near or in genes involved in various circulatory and blood related functions, including blood group determination, regulation of blood pressure and/or circulation, vascular growth factors. Our strongest association was with SNP rs507666 (p=5.2e-20), found in ABO, which encodes a protein defining the ABO blood group. Variation in SNP rs507666 determines A1 blood subgroup status. This finding is consistent with a previous study that reported the association between the A blood group and varicose veins. It also identified an additional SNP, rs696562 (p=1.8e-16), near XR5 associated in Kell blood group determination. Our study also identified SNPs in genes which have previously been characterized as players in the regulation of blood pressure or circulation, including SNP rs11121615 (8.1e-12) in CASZ1, associated with blood pressure variation and SNP rs7111987 (2.2e-10) near ADM which encodes a peptide that may function as a hormone in circulation control. Finally, the GWAS of varicose veins in the 23andMe participant cohort is enriched with genes involved in cell migration and movement. For example gene rs1433196 (5.2e-11) is found in ANGPT1 that encodes a type of angiopoietin, a group of proteins with important roles in blood development; SNP rs696562 (mentioned above) near ANGPT2, that encodes an angiopoietin, is associated with movement in cell migration.

A protein whose effects include angiogenesis, vasculogenesis and endothelial cell migration has been found to be enriched in the cKO mice without an increase of other cell types including inflammatory cells. De novo motif prediction has identified enrichment of EGR2, NR5A2, MYC and several others. This is the first report of gaining circadian oscillation in a genetic mutant mice. We are currently confirming and studying the molecular mechanism of this phenotype.
2083S
Carotid plaque is a measure of subclinical cardiovascular disease and is a risk factor for stroke and heart attack. The genetic underpinnings of plaque are not known. The study goal was to identify risk variants for plaque using an extreme phenotype approach in the analysis of our genome-wide association study. In individuals from the population-based Northern Manhattan Study, plaque was detected by high-resolution B-mode ultrasound and plaque volume was measured by an image analysis software and expressed as total plaque area, a sum of all plaque areas within an individual. Genotyping was done with the Affymetrix 6.0 SNP array. Among the 908 Hispanics, we first calculated the residual score for each participant by regressing plaque area on significant known risk factors (age, sex, packyears of smoking, SBP, diabetes, LDL-HDL ratio, homocysteine levels, high school completion, LDL, lipid lowering medication, and WBC count). To enrich the genetic effects and increase our power to detect associations, we next identified the individuals in the extreme 10% and 20% of the residual distribution, for a total of 90 and 200 individuals on each side respectively, thus ensuring that the phenotypic extremes were largely unexplained by these known risk factors. For each of the extreme thresholds, we performed logistic regression analysis on 877K SNPs, while controlling for the top 3 PCs. A gene based analysis was then performed based on the SNP results using VEGAS. The top gene from the 10% threshold (p=3.80E-05) was member of the oncocogene family (RAS22B). Members of the RAS family are thought to play a crucial role in endotheial function. The top gene from the 20% threshold (p=2.30E-09) was the retinal pigment epithelial transcription factor (RPE65). In addition to the gene based analysis, we also conducted a pathway based analysis using genes with a p<0.05 as input into WebGestalt. After adjusting for multiple testing, several pathways of interest were identified. In particular for the 10% threshold, the transcription pathway was significant (p=5.0E-04). An analysis of our diet data found differences in the intake of vegetables (but not fruit) between the lower 10% and upper 10%, as well as the lower and upper 20%, with individuals with smaller than expected plaque area consuming more vegetables. These results suggest that genetic variants affecting taste could contribute to plaque development via influencing vegetable consumption.

2085S
The Kaiser Permanente/UCSF Genetic Epidemiology Research study on Adult Health and Aging: A blood pressure genome-scan in ~100,000 Subjects. T. Hofmann1, A. Chakravarti2, G. Ehret2,3, C. Iribarren4, Y. Banda, E. Jorgenson3, M.N. Kvalem1, C. Schaefer4, N. Rischa1. 1) University of California San Francisco, San Francisco, CA; 2) John’s Hopkins University, Baltimore, MD; 3) Geneva University Hospitals, Switzerland; 4) Division of Research Information, Kaiser Permanente, Northern California, Oakland, CA.
Blood pressure (BP) is quantitatively the major risk factor for cardiovascular disease. The Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort has detailed data on longitudinal blood pressure, prescription medication, and genotypes on ~675,000 markers for nearly all of its 100,000 participants (81% white, 7.5% Asian, 7% Latino, and 3.5% African American). Approximately 1.3 million systolic (SBP) and diastolic (DBP) blood pressure measures from electronic health records were used to derive long term average phenotypes and association tests with genotyped variants were run based on the directly genotyped markers. We conducted a genome-wide association study across the subjects of European ancestry. All analyses were adjusted for age, age2, sex, BMI, and the first ten principal components. Follow-up analysis was conducted in the other races/ethnicities. We compared the results of our association study to previously published GWAS hits from the NHGRI GWAS catalog, and found that, at a p<0.0016 (0.05/32 SNPs), 23 SNPs replicated, 8 did not replicate, and 1 was not informative. Further we identified 3 novel loci at genome-wide significance (p<5.0×10^-8). When comparing to DBP, we found that, at a p<0.0016, 23 SNPs replicated, 6 did not replicate, and 1 was not informative. Further we identified 7 novel loci at genome-wide significance. In summary, we replicate the great majority of BP SNPs listed in the NHGRI GWAS catalog, and identify 10 new common BP loci.

2086M
Variation in LECR1 and GALNT10 modulate sex-difference in carotid intima-media thickness: A genome-wide interaction study. C. Dong1, A. Beecham2, L. Wang2, D. Cabral2, S.H. Blanton2, R.L. Sacco1,2,3, T. Rundek1,2,4. 1) Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL; 2) John T. McDonald Department of Human Genetics, John P Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL; 3) Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL; 4) Division of Neurology, John’s Hopkins University School of Medicine, Baltimore, MD; 5) Department of Human Genetics, MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom; 6) National Heart, Lung, and Blood Institute's Framingham Heart Study and Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA.
Fibrinogen, coagulation factors VII (FVII) and VIII (FVIII), and von Willibrand factor (vWF) occupy key roles in blood coagulation. Plasma levels are associated with risk of arterial and venous thrombosis and fibrinogen is also a marker of inflammation. The CHARGE Hemostasis Working Group previously identified common variants associated with plasma levels of these factors. We now explore associations with rare and functionally relevant variants.
Individuals of European, African, East Asian and Hispanic ancestries were genotyped using the Illumina “Exome Chip” genotyping array and variants called using a common algorithm across cohorts. Association of rare variants was assessed using sex difference in carotid artery intima-media thickness (IMT), a recognized marker of atherosclerosis. However, the genetic underpinnings that modulate sex-differences in gene-IMT associations are largely unknown. With a multistage design using 722,379 single nucleotide polymorphisms (SNPs) genotyped with the Affymetrix 6.0 chip, we performed a genome-wide sex-by-SNP interaction study in 931 Hispanics (mean age: 69±8 years, 61% women, mean IMT: 0.70±0.06 mm), followed by replication in 257 non-Hispanic blacks (mean age: 73±9 years, 62% women, mean IMT: 0.74±0.10 mm) and 153 non-Hispanic whites (mean age: 73±9 years, 50% women, mean IMT: 0.76±0.10 mm). Assuming an additive genetic effect for each SNP based on the minor allele number, we performed multiple linear regression analysis to test for sex-by-SNP interaction on the IMT while controlling for age and the top 3 principal components estimated to capture ancestry by EIGENSTRAT. Among 14 SNPs with an interaction of p<5.0E-6 in Hispanic discovery sample, replicated interactions were found for four SNPs in leucine, glutamate and lysine rich 1 (LEKR1) gene in non-Hispanic whites (p=5.2E-3 for or SNPs), leucine rich polypeptide N-acetylgalactosaminyltransferase 10 (GALNT10) gene in non-Hispanic blacks (p=1.3E-3). Specifically, for the top SNP rs7616559 in LEKR1 gene, the adjusted mean difference between men and women was 0.003 mm (p=0.67) for AA-carriers, 0.044 mm (p=2.1E-9) for AG-carriers, and 0.064 mm (p=3.0E-6) for GG-carriers, compared to the control sample. For GALNT10, the adjusted mean difference between men and women was -0.017 mm (p=0.13) for TT-carriers, 0.022 mm (p=0.001) for TC-carriers, and 0.051 mm (p=0.002) for CC-carriers in the combined sample. Genetic variants near LEKR1 gene had significant associations with measures of newborns and variants near GALNT10 gene have been associated with body mass index. Given the consistent findings across different-ethnic groups, further studies are warranted to perform in-depth investigations of functional genetic variants in these regions.
2087S

Carotid intima-media thickness: a genome-wide association analysis among African Americans. S.M. Tajuddin1, M.A. Nalis1, M.K. Keller2, A.B. Zonderman1, M.C. Evans1. 1) Laboratory of Epidemiology and Population Sciences, National Institutes of Health, Baltimore, MD; 2) Laboratory of Neurogenetics, National Institutes of Health, Bethesda, MD.

Although the age-adjusted death rate from coronary heart disease (CHD) continues to decline in the population overall, there remains significant morbidity and mortality disparities for African Americans (AAs). The atherosclerotic process of CHD leads to thickening of the intimal and medial layers of the common carotid artery. Carotid intima-media thickness (CIMT), measured on ultrasound, is a known surrogate marker of susceptibility in AAs. Replication studies in other AA cohorts will help confirm the findings that may shed light on atherosclerosis development and provide a biomarker of susceptibility in AAs. Replication studies in other AA cohorts will be conducted to validate these findings.

2088M

Genetic determinants underlying hypertension in multi-ethnic populations. N. Vasudeva1, L. Wang1–2, Z. Liu1, P. Goldschmidt1, M. Cicek-Vance1–3, D. Seo1, G. Beecham1–2. 1) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 2) Division of Cardiology, Miller School of Medicine, University of Miami, Miami, FL.

Hypertension is a major cardiovascular health risk that is well-established as a heritable trait. The environmental factors influencing it, as well as their complex interactions make it challenging to discover the genetic determinants affecting the trait. A cumulative effect of multiple variants has been believed to contribute to this polygenic trait. To dissect the genetic basis of this trait, we analyzed European, Hispanic and African ancestry samples from the 1000 Genomes Project imputed SNPs that passed stringent quality control criteria in 584 AA (55% men). We estimated beta coefficients and p values using linear regression models adjusted for age, sex and the 10 principal components under additive genetic effect model. We identified eight single nucleotide polymorphisms (SNPs) in AURKAIP1 (p=5.39E-06), FAM53C (p=6.59E-06), CAMTA1 (p=2.7x10^-8), and LINC00229 and CAMTA1 associated 3 algorithms to make the CIMT at genome-wide significance level of P<5 x 10^-8. The lead SNPs were MMAA-rs142277468 (P=2.6 x 10^-10), and CAMTA1-rs72683082 (P=2.8 x 10^-10). We identified three genetic regions that are associated with CIMT in this African ancestry population. SNPs in CAMTA1 and LINC00229 have been implicated in plasma fatty acid levels and obesity related traits, respectively. MAA4 is involved in calbinamin processing and transportation into the mitochondria and may have an effect on atherosclerosis through homocysteine metabolism. However, our findings may shed light on atherosclerosis development and provide a biomarker of susceptibility in AAs. Replication studies in other AA cohorts will be conducted to validate these findings.

2089S

Contribution of Global Copy Number Variants to Down Syndrome-associated Atrioventricular Septal Defects. D. Ramachandran1, J.G. Maffe1, E.A. Locke1, L.J. Bean1, T.C. Rosser1, P. Bose1, K.J. Dooley4, C.L. Cua1, G.T. Capone1, R.H. Reeves1, C.L. Maslen6, D.J. Cutler1, S.L. Sherman1, M.E. Zwicky1. 1) Department of Human Genetics, Emory University, Atlanta, GA; 2) Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA; 3) Center for Statistical Genetics and Department of Biostatistics, University of Michigan School of Public Health, MI; 4) Sidney Heart Center Cardiology, Children’s Hospital of Atlanta, GA; 5) Heart Center, Nationwide Children’s Hospital, Columbus, OH; 6) Down Syndrome Clinic and Research Center, Kennedy Krieger Institute, Baltimore, MD; 7) Department of Physiology, McKusick Nanthans Institute for Genetic Medicine, School Of Medicine, Johns Hopkins University, MD; 8) Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR.

Atrioventricular septal defects (AVSD), a severe congenital heart defect, occur in the general population in ~1 in 10,000 births. Nearly 20% of infants with Down Syndrome (DS) have an AVSD, representing an approximately 2000-fold increased risk compared to the euploid population. Our central hypothesis was that in the presence of an extra chromosome 21, otherwise benign copy number variants (CNVs) act to increase the risk of AVSD in individuals with DS. Here we sought to test two specific questions: (1) Do common CNVs of large effect contribute to the elevated risk of AVSD in the DS population? (2) Do rare CNVs increase the risk of AVSD in the DS population? We used the Affymetrix SNP 6.0 genotyping platform to comprehensively characterize CNVs in 452 ethnically matched individuals with DS, comprising of 210 cases (DS + complete AVSD) and 242 controls with a structurally normal heart (DS + NH). We implemented strict quality control filters to minimize false positive calls, including 3 algorithms to make the CNV calls (BEAST, GADA and GLAD). We also required each putative CNV call to contain > 20 SNPs within the interval. After excluding CNVs overlapping centromeres, we identified 541 deletions (253 in cases and 288 in controls). Results from burden and region-wise analyses using PLINK revealed that despite the 2000 fold elevated risk, common CNVs of large effect (OR > 2.0) do not account for the increased risk observed in DS-associated AVSD. In contrast, rare CNVs of large effect (OR > 100kb, > 1% frequency) (p < 0.01) and case deletions intersect genes more often than those observed in controls (p < 0.007). Gene enrichment analysis showed a trend for enrichment among deletions impacting tricileogue pathways in cases compared to controls. No significant differences were observed for large rare duplications between cases and controls. Our findings suggest that the etiology of AVSD is highly complex and does not arise from the action of a few common variants of large effect. Instead, our data support a multifactorial model, wherein large rare deletions play a significant role in elevating the risk of AVSD in a trisomic background. Our study further suggests an important role for ciliome genes in AVSD.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Identification of Loci associated with Bicuspid Aortic Valve (BAV). S.C. Body, M. Heydarpour, J.G. Seldman, S. Prakash, D. Milewicz, Y. Bousse, G. Limongelli, the Bicuspid Aortic Valve Consortium (BAVCoin), 1) Anesthesia, Brigham & Womens Hosp, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) University of Texas, Houston, TX; 4) University of Laval, Quebec, Canada; 5) Monaldi Hospital, Naples, Italy.

OBJECTIVE: Bicuspid aortic valve (BAV) is observed in 0.6%-1% of humans and is associated with premature aortic stenosis, aortic aneurysm and other congenital cardiac anomalies. BAV has been associated with high-penetration rare mutations, mainly in NOTCH1, in a few families, but the majority of BAV is sporadic with unidentified intronic and intergenic polymorphisms in mechanisms causing BAV are unknown. The objective of this study was to identify pathway/genes associated with sporadic BAV. METHODS: We genotyped 456 Caucasian cases (339 males, 117 females) using Illumina Omni12.8-8 BeadPacs genotyped with either Omni2.5 or Omnis arrays containing 1,869, 897, and 2,450 Caucasian controls respectively. Quality control and population stratification of the genotype data were performed in PLINK. An additive logistic regression model was performed for association analysis across all three GWAS studies accounting for gender and variants associated with aortic stenosis and aortic aneurysm in order to remove confounding by presentation. We performed three analyses, one for each control group versus the cases, then pooled significant P-values (<1*10^-5) for combined SNPs (>5%) with the same direction for genetic marker effects (ORs). RESULTS: 22 SNPs in 12 regions with P<1*10^-5 in all three GWA studies remained significant after correction for multiple testing. Five regions (1p21, 4q32, 5q14, 8p23, 12q22) contained SNPs of genome-wide significance (P<5*10^-8) with odds ratios of 1.5-2.6. The most significant associations were in 1q (p=1.8*10^-11) and 10q (p=2.9*10^-8). CONCLUSION: Our findings suggest that several polymorphisms could explain the higher predisposition of BAV in Caucasians. Novel regions were identified on chromosomes 1, 4, 5, 8, 10, 12, 16, and 18, which have important roles in heart development and can be good candidates for fine mapping and replication in other cohorts. In addition, prior work has demonstrated BAV-linked genes involved in cardiac development of other vertebrates (e.g. Notch1-deficient-mice, Gata5-null-mice). Such studies combined with our ability to delete genes in specific cells at specific developmental stages in animal models will undoubtedly unravel numerous candidate Coronary Heart Disease causing genes that can be directly tested in human genetic studies. Conversely loci identify in human cohorts can be tested in animal models to confirm (or not) their causative link to disease. These findings are being replicated.

Relationship Between Plasma Betaine Levels and Cardiovascular Disease: Results of a Genome-Wide Association Study. J. Hartila, W.H. Tang, Z. Wang, S.L Hazen, H. Allayee. 1) Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90033; 2) Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, OH 44195.

We recently identified several metabolites generated through gut microbiome and hepatic-mediated metabolism of dietary choline and L-carnitine as novel risk factors for cardiovascular disease (CVD). Elevated levels of trimethylamine-N-oxide (TMAO) in particular were found to be pro-atherogenic in mice and humans but the role of betaine, which can also be generated from choline, in CVD is less clear. Following our recently published comparative genomics analysis of TMAO, we sought to identify the genetic determinants of plasma betaine levels. In stage 1, a GWAS with ~1000 subjects from the GeneBank study identified 4 loci on chromosome (chr) 1q32, 2q34, 5q14, and 16q24 that were significantly or suggestively associated with betaine levels. In stage 2 with ~2000 additional GeneBank subjects, 2 independent SNPs on chr 5q14 (rs617219 and rs18676503) and rs715 on chr 2q34 demonstrated association with betaine, which became even more significant in ~4000 subjects (p=6.0E-9 - 9.0E-13). A weighted genetic risk score with these 3 SNPS also yielded highly significant association with betaine (p=1.1E-23). The SNPs on chr 5q14 are located near the BHMt gene (rs617219), which catalyzes the simultaneous conversion of betaine and homocysteine to dimethylglycine (DMG) and methionine, respectively, and the DMGDH gene (rs18676503), which further metabolizes DMG to sarcosine and subsequently glycine. Of these additional metabolites, rs517911a modestly increased methionine (p=0.04). By comparison, rs715 on chr 2q34 is located within the CPS1 gene, which is the rate-limiting enzyme in the urea cycle. In addition to the association with betaine, rs715 increased ornithine levels (p=9.0E-3) and decreased citrulline (p=3.0E-4), arginine (p=4.0E-3), and urea (p=0.07). These results suggest that rs715 reduces flux down the urea cycle by decreasing CPS1 activity and/or expression since the strongest effects were on the most proximal urea cycle metabolite (citrulline) with increasingly weaker effects on more distal metabolites (i.e. arginine and urea). Lastly, of the betaine-associated SNPs on chr 2q34 and 5q14, rs715 was significantly associated with decreased risk of CVD (OR=0.71, 95% CI 0.58-0.87; p=0.004) in all ~10,000 GeneBank subjects. Taken together, these studies provide evidence for a genetically mechanistic link between betaine metabolism, the urea cycle, and CVD in humans.

Genome-wide association study identifies novel susceptibility loci for venous thromboembolism in African Americans. W. Hernandez, E.R. Gamazon, A. Konkashbaev, R.A. Kittles, L.H. Cavallari, M.A. Perera. 1) The University of Chicago, Department of Medicine, Section of Genetic Medicine, Chicago, IL; 2) University of Illinois, Department of Medicine, Institute of Human Genetics, Chicago, IL; 3) University of Florida, Department of Pharmacology and Translational Research, College of Pharmacy, Gainesville, FL.

Venous thromboembolism (VTE) is a chronic multifactorial disease encompassing deep vein thrombosis (DVT) and pulmonary embolism (PE). It is a major public health burden and the third most common life-threatening cardiovascular condition resulting in high rates of hospitalization and mortality. In the US, African Americans (AAs) have the highest incidence and mortality rates. Although studies suggest that VTE is highly heritable, few associations between VTE risk and genetic variants have been established. Furthermore, the role of these polymorphisms in the risk of VTE among AAs is conflicting at best. The lack of genetic susceptibility to VTE among AAs, in great part, may be due a combination of identifying susceptibility genetic variants in Caucasian cohorts and subsequently testing these variants in AAs as well, as the very low frequency of these variants among AAs. In this study, we present preliminary findings from the first genome-wide association study (GWAS) and risk of VTE conducted exclusively on AAs (137 cases and 437 control). Our results revealed 15 independent signals that reached a genome-wide significance level of 5.0*10^-8. Five variants in PTGER3 (rs570021), PPP1R12B (rs705744), SLC2A23 (rs9364552), SLC9A11 (rs951284), and MGST1 (rs2975139) and eight intergenic variants (rs7503127, rs7668402, rs7729075, rs9276835, rs121319, rs2749490, rs711931) were associated with increased risk of VTE (ORs=2.2 - 2.6). Two variants, rs10234060 and rs711831, were found to decrease risk of VTE (OR=0.4 and 0.44, respectively). These ORs are larger than those previously identified in GWASs on VTE risk among Caucasians. In addition, we conducted bioinformatics analysis of the top signals to identify potential expression quantitative trait loci (eQTLs) utilizing whole blood from AAs. The risk variant, rs9364552 located in SLC22A3 (OR=2.3, P=4.30*10^-8), was found to be cis-eQTL (dQTL) to SLC22A3, with decreasing downregulation of SLC22A3 in individuals carrying the rs9364552 minor allele would express lower levels of PLG, consequently increasing their risk of VTE. Furthermore, SLC22A3 and PLG have been shown to affect lipoprotein(a) levels and interact as independent risk factors for VTE. Currently, we are in the process of validating these finding in an independent AA cohort of VTE. Our study provides new molecular insight into the underlying mechanism that may regulate VTE susceptibility in AAs.
Nervous system-related loci determining sex-difference in blood pressure reactivity to cold stress in both Chinese and Whites. Q. Zhao, X. Kong,1,3 T.N. Kelly,1 C. Li,1 D. Gu,3 J. He.1 1Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA; 2Department of Endocrinology, China-Japan Friendship Hospital, Beijing, China; 3State Key Laboratory of Cardiovascular Disease, Fujui Hospital, National Center of Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

Exaggerated blood pressure (BP) response during the cold pressor test (CPT) has been suggested as a risk factor for developing cardiovascular disease. However, little is known about the genetic basis for sex-related differences in BP reactivity to cold stress. The discovery sample was from the Genetic Epidemiology Network of Salt Sensitivity (GENSalt), which included 3,495 participants. During the CPT, BP was measured prior to and after the participants immersed their right hand in ice water for 1 minute. A total of 1,881 GenSalt participants completed the CPT and were genotyped using the Affymetrix SNP array 6.0. A genome-wide association analysis was conducted to examine the interaction effect of SNPs and sex on BP reactivity during the CPT. A total of 13 loci showed potential sex difference in the association with BP reactivity variables (P for interaction < 1.0 x 10^{-5}). These loci were further tested among 1,448 participants of European ancestry from the Coronary Artery Risk Development in Young Adults (CARDIA) study, in which a similar cold stress test was conducted. Five of the 13 loci (1p32.2, 2q33.1, 5q23.1, 5q15, and 13q33.3) identified in the GenSalt study showed potential transethnic replication in the CARDIA study (P for interaction < 0.1). Four of the five loci (2q33.1, 5q23.1, 5q15, and 13q33.3) identified in the GenSalt study showed potential replication among multiple ethnic groups, which included European, African, and East Asian in the CARDIA study. For example, the index SNP rs3766457 at the locus 1p32.2 was associated with lower pulse pressure (PP) response in men but greater PP response in women (-2.64 vs. 1.48 mm Hg for each G allele) in the CARDIA study. In conclusion, our study provides the first evidence for the transethnic replication of sex-specific genetic effects for BP response to cold stress and implicates multiple potential nervous system-related genes in determining the sex difference in BP reactivity to stress.

Finding low frequency causal genetic variants for lipids by genotyping subjects with extreme HDL-c levels. W. Zhou1, OL. Holmen2,5, H. Zhang, J. Chen,1 M. Boehnke,4 GR. Abecasis,4 K. Hveem2,6,7, CJ. Miller1,6,7,1 1Department of Internal Medicine, Division of Cardiology, University of Michigan School of Medicine, Ann Arbor, Michigan, 48109, United States of America; 2HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, 7600 Trondheim, Norway; 3St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway; 4Department of Epidemiology, School of Public Health, University of Michigan School of Public Health, Ann Arbor, Michigan, 48109, United States of America; 5Department of Medicine, Leverkusen Hospital, Nordrhein-Westfalen Health Trust, 7600 Leverkusen, Norway; 6Department of Computational Medical and Bioinformatics, University of Michigan, Ann Arbor, Michigan, 48109, United States of America; 7Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan, 48109, United States of America.

Genome-wide association studies have identified more than 157 loci associated with lipid levels, which are treatable and heritable risk factors for cardiovascular diseases. Most of these variants are common with modest effect sizes. Missing heritability of blood lipid levels suggests less frequent (1 - 5% minor allele frequency) or rare (< 1% minor allele frequency) coding variants may exist for lipids. To detect low-frequency causal variants, samples with extreme phenotypes likely have increased power compared to population-based. We successfully genotyped 1,751 non-MI samples with extremely high HDL-c levels, 1,744 non-MI samples with extremely low HDL-c levels from the Norwegian HUNT-study using the Illumina HumanExome Beadchip, allowing for direct genotyping of 72,399 polymorphic coding variants. We evaluated the power of the extreme HDL samples (N=3,495) in comparison to a genotyped set of non-extreme samples (N = 3,493, from the same cohort) to detect the known HDL variants and genes. Single SNP association results for the 50 HDL known variants are more significant using extreme HDL samples than population-based samples (Pwilcox = 9.3 x 10^{-6}), suggesting that even with extreme phenotypes, larger sample size or more complete genotypes from sequencing may be needed to detect novel loci.

Identification of blood pressure related genes by population-based transcriptome analyses. C. Müller1,4, K. Schramm1,4, C. Schumann5,4, S. Völker4,6, A. Hoffmann6, H. Wahl2, J. Winter6, H. Grallert1, T. Illig11,12, A. Peters11,13, M. Dör2,4,5, X. Guo7, W. Palmas18, T. Melting1,4,12, A. Teumer12,5, M. Cartensen9,10, P.S. Wild20,21, H. Völker4, K. Kühn7, M. Roden10, D.M. Harrington2,5, U. Volker7, A. Ziegler12,4, V. Lühn2, T. Zeller12,5, H. Blankenberg4, H. Preuss5,4, S.B. Felix14,15,1,6,11 1Clinic for general and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany; 2DZHK (German Centre for Cardiovascular Research), partner site Hamburg/Kiel/Lübeck; 3Institute of Human Genetics, Helmholtz Center Munich, German Research Center for Environmental Health, Neuherberg, Germany; 4Institute of Human Genetics, Technical University Munich, München, Germany; 5Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany; 6The Genetics of Obesity and Related Metabolic Traits Program, The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 7Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA; 8Department of Biometry and Statistics, University of Lübeck, UNIVERSITY HOSPITAL SCHLESWIG-HOLSTEIN, Lübeck, Germany; 9Institute for Clinical Diabetology, German Diabetes Center, Leibniz Institute for Diabetes Research at Heinrich Heine University Düsseldorf, Germany; 10German Center for Cardiovascular Research, partner site Munich, Munich, Germany; 11Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 12Medical School Hannover, Hannover; 13Institute of Genetic Epidemiology, University Medicine, Hannover, Germany; 14Department of Internal Medicine II, University Medicine Greifswald, Greifswald, Germany; 15DZHK (German Centre for Cardiovascular Research), partner site Greifswald, Greifswald, Germany; 16DZHK (German Centre for Cardiovascular Research), partner site Düsseldorf, 11 Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 17Institute for General Medicine, University Heart Center Hamburg, Hamburg, Germany; 18Institute for Clinical and Experimental Medicine, Prague, Czech Republic; 19Institute for Epidemiology, Technical University Munich, München, Germany; 20Department of Health Sciences and Population Studies, University of Southern Denmark, Odense, Denmark; 21Cardiovascular Medicine, University of Copenhagen, Copenhagen, Denmark; 22Institute for Community Medicine, University Medicine Göttingen, Göttingen, Germany; 23Section on Cardiology, Department of Internal Medicine, Wake Forest School of Medicine, Winston Salem, North Carolina, USA; 24Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University, Winston-Salem, North Carolina, USA.

High blood pressure (BP) is a major risk factor for cardiovascular diseases. We analyzed associations between the blood transcriptome and BP traits within large cohorts of the MetaXpressor consortium by applying standardized protocols. The study population included 4,533 individuals with available transcriptomic data from three German cohorts: GenSalt (GHS), Cooperative Health Research in the Region of Augsburg (KORA F4) and Study of Health in Pomerania (SHIP-TREND) and one US cohort: Multi-Ethnic Study of Atherosclerosis (MESA). Expression levels were measured in 12,539 blood samples by using the Illumina HumanHT-12 BeadChip arrays. Associations with systolic BP (SBP), diastolic BP (DBP) and pulse pressure (PP) were computed by linear regression models adjusted for sex, age, BMI and technical covariates within each study. A meta-analysis was conducted within GHS and MESA using the inverse variance method. Significant associations (FDR<0.05) were selected for replication in KORA F4 and SHIP-TREND. Genes with consistent effect directions in all four studies and p-values ≤ 0.05 in the replication cohorts were reported as candidate genes. In total, transcripts of eight distinct genes were consistently associated with at least one of the traits SBP, DBP or PP in discovery and replication steps: CBPBA, CRIP1, F12, LMNA, MYADM, TIPARP, TTPPS3 and TSC22D3. Effect sizes were comparable between associations in monozygotic and whole blood. In total, the candidate genes explained between 4-13%, 4-6% and 2-8% of inter-individual variance of SBP, DBP and PP, respectively.

This is the first study investigating the associations between BP traits and whole transcriptomes across different blood cell populations on a large scale. The comprehensive analyses highlight eight genes, correlated with BP.
**Posters: Cardiovascular Genetics**

**2096M**
Genetics of Plasma Lactate. P. Balakrishnan1, A. Tin1, J. Pankow2, E. Boenwinkel1, R. Hoogeveen2, J.H. Young1, WH.L. Kao1, 1) Epidemiology, Johns Hopkins University, Baltimore, MD; 2) University of Minnesota School of Public Health, Minneapolis, MN 55454; 2) Human Genetics Center, University of Texas School of Public Health, Houston, TX 77030; Department of Medicine, Baylor College of Medicine, Baylor College of Medicine, Houston, TX 77030; 5) Department of Medicine, The Johns Hopkins University, School of Medicine, Baltimore, MD 21205.

**Background:** Type 2 diabetes affects 25.8 million of the U.S. population. Recent studies have shown that increased plasma lactate levels are associated with insulin resistance and type 2 diabetes. Plasma lactate is a measure of oxidative capacity, the ability to meet increased energy demand. The genetics of plasma lactate have not been well-characterized.

**Materials and Methods:** We performed a genome-wide association study (GWAS) in the Atherosclerosis Risk in Communities (ARIC) study - an ongoing prospective epidemiological study conducted in four U.S. communities. Plasma lactate (mg/dl) was measured from blood collected at Visit 4, at which the participants’ age ranged between 54-73. Single nucleotide polymorphisms (SNPs) were also genotyped from ARIC blood samples and imputation was done using the 1000 Genome (NCBI build 37 - hg 19) reference. Linear regression models of single nucleotide polymorphisms (SNPs) were assessed for association with log transformed plasma lactate, separately in European Americans and African Americans. Models were additionally adjusted for age at Visit 4, sex, ARIC center, body mass index, waist circumference and significant principal components of ancestry. The p-values from the European American GWAS SNPs with minor allele frequency >0.01 were used as an input for pathway based analysis using MAGENTA.

**Results:** Four SNPs reached genome-wide significance (P < 5*10^-8) in the European Americans and were also associated in the African Americans based on locus specific threshold. The four SNPs are located within GCKR gene, involved in glucokinase inhibition, and PPP1R3B gene, involved in glycogen synthase inhibition. The vasoconstriction and vasodilation regulation pathway was the pathway most statistically significantly associated with plasma lactate.

**Conclusions:** The genetics of plasma lactate may provide important clues regarding the pathogenesis of type 2 diabetes.

**2097S**
Identification of three novel genetic variants associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. J.E. Lim1, R.W. Hong1, J.W. Kim2, Y. Tabara3, H. Ueshima4-5, T. Miki5, F. Matsuda5, Y.S. Cho5, Y. Kim5, B. Oh1, 1) Kyung Hee University, Seoul, South Korea; 2) Center for Genome Science, Korea Centers for Disease Control & Prevention, Division of Epidemiology and Health Index, Chungcheongbuk-do, South Korea; 3) Department of Internal Medicine, Inje University Ilsan Paik Hospital, Gyeonggi-do, South Korea; 4) Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; 5) Center for Epidemiologic Research in Asia, Shiga University of Medical Science, Otsu, Japan; 6) Department of Health Science, Shiga University of Medical Science, Otsu, Japan; 7) Department of Geriatric Medicine, Ehime University Graduate School of Medicine, Toon, Japan; 8) Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, South Korea.

The electrocardiogram has several advantages in detecting cardiac arrhythmia - it is readily available, noninvasive, and cost-efficient. Recent genome-wide association studies have identified single nucleotide polymorphisms that are associated with electrocardiogram measures. We performed a genome-wide association study using Korea Association Resource data for the discovery phase (phase 1, n = 6,805) and 2 consecutive replication studies in Japanese populations (phase 2, n = 2,285; phase 3, n = 5,010) for QRS duration and PR interval. Three novel loci were identified: rs2483280 (PRDM16 locus) and rs335206 (PRDM6 locus) were associated with QRS duration, and rs17026156 (SLC8A1 locus) correlated with PR interval. PRDM16 was recently identified as a causative gene of left ventricular noncompaction and dilated cardiomyopathy in 1p36 deletion syndrome, which is characterized by heart failure, arrhythmia, and sudden cardiac death. Thus, finding that a PRDM16 SNP is linked to QRS duration strongly implicates PRDM16 in cardiac function. In addition, C allele of rs17026156 increases PR interval (beta±sd, 2.39±0.40 ms) and exist far more frequently in East Asians (0.46) than in Europeans and Africans (0.05 and 0.08, respectively).

**2098M**
Identify genetic risk factors for coronary collaterals in a genetically diverse population. Z. Liu1, L. Wang2-3, N. Vasudev4, P.J. Goldschmidt-Clermont1, M.A. Pericak-Vance1,2, D.M. Seo2, G.W. Beecham1,2, 1) Human Genetics, John T. Macdonald Foundation, University of Miami, Miami, FL; 2) John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL; 3) Division of Cardiology, Miller School of Medicine, University of Miami, Miami, FL.

Coronary artery disease (CAD) is the leading cause of death worldwide. Important risk factors for CAD have been identified, but they fail to explain why some occlusive CAD patients with fewer or no natural bypass collateral vessels have a lower rate of survival compared with patients with greater number of such vessels. Genetic components in combination with multiple other factors may play roles, but the genetic pre-determinant for developing coronary collateral vessels in patients is unclear, mostly due to study populations too small to represent all CAD patients. Under-represented populations, especially Hispanics, will provide additional information to study the genetic contribution to coronary collaterals formation. Toward this end, we ascertained 879 occlusive CAD patients from South Florida region and conducted a genome-wide association study. These patient samples represented the large diversity of populations in South Florida region, including Hispanic, African American and Caucasian populations. Each patient had coronary stenosis greater than 50% of at least one coronary artery. Among these occlusive CAD patients, 473 had collaterals and 406 did not. We genotyped patient samples on Affymetrix 6.0 platform and performed standard QC, followed by single genetic variant test using covariates of gender and principal components as previously identified. For single genetic variant test, SNPs rs16883071 (p = 5.58E-06), rs17495254 (p = 8.75E-06), rs10894957 (p = 8.86E-06) showed top association with collateral binary trait. Furthermore, we performed gene-based test with VEGAS, and the top associated gene was C1QTNF9B (p = 2.00E-04), which passed Bonferroni corrected genome-wide significant threshold (2.8E-06); other top associated genes were SNAI3 (p = 5.31E-05), CTU2 (p = 1.53E-04), RNF166 (p = 1.56E-04) and PIEZ1 (p = 2.2E-04). In conclusion, we identified potential common SNPs and genes that are associated with collaterals in a genetically diverse population. We will further investigate genetic components of rare variants in coronary collateral vessels. Understanding the genetic components of coronary collaterals allows us to understand why there are differences between occlusive CAD patients of developing a sufficient collateral circulation, therefore make it possible to develop adjuvant treatment to increase CAD patient survival rate through manipulating collateral process.
2099S

Meta-analysis of variants on the Exome Chip in 120,700 individuals of European ancestry identifies multiple rare and common loci for blood pressure and nicotine dependence.

A.Y. Chu1, A.C. Morrison2, Y. Lu1, S. Weiss3, P.L. Auer4, X. Guo5, A.Y. Chu1, J. Jakobsdottir6, J.C. Bis7, W. Zhao8, K. Tsiose9, N. Anjin10, H. Mei11, C. Newton-Chen12, W. Palmas13, Y. Liu14, R.J. Loos15, T.L. Edwards16, U. Volke7, M. Frommer15, C.M. Djuric17,3,5,6, S.B. Semsarian17,18,5,6, E. Ehrat19,20, V. Guðnason21,22, D. Chasan-Taberle17,18, D. Levy1, CHARGE Plus Exome Chip Blood Pressure Consortium. 1) Framingham Heart Study, NHLBI/NIH, Framingham, MA, 01702, USA; 2) Division of Statistical Genomics, Department of Genomic Medicine and Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63110, USA; 3) Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA; 4) Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98195, USA; 5) Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; 6) Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, TX 77030, USA; 7) The Charles Bronfman Institute for Personalized Medicine, Icahn Sinai School of Medicine at Mount Sinai, New York, NY 10029, USA; 8) Ernst-Moritz-Arndt-University of Greifswald, Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, Greifswald, Germany; 9) School of Public Health, University of Wisconsin-Milwaukee, Milwaukee WI 53211, USA; 10) Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA; 11) Division of Preventive Medicine, Brigham and Women’s Hospital, Boston, MA 02215, USA; 12) Center for Human Genetics Research, Vanderbilt University, Nashville, TN 37212, USA; 13) Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, 3015 CN Rotterdam, the Netherlands; 14) Broad Institute of Harvard and MIT, Massachusetts General Hospital, Boston, MA 02124, USA; 15) Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; 16) Institute of Molecular Medicine and School of Public Health Division of Epidemiology Human Genetics Research, Vanderbilt Institute of Imaging Science and Technology, University of Tennessee, Nashville, TN 37212, USA; 17) Center of Biostatistics & Bioinformatics, Department of Medicine, University of Mississippi Medical Center, MS 39216, USA; 18) Icelandic Heart Association, Kopavogur, Iceland; 19) University of Southern California School of Medicine, University of Southern California School of Medicine at Los Angeles, CA 90048, USA; 20) Cardiology, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil, 4, 1211 Genève 14 Switzerland; 21) University of Iceland, Reykjavik, Iceland.

Introduction: High blood pressure (BP) or hypertension is a major risk factor for cardiovascular disease. Previous genome-wide association studies (GWAS) have identified >60 genetic loci for BP. However, most of the published variants are non-coding and explain a small proportion of BP heritability. Goal: To investigate the role of coding variation on interindividual variability in BP (systolic, diastolic, mean arterial, and pulse blood pressure).

Methods: We genotyped ~250,000 mostly coding variants in 120,700 individuals of European ancestry using the Infinium HumanExome BeadChip. Association analyses were conducted using the seqMeta software in each of 15 participating studies and results were combined by fixed effect meta-analysis. Array-wide statistical significance was defined as p < 5 x 10-7. Conditional analyses were performed to identify independent signals at known or novel loci. Our aim was to identify larger effect size rare variants (minor allele frequency < 1% and frequency < 0.05%) than those identified in GWAS. Significant novel associations (p < 5 x 10-7) were discovered for common variants in several genes with influences on kidney function, erythrocyte measures, and cardiomyopathy: RGL3 (missense), PCNX3 (intronic), TBX2 (intronic), TNNT3 (intronic), PHF19 (upstream), RABEPK (missense), and PRDM16 (missense). Among low-frequency and rare variants, novel associations were found at NPR1 (MAF 0.013, missense, p = 7 × 10-7 MAP), IVL (4,002), we identified a novel locus represented by rs2282679 at the GC gene (4q12.13) with (β = -0.13, p = 3.0 × 10-2) in Sikhs. The IVL gene, located on chromosome 1q21.3, encodes involucrin. Both IVL and GC genes are implicated in the synthesis of vitamin D. Genetic variation in the IVL gene is associated with 25(OH)D levels. Our results also confirmed a previously reported association with 25(OH)D levels (β = 0.10, p = 3.1 x 10-4) after adjustment with age, sex and type 2 diabetes status. These findings are currently being replicated in other independent larger datasets. Our results also confirmed a previously reported association with 25(OH)D deficiency with the risk of PAD up to 10 fold and there is evidence that there may be different pathways contributing to the disease in smokers compared to non-smokers, in whom the major risk factor is diabetes. There are two established loci for PAD, a genome wide significant association on 9p21 and a single variant with borderline significance (p=5e-8) signal near CHRNA3 and another signal (p=5e-7) in the 9p21.3 region. The lack of additional associations may be due to the highly polygenic architecture of PAD. To identify variants specifically associated with PAD in current-smokers, we conducted a smoking interaction meta-analysis and also performed a smoking interaction analysis to identify variants that may interact with smoking status to modify the risk of PAD. We combined summary statistics in a fixed effects meta-analysis for 2,356,286 SNPs in 4,544 PAD cases and 30,404 PAD controls - of whom 3,355 cases and 20,212 controls were smokers. We also performed an interaction analysis of allelic effects in smokers vs. non-smokers. The top SNP in smokers was rs1051730 (p=7.2e-7), near CHRNA3, that is established for PAD and is associated with nicotine dependence and smoking quantity. The top SNP was reported by Liu et al. (2013) and was associated with PAD in smokers (OR=1.18, p=2.5e-6). The smoking interaction analysis highlighted signals that had large effect sizes in smokers but little or no effect in non-smokers, rs2076516 (OR=4.6, p=1.4e-7, phet=3.5e-4), near BDNF, rs2076516 (OR=4.5, p=7.8e-7, phet=1.3e-5), near ADAMTS17 and rs11214800 (OR=1.3, p=2.6e-4, phet=2.7e-5), near HTRA3 were associated in PAD in smokers and showed evidence for interaction with smoking status. The expression of BDNF is directly affected by smoking and cigarette smoking, while SNPs near ADAMTS17 have been shown to interact with smoking status to affect blood pressure and SNPs near HTRA3 have been associated with smoking quantity. Our analyses stratified on smoking status highlight biologically plausible signals related to smoking and suggests that cigarette smoking may help to elucidate differences in disease aetiology in smokers compared to non-smokers.
2102M
A Comprehensive 1000 Genomes-based GWAS of Coronary Artery Disease. H.-H. Won1,2,3, K. Mohlke1,2, K. Xu1,2, S. Ganesh1,2,3, C.J. Willer1,2,4, 1 Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, Michigan, USA; 2) First Hospital, Peking University, Beijing, China; 3) HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway; 4) Centre for Genomic Sciences, Jockey Club Building for Interdisciplinary Research, State Key Laboratory of Brain and Cognitive Sciences, and Department of Psychiatry, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China; 5) Department of Medicine, The University of Hong Kong, Hong Kong SAR, China; 6) Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 7) Third Hospital, Peking University, Beijing, China; 8) Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA; 9) Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA.

Cardiovascular disease (CVD) is the leading cause of death globally. Circulating blood lipids are heritable, treatable, risk factors for cardiovascular disease. Over the past six years, genome-wide association studies (GWAS) have identified 157 loci associated with plasma lipid levels. However, most of the GWAS were conducted on individuals of European descent, and so the degree to which knowledge gained from these studies is applicable to other populations has not been systematically investigated.

To have a better understanding of the global genetic variants associated with blood lipid levels, we conducted an exome-wide association study in samples from the Chinese population. In this study, we tried to answer three questions. (1) Whether the known lipids associated genes/variants were associated with blood lipids levels and have similar effect size in Asian samples? (2) Are there any novel genes/variants, especially low-frequency and rare variants, that are associated with blood lipid levels? (3) Whether the lipids-associated genes were affected by positive selection during evolution?

In the analysis, many of the known lipids associated loci/variants were also associated with blood lipids in our samples, and have similar effect size. Heritability analysis indicated that low-frequency, high-penetrance variants contribute a considerable proportion of the heritability, however, we could not identify any novel genes or variants that reached exome-wide significance. The rare variant, p.Leu3548Ile in DNAH17 approached exome-wide significance (P = 3×10^{-7}) but we could not identify sufficient copies of this variant in three additional East Asian studies to assess association in an independent cohort. We also identified several Asian-specific variants in known lipids associated loci (APOA5, APOB and PCSK9) that also reached genome-wide significant level. Using data from the 1000 Genomes Project, we found that lipid genes have a higher Fst between Asian and European populations than expected by chance. By estimating Fay and Wu’s H, we also find these lipid genes tend to have a higher proportion of derived alleles in the Asian population. These suggested some lipids genes diverged quickly in Asian populations and may be driven by positive selection during evolution.
Chronic venous disease (CVD) is one of the most common vascular abnormalities, especially in Northern and Western Europe, with a prevalence of up to 20% in Northern and Western Europe. According to the CEAP guidelines, CVD comprises those clinical entities that are characterized by visible venous ectasies but which are not associated with an identifiable mechanism of venous dysfunction. CVD represents a summary term comprising superficial venous malformations, especially in Northern and Western Europe, with a prevalence of up to 20% in Northern and Western Europe. According to the CEAP guidelines, CVD comprises those clinical entities that are characterized by visible venous ectasies but which are not associated with an identifiable mechanism of venous dysfunction. CVD represents a summary term comprising superficial venous malformations, especially in Northern and Western Europe, with a prevalence of up to 20% in Northern and Western Europe. According to the CEAP guidelines, CVD comprises those clinical entities that are characterized by visible venous ectasies but which are not associated with an identifiable mechanism of venous dysfunction. CVD represents a summary term comprising superficial venous malformations, especially in Northern and Western Europe, with a prevalence of up to 20% in Northern and Western Europe.

Methods and Results: Participants from the Northern Swedish GLACIER Study were genotyped with the MetaboChip array. Main effects of variants mapping to previously established lipid loci (n=3,898 SNPs) were estimated with linear regression models by fitting the lipid levels at follow-up as the dependent variables and conditioning on the corresponding baseline lipid measure (N=3,492 for total cholesterol change (ΔTC) and N=2,209 for triglyceride change (ΔTG)). We sought replication in three other Swedish studies (MDC, PIVUS, ULSAM; total N~4,000) and conducted an in silico look-up for the top ranking lipid-change associated variants in the CARDIoGRAMplusC4D Consortium (N~190,000) to explore whether these variants also associate with coronary artery disease (CAD). In total, 227 variants (in 11 loci) had statistically significant (P<0.05) pooled effects for lipid changes after meta-analyzing the 4 cohorts’ effect estimates; seven of these loci have not been discovered in previous cross-sectional meta-analyses. A novel genetic association with CAD was observed for rs2000999 in haptoglobin-related protein (HPR), which was associated with ΔTC (P=5.1×10⁻⁸) and CAD (P=3.6×10⁻⁸). Conclusion: In the GLACIER Study, we found 4 genome-wide significant hits for TG changes. Furthermore, through meta-analysis, we identified 11 regions associated with TC and TG changes and one novel locus, HPR, for CAD.
Familial aggregation of CCC suggests that there might be a genetic compon-ent to disease susceptibility. We hypothesized that there are immunoglobulins, cytokines, chemokines, integrins. There are genetically relevant genes and proteins in the myocardial tissue of CCC patients and controls. There are good relationships between miRNA expression levels and protein expression data of early expressing cardiac genes were obtained in 6 Caucasian families, with 14 affected members with CHD.

Stroke/Richard Lewar Centre of Excellence, University of Toronto. University of Toronto, Toronto, Ontario, M5S 1A8, Canada; 9) Heart & Stroke/Richard Lewar Centre of Excellence, University of Toronto.

The genetic etiology is known in less than 20% of congenital heart disease (CHD). The objective of our study was to identify genes enriched for novel and rare damaging variants in pathways crucial for cardiovascular development. We performed whole exome sequencing (Agilent SureSelect V3, Illumina HiSeq) in 6 Caucasian families, with 14 affected members with CHD, 9 non-affected members and 147 unrelated Caucasian CHD subjects. Gene and protein expression data of early expressing cardiac genes were obtained and genes with novel/rare variants on highly conserved locations in these genes that were predicted damaging were prioritized. Analysis of these 56 prioritized genes using DAVID (GO-term enrichment analysis) identified strong functional interactions amongst 9 of these genes for extra-cellular matrix (ECM) function. Novel/rare pathogenic variants, inherited and de novo, in these 9 genes were seen in 12.7% of the CHD cases, 3 cases showed co-segregation of mutations in more than one gene. Expression knockdown of individual genes in zebrafish embryos caused abnormal cardiac phenotypes. More severe phenotypes were seen with concurrent knockdown of more than one gene supporting a multigenic etiology for CHD. Our study identifies novel variants in functionally interacting ECM genes that contribute individually or in combination to CHD.

Chagas disease, due to Trypanosoma cruzi, occurs exclusively in the Americas, particularly in poor, rural areas of Central America, and South America. An estimated 300,000 new cases and 50,000 fatalities occur per year. Chronic Chagas disease cardiomyopathy (CCC) is an inflammatory cardiomyopathy that affects approximately 30% of infected individuals. Familial aggregation of CCC suggests that there might be a genetic component to disease susceptibility. We hypothesized that there are immunoglobulins, cytokines, chemokines, integrins. There are genetically relevant genes and proteins in the myocardial tissue of CCC patients and controls. There are good relationships between miRNA expression levels and protein expression data of early expressing cardiac genes were obtained in 6 Caucasian families, with 14 affected members with CHD, 9 non-affected members and 147 unrelated Caucasian CHD subjects. Gene and protein expression data of early expressing cardiac genes were obtained and genes with novel/rare variants on highly conserved locations in these genes that were predicted damaging were prioritized. Analysis of these 56 prioritized genes using DAVID (GO-term enrichment analysis) identified strong functional interactions amongst 9 of these genes for extra-cellular matrix (ECM) function. Novel/rare pathogenic variants, inherited and de novo, in these 9 genes were seen in 12.7% of the CHD cases, 3 cases showed co-segregation of mutations in more than one gene. Expression knockdown of individual genes in zebrafish embryos caused abnormal cardiac phenotypes. More severe phenotypes were seen with concurrent knockdown of more than one gene supporting a multigenic etiology for CHD. Our study identifies novel variants in functionally interacting ECM genes that contribute individually or in combination to CHD.
Identification of pathwaygenes associated with Bicuspid Aortic Valve in three Caucasian Cohorts. M. Heydarpour1, J.G. Seidman2, S. Prakash2, D. Milewicz2, Y. Bosse3, G. Limongelli3, S.C. Body3, the Bicuspid Aortic Valve Consortium (BAVCOn). 1) Brigham and Women’s Hospital, Boston, MA; 2) Harvard Medical School, Boston, MA; 3) University of Texas, Houston, TX; 4) University of Laval, Quebec, Canada; 5) Monaldi Hospital, Naples, Italy.

OBJECTIVE: Bicuspid aortic valve (BAV) is observed in 0.6-1% of humans and is associated with premature aortic stenosis, aortic aneurysm and other congenital cardiac anomalies. BAV has been associated with high-penetration rare variants in a few families, but the majority of BAV is sporadic with unidentified inheritance. The embryologic molecular mechanisms causing BAV are unknown. The objective of this study was to identify pathways associated with sporadic BAV. METHODS: We genotyped 456 adult US Caucasians with BAV (339 males, 117 females) using Illumina Omni2.5 5,216 Caucasians from three dbGaP cohorts were used as controls and typed with either Omni2.5 or Omni5 arrays. Quality control of the genotype data were performed using PLINK. After QC, we examined 603,499 SNPs mapped to 500kbp regions covering 16,775 genes. We used i-GSEA4GWAS and i-CSNPathway software (Zhang et al. 2010, 2011) to identify gene sets associated with BAV. This method integrates gene set enrichment, linkage disequilibrium analysis, functional SNP annotation and pathway-based analysis. Permutation test and False Discovery Rate (FDR) were used for multiple testing corrections. RESULTS: Of 1,843 gene sets selected using canonical pathways, 45 were associated with BAV (FDR<0.05). Four pathways - Small GTPase mediated signal transduction (SGMST), Ras protein signal transduction (RPST), Cell recognition (CR), and Regulation of G-protein coupled receptor signaling (RGPCRPS), were identified as candidate pathways for BAV. The top pathway SGMST contained 44 significant genes (FDR<0.001) of which, the SNP rs2842895 in the regulatory region (promoter) had high-LD (r>0.8) with rs4585612. Both SNPs were present in the region of gene RREB1. Two other SNPs (rs3360 and rs8096892) near OPCML and RG511 were also identified as candidate causal SNPs, respectively. CONCLUSION: We identified potential candidate genes for BAV in the SGMST, RPST, CR, and RGPCRPS pathways. The top candidate gene RREB1 is a zinc finger protein involved in Ras/Raf-mediated cell differentiation by enhancing calcitonin expression. RREB1 is essential to reduce cell-cell adhesion when epithelial cells are undergoing the dynamic changes for cell shape (Melani et al. 2008). RREB1 may be essential to reduce cell-cell adhesion when epithelial cells are undergoing the dynamic changes for cell shape (Melani et al. 2008). RREB1 may be essential to reduce cell-cell adhesion when epithelial cells are undergoing the dynamic changes for cell shape (Melani et al. 2008). RREB1 may be essential to reduce cell-cell adhesion when epithelial cells are undergoing the dynamic changes for cell shape (Melani et al. 2008). RREB1 may be essential to reduce cell-cell adhesion when epithelial cells are undergoing the dynamic changes for cell shape (Melani et al. 2008).
2112M A Gene Network approach to Rare Variant analysis. T.G. Richardson1, N.J. Timpson1, C. Campbell1, T.R. Gaunt1. 1) MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol; Bristol, United Kingdom; 2) Intelligent Systems Laboratory, University of Bristol, Bristol, United Kingdom.

Background: Current endeavours in rare variant analysis are typically undertaken after mutations are found in individual genes. However, we undertook a novel approach to rare variant analysis by utilising knowledge regarding interactions and relationships between sets of genes. Methods: Using the resource STRINGdb we identified 263,666 human protein-protein interactions and assembled evidence for 247,132 of these interactions to construct the 2.099 Gene Networks. The variant effect predictor (VEP) was used enrich our analysis by identifying 116,621 unique variants in the UK10K sequence data predicted to be nonsynonymous. Variants within the same network were collapsed together and analysed with lipid traits using the sequence kernel association test (SKAT) after applying a threshold of 5% and 1% minor allele frequency. Results: Three networks provided evidence of association after correcting for multiple comparisons and were evaluated further using single gene analyses and permutation tests. The most statistically robust evidence found that variants from within 3 genes (PDE5A, MRV11 and NRP1), whose products were known to interact, collectively contributed to an observed signal with LDLc levels (P = 5.30 × 10-6). Conclusion: We have undertaken a novel approach to rare variant analysis which examines the combined effect of variants from across genes which have been shown to interact according to experimental evidence. We identified networks which provide evidence of association with lipid traits which suggest further analysis of polygenic effects across networks and pathways should be prioritised in better understanding the genetic architecture of complex disease.

2113S WES reporting of mutations from cardiovascular "actionable" genes in clinical practice: a look for key knowledge gaps. A. PINARD1, N. HANNA2,3, C. GUEN3,2, L. FAIVRE2,3, G. JONDEAU2,4, C. BOILEAU2,4, C. BEROUD2,5,6, G. COLLOD-BEROUD2,5,6. 1) Inserm UMR_S910, Marseille, France; 2) Aix Marseille Université, GMGF, 13385, Marseille, France; 3) Department of Genetics, Bichat University hospital, 46 rue Henri Huchard, 75018 Paris France; 4) Inserm U1148, Bichat University hospital, 46 rue Henri Huchard, 75018 Paris; 5) Centre de Référence Anomalies du Développement et Syndromes Malformatifs et FHU TRANSLAD, CHU Dijon, Dijon, France; 6) EA 6271 Génétique des Anomalies du Développement, Université de Bourgogne, Dijon, France; 7) Cardiology Department, Bichat University Hospital, 46 rue Henri Huchard, 75018 Paris; 8) APHM, Hôpital d’Enfants de la Timone, Département de Génétique Médicale et de Biologie Cellulaire, 13385, Marseille, France.

High throughput next generation sequencing (NGS) such as whole exome sequencing (WES) are being rapidly integrated into clinical practice. The use of these techniques leads to unexpected findings for which decisions about the implications of the identified variant need to be evaluated. In the next generation of cardiac genetic studies, we hypothesize that a systematic database of validated cardiovascular genes can improve survival and quality of life. In an effort to standardize information regarding mutations in the FBNI gene, we developed in 1995 a locus specific database with the generic system called Universal Mutation Database (UMD). Subsequently, databases for TGFBR2 and FBNI were created. All are now recognized as international references. To be exhaustive and facilitate NGS analysis, we have now developed databases for the ACTA2, SMA3D, MYH11 and MYLK genes. They contain all known mutations collected from literature and through direct collaborations with diagnostic laboratories. For FBNI (320 entrees), TGFBR2 (300 entrees), and the new TGFBR1 (125 entrees), ACTA2 (203 entrees), SMA3D (61 entrees), MYH11 (44 entries) and MYLK (13 entries). Each mutation is annotated at gene, protein and clinical levels. Several tools are also available for studying and extracting data of interest as well as complex algorithms to predict pathogenicity of missense variations; consequences of variations on splicing signals, or to search for genotype-phenotype correlations. These databases are updated regularly and curated by experts. They are accessible at: http:// www.umd.be/. UMD and TGFBR2 databases are available at University of Michigan and UCSD libraries: (300 entries) and the new database with the generic system called Universal Mutation Database (UMD) are now recognized as international references. To be exhaustive and facilitate NGS analysis, we have now developed databases for the ACTA2, SMA3D, MYH11 and MYLK genes. They contain all known mutations collected from literature and through direct collaborations with diagnostic laboratories. For FBNI (320 entries), TGFBR2 (300 entries), and the new TGFBR1 (125 entries), ACTA2 (203 entries), SMA3D (61 entries), MYH11 (44 entries) and MYLK (13 entries). Each mutation is annotated at gene, protein and clinical levels. Several tools are also available for studying and extracting data of interest as well as complex algorithms to predict pathogenicity of missense variations; consequences of variations on splicing signals, or to search for genotype-phenotype correlations. These databases are updated regularly and curated by experts. They are accessible at: http://www.umd.be/. UMD and TGFBR2 databases are available at University of Michigan and UCSD libraries (300 entries) and the new database with the generic system called Universal Mutation Database (UMD) are now recognized as international references.

2114M Sequencing of 2,000 Norwegians to evaluate genetic architecture of lipid and MI-associated variants. C. Willer1,2, W. Zhu1, O. Holmen1, H. Zhang1, J. Chen1, D. Hovelson1, M. Boehnke1, G. Abecasis2, K. Hveem1. 1) Int Med & Human Genetics, University of Michigan, Ann Arbor, MI; 2) Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 3) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 4) Center for Cardiovascular Research, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway; 5) Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI.

Genome-wide studies have been successful at identifying novel loci and genes associated with blood lipid levels and coronary artery disease. However, these approaches have not adequately tested the role of low frequency variants. To attempt to identify low frequency variants with large effect on risk of myocardial infarction or impact on blood lipid levels, we performed low-pass whole genome sequencing of 2,000 individuals from Nord-Trondelag, Norway. We selected individuals with medical-record confirmed MI with early-onset (< 58 years for men and < 68 years for women). For each MI-case, we selected one sex- and birth-year matched healthy control from 70,300 potential controls by excluding individuals with a variety of cardiovascular and metabolic conditions. We opted to perform lower-pass whole genome sequencing for the following reasons: i. increased power relative to higher-depth whole genome sequencing in fewer samples; ii. utility of whole genome sequence data for imputation into GWAS samples relative to exome-sequence data and iii. to assess low frequency non-coding variation. Our first data freeze of 1,237 individuals (626 MI cases and 611 controls) did not identify any novel genes or genetic variants associated with MI, but did identify association at several known loci. Furthermore, we identified eight low-frequency non-coding variants (freq < 5%) with P < 5×10-7 (including chr4:76188636 near PAR1 with P = 7×10-16 with total cholesterol) that were not previously evaluated in large-scale GWAS studies and may represent novel low frequency large-effect loci.

2115S Whole exome sequencing highlights the importance of the CRELD1 interactome in atrioventricular septal defect in Down syndrome. C.M. Ackerman1, H. Li2, D. Klinedinst1, R. Polik1, S. Blackshaw1, H. Corbitt1, R.H. Reeves2, C.L. Maslen1. 1) Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR; 2) Department of Physiology and the Institute for Genetic Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD.

Atrioventricular septal defect (AVSD) is a congenital heart defect (CHD) frequently associated with Down syndrome (DS). Children with DS account for 65% of all cases of AVSD and have a 2000-fold increased risk of the defect. However, since no specific genes nor proteins have been identified that interact with CRELD1 contribute to the risk of AVSD in DS. The study cohort was comprised of individuals with DS and a complete AVSD cases (cases) and individuals with DS and a normal echocardiogram (controls), including individuals of European Ancestry and African Americans. Missense mutations were characterized as damaging, benign, or unknown based on PolyPhen2 and MutPred analyses.

Of the 33 genes we evaluated, 18 genes had at least one damaging mutation. The majority of these genes have not been implicated in the pathogenesis of CHD, with the exception of PTPN11, a gene in which we previously reported mutations in individuals with non-syndromic AVSD. We detected 23 damaging variants in cases and only 3 in controls (p=0.0001), with the majority being missense variants. Other case-specific variants identified included one recurrent frameshift mutation in RAB26 and a premature termination codon in EDIL3. There was no significant skewing of variants between genders, and inclusion of African American-specific variants did not falsely elevate the percentage of case-specific deleterious variants. We conclude that further studies of DS cases with damaging variant in a CRELD1 pathway gene. The fact that we see mutations in the same genes in both euploid and DS with AVSD supports our hypothesis that the same genes involved in the pathogenesis of AVSD in DS, also contribute to AVSD in the euploid population. Further functional analysis of these CRELD1 pathway variants will provide insight into the mechanisms that underlie heart defects during development.
2116M
Gene-centric association tests applied to cardiovascular disease using whole genome sequencing. M.A. Almeida1, J. Peralta1,2, J.W. Kent1, T.M. Teslovich3, G. Jun3, C. Fuchsberger3, A. Wood2, A. Manning2, T.M. Frayling1, P. Cingolani3, T.W. Blackwell3, R. Sladek6, T.D. Dyer3, A.G. Com- muzie1, H.H.H. Goring1, L. Almasy1, M.C. Mahaney1, D.M. Lehman3, J.E. Curran1, G. Abecasis1, R. Duggirala1, J. Blangero1. 1) Genetics, Texas Biomedical Research Institute, San Antonio, TX; 2) Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, WA, Australia; 3) University of Michigan, Ann Arbor, MI, USA; 4) University of Exeter, Exeter, United Kingdom; 5) Broad Institute, Boston, MA, USA; 6) McGill University, Montreal, Canada; 7) Montreal Diabetes Research Institute, Montreal, Canada; 8) University of Texas Health Center at San Antonio, San Antonio, TX, USA.

The advent of whole genome sequencing provides a unique opportunity to improve our understanding of the genetic variation underlying complex diseases. As part of the T2D-GENES Consortium, we have directly sequenced 590 individuals (and accurately imputed another 448 members) from 20 large Mexican American pedigrees aiming the investigation of rare genetic variants contribution to the type 2 diabetes development. Those individuals are part of the SAFS (San Antonio Family Study) and have been extensively phenotyped during the 25 years of this project. The immense number of identified SNVs imposes new statistical and analytical barriers that require the development of alternative approaches for screening of potential causal variants and pathways. We employed a variance component-based single degree-of-freedom test using an empirical gene-specific genetic relationship matrix (GRM) as the focal covariance kernel. The empirical gene-specific GRM (the GSGRM) utilizes any set of chosen variants identified in a gene by gene pathways of interest and provides the correlation of the dosage vectors between individuals. The efficiency of this new focal kernel for explaining phenotypic variance is tested by the use of a single-degree freedom likelihood ratio test. Gene pathway definitions were obtained from the latest CytoScape database release and a GSGRM was estimated for each one of 1191 genes that constitute 9 CVD (Cardiovascular Diseases) related pathways. Each empirical kernels was tested against a set of CVD-related traits collected in the SAFS cohort. As might be expected, our top association was for the coronary artery disease of the apolipoprotein B gene (this gene (p-value = 2.1×10^-5); this is known case of cys-regulation and serves as a positive control of our approach. We also observed a strong association between the gene PCSK9 and CVD itself (p-value = 8.2×10^-5). The PCSK9 gene product plays a central role in cholesterol metabolism and is a major target for the new generation of anti-ldip drugs. Each variant in the gene was independently tested and two non-coding variants achieved genome-wide significance and a non-synonymous variant, rs11583680, located in first exon of this gene (p-value = 1.6×10^-4) was also identified. All non-synonymous variants achieved genome-wide significance and a non-synonymous variant, rs11583680, located in first exon of this gene (p-value = 1.6×10^-4), was also selected for further analysis. CM most often occurs as an isolated and sporadic feature. Syndromic forms include Sturge-Weber syndrome, Klippel-Trenaunay syndrome and Parkes Weber syndrome. There is often associated soft tissue and/or bony overgrowth. In the autosomal dominant capillary malformation-arteriovenous malformation (CM-AVM), CMs increase in number with age. A third of the CM-AVM patients have an associated fast-flow anomaly, most often located in the head and neck region. CM-AVM is caused by haploinsufficiency, most likely combined with a missense second-hit, p.120-RasGAP, the protein product of RASA1, a series of Sturge-Weber syndrome and sporadic CMs were shown to harbor a non-synonymous somatic single-nucleotide variant in GNAQ, encoding guanine nucleotide binding protein. We assessed 33 fresh-frozen lesions by both Sanger sequencing on cDNA and allelic specific PCR on genomic DNA, for the presence of the c.548G>A (p. Arg183Gln) hotspot mutation. In total, 11 lesions with mutations were detected in 9 out of 22 patients was analyzed by NGS, searching for genes of predisposition. DNA from blood samples of 22 patients with Balkan Endemic Nephropathy (BEN) were found further evidence for the presence of the GNAQ hotspot mutation. Targeted massive parallel sequencing of the complete coding sequences of GNAQ and RASA1 was therefore performed on the 16 negative lesions using Ion AmpliSeq Panel on PGM. Vertical coverage varied from 400x to 1550x. We did not find additional mutations. Six tissues with CM-AVM (16,54,5%); 5/5 (100%) of Sturge-Weber Syndrome lesions and 13/16 (81.2%) CMs. This mutation was not detected in the fifteen other lesions with other CM phenotypes (45.5%). This could be explained by the fact that normal CMs with low-mitogenic mutator cells in the lesionated region. Somatic Activating GNAQ Mutations are Frequent in Capillary Malformation (Port-Wine Stains). M. Amyere1, N. Revencu2, A. Dompmartin2, R. Healer1, L. Boon1, M. Vlkulla1. 1) Laboratory of Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Centre for Human Genetics, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium; 3) Department of Pathology, Université de Caen Basse Normandie, CHU Caen, France; 4) Centre for Vascular Anomalies, Division of Plastic Surgery, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium.

Capillary malformations are localized defects of vascular development that are classified into capillary, venous, arterial, lymphatic, and combined anomalies. Capillary Malformation (CM) is the most common affecting cutaneous capillary vessels in 0.3% of normals. These lesions can be localized or diffuse. CM most often occurs as an isolated and sporadic feature. Syndromic forms include Sturge-Weber syndrome, Klippel-Trenaunay syndrome and Parkes Weber syndrome. There is often associated soft tissue and/or bony overgrowth. In the autosomal dominant capillary malformation-arteriovenous malformation (CM-AVM), CMs increase in number with age. A third of the CM-AVM patients have an associated fast-flow anomaly, most often located in the head and neck region. CM-AVM is caused by haploinsufficiency, most likely combined with a missense second-hit, p.120-RasGAP, the protein product of RASA1. A series of Sturge-Weber syndrome and sporadic CMs were shown to harbor a non-synonymous somatic single-nucleotide variant in GNAQ, encoding guanine nucleotide binding protein. We assessed 33 fresh-frozen lesions by both Sanger sequencing on cDNA and allelic specific PCR on genomic DNA, for the presence of the c.548G>A (p. Arg183Gln) hotspot mutation. In total, 11 lesions with mutations were detected in 9 out of 22 patients was analyzed by NGS, searching for genes of predisposition. DNA from blood samples of 22 patients with Balkan Endemic Nephropathy (BEN) were found further evidence for the presence of the GNAQ hotspot mutation. Targeted massive parallel sequencing of the complete coding sequences of GNAQ and RASA1 was therefore performed on the 16 negative lesions using Ion AmpliSeq Panel on PGM. Vertical coverage varied from 400x to 1550x. We did not find additional mutations. Six tissues with CM-AVM (16,54,5%); 5/5 (100%) of Sturge-Weber Syndrome lesions and 13/16 (81.2%) CMs. This mutation was not detected in the fifteen other lesions with other CM phenotypes (45.5%). This could be explained by the fact that normal CMs with low-mitogenic mutator cells in the lesionated region. Somatic Activating GNAQ Mutations are Frequent in Capillary Malformation (Port-Wine Stains). M. Amyere1, N. Revencu2, A. Dompmartin2, R. Healer1, L. Boon1, M. Vlkulla1. 1) Laboratory of Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Centre for Human Genetics, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium; 3) Department of Pathology, Université de Caen Basse Normandie, CHU Caen, France; 4) Centre for Vascular Anomalies, Division of Plastic Surgery, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium.

Somatic Activating GNAQ Mutations are Frequent in Capillary Malformations (Port-Wine Stains). M. Amyere1, N. Revencu2, A. Dompmartin2, R. Healer1, L. Boon1, M. Vlkulla1. 1) Laboratory of Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium; 2) Centre for Human Genetics, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium; 3) Department of Pathology, Université de Caen Basse Normandie, CHU Caen, France; 4) Centre for Vascular Anomalies, Division of Plastic Surgery, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium.

2117S
Molecular-genetic factors in normal and pathological angiogenesis. I.I. Dimova1, V.G. Ojcov2, A. Makanya3, R. Hlushchuk4, M. Polenakovic5, M.G. Mihailova-Hristova1, R. Vazharova5, L. Balabanski2, S. Ivanov2, D.I. Toncheva1. 1) Medical University Sofia, Sofia, Bulgaria; 2) Institute of Anatomy, University of Bern, Switzerland; 3) Faculty of Medicine, University of Nis; 4) Macedonian Academy of Sciences and Arts; 5) Laboratory Genomics, Malinov Hospital.

Angiogenesis is involved in many physiological processes, but also as a hallmark in the pathology of many diseases (cancer, ischemia, atherosclerosis, inflammatory diseases), in wound healing and in tissue regeneration. We aimed in studying the molecular-genetic mechanisms of angiogenesis in normal and pathological conditions. Chicken area vasculosa was used as a model for investigation of essential molecular pathways - Notch, EphrinB2/ EphB4 and SDF-1/CXCR4 signaling. Glomerular sclerosis and interstitial fibrosis are characteristic angiogenetic pathological changes in kidneys of patients with Balkan Endemic Nephropathy (BEN). DNA from blood samples of 22 patients was analyzed by NGS, searching for genes of predisposition. We demonstrated that Notch inhibition disturbed vessel stability and led to pericyte detachment followed by extravasation of mononuclear cells. The mononuclear cells contributed to formation of transmural pillars with sustained intussusceptive angiogenesis resulting in a dense vascular plexus without concomitant vascular remodeling and maturation. Inhibition of ephrinB2 or EphB4 signaling induced some pericyte detachment and resulted in up-regulation of VEGF-Rs but with neither an angiogenic response nor recruitment of mononuclear cells. Notably, Tie-2 receptor was downregulated, and the chemotactic factors SDF-1/CXCR4 were up-regulated only due to the Notch inhibition. Using NGS we found additional patients with deleterious/damaging effects in three genes - CELA1, HSPG2 and KCNK5. These mutant genes in BEN patients encode proteins involved in basement membrane/extracellular matrix and vascular tone, tightly connected to processes of angiogenesis. The studies contributed to elucidation of the molecular mechanisms of normal angiogenesis and genetic factors predisposing to abnormal angiogenesis in glomerular sclerosis.
Abnormal lipid-lipoprotein levels are known to be the risk factors for cardiovascular disease. One of the genes involved in lipid metabolism is scavenger receptor class B type I (SCARB1), which has the main function in a selective uptake of cholesteryl esters in reverse cholesterol transport. Using Sanger sequencing, we examined 13 exons and exon-intron boundaries of SCARB1 in 95 African Blacks from Nigeria having extreme high-density lipoprotein cholesterol (HDL-C) levels. Sequencing analysis revealed a total of 83 variants (MAF ≤5%, n = 32; MAF >5%, n = 51). Common tagSNPs (MAF ≥5%) and uncommon/rare variants identified by sequencing (n = 78), plus 69 additional common tagSNPs from HapMap project covering the entire gene and 2 additional relevant variants from the literature were genotyped in the entire Nigerian sample (n = 789). A total of 137 successfully genotyped variants were then analyzed for their associations with various lipid traits. The most significant associations were observed with HDL-C levels. Single-site analysis identified 20 common SCARB1 variants significantly (P < 0.05) associated with one or more lipid traits, of which the most significant was the rs11057851 SNP associated with HDL-C levels (P = 0.0043). Optimal sequence kernel association test (SKAT-O) revealed also a nominally significant association of rare variants (MAF ≤1%, n = 23) with HDL-C levels (P = 0.0475). In summary, our results suggest the genetic contribution of SCARB1, both common and rare variants, in the lipid metabolism in humans.

**2120M**

Identification of novel genetic mutations causing familial hypercholesterolaemia among Saudi Arabian population. F.A. AL-ALLAF1,2, M. Athar1,2, Z. Abduljaleel1,2, A. Bouazzouali1,2, M.M. Taher1,2, R. Own1, A.F. AL-ALLAF1,2, I. AboMansoor1, Z. Azhar1, F.A. BA HAMMAM1,2, H. Abalkhair1, A. Alashwa1, S.S. Siddiqui1. 1) DEPARTMENT OF MEDICAL GENETICS, FACULTY OF MEDICINE, UMM AL-QURA UNIVERSITY, MAKKAH, Saudi Arabia; 2) Science and Technology Unit, Umm Al-Qura University, Makkah, Saudi Arabia; 3) Faculty of Medicine, Alfaisal University, Riyadh, Saudi Arabia; 4) King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia; 5) FACULTY OF DENTISTRY, UMM AL-QURA UNIVERSITY, MAKKAH, Saudi Arabia.

Familial hypercholesterolemia (FH) is a major risk factor for the development of Coronary heart diseases (CHD). The disease is hereditary in an autosomal dominant manner. Mutation(s) and/or deletion(s) in the LDL-receptor (LDLR) gene, in the apolipoprotein B-100 (ApoB) gene, or in the proprotein convertase subtilisin kexine 9 (PCSK9) gene are the most known causative mutations. The most accurate and unequivocal method for FH diagnosis is by molecular genetic testing of suspected cases and therefore this study aims to identify the genetic defects causing FH in Saudi population and to develop a diagnostic test for detecting such mutations. Forty four subjects were genotyped screened for mutations in the LDLR, ApoB and PCSK9 genes through direct PCR-sequencing using ABI 3500 genetic analyzer. We identified four common mutations in coding sequences of the LDLR gene, one mutation in ApoB gene and three mutations in PCSK9 gene. Among the identified LDLR gene mutations, two have been reported previously and two are novel. In addition, a base substitution in the splice acceptor site of LDLR intron 11 and a second mutation was also observed in LDLR intron 11. We have also identified a single previously reported heterozygous mutation in ApoB gene. Similarly we found three mutations in PCSK9 gene which are also reported earlier. Moreover, two heterozygous mutations c.658-7C>T and c.799+3A>G in the PCSK9 gene intron 4 and 5 respectively, have also been identified. This knowledge is important for optimizing cholesterol lowering therapies and mutation analysis diagnostic test. In addition, these data contribute to the understanding of the molecular basis of FH in Saudi Arabia.
Low-frequency coding variation in DNAH11 is associated with Sudden Cardiac Arrest among African Americans. J.A. Brody1, F. Asher2, N. Bihlmeyer2, X. Zhao1, A. Mak1, A.S. McCallion2, T. Lumley2, C.M. Sitlani1, K. Rice3, B.M. Psaty1,7,8, R. Lemaitre1, P.Y. Kwok1, F.D. Kolodgie3, D.S. Siscovick1,7,9, N. Sotoodehnia1,10. 1) Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA; 2) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 3) CVPath Institute, Inc., Gaithersburg, MD; 4) Cardiovascular Research Institute and Institute for Human Genetics, University of California, San Francisco, CA; 5) Department of Statistics, University of Auckland, Auckland, New Zealand; 6) Department of Biostatistics, University of Washington, Seattle, WA; 7) Departments of Medicine and Epidemiology, University of Washington, Seattle, WA; 8) Group Health Research Institute, Seattle, WA; 9) New York Academy of Medicine, New York, NY; 10) Division of Cardiology, University of Washington, Seattle, WA.

Background: Sudden cardiac arrest (SCA) is a major public health concern, particularly among African Americans where risk of cardiac arrest is higher than that of the general population, and survival is poor. While environmental factors clearly contribute to SCA risk, familial aggregation studies and molecular genetics studies of inherited arrhythmias suggest that genetic factors confer susceptibility to SCA in the general population.

Methods: We genotyped >240,000 common and low-frequency, nonsynonymous and splice-site variants with the Illumina HumanExome chip. Our study sample of African Americans consisted of 402 SCA cases and 4673 controls from two cohort studies, Atherosclerosis Risk in Communities study and the Cardiovascular Health Study, and one case-control study of SCA: the Cardiac Arrest Blood Study (CABS). Associations were modeled using Cox-proportional hazards models for the cohort studies or logistic regression for the case-control study and were combined using fixed effects meta-analysis. Common variants (minor allele frequency [MAF] ≥1%) were tested individually and rare and common variants were jointly modeled within gene using the Sequence Kernel Association Test (SKAT). We then examined our findings among 2405 cases and 2431 controls of European ancestry from the CABS study.

Results: Gene-based SKAT analyses identified Dynemin, axonemal, heavy chain 11 (DNAH11) to be a novel gene associated with SCA risk among African Americans. The findings for DNAH11 (p = 1.16E-6, cumulative MAF = 249%) reached significance after Bonferroni correction for the number of genes examined (N=17,574). The gene test result was primarily driven by two missense variants (rs147478795, His2795Asp, MAF = 1.5% and rs74667831, Cys2763Arg, MAF = 2.0%) in high linkage disequilibrium (r2 > 0.8). Rs147478795 (OR=4.34; 95% CI=2.43-6.25; p=7.95E-11) is likely to impact protein function as it was predicted to be deleterious by all four bioinformatic prediction algorithms examined. Among European descent individuals, rs147478795 is rare and was found in 1 case and no controls. The gene-based SKAT result showed no evidence of association in those of European ancestry (p=0.84, cumulative MAF=238%). Other mutations in DNAH11 cause ciliary dyskinesia type 7, which is characterized by situs abnormalities including dextrocardia. Conclusions: Through analyses of low-frequency coding variation, we have discovered a novel gene associated with SCA among African Americans.
2124M

Rare variant association analysis for plasma lipids in coronary artery disease susceptibility loci. H.K. Chheda1, P.P. Paatta1, E.T. Tikkanen1, S.M. McCarthy2, V.S. Salomaa3, R.D. Durbin3, A.P. Palotie1, T.A. Alttoa, S.R. Ripatha,1,2,6 1) Institute for Molecular Medicine, Finland, Helsinki, Uusimaa, Finland; 2) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom; 3) National Institute for Health and Welfare, Department of Chronic Disease Prevention, Helsinki, Finland; 4) Program in Medical and Population Genetics and Genetic Analysis Platform, The Broad Institute of MIT and Harvard, Cambridge, USA; 5) Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland; 6) Hjelt Institute, University of Helsinki, Helsinki, Finland.

Coronary artery disease (CAD) is the leading cause of deaths worldwide. Genome-wide association studies have identified multiple common variants associated with CAD and related traits. Some of these loci mediate their effects via risk factors such as lipids or hypertension. However, most of the common loci identified so far have a relatively small effect on the disease susceptibility. The exact molecular mechanisms behind most of these loci are also unknown. Therefore, we aimed at identifying rare and low frequency variants associated with plasma lipids and other metabolites in the known CAD loci. We analyzed low pass whole genome sequenced 1217 Finnish individuals from the FINRISK cohort with deep phenotypic data for CAD risk factors including metabolic profiles using different platforms. We undertook the analysis using sequence kernel association test (SKAT) for variants with minor allele frequency less than 1% or 5%. We performed SKAT using two approaches; window-based (20 or 50 variants in each window) and genomic region-based (3kb regions). Our preliminary results show a higher burden of CAD loci on plasma lipids and other lipids in the known CAD loci. 2125S

Identification of potentially pathogenic mutations in 9 candidate genes for bicuspid aortic valve using next-generation sequencing. N. Dargis1, M. Lamontagne1, N. Gaudreault1, L. Sbars2, C. Henry1, P. Pibarot1, P. Mathieu1, Y. Bosse1,2,1) Quebec Heart and Lung Institute, Quebec City, Quebec, Canada; 2) Department of molecular medicine, Laval University.

BACKGROUND: Bicuspid aortic valve (BAV) is the most frequent congenital heart disease. Affected individuals are at greater risk for developing aortic valve stenosis as well as other valvuloaortopathies. Despite the high prevalence of BAV, its etiology and genetic origins remain elusive. To improve our understanding of the genetic components contributing to the development and progression of BAV, we sought to identify and characterize the pathogenicity of all genetic variants located in 9 candidate genes for BAV (AXIN1, EGFRI, ENG, GATA5, NKX2-5, NOS3, NOTCH1, PDA12 and TGFBR2) in affected French Canadians using next-generation sequencing approaches.

METHODS: Genomic DNA from 48 French Canadians affected with BAV confirmed at surgery and having no concomitant aortic insufficiency and/or aortic aneurysm was extracted from blood buffy coat. Targeted DNA sequencing was performed using the Ion Torrent™ Personal Genome Machine™ (PGM™) from Life Technologies™. We designed an Ion AmpliSeq™ custom panel of primers covering 1,000 bp of the promoter region, the entire coding sequence and both 3' and 5' untranslated regions (UTR) for each of the nine genes. Novel variants were validated by conventional Sanger sequencing and the pathogenicity of all variants was evaluated with the Combined Annotation Dependant Depletion (CADD) method.

RESULTS: Sequencing revealed a total of 217 genetic variants, 38 of which were novel. Ten new variants were detected in the NOTCH1 gene, including one novel missense mutation present in one patient (p.G152S, located in exon 4). This variant was found to be pathogenic by functional testing using both approaches. In the known CAD loci, we observed a combined proportion of 20-25% windows associated with high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), total cholesterol (TC) and triglycerides (TG). One such region upstream of NID1 gene on chromosome 7 showed significant association with TG (window-base approach: P<9x10-5, genomic region-based approach: P=2x10-5). Fine-mapping of this region identified 3 independent low frequency variants showing association with TG: rs150895469 (frequency: 0.8%, P=2x10-5); var_7_18328999 (frequency: 0.2%, P=7x10-4); rs117533313 (frequency: 0.1%, P=4x10-4). We performed whole exome sequencing on 48 French Canadians to identify and assess the frequencies of these variants in French Canadians and other populations.

CONCLUSION: We identified 38 novel genetic variants within 9 candidate genes for BAV in 48 subjects affected with the disease. One novel mutation in NOTCH1 is predicted to be damaging and deleterious. Further functional studies will be required to fully evaluate the pathogenicity of these variants in French Canadians and other populations. Identifying causal mutations is essential to increase our knowledge of the molecular mechanisms underlying this frequent cardiac pathology, which could ultimately translate into earlier diagnosis, screening for families at risk, predicting disease progression and personalizing treatments.

2126M

Whole Exome Sequencing Reveals Rare, Truncating Variants in Nuclear Envelope Genes are Present in a Large Subset of Cardiovascular Patients. G.T. Haskell1, B.C. Jensen2, D.S. Marchuk1, C. Skrzynia1, C. Bizon2, K. Wilhelmsen1, K.E. Weck1, J.P. Evans1, J.S. Berg1, 1) Department of Genetics, UNC-Chapel Hill, Chapel Hill, NC; 2) Renaissance Computing Institute, Chapel Hill, NC; 3) Division of Cardiology, UNC-Chapel Hill, Chapel Hill, NC; 4) McAllister Heart Institute, UNC-Chapel Hill, Chapel Hill, NC; 5) Department of Pathology and Laboratory Medicine, UNC-Chapel Hill, Chapel Hill, NC.

The NC4GENES (North Carolina Genomic Evaluation by Next-Generation Exome Sequencing) clinical trial is applying whole exome sequencing (WES) in a bedside to bench approach, in order to evaluate the use of genome-scale sequencing to identify genomic determinants of a variety of genetic disorders, including hereditary cardiac conditions. Early pilot studies suggested of having a genetic form of cardiac disease have been enrolled, and WES sequence variants analyzed against a diagnostic list of ~75 known cardiac disease genes. The diagnostic yield for clearly positive results in this cohort was 16.7 percent. In patients with negative or equivocal diagnostic results, we analyzed other rare, predicted deleterious variants that might account for their conditions. Given that mutations in the nuclear envelope (NE) gene LMNA are a major genetic risk factor for severe cardiac disease, we were particularly interested in evaluating whether deleterious variants in other NE genes might contribute to cardiac disease as well. We used bioinformatics tools to systematically annotate a number of variant characteristics, including frequency in population databases and our own in-house WES database as well as in silico pathogenicity predictions. We identified novel or extremely rare truncating variants in NE genes in 5 out of 14 patients.

Three of these are nonsense or splice site variants in genes encoding nucleoporins, and one alters the splice donor site between intron 43 and exon 44 in SYNE1, a gene associated with Emery-Dreifuss Muscular Dystrophy, that is also important for mechanotransduction and heart function in mice. The SYNE1 variant was identified in a patient with dilated cardiomyopathy (DCM) who underwent heart transplant at age 15, and is present in his father, who also has advanced DCM. Many of the WES-identified NE genes are expressed in the developing as well as the adult heart, and we are evaluating whether these NE genes play a role in early cardiac development. Together, these data suggest the hypothesis that integrity of the nuclear envelope is particularly important for cardiac function and that alterations in this class of genes may be an important cause of genetic cardiac disease. Genome-scale sequencing technologies provide an especially attractive strategy for surveying clinically relevant variants in cardiac disease, and provide a rich source of data for discovery-based investigations aimed at identifying novel disease-associated variants.
2127S Initial analysis of exome sequence from 410 individuals with familial or simplex dilated cardiomyopathy. D.J Hedges1,2, A. Morales1,2, D. Kinnaman1,2, D. Wheeler1,2, J. Shendure1, M. Bamshad1, D. Nickerson1, R. Hershberger1,2,3. 1) University of Washington Center for Mendelian Genomics. 2) Department of Medicine, University of Washington. 3) Department of Pediatrics, University of Washington.

Coronary artery disease (CAD) is the leading cause of mortality in the world and it is expected to be the first cause of death worldwide by 2020. Genome-wide association studies (GWAS) have identified several genetic variants associated with CAD in LDLR and PCSK9 genes. This study was evaluated the possible association of common polymorphisms at LDLR and PCSK9 genes with the risk and severity of CAD in the Iranian patients. Sequencing of 18 exons of the LDLR gene and promoter region of the PCSK9 gene was performed in 202 Iranian patients angiographically confirmed CAD and 150 healthy controls. Sullivan’s scoring system was used for determining the severity of CAD in cases. Our results showed that homozygote genotypes of rs1122608 (P<0.001), rs4300767 (P<0.005) and rs10417578 (p<0.007) SNPs have strong protective effects on the CAD. In addition, we found that rs1122608 (GT or TT) was at higher risk of three vessel involvement compared to single vessels affecting (P<0.01). The only change in the coding region was c.1620 G>A in exon 11 of LDLR gene, which leads to a change in amino acid production (Gly->Glu). Moreover, according to the known databases this alteration probably is dangerous and benign, respectively. The results from our case-control study and other studies might be explained by genetic heterogeneity in the susceptibility of CAD and ethnic differences between Asians and Caucasians. Further investigation on other SNPs in these genes is warranted to validate our findings in Iranian population.
2130M
Association of rare variants in LDLR, HMGC, NAT2, ABCA1, and APOA1 with plasma lipid levels; initial results from the eMERGE PGx project. I.J. Kullo1, S.J. Bielinski2, D.S. Carrell3, S. Stallings4, J. Pathak5, A. Gordon3, L. Rasmussen-Torvik3, K.F. Dohney3, S. Volp3, M.D. Ritchie3, D.C. Denny3, C.G. Chute2, D.M. Roden4, D.A. Nickerson5, G.P. Jarvick7, D.R. Crosslin8, 1) Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, MN; 2) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 3) Group Health Research Institute, Center for Health Studies, Seattle, WA; 4) Vanderbilt University Medical Center, Nashville, TN; 5) University of Washington, Seattle, WA; 6) Northwestern University Feinberg School of Medicine, Chicago, IL; 7) Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore MD; 8) National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 9) Center for Systems Genomics, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA; 10) Departments of Biomedical Informatics and Medicine, Vanderbilt University, Nashville, TN; 11) Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA.

Linking genome sequence data to the broad range of phenotypes in the electronic medical record (EMR) is a potential means of obtaining insights into the pathogenicity of rare variants. The Electronic Medical Records and Genomics (eMERGE) pharmacogenomics (PGx) project is using the next-generation PGRN-Seq capture reagent to sequence 84 drug response genes in 9,000 individuals across the network. In a preliminary analysis of the first 894 participants, we examined the association between rare variants in five candidate genes (LDLR, HMGC, NAT2, ABCA1, and APOA1) that are implicated in lipid metabolism and plasma lipid levels (total cholesterol, LDL-C, HDL-C, and triglycerides) abstracted from the EMR at Group Health. For each trait, we ascertained the maximum level in the EMR for each participant. Of the 354 variants found in these 5 genes we identified 24 variants with evidence of high or moderate effects through in silico analyses using SnpEff and SeattleSeq annotation tools. For each functional variant, we assessed the differences in lipid levels between genotype groups. We describe the effect of one variant (C2177T) in LDLR on LDL-C levels in this cohort. The variant (rs45508991) leads to a THR/ILE amino acid change and was present in 2% of the Genotype - Phenotype (GxP) cases, with LDL-C in the highest tertile of the cohort. The LDLR mutation carriers have LDL-C levels below 75th percentile. The remaining 24 variants had no effect on lipid levels. Preliminary findings suggest that EMR-based studies may be helpful in confirming/assigning pathogenicity to rare variants identified by genome sequencing.

2131S
TGFB3 pathogenic mutations cause MFS/LDS phenotypes and aortic aneurysms in 3 Japanese families. H. Morisaki1, S. Ono1, I. Yamana1, R. Sultana1, T. Oda2, H. Tanaka2, H. Sasaki2, K. Minatoya2, R. Matsukawa2, T. Tsukube4, N. Kubo2, T. Morisaki1, 1) 1) Bioscience and Genetics, National Cerebral and Cardiovascular Center Research Institute, Suita, OSAKA, Japan; 2) Medical Genetics, National Cerebral and Cardiovascular Center, Suita, OSAKA, Japan; 3) Vascular Surgery, National Cerebral and Cardiovascular Center, Suita, OSAKA, Japan; 4) Cardiovascular Surgery, Japanese Red Cross Kobe Hospital, Kobe, HYOGO, Japan; 5) Pediatrics, Urakawa Red Cross Hospital, Urakawa, HOKKAIDO, Japan.

To date, several genes have been identified to be responsible for hereditary aortopathies including systemic connective tissue disorders such as Marfan syndrome (MFS) and Loeys-Dietz syndrome (LDS). Some of these genetic mutations were found to cause high transforming growth factor (TGF)-β signaling. To identify new genes responsible for these diseases, 292 probands with MFS/LDS-like features negative in FB1N, TGFB1R, TGFB2R, TGFB2, or SMAD3 mutations were subjected to exome sequencing. Three probands with family history of skeletal and/or aortic features resembling MFS/LDS were identified to have three novel mutations (p.N235fs, p.I322T, or p.R300W) in TGFB3 gene. Case 1 was a 43y tall male with scoliosis, strabismus and a history of thoracic aortic dissection (Stanford A) at 41y with aortic root dilatation. He was treated with emergency total arch replacement followed by aortic root repair (David operation) at 43y. His mother was tall and slender, and diagnosed as MFS with a history of thoracic aortic dissection at 59y. Case 2 was a 56y tall male who had a Bentall operation for severe aortic root dilatation. His daughter and son were both tall and had chest deformities along with mild aortic root dilatation. Case 3 was a 3y female showing LDS features with arrhythmogenic right ventricular dysplasia and aortic root dilatation. Her father also had tall stature, hypertelorism, bifid uvula and translucent skin. Genetic analysis of family members revealed the TGFB3 mutations segregated with phenotypes in all three cases. It was reported that TGFB3 mutations cause familial arrhythmogenic right ventricular dysplasia and that TGFB3 gene has an association with cleft lip and palate, hypertension, or ossification of the posterior longitudinal ligament of the spine. Murine model revealed that Tgfb3 gene plays an important role in embryogenesis and pathogenesis. In addition, recently, it was reported that the patients with de novo TGFB3 mutations showed MFS/LDS-like features with or without aortic involvement. This is the first report of TGFB3 mutations found in familial cases with MFS/LDS-like features, and three of our patients required aortic root replacement. Based on these results, we conclude that TGFB3 is essential not only for palatogenesis but also for systemic connective tissue disorders similar to MFS/LDS. Aortic aneurysms and dissections should be the important features of patients with TGFB3 mutations.
Posters: Cardiovascular Genetics

505

2132M

2133S

Exome sequences and rare variant association for plasma lipids in
over 18,900 individuals. G.M. Peloso 1,2,3,4, J.A. Brody 5, J.R. Crosby 6,7,
A. Isaacs 8 , Y. Zhou 9,10 , L.A. Lange 11 , J.C. Bis 5 , A.C. Morrison 6 , A.
Reiner 12,13, C. Kooperberg 12, C. van Duijn 8, S.S. Rich 14, C. Willer 15,16,17,
C.J. O'Donnell 3,10,18, B.M. Psaty 5,13,19,20,21, J.G. Wilson 22, L.A. Cupples 9,10,
S. Kathiresan 1-4, E. Boerwinkle 6,23 for the Lipids RVAS Working Group.
1) Center for Human Genetic Research, Massachusetts General Hospital,
Boston, MA; 2) Cardiovascular Research Center, Massachusetts General
Hospital, Boston, MA; 3) Department of Medicine, Harvard Medical School,
Boston, MA; 4) Program in Medical and Population Genetics, Broad Institute,
Cambridge, MA; 5) Cardiovascular Health Research Unit, Department of
Medicine, University of Washington, Seattle, WA; 6) Human Genetics Center, School of Public Health, University of Texas Health Science Center at
Houston, Houston, TX; 7) Department of Biostatistics, Bioinformatics, and
Systems Biology, The University of Texas Graduate School of Biomedical
Sciences at Houston, Houston, TX; 8) Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the
Netherlands; 9) Department of Biostatistics, Boston University School of
Public Health, Boston, MA; 10) National Heart, Lung, and Blood Institute
(NHLBI) Framingham Heart Study, Framingham, MA; 11) Department of
Genetics, University of North Carolina, Chapel Hill, NC; 12) Public Health
Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA;
13) Department of Epidemiology, University of Washington, Seattle, WA;
14) Center for Public Health Genomics and Department of Public Health
Sciences, University of Virginia, Charlottesville, VA; 15) Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan,
Ann Arbor, MI; 16) Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; 17) Department of Human
Genetics, University of Michigan, Ann Arbor, MI; 18) Division of Intramural
Research, NHLBI, US National Institutes of Health, Bethesda, MD; 19)
Department of Medicine, University of Washington Medical Center, Seattle,
WA; 20) Group Health Research Institute, Group Health Cooperative, Seattle, WA; 21) Department of Health Services, University of Washington, Seattle, WA; 22) Department of Physiology and Biophysics, the University of
Mississippi Medical Center, Jackson, MS; 23) Human Genome Sequencing
Center, Baylor College of Medicine, Houston, TX.
Genetic discovery using screens of common DNA sequence variants has
identified many novel gene regions associated with plasma lipids but it
has been challenging to pinpoint specific causal genes. Discovery using
sequencing of protein-coding regions across the genome may overcome
this barrier. Here, we test the hypothesis that exome sequencing in the
population can identify novel genes associated with plasma levels of lowdensity lipoprotein cholesterol (LDL-C), triglycerides, high-density lipoprotein
cholesterol (HDL-C), and/or total cholesterol. We sequenced the proteincoding regions (exome) of >17,000 genes from ~15,000 European ancestry
(EA) and ~3,900 African American (AA) individuals and tested the association of plasma lipid levels with single rare coding variants or those aggregated
within a gene. For gene-based tests, we implemented two methods Sequence Kernel Association Test (SKAT) and burden - with various minor
allele frequency thresholds. Functional classes of variants were aggregated
into categories: loss-of-function (nonsense, splice-site and frameshift) and/
or those predicted to be deleterious. seqMeta software was used for efficient
gene-based meta-analysis, where each study computes association-specific
statistics for each variant and genotypic covariance matrices within predefined gene regions. Score statistics and genotypic covariance matrices are
combined across studies and used to construct the gene-based tests. With
a sample size of 15,000 individuals, we had > 80% power to detect a gene
with a cumulative minor allele frequency of 0.5% and a ½-SD unit effect at
exome-wide statistical significance (P < 2.5×10 -7). We found aggregated
rare variant associations in many genes previously known to be associated
with plasma lipids. For example, in EA individuals, LDL-C was associated
with LDLR (beta=15 mg/dl; p=2.7×10 -22 with Burden 0.1%), HDL-C was
associated with ABCA1 (beta=-4%; p=2.4×10 -7 with Burden 1%), and triglycerides were associated with APOC3 (beta=-37%; p=6.4×10 -23 with Burden
1%). In AA individuals, LDL-C was associated with PCSK9 (p=6.7×10 -9 with
SKAT 1%) and HDL-C was associated with ABCA1 (beta=-8%; p=4.8×10 -8
with Burden 1%). Surprisingly, despite analysis in over 18,000 individuals
with exome sequences, our gene-based association analysis did not reveal
any novel genes at exome-wide significance. These data highlight the challenges of rare variant association to identify novel genes for complex traits.

Rare mutations in cardiomyopathy genes are associated with takotsubo cardiomyopathy. A.L. Siniard 1, P. Nakaji 2, J.J. Corneveaux 1, R.F.
Richholt 1, R.P. Bruhns 1, E. Carlson 1, A. Courtright 1, K. Van Keuren-Jensen 1, J. Forseth 3, J.M. Zabramski 2, R.F. Spetzler 2 , M.J. Huentelman 1,
M.Y.S. Kalani 2 . 1) Neurogenomics Division, Translational Genomics
Research Institute, Phoenix, AZ; 2) Division of Neurological Surgery, Barrow
Neurological Institute/Saint Joseph’s Hospital and Medical Center, Phoenix,
Arizona; 3) Arizona Pulmonary Specialists, Phoenix, Arizona.
Takotsubo cardiomyopathy (TC), defined as transient cardiac malfunction
with characteristic left ventricular (LV) wall motion abnormalities in the setting
of normal coronary vasculature, is a poorly understood complication of
cerebral aneurysm rupture that most commonly affects post-menopausal
women. Little is known about the etiology of this condition; however, the
clinical approaches for management of TC are largely supportive and reactive to the presentation of symptoms. In this study we utilized next generation
exome sequencing analysis in an effort to define genetic variation associated
with and responsible for this phenomenon. We performed exome sequencing
on eight female patients diagnosed with TC after aneurysmal subarachnoid
hemorrhage. Unless the variant was already present in ClinVar, filtering of
the identified variants was performed to limit our investigation to those
candidates that passed freebayes-derived haplotypic scoring, were rare
(<0.5% MAF in NHLBI GO ESP) or novel (not observed in publicly available
databases including 1000 genomes, NHLBI GO ESP, and an internal TGen
database), were determined algorithmically to have high potential impact on
the associated protein function (Combined Annotation Dependent Depletion,
CADD, score >15), were within regions of high species conservation (PhyloP
> 50%), and exist in genomic loci previously associated with cardiac dysfunction. Additionally, all candidate variants were verified to be present using
Sanger sequencing chemistry. Exome sequencing and analysis revealed
that 7 out of 8 patients with TC after aneurysmal subarachnoid hemorrhage
were heterozygote carriers of predicted deleterious mutations in established
cardiomyopathy genes including MYOZ2, SDHA, ANKRD23, MYLK2, PKP2,
DSP, and TTR. Our study suggests that patients with TC following cerebral
aneurysm rupture harbor deleterious mutations in established cardiomyopathy genes. We postulate that these patients likely live in a compensated
state of cardiac dysfunction that manifests only after the myocardium is
stressed by the adrenergic surge associated with aneurysm rupture or the
stress associated with the aggressive hemodynamic treatment of cerebral
vasospasm. It is possible that the clinical knowledge of the presence of these
mutations would change the management of patients presenting with TC.

2134M
Screening of the PRKG1 gene in a British cohort of Thoracic Aortic
Aneurysm and Dissection (TAAD) patients. Y.B.A. Wan 1, J.A. AragonMartin 1, L.J. Collins 1, D.C. Guo 4, A. Sagger 2, M. Jahangiri 3, D. Milewicz 4,
A.H. Child 1 . 1) Cardiovascular & Cell Sciences Research Institute, St.
George's Hospital, University of London, SW170RE, United Kingdom; 2)
Clinical Genetics Unit, St George's University of London, Cranmer Terrace,
SW170RE, London, UK; 3) Department of Cardiothoracic Surgery, St
George's Healthcare NHS Trust, London, UK; 4) Department of Internal
Medicine, University of Texas, USA.
Background: A mutation in the PRKG1 gene, which encodes for the
type I cGMP-dependent protein kinase (PKG-I, [OMIM 176894]) that was
published in 2013, causes decreased contraction of the vascular smooth
muscle cells. This in turn can increase risk of TAAD.
Method: A total of 101 UK patients (80M: 21F, mean age 53±13) with
known TAAD who did not fulfill the revised Ghent criteria for Marfan syndrome, and with no demonstrable mutations in FBN1, TGFB2, TGFBR1/2,
ACTA2 and SMAD3 genes were recruited to this study. These patients were
screened for the whole gene and in particular the published variant c.530G>A
(p.R177Q) in exon 3 of PRKG1, by bi-directional Sanger sequencing. The
methodology was taken from the original paper (Recurrent gain-of-function
Results: Our TAAD cohort of 101 British samples did not carry the
c.530G>A (p.R177Q) mutation in exon 3 of the PRKG1 gene. However, 2
heterozygous synonymous mutations (c.477C>T/p.T159T and c.993T>C/
p.V331V; NM_001098512) were found in 2 unrelated probands. These mutations were not reported on genome databases (1000 genomes and Exome
Variant Server). In-silico (mutation t@ster) predicted these to be disease
causing. Splice site tools (BDGP and ESE finder) were used to assess the
pathogenicity of these mutations and the causativity remains to be proven.
Conclusion: This study was negative for the recurring mutation found in
3 families from a USA cohort. Family segregation studies will be performed
for the two synonymous mutations in the two families to demonstrate that
they could be linked to the disorder. Samples should then assayed to determine if it increases the activity of the kinase before the variant can be
considered to be disease-causing. Further screening of novel single nucleotide variations through next generation exome sequencing data in our British
cohort of TAAD patients is currently underway.
You may contact the author at P0900001@sgul.ac.uk.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.


2135S
Exome sequencing identifies a homozygous splice site variation in ZBTB17 in arrhythmogenic ventricular dysplasia. J. Wang1, 2, H. Lee3, C. Silverstein4, S. Nelson1, 2, H. Accurso1, 2, D. Wood3, S. Nieswandt1, 2, M. Stein4, H. Luo3, R. Aggarwal3, 2, D. S. Chinnery1, 2, 3, 4.
Background: Arrhythmogenic right ventricular dysplasia cardiomyopathy (ARVD/C) is an inherited genetic disorder of the myocardium characterized by progressive fibrofatty replacement that predisposes to ventricular tachycardia and sudden death. Although it classically affects the right ventricle, left and biventricular forms are increasingly recognized. A mutation in one of the 8 known genes, primarily encoding desmosomal proteins, is found in fewer than half of patients. We identified a family with onset of dilated cardiomyopathy in the 2nd or 3rd decade in four out of four siblings with progression to cardiac transplantation so far in three. Gross and histopathological examination of the three explanted hearts demonstrated the characteristic fibrofatty replacement of ARVC/D involving predominantly the left ventricle. The family reported a high degree of consanguinity. Methods: To identify the molecular basis of these findings in the family, Clinical Exome Sequencing was carried out in the index patient. Exome variants were filtered for protein-altering rare heterozygous variants and homozygous variants to identify causative mutations. Mean oversampling of bases was 107X and at least 93.9% of the 33.4 Mb of protein coding sequence defined by RefSeq was read at least 10X depth. HumanConeindex BeadChip assay was performed on the living affected siblings to filter the candidate variants. Results: Within the shared linkage regions, three rare homozygous variants, predicted to cause missense, essential splice site, and premature stop codon changes, were identified in the index case and confirmed to be shared among the three affected siblings. Among the candidate variants, ZBTB17 resides in a region that was previously implicated in dilated cardiomyopathy by association studies in European and Han Chinese populations. Conclusions: ZBTB17 (zinc finger and BTB domain containing 17), also known as Miz-1, encodes a zinc finger protein involved in the regulation of c-myc. Our findings implicate ZBTB17 as a new gene for ARVD/C and provide new evidence of a critical role for ZBTB17 in cardiac function. The significance of the three variants shared by the affected siblings is to be established experimentally using animal models.

2136M
Identification of Novel Genes and Variants Associated with Hypertrophic Cardiomyopathy in Panel Negative Patients Through Targeted Sequencing of 450 Genes and Rare Variant Association Testing. M.T. Wheeler1, 2, 3, J.R. Homburger1, 4, F.E. Dewey1, 2, 3, A. Pavlovic1, 2, 3, H. Chabi1, 4, J.R. Priest1, 2, 3, R.L. Goldfeld1, 4, D.M. Waggott1, O. Soyinka1, E.A. Ashley1, 2.
1) Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA; 2) Stanford Center for Inherited Cardiovascular Disease, Stanford University School of Medicine, Stanford, CA; 3) Cardiovascular Institute, Stanford, CA; 4) Department of Genetics, Stanford University School of Medicine, Stanford, CA; 5) Program in Biomedical Informatics, Stanford University School of Medicine, Stanford, CA.
Hypertrophic cardiomyopathy (HCM) is a genetic disorder of the heart muscle leading to increased heart wall thickness and risk of sudden death. HCM is typically inherited in an autosomal dominant fashion. Clinical genetic testing for HCM patients sequences 8 to 17 genes to identify a causal variant. In about half of HCM patients who undergo clinical genetic testing, no known mutation is identified. In the remaining patients, it is estimated that a subset of HCM in these panel-negative patients, we performed targeted sequencing of 450 cardiac genes covering 3.3 Mbp in 58 unrelated HCM patients without known causal variants. We used Real Time Genomics (RTG) alignment and variant calling pipeline together with sequence to medical phenotypes (STMP) to align, call, and prioritize variants by comparison to publically available and cardiac-specific annotation data. We identified 126 extremely rare or novel exonic nonsynonymous and premature termination variation variants. Further sequencing revealed novel stop variants in GJAS, CSRP3, and ATP1A4. To prioritize novel additional genes that may contribute to development of HCM, we performed gene-delimited rare variant association testing using the SKAT-O test. We compared our HCM patients against a cohort of 424 exome sequenced individuals without HCM. We identified new genes significantly associated with HCM case status including ILK (p = 9.963 x 10^-18) and ADAMST9 (p = 1.857 x 10^-5). To control for the potential confounding effect of ancestry, the results were replicated using a subset of the data set that contained only individuals of European ancestry. ILK has previously been implicated in dilated cardiomyopathy and identified that it is a strong biological candidate for affecting HCM. Through both functional annotation and burden testing we identify new candidate genes and variants that may play an important role in the genetic basis of HCM.

2137S
Whole exome sequencing of a family trio to identify potential genetic modifiers of LMNA arrhythmia and cardiomyopathy. M. Zaragoza, S. Pikakis, V. Fischer, D. Pediatric Genetics & Genomics, University of California, Irvine, School of Medicine, Irvine, CA.
We identified a novel LMNA splice site mutation (c.357-2A>G) as the primary mutation associated with Sick Sinus Syndrome (SSS), Dilated Cardiomyopathy (DCM), and Sudden Cardiac Death (SCD) in a large family. To identify potential secondary factors, genetic “suppressors” or “enhancers” that may result in clinical variability within this family, we then focused on two affected members: a 71-year-old female and her 39-year-old daughter. The mother had bradycardia and atrial fibrillation at 49, sinus arrest at 52, and DCM at 61 while her daughter presented at 36 with bradycardia and ventricular arrhythmias. To identify potential modifiers associated with this phenotypic range, we conducted whole exome sequencing (WES) on the affected mother and daughter and the unaffected father. This produced 42 Gb of sequencing data in which we identified ~1 million unique variants for each parent and ~700,000 for the daughter. We then focused on the variants that were not shared by mother and daughter located in over 700 genes for ion channels, cardiac muscle contraction, LMNA-associated proteins (nuclear envelope proteins, transcription factors, chromatin-associated proteins), and LMNA-associated signal transduction (MAPK, mTOR, TGF Beta pathways). We validated the potential modifier variants by Sanger sequencing. We found seven potential “suppressors” including rare missense variants in the Lysosomal-associated membrane protein 2 (LAMP2), Mitogen-activated protein kinase 3 (MAPK3), and Retinoblastoma-like 2 (RB2) genes. We found six potential genetic “enhancers” including a novel missense variant in the Unc-51-like kinase 3 (ULK3) gene and rare missense variants in SHMT2. Among the motifs protected 20 (RBM20) and Spectrin repeat containing nuclear envelope 1 (SYNE1) genes. In conclusion, we used WES to identify novel and rare variants in LMNA pathway-associated genes that serve as potential modifiers of SSS, DCM, and SCD. Gene expression and functional studies are planned. These studies may provide insight into the molecular mechanisms for the intrafamilial variability in LMNA-associated arrhythmia and cardiomyopathy. By taking advantage of the endogenous methods to alter disease; we hope that molecular targets may be discovered to develop new preventive strategies and treatments.

2138M
Timothy syndrome type 2 associated CACNA1C mutation G402S in a teenager with normal development presenting with ventricular fibrillation. T.P. Alastalo1, A. Hirppula1, J. Taillit2, S. Myllykangas3, JW. Kossen-vuur4, 1) Padiatric Cardiology, University Hospital Helsinki, Helsinki, Finland; 2) Blueprint Genetics, Helsinki, Finland; 3) Institute of Biomedicine, University of Helsinki; 4) Research Center of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland.
Timothy syndrome (TS) is a multigorgan disorder with prolonged QTc interval, congenital heart defects, syndactyly, typical facial features and neurodevelopmental problems. Ventricular tachyarrhythmia is the leading cause of sudden death. From diagnostic point of view the most characteristic features of TS are craniofacial dysmorphism, calcified arteries, myoclonus, and developmental delay. Classical TS type I results from a recurrent de novo CACNA1C mutation, G406R in exon 8A. An atypical form of TS, type 2, is caused by G406R and G402S mutations in the alternatively spliced exon 8B. Only one individual for each exon 8 mutations has been described to date. In contrast to multiorgan disease caused by G406R either in exon 8A or 8B, the G402S carrier manifested only an isolated cardiac phenotype with LQT and cardiac arrest. This isolated phenotype was suggested to result from somatic mosaicism. We describe a teenager patient resuscitated from ventricular fibrillation and treated with ICD. She has no other organ manifestations, no syndactyly, normal neurodevelopment and her QTc is <480ms. There is no family history of arrhythmias or sudden deaths. Targeted oligonucleotide-selective sequencing (OS-Seq) of channelopathy genes revealed a de novo substitution, G402S in exon 8 of CACNA1C. This is the third reported case of TS type 2. Direct sequencing of blood and oral mucosa derived DNA showed an identical mutation peak suggesting ubiquitous expression in different tissues. The phenotype of our patient and the previously described patient show an isolated arrhythmia disease with no other clinical manifestations of classical TS.
Rare potentially pathogenic variants in the congenital arrhythmia syndrome genes SCN5A and KCNH2 are detected frequently but rarely associated with arrhythmia phenotypes in electronic health records. A highly conservative approach to returning results of uncertain effect as normal corrected QT interval on many other ECGs. Four individuals had the 48 had a family history of arrhythmia in the EHR. One individual with rs150264233 (SCN5A S1904L) had multiple instances as pathogenic for cardiac arrhythmia phenotypes in the ClinVar database.

The 48 had a family history of arrhythmia in the EHR. One individual with rs199473618 (SCN5A V1523I) had atrial fibrillation, a common arrhythmia. One individual with rs150264233 (SCN5A S1904L) had multiple instances of prolonged corrected QT interval (471-540 ms, normal < 440 ms), as well as normal corrected QT interval on many other ECGs. Four individuals had evidence of bundle branch block. No additional arrhythmia phenotypes were identified as pathogenic or likely pathogenic by ≥ 1 lab. 9 by > 2 labs, and only 4 by all 3 labs (rs199473166, rs137854618, rs199473603, and rs137854602, causing SCN5A 1848F, D1275N, T1304M, and R1512W, respectively). To determine arrhythmia phenotypes for the 48/2/022 individuals with ≥1 of the 40 putatively-pathogenic variants, we reviewed electronic health record (EHR) data and ECG tracings (available in 33/48). None of the 48 the majority of patients have a similar deletion on chromosome 22q11.2. Approximately 60% of patients have varied conotruncal heart defects (CTDs). While 40% have normal cardiac anatomy. We hypothesized that novel (not previously reported) or rare (<1%) functional coding variants, could serve as genetic modifiers contributing to the risk of CTDs in 22q11DS patients. Genomic DNA samples from 186 22q11DS subjects including 91 cases with CTDs and 95 controls without CTDs were selected for whole exome sequence (WES) capture using the NimbleGen SeqCap EZ Exome Library V3.0, followed by Illumina HiSeq 2000 sequencing yielding 100-bp paired-end reads (NHBB; Nickerson, U. Washington). A total of 77,836 coding variants were identified from the WES data. Among these, 10,621 were missense or frameshift mutations not observed in the 1000 Genomes project and the Exome Aggregation Consortium (ExAC). Five (not previously reported) or rare (<1%) functional coding variants, supporting the idea that missense mutations in histone demethylases may contribute to CTDs in 22q11DS. Interestingly, de novo rare functional variants in histone modifying genes were found in WES data in the NIH congenital heart disease, PCGC cohort (PMID: 23665959). Among the 26 genes reported from the PCCG, we found a mutation in RNF20 (E3 ubiquitin ligase; histone H2B) in one of our cases. Overall, the two studies implicate histone-modifying genes either as causative or as modifiers in the etiology of CTDs. In addition, the number of such genes that may be important for heart development in humans.
Aortic valve stenosis (AS) is a cardiovascular disease that can be fatal in the absence of treatment. Individuals with a bicuspid aortic valve (BAV), the most common cardiovascular congenital abnormality, develop symptoms of AS 10 to 15 years earlier compared to those with a tricuspid valve. The molecular mechanisms leading to the premature development of AS in BAV patients are unknown. The objective of this study was to identify genes differentially expressed between calcified BAV and tricuspid valves with (TAVc) and without (TAVn) calcification using RNA sequencing. Ten human calcified BAV and nine TAVc were collected from male patients who underwent aortic valve replacement at the IUCPQ. All valves had the same degree of fibro-calcific remodeling. Ten TAVn were obtained from male patients who underwent heart transplantation. mRNA levels were measured using the Illumina HiSeq 2000 system. The alignment of paired-end reads and the comparison of gene expression among three groups of valves were performed using the TopHat-Cufflinks protocol. A q-value <0.05 was used as a significant threshold. Genes with FPKM <10 in two tissues compared were excluded. IPA was used to identify molecular functions enriched for genes differentially expressed. Thirty-one genes were up-regulated and 33 were down-regulated in BAV compared to TAVc, including 248 genes with fold change >2. The atherosclerosis (q=1.95E-06), and phospholipases (q=7.76E-03) pathways were significantly enriched with differentially expressed genes. A total of 283 genes were up-regulated and 340 were down-regulated in BAV compared to TAVn, counting 288 genes with fold change >2. Compared to TAVn, 189 and 233 genes were up- and down-regulated in TAVc, respectively, including 248 genes with fold change >2. Tens of genes were newly identified to be differentially expressed among these groups of aortic valves. This is the first study to analyze the transcriptome of aortic valves by using RNA sequencing. AS patients showed significant changes in the expression profile of BAV and TAVc compared to normal valves. Pathways biologically relevant for AS were significantly enriched with differentially expressed genes. The role of these genes and pathways in the pathogenesis of AS in BAV patients requires further investigation. These results contribute to the search of new therapeutic targets with the potential to avoid or slow the development and/or progression of AS in patients with BAV and TAV.
Whole genome sequencing association study of ECG conduction traits. B.P. Prins¹, F. Petropoulos¹, E. Zeggini², N. Soranzo³, T.D. Spector⁴, Y. Jamshidi⁵, UK10K Consortium Cohorts Group. 1) Cardiogenetics Lab, Cardiovascular and Cell Sciences Institute, St George's University of London, London, London, United Kingdom; 2) Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom; 3) Department of Twin Research & Genetic Epidemiology, King's College London, London, United Kingdom.

The UK10K project, funded by the Wellcome Trust, has completed whole genome sequencing (WGS) to ~6.7x coverage of 1,754 individuals from the TwinsUK study. Variants discovered through WGS, along with those from 1,000 Genomes were imputed into the full cohort with genome-wide genotype data, increasing the sample size to 4,329, for which for 2,327 samples ECG measurements were available. We performed discovery analyses testing 9.2 million variants (MAF>1%) in a meta-analysis of sequenced and imputed variants for association with ECG conduction traits (PR interval, QT interval, QRS duration, and heart rate) and with the extremes of these traits. We identified 2 variants (p < 5 × 10⁻⁸) in known ECG trait-loci; a novel variant for PR interval and SCN10A (rs6598255, intronic, MAF 41.6%), and a known variant for QT interval and NOS1AP (rs12143842, 6 kb upstream, MAF 24.6%). We also identified 133 variants across 131 known and novel loci suggestively associated (p < 1×10⁻⁵) across the four ECG traits, some of which appear to be pleiotropic in nature.

We found significant enrichment (p < 5×10⁻²) for genes involved in sarcolemma function, calcium binding, voltage-gated channel activity and cytoskeletal protein binding. Using publicly available data from the ENCODE project on chromatin states, epigenetic modifications and sequence motifs, we show that some of these variants are likely to have functional consequences affecting regulatory elements. Currently rare variant-, and combined common- and rare variant analyses are being performed using SKAT, where we collapsed variants in 57,820 gene sets based on GENCODE v19 definitions to test whether rare variation may be associated with ECG traits, and to dissect the contribution of common and rare variants to each ECG trait. These results will provide insights into how both rare and common variation are likely to contribute to the understanding of the genetic architecture of ECG traits, and are likely to explain a substantially larger fraction of the missing heritability.
2146M

Replication of a single nucleotide polymorphism variant in CETP gene associated with HDL level in the ClinSeq® Study. H. Sung1, T. Schwanetz-An1, B. Suktitip2, K. Lewis2, D. Ng2, S.G. Goncalves2, J.C. Mullikin3,4, L.G. Biesecker2, A.F. Wilson1. National Institutes of Health Intramural Sequencing Center (NISC), 1) Genometrics Section, Computational and Statistical Genomics Branch, NHGRI, NIH, Baltimore, MD; 2) Medical Genomics and Metabolic Genetics Branch, NHGRI, NIH, Bethesda, MD; 3) Comparative Genomics Analysis Unit, Cancer Genetics and Comparative Genomics Branch, NHGRI, NIH, Bethesda, MD; 4) National Institutes of Health Intramural Sequencing Center (NISC), NHGRI, NIH, Bethesda, MD.

ClinSeq® is a large-scale medical sequencing study designed to investigate associations of rare sequence variants with traits related to Coronary Artery Disease (CAD). The study currently includes more than 1000 non-smoking participants between the ages of 45 to 65 with normal to severe coronary artery calcification scores. About 200 CAD-related traits were measured at the NIH Clinical Research Center in Bethesda, MD. Whole-exome sequencing was performed with the Agilent SureSelect 38Mb and 50Mb capture kits for 387 and 325 individuals, respectively, at the NIH Intramural Sequencing Center. Single nucleotide variants (SNVs) common to both capture regions with a call rate >98% and minor allele frequency (MAF) >1% were used to check for cryptic relatedness and for misspecified population stratification by multidimensional scaling analysis - 635 unrelated European Americans (EAs) remained. For each capture kit with EAs only, SNVs with at least one homozygote and a call rate greater or equal to 50% were included. The two capture regions with only SNVs in common were merged, yielding 439,807 SNVs. Of these SNVs, 74% and 46% had MAFs < 0.01 and < 0.001, respectively. The SNVs with MAF < 0.01 were collapsed into a single derived variant for each genomic region defined by hotspot blocks. Collapsed variants were coded as the proportion of minor allele occurring within each region; common variants were coded as the number of minor alleles (scaled from 0 to 1). Tests of association of each SNV with High Density Lipoprotein (HDL) level was performed on untransformed, log-transformed and rank-inverse transformed HDL traits with simple linear regression, adjusting for age, sex, BMI and medication. The SNV rs1532625 in the intron of CETP gene was found to have a significant association for all untransformed and transformed HDL traits at the level of 1e-05 and one SNV in the exon of PNMAL1 (rs111356009) at the level of 1e-04. CETP and RBBP6 genes are known to be associated with HDL and PNMAL1 gene with triglycerides.

2147S

Assessment of the implication of rare coding variants in familial and sporadic Fibromuscular Dysplasia. N. Bouatia-Naji1,2, SR. Kiando1,2, PF. Plouin1,2, MD. Barlassina3, D. Cusi5, P. Galan4, M. Lathrop5, X. Jeunemaitre1,2,3. 1) INSERM UMR970 PARCC, Paris, FRANCE; 2) Université Paris Descartes, PRES Sorbonne Paris Cité, Paris FRANCE; 3) AP-HP. Department of Hypertension, Hôpital Européen Georges Pompidou, Paris, FRANCE; 4) Filaret Foundation, Genomic and Bioinformatics Unit ; Dept. of Medicine, Surgery and Dentistry, University of Milano, ITALY; 5) Graduate School of Nephrology, University of Milano, Division of Nephrology, San Paolo Hospital, Milano, ITALY; 6) Unité mixte de recherche 557 Inserm/Inra/Cnam/Paris13, FRANCE; 7) McGill University and Génome Québec Innovation Centre, Montréal, Québec, CANADA; 8) AP-HP. Referral Center for Rare Vascular Diseases, Hôpital Européen Georges Pompidou, Paris, FRANCE.

Fibromuscular dysplasia (FMD) is an arterial disease characterized by nonatherosclerotic stenosis reported mostly in renal and extra-cranial carotid arteries. FMD predisposes to hypertension, renal ischemia and stroke, the first cause of disability worldwide. The genetics of fibromuscular dysplasia is under-investigated because of the lack of large families and cohorts due to the rarity of the disease and the complexity of the imaging-based diagnosis. Causes of FMD are unknown and it occurs predominantly in females with a prevalence of ~4/1000 for the clinical forms. There are strong arguments in favour of the genetic origin of FMD, based on documented and reported family history, although the precise estimation of its heritability is missing. FMD is probably a typical complex genetic disease and it is challenging to investigate because it is also a rare disease. Exome sequencing in 16 familial cases of FMD (five sibs and two sib-trios) generated 3,971 SNPs with at least one rare (MAF<0.01) and predicted functional variants. Inter-families analyses showed that no gene could explain the shared FMD status in at least three out of seven families. We also analysed 15 known causative genes of vascular diseases and/or connective tissue syndromes (e.g. FBN1, TGFB2 and COL3A1). None of these genes presented rare coding and putatively functional variants that could explain the FMD status in one of the seven families (intra-family analyses). Then, we aimed to follow-up 22 genes that were identified as harboring at least four rare variants and the 15 candidate genes in a larger sample of 259 FMD unrelated cases and 698 controls using genotyping data generated by the exome chip, an array enriched for rare and predicted functional variants. Neither gene-based association analyses of rare variants (MAF<0.01) using SKAT-O nor single SNP association of common SNPs by logistic regression (MAF>0.05) support a role of the 37 genes tested (Ntotal=982 SNPs, including 264 common SNPs). These findings support strong genetic heterogeneity for FMD and encourage more powerful and comprehensive genomic approaches, such as genome-wide association studies, to decipher the genetic basis of FMD.
2148M
High Throughput Sequencing and Bioinformatic Analysis in Familial Congenital Heart Disease. D. Corsmeier1, S. Fitzgerald-Butt1,2, G. Zender2, M. Mori3, K. Kelly2, K. Walters2, V. Vielant1, H. El Houdi3,5, V. Garg1,2, K. McBride1,2, P. White1,2.
1) The Biomedical Genomics Core & The Center for Microbial Pathogenesis, The Research Institute at Nationwide Children's Hospital, Columbus, OH; 2) The Center for Cardiovascular and Pulmonary Research, The Research Institute at Nationwide Children's Hospital, Columbus, OH; 3) The Center for Molecular & Human Genetics, The Research Institute at Nationwide Children's Hospital, Columbus, OH; 4) Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital, Columbus, OH; 5) Department of Pediatrics, College of Medicine, The Ohio State University, Columbus, OH.

Congenital heart disease (CHD) is the most common type of birth defect, appearing in eight of one thousand live births. CHDs are a leading cause of death in infants and represent an increasing cause of adult morbidity and mortality. There is strong evidence for a large genetic component in the etiology of non-syndromic CHDs. As such, whole exome and whole genome sequencing will be invaluable in elucidating the underlying causal variants which lead to disease. In this study, high throughput sequencing approaches were utilized to identify genetic variants causing non-syndromic CHDs in families exhibiting Mendelian segregation of CHDs. Our subject cohort consists of over 800 affected cases and includes over 100 multiplex families. To date whole exome sequencing has been performed on several hundred of these samples using Agilent SureSelect Human All Exon V5 kits and the Illumina HiSeq platform. In selected cases for which preliminary results yielded no disease candidate mutations, more extensive sequencing methods were utilized including paired end and mate-paired whole genome sequencing. Bioinformatic analysis was performed using Churchill, the pipeline we developed for the discovery of human genetic variation. Churchill utilizes the Burrows Wheeler Aligner (BWA) and the Genome Analysis Toolkit (GATK) to perform sequence alignment and variant calling. The initial filtering and prioritization of variants, we used traditional heuristic methods based on existing variant databases (dbSNP, 1000 Genomes, NHLBI Exome Sequencing Project, HGMD). For bioinformatic and statistical validation, we explored other widely-used methods. These different methods yield highly consistent results, as we have been able to identify putative causal variants in these families and mark them for further validation and functional studies. These findings will provide a better understanding in the pathogenesis of CHD and can lead to improved diagnosis and novel preventive and therapeutic approaches.

2148S
Exome Sequencing Reveals Candidate Genes for Spontaneous Coronary Artery Dissection. J.L. Theis, M.S. Tweet, J.M. Evans, R. Gulati, S.N. Hayes, T.M. Olson. Mayo Clinic, Rochester, MN.

Spontaneous coronary artery dissection (SCAD) is an uncommon, idiopathic disorder that primarily affects young, otherwise healthy women. The pathophysiologic process involves an intimal dissection or intramural hematoma, which can obstruct blood flow and cause myocardial infarction and sudden death. A majority of patients have concomitant fibromuscular dysplasia (FMD), implicating a generalized arteriopathy. Postulating a genetic basis for sporadic SCAD, whole exome sequencing and comparative variant filtering of proband-parent trios was performed. The cohort for this pilot study was comprised of 28 unrelated patients with angiographically confirmed SCAD in the absence of a systemic connective tissue disorder (93% female; mean age 41 years). Exomes of each trio were uploaded into Ingenuity Variant Analysis (IVA) software and filtered for rare (<1% minor allele frequency), predicted-deleterious coding variants. Modeling for de novo and recessive modes of inheritance yielded 150 unique candidate genes, none of which comprise commercial gene testing panels for thoracic aortic aneurysm and dissection. A short list of 8 prioritized candidate genes was generated based on presence in >2 trios (n=1), known coronary artery expression (n=2), or abnormal vascular phenotype associated with disruption of the murine ortholog (n=5). Each candidate gene harbored missense variants arising de novo (n=1) or inherited as recessive alleles (5, compound heterozygous; 2, homozygous). Two of these genes, each linked to SCAD associated with FMD, have been selected for mutation scanning of a cohort of 270 additional unrelated SCAD subjects by high throughput heteroduplex analysis. While our preliminary investigation has ruled out a common disease gene for SCAD, whole exome sequencing has identified plausible candidate genes, a first step toward deciphering molecular underpinnings of a disorder that may account for 40% of myocardial infarctions in women <50 years.

2150M
Feasibility of a Targeted Ion Torrent Next Generation Sequencing platform for Identification of Mutations Associated with Inherited Cardiac Arrhythmia Syndromes. E. Burashnikov, R. Pfeiffer, G. Caceres, Y. Wu, C. Antzelevitch. Molecular Genetics, Masonic Medical Research Laboratory, Utica, NY.

Introduction: Many unexplained sudden deaths are due to inherited cardiac arrhythmia syndromes such as Brugada Syndrome (BrS), Early Repolarization Syndrome (ERS), Short QT Syndrome (SQT) and idioventricular fibrillation (IVF). We sought to determine the suitability of screening a panel of 30 genes encoding sodium, calcium and potassium channels genes, including ATP-sensitive potassium channels, glycerol-3-phosphate dehydrogenase 1-like (GPD1L), genes previously associated with arrhythmia syndromes. We also included ankyrin 2 (ANK2), ryanodine receptor 2 (RYR2), ATP-binding cassette transporters (ABCC8, ABCB9), caveolin 3 (CAV3) and h-es related family gene (HEY2), as a potential causative contributors to these diseases. Methods: We employed an Ion Torrent Personal Genome Machine (Life Technologies) for Next Generation Sequencing and analyzed data with Ion Reporter Software (Life Technologies). Variants were annotated by Ingenuity Variant Analysis (Qiagen) and were confirmed by Sanger method on 3730 Automatic DNA Analyzer. Any variation with a frequency less than or equal to 0.5% was considered as a mutation, based on 1000 Genome Frequency. Libraries were constructed using Custom Ion Ampliseq to amplify all exons and intron borders of the targeted genes. Proband included in the study were predominantly SCNSA-negative. Results: Total of 121 probands were screened, among them 50 BrS, 42 ERS, 17 SQT and 12 cases IVF/VF. We identified mutations in 33 of 50 BrS probands screened, for a yield of 66%. Mutations in ANK2, calcium channel genes, SCN1A, SCN5A and I K-ATP genes were found in 17, 13, 14, 7 and 9% of probands, respectively. In the ERS group, we identified mutations in 21 of 42 probands screened, for a yield of 50%. Mutations in ANK2, calcium channel genes, SCN10A, SCN5A, and I K-ATP genes were found in 17, 13, 14, 7 and 9% of probands, respectively. The IVF group, we uncovered mutations in 6 of 17 probands screened, for a yield of 35%. Mutations in ANK2, SCN1A, SCN1B, RYR2, GPD1L and I K-ATP genes were 8% for each. Conclusions: Our study shows the feasibility of using a targeted Next Generation Sequencing approach for high throughput screening of SCD mutations. The study shows a yield of 50 to 76% for BrS, ERS, SQT and IVF probands.
Thrombotic storm (TS) is a devastating and extremely rare disorder that occurs in a small subset of patients with venous thromboembolic disease. TS is characterized by patients exhibiting ≥ 2 of the following in a short period: 1) more than two acute arterial/venous thromboemboli, or thrombotic microangiopathy, 2) unusual location, 3) progressive/recent unexplained recurrence, or 4) refractory/atypical response to therapy. The causes of TS are currently unknown, but we hypothesized that rare genetic variants explain a proportion of the risk for TS. To identify candidate genetic variants we performed whole-exome sequencing (WES) on the only known TS multiplex family existing in the world. Within our multiplex TS family, 10 family members were WES including four sisters, 2 of which were TS affected and the other 2 unaffected. DNA was captured using the Agilent 50mb kit and sequenced using Illumina HiSeq2000. Alignment and genotype calling were performed with Burrows-Wheeler Aligner and Genome Analysis ToolKit. Using this WES data we filtered variants to identify those inherited under 3 different inheritance models; dominant, recessive, and co-dominant. For all analyses we included only those variants that met the following criteria; 1) inheritance models (dominant, recessive, and co-dominant), 2) variant allele frequency (MAF)<5% and either missense, nonsense or splice site variants, 3) variant with alternative variants in multiple families. There is intense discussion on the role of rare and common genetic variants in the etiology of complex diseases. Many studies attempting to identify the rare variants associated with the disease are large population based association studies whereas our study is designed to recognize inherited risk alleles. The discovered novel variants are located on promising genes and a more detailed analysis might provide new insights to molecular pathways responsible for the pathogenesis of CHD. Our family-data based findings are now being confirmed by Sanger sequencing in the remaining study population and also in a large Northern Finland Birth Cohort.
2153S
Retrotransposons in Nonsyndromic Conotruncal Heart Defects. D. Webber1,2, S. MacLeod1,2, S. Erickson1,2, M. Li1,2, L. Murphy1,2, C. Hobbs1,2, 1 Pediatrics Department, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR; 2) Arkansas Children’s Hospital Research Institute, Little Rock, AR.

Introduction: Congenital heart defects (CHDs) are a leading cause of infant morbidity and mortality. Research on CHDs has focused primarily on genomic variants within coding regions and has yet to investigate DNA transposable elements, which occupy nearly half of the human genome and are implicated in a wide range of structural and developmental conditions. We identified and characterized retrotransposon insertion polymorphisms (RIPs) among families affected by nonsyndromic conotruncal heart defects (CTDs).

Methods: Participants were enrolled in the National Birth Defects Prevention Study at the Arkansas site between 1997 and 2007. We used probe hybridization enrichment and sequencing to map the position of human specific retrotransposons, including L1Hs and AluYb9, throughout the genomes of 88 case families (n = 263) affected by conotruncal heart defects and 40 unaffected control families (n = 120). Paired-end sequencing data was aligned to the reference genome with the Burrows-Wheeler Aligner (BWA), and non-reference RIPs were identified using RetroSeq. RIPs reported in dbRIP or 1000 Genomes were filtered out of the dataset in order to enrich for novel retrotransposon insertions, likely to be under selective pressure or to be a product of recent retrotransposition. Results: A similar proportion of novel RIPs were identified in case families compared to controls. Among novel cases, there were 26 novel RIPs in mothers, 37 in fathers, and 30 in infants. Nearly half (14/30) of the RIPs affecting case infants were located within introns, and the majority (21/30) were members of the LINE-1 family. Eight RIPs were observed in more than one case family, including an Alu retrotransposon within the GLUD1 intron that was present in 5 case families and in no controls. Modeling suggest that this novel insertion introduces two canonical splice acceptor sites that are poised to disrupt the adjacent exon, which encodes a GTP binding domain previously implicated in familial hyperinsulinemic hypoglycemia 6 (HHF6; OMIM #606762). These findings are strengthened by gene ontology analysis among case infants, which showed significant enrichment (P = .003) of intronic RIPs in genes (GO: 0006952) involved in “regulation of insulin secretion.”

Conclusion: Proper glucose homeostasis is critical to cardiogenesis, and maternat diabetes and obesity are well-established risk factors for CHD. Results suggest that rare RIPs may influence glucose metabolism and risk of CTDs.

2155S
Genotype and Phenotype Characteristics of Filipino Families with Familial Hypercholesterolemia and Novel LDL-R Gene Mutation. E.C. Cutiongco de la Paz1, R.G. Sy2, K.N. Hernandez3, E.B. Llanes2, F.E. Punzalan1,2, C.P. Cordero3. 1) Institute of Human Genetics, National Institutes of Health, University of the Philippines Manila, Philippines; 2) Section of Cardiology, Department of Internal Medicine, College of Medicine-Philippine General Hospital, University of the Philippines Manila; 3) Institute of Clinical Epidemiology, College of Medicine, University of the Philippines Manila.

Familial hypercholesterolemia (FH) is a disorder of lipoprotein metabolism and is one of most common inherited disorders worldwide. A diagnosis of FH is typically based on high plasma levels of total and LDL cholesterol, a family history of childhood-onset hypercholesterolemia, cholesterol deposition in extracellular tissues (e.g. tendon xanthomas, corneal arcus), and a personal and family history of premature CVD. However, there are no absolute predictive clinical criteria for FH diagnosis. Genetic testing is recommended in combination with clinical assessment to provide a more definitive diagnostic process. The LDL-R gene codes for the LDL receptor protein which is responsible for the uptake (and subsequent degradation) of serum LDL cholesterol via receptor-mediated endocytosis. Mutations in the LDL-R gene, can impair receptor synthesis or lead to the production of defective receptors. To date, 1,122 unique genetic variations have already been reported for the LDL-R gene and may be found in the online registry maintained by the Leiden Open (source) Variation Database (LOVD). Included here are the five novel mutations found by study involving selected Filipino FH patients: exon 3 G50R, exon 10 D147N, exon 13 I602V, intron 2 190+4 A>T 5’ splice donor, and intron 8 1187-10 G>A 3’ splice acceptor. Objectives: 1) To determine the presence of LDL-R gene mutation among first degree relatives of FH patients: exon 3 G50R, exon 10 D147N, exon 13 I602V, intron 2 190+4 A>T 5’ splice donor, and intron 8 1187-10 G>A 3’ splice acceptor. Objectives: 1) To determine the presence of LDL-R gene mutation among first degree relatives of FH patients. 2) To describe the phenotype among these family members. Methodology: First degree relatives of the index patients identified with novel LDL-R mutations in the 2005 study of Punzalan et al. were recruited for the study. For each participant, clinical history and lipid profile were obtained and the presence/absence of the proband mutation determined via direct sequencing. The p values were calculated using the Kruskall-Wallis test. Results: A total of seven participants from four families were included, two (28.6%) males and five females (71.4%) with ages ranging from 16 to 60 years old and having a mean LDL cholesterol level of 4.874 mmol/L. Calculated p values show significant association of elevated LDL- and total cholesterol levels as well as HDL ratio with the presence of an LDL-R mutation. Conclusion: Consistent with the known etiology of FH, elevated LDL cholesterol levels was found to be associated with mutations in the LDL-R gene.

2155M
Cardiovascular Disease in Pediatric Patients with Ciliopathies. S. Bowdin1,2, N. Karp1. 1) Hospital for Sick Children, Clinical Genetics, Hospital for Sick Children, Toronto, Ontario, Canada; 2) Department of Pediatrics, University of Toronto, Canada.

Ciliopathies are a heterogeneous group of disorders caused by mutations in single genes, which alter proteins of the cilium-centrosome complex (CCC). As the primary cilia can play multiple roles and are ubiquitously distributed throughout the body, abnormal ciliary proteins can lead to multiple organ malformations and dysfunctions, including congenital heart disease. Published information on cardiovascular disease in patients with ciliopathies remains scarce, although the presence of significant cardiac disease is of cardinal importance to the clinical outcome. We conducted a retrospective review of the hospital charts and echocardiograms/electrocardiograms of 72 pediatric patients diagnosed with a ciliopathy syndrome, all of whom were cared for at the Hospital for Sick Children, in Toronto, Canada between 1983-2013. Each diagnosis of a ciliopathy was made on the basis of clinical phenotype, when possible confirmed by molecular genetic testing. A total of 23 out of 72 ciliopathy patients (32%) were found to have an abnormal echocardiogram. Five out of six patients with Alstrom syndrome had dilated cardiomyopathy as well as other structural cardiac anomalies and ECG findings. Six out of 39 patients with Bardet-Biedl syndrome, and 7 out of 19 patients with Joubert syndrome and related disorders, had structural cardiac anomalies and/or rhythm disturbances. Three out of 6 patients with nephronophthisis had hypertension but none had structural cardiac defects. Seven patients with Leber Congenital Amaurosis, 1 patient with Senior-Loken syndrome, and 1 patient with Oral-Facial-Digital Syndrome type 1 had no documented cardiovascular disease. This data suggests that cardiovascular evaluation should be given to a cardiovascular evaluation for all patients with a known or suspected ciliopathy. In the future, collection of prospective data may enable screening recommendations to be tailored to the patient’s specific ciliopathy diagnosis.
2156M
Implantable cardioverter defibrillator (ICD) therapy in two founder arrhythmmogenic right ventricular cardiomyopathy (ARVC) populations with mutations in TMEM43 (p.S358L) and PKP-2 (p.Q378X). When genotype matters. K. Hodgkinson1,3, P. Bolland2, A. Bullen1, F. Curtis2, A. Collier3, S. Stuckless1, T.-L. Young3, B. Fernandez3, A.J. Howes2, S. Connors1. 1) Clinical Epidemiology, Faculty of Medicine, Memorial University, Health Sciences Centre, St John’s, NL, Canada; 2) Division of Cardiology, Faculty of Medicine, Memorial University, Health Sciences Centre, St. John’s, NL, Canada; 3) Discipline of Genetics, Faculty of Medicine Memorial University, Health Sciences Centre, St. John’s, NL, Canada.
Background: Arrhythmmogenic right ventricular cardiomyopathy (ARVC) may lead to sudden cardiac death (SCD) due to ventricular tachycardia (VT) for which implantable cardioverter defibrillator (ICD) therapy is warranted. Determining when to provide primary ICD therapy in asymptomatic individuals is difficult. We assessed the efficacy of ICD therapy in two homogeneous founder cohorts with disease-causing mutations in PKP-2 (p.Q378X) and TMEM43 (p.S358L), the latter known for the significant difference in survival between mutation positive and negative individuals and the significantly worse disease in males than females. Affected status comprised mutation positive and/or clinically positive individuals. Unaffected status could only be defined as mutation negative. Methods: Genetic and clinical information was obtained from an existing database. Of 867 individuals at a priori 50% risk for TMEM43 (p.S358L), mutation status was known in 627 (398 affected). Of 117 individuals at a priori 50% risk for PKP-2 (p.Q378X), disease status was known in 61 (42 affected). The natural history to death in the PKP2 cohort was defined and outcomes following ICD in both cohorts assessed. Results: In 129 TMEM43 (p.S358L) mutation positive individuals with an ICD for primary prophylaxis, survival was significantly improved in both males (p<0.001, RR 9.31, 95% CI 3.34-26.00) and females (p<0.001, RR 3.57, 95% CI 1.34-9.54). In the PKP2 cohort, there was no survival difference between affected and unaffected individuals (p=0.533). Ten affected individuals had ICDs, nine for primary prophylaxis. One had VT requiring defibrillation. One SCD occurred after ICD explantation (implant infection). Six (67%) individuals had inappropriate ICD therapy encompassing 92 defibrillations. Discussion: In the missense TMEM43 (p.S358L) cohort there was a significant reduction in all-cause mortality in mutation positive individuals with the nonsense PKP-2 (p.Q378X) mutation had no difference in survival and thus no benefit from the ICD. However, they had an increased burden of inappropriate defibrillation. Conclusion: ICD therapy for individuals with the PKP-2 (p.Q378X) mutation should be considered as primary VT whereas primary ICD therapy is not indicated for individuals with TMEM43 (p.S358L). This translational research supports the necessity for mutation-specific population studies to provide accurate natural history and clinical course information to help direct costs and lower patient burden.

2157S
Elevated risk of abnormal arterial remodeling in relatives of individuals with fibromuscular dysplasia. A. Katz1,2, M. Oerderoster2, S. Blackburn2, D. Coleman1, J. Stanley2, A. Ruesch2, J. Douglas3, J. Li2, S. Ganesh1. 1) Cardiovascular Medicine, University of Michigan, Ann Arbor, MI; 2) Vascular Surgery, University of Michigan, Ann Arbor, MI; 3) Human Genetics, University of Michigan, Ann Arbor, MI. Fibromuscular dysplasia (FMD) is a rare disease of abnormal arterial remodeling, predominantly occurring in women (90%). While prior research has suggested a genetic etiology of FMD, little is known of vascular risk to family members of affected individuals. We hypothesized that compared to the general population, first degree relatives of individuals with FMD have a higher rate of diseases of abnormal arterial remodeling, manifesting as FMD with multi- or uni-focal stenosis, arterial aneurysm, and arterial dissection. In individuals with FMD were enrolled through care received at the University of Michigan and through self-referral to our center. Probands (N = 62) all had physician diagnosed FMD. Medical history of the family members was assessed for: (1) physician diagnosed FMD, (2) clinical signs of suspected but not formally diagnosed FMD, (3) aneurysm without evidence of FMD. Clinical signs of suspected FMD were stroke before the age of 55, hypertension diagnosed before age 30, renal artery stenosis, or sudden death without clear alternative etiology. Familial relative risk was calculated by comparing observed number of cases to expected number of cases among relatives of probands. Among all first degree relatives for which a medical history could be obtained (373; 120 parents, 153 full siblings, and 100 offspring), 2.7% had physician diagnosed FMD, 4.8% had clinically suspected FMD, and 3.5% had an aneurysm. Among female relatives, 5.5% (10/181) had physician diagnosed FMD, and 9.9% (18/181) had physician suspected FMD, while 11.7% (20/177) had an aneurysm with a physician diagnosed FMD. Using an estimate of 0.1% population prevalence of FMD, these data represent a crude unadjusted relative risk of 55 for female first degree relatives and 64 for sisters of FMD affecteds. Among proband fathers, 11.7% (7/60) had been diagnosed with an abdominal or lower extremity aneurysm (AAA). Using an estimate of 6% prevalence of AAA among men over 65 (from population screening studies), this represents a relative risk of 1.9. Our data indicate an elevated risk of arterial dysplasia in relatives of individuals with FMD, which supports the hypothesis that FMD is a heritable, systemic arteriopathy with variable clinical manifestations that may be specific to gender. These findings have potential clinical screening implications for family members of individuals with FMD. Further studies are needed to delineate the genetic basis and clinical presentations of FMD.

2158M
First 2 years of experience of an integrated multidisciplinary clinic for adults with aortopathies in a Canadian context. A.M. Laberge1,2, I. El Hamamsy3, L. Robb1, N. Poirier2, P. Demers2, F.P. Mongeon2, 1) Division of Medical Genetics, Dept of Pediatrics, CHU Sainte-Justine, Montreal, Canada; 2) Dept of Heart Surgery, Montreal Heart Institute, Montreal, Canada; 3) Dept of Medicine, Montreal Heart Institute, Montreal, Canada.
In 2012, the Montreal Heart Institute started an integrated multidisciplinary clinic for adults referred for suspicion of Marfan syndrome or other connective tissue disorders at risk of aortic disease. In this clinic, a specialized heart team (cardiologist specialized in adult congenital heart disease and heart surgeons specialized in aortic surgery) and a genetics team (clinical geneticist, genetic counselor) see patients side-by-side. Patients with a family history of aortic disease or systemic features of Marfan syndrome are seen initially by both teams, unless a genetic diagnosis had previously been confirmed. Patients with presumed isolated aortic disease are seen first by the heart team, who determines if evaluation by the genetics team is needed. Special consultation corridors were established with a specialized orthopedist, ophthalmologist and physiatrist. A patient-centered, holistic approach is our standard of care. We report here our experience with this integrated multidisciplinary clinic over our first two years of activities. Between May 2012 and April 2014, 183 new patients were seen, from 146 different families. Reasons for referral included suspicion of Marfan syndrome (72), TAAD (56), suspicion of Loey-Dietz syndrome (15), suspicion of Ehlers-Danlos syndrome (8), sudden death in the family (6) and various other reasons (26). All were seen by the heart team, 70 were seen by the geneticist for a dysmorphological exam. All had dedicated cardiovascular imaging in our center. Genetic tests were ordered for 35 patients. To date, genetic testing led to identification of mutations in FBN1 in eight patients with suspected Marfan syndrome and ACTA2 mutations in two families. Close links with the equivalent pediatric clinic and the prenatal genetics clinic at the affiliated mother and child university health center facilitated the evaluation of eight children of our adult patients. Most importantly, these close links allowed for the rapid evaluation of three pregnant women at risk of aortic disease and eight affected parents who had no active follow-up at the time their child was seen in the pediatric clinic. Our integrated multidisciplinary approach results in efficient access to specialized cardiac and genetic assessment and rapid treatment when required. Further follow-up is required to ascertain the value of such dedicated clinics in altering the natural history of these conditions.
2159S Patterns of discordant phenotypes in familial congenital heart disease. L.A. Larsen1, S.G. Ellesæe2, C.T. Workman2, P. Bouvagnet2, K. van Engelen3, R. Hinton4, V. Hoylaerts5, C.A. Mulder6, K. McBride3, B.J.M. Mulder10, A. Postma6, L. Søndergaard12, S. Brunak6,11. 1) Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark; 2) NNF Center for Protein Research, University of Copenhagen, Copenhagen, Denmark; 3) Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; 4) Laboratoire Cardiogénétique, Hospices Civils de Lyon, Groupe Hospitalier Est, Bron, France; 5) Department of Clinical Genetics, Academic Medical Centre, Amsterdam, The Netherlands; 6) Division of Cardiology, The Heart Institute, Cincinnati Children’s Hospital Medical Center; 7) Department of Cardiothoracic Surgery, Aarhus University Hospital, Skejby, Denmark; 8) Department of Oncology, Lombardi Cancer Center, Georgetown University, Washington, DC; 9) Center for Cardiovascular and Pulmonary Research, Nationwide Children’s Hospital, Department of Pediatrics, Ohio State University, Columbus OH; 10) Department of Cardiology, Academic Medical Centre, Amsterdam, The Netherlands; 11) Department of Anatomy, Embryology & Physiology, Academic Medical Centre, Amsterdam, The Netherlands; 12) Department of Cardiology, Rigshospitalet, University Hospital of Copenhagen, Denmark.

Congenital heart disease (CHD) affects approximately 1% of live births. CHD may be caused by both genetic and environmental factors, but much of the molecular pathology remains to be elucidated. Observations of co-occurrence of specific heart defects in CHD families and in specific genetic syndromes suggest that some discordant heart defects may share susceptibility genes or pathways. Here, we have investigated the occurrence of concordant and discordant heart defects in a dataset of 4004 clinically diagnosed CHD cases from 1175 families with at least two affected individuals per family (663 unpublished families, 512 previously published families). To investigate if specific discordant heart defects co-occur in families, we arranged the dataset into 62 anatomical subgroups of heart defects. Next, to determine the strength of association between pairs of anatomical subgroups, we calculated the log odds ratio of familial occurrence for each pair. Hierarchical clustering of the observed log odds ratios revealed specific patterns of occurrence of discordant heart defects within families, supporting a genetic model of CHD susceptibility genes. The analysis revealed correlation between phenotype-pairs with high familial occurrence and phenotype-pairs with a large number of overlapping susceptibility genes in mice. For example; ventricular septal defect pairs with high familial occurrence and phenotype-pairs with a large number of overlapping ventricular septal defects rarely co-occur with right-sided defects. Targeted deletion of specific genes in mouse models often results in more than one type of heart defects typical for heterotaxy co-occur in families and suggest that left-sided defects rarely co-occur with right-sided defects. Targeted deletion of specific genes in mouse models often results in more than one type of heart defect, and for a specific subgroup of heart defects several overlapping susceptibility genes may exist in mouse models. We used generated phenotype-data from mouse models to investigate if the observed patterns of occurrence in families could be explained by shared susceptibility genes; we compared the pairwise discordance log odds ratios in families with pairwise subgroups of heart defects observed in mouse models, and we calculated the frequency of overlapping susceptibility genes between the pairs of phenotypes. The analysis revealed correlation between phenotype-pairs with high familial occurrence and phenotype-pairs with a large number of overlapping susceptibility genes in mice. For example; ventricular septal defects and interrupted aortic arch show high familial occurrence and in mouse models 14 susceptibility genes are overlapping between these two groups. In summary, we observed significant clustering of specific discordant CHD phenotypes within families and our data suggest that shared susceptibility genes may be an underlying mechanism.

2160M Titin As a Gene for Conduction Defects With and Without Cardiomyopathy. E. Smith1, A. Mansi2,3. 1) Program for Cardiovascular Genetics, Department of Internal Medicine/Section of Cardiology, Yale University School of Medicine, New Haven, CT; 2) Department of Genetics, Yale University School of Medicine, New Haven, CT.

Radical mutations in the gene Titin (TTN) have been associated with dilated cardiomyopathy. Mutations in this gene are also reported to be associated with arrhythmogenic right ventricular cardiomyopathy, but not with isolated arrhythmia. We report three cases of sudden cardiac death due to ventricular fibrillation, where the underlying genetic causes have been narrowed down to a radical, protein altering TTN mutation. Case one is a young woman who had cardiac arrest due to ventricular fibrillation while driving a car. After resuscitation she underwent a single chamber implantable cardioverter defibrillator (ICD) placement. Post-arrest, she had several runs of non-sustained ventricular tachycardia and frequent bouts of premature ventricular excitation. Genetic testing by whole exome sequencing (WES) identified a TTN frameshift (p.86985fs) mutation. She was later found to have mildly dilated left ventricle by echocardiography examination. The subject had family history of unspecified sudden death at young ages. Case two is a gentleman who had ventricular fibrillation at home. He had a history of paroxysmal atrial fibrillation. He too had a single chamber ICD implantation post-arrest. There is no history of cardiomyopathy or sudden death in the family. His genetic testing by WES identified a TTN nonsense mutation (p.I81966X). He did not meet the criteria for dilated cardiomyopathy and instead had mild left ventricular hypertrophy. Case three is a woman who also had ventricular fibrillation while out with friends. Though she carried a diagnosis of hypertrophic cardiomyopathy post-arrest, serial echocardiograms did not corroborate this finding. An automatic implantable cardioverter defibrillator (AICD) was implanted post-arrest. Her family history was significant for young death in her maternal grandparents and great-grandfather. Genetic testing revealed a splice site substitution (c.2809-2A>C). These three isolated cases raise the possibility that radical TTN mutations may underlie malignant ventricular arrhythmia with and without cardiomyopathy, which constitutes a new phenotype for this gene. Further studies are needed to establish the association and to determine the mechanism.

2161S Cystatin C and cardiovascular disease: a Mendelian randomization study. S.W. van der Laan1’, T. Fall2, J. van Setten3,4, P.I.W. de Bakker3,4, G. Pastorik3, J. Amlöv5, M.V. Holmes6, F.W. Asselbergs7,8,10 on behalf of the Cystatin C MR Consortium. 1) Experimental Cardiology, UMC Utrecht, Utrecht, the Netherlands; 2) Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 3) Department of Medical Genetics, UMC Utrecht, Utrecht, the Netherlands; 4) Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, the Netherlands; 5) Penn Medicine, University of Pennsylvania Health System, United States of America; 6) Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom; 7) Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, the Netherlands; 8) Department of Cardiology, UMC Utrecht, Utrecht, the Netherlands.

Observational studies show a strong dose-response association between circulating Cystatin C (encoded by CST3) and incident coronary heart disease (CHD), independent of traditional risk factors and renal function. This supports the hypothesis that circulating Cystatin C could represent a causal factor for CHD. However, residual confounding and reverse causality could be alternative explanations that are difficult to tease from observational studies. We sought to investigate the causal role of Cystatin C in CHD development by conducting a Mendelian randomization (MR) analysis using a common variant in the CST3 locus.

We conducted a MR analysis in 43 studies including 219,219 individuals with 37,321 measures of Cystatin C and 41,162 CHD events. We used rs911119 (or a proxy, r^2>0.90) in the CST3 locus (identified previously by GWAS) as a genetic instrument for MR analysis.

Cystatin C associated with risk of CHD in an observational analysis adjusted for age and sex (odds ratio [OR] 1.20; 95% confidence interval [CI]: 1.10, 1.31) per doubling of Cystatin C concentration; p=8.87×10^-4; additional adjustment for confounders (smoking, HDL-cholesterol, BMI, CKD-EPI, and systolic blood pressure) diminished the association (OR 1.10; 95%CI 1.00, 1.21) per doubling of Cystatin C concentration; p=0.19). Rs911119 had a strong effect on circulating Cystatin C concentration (95%CI 0.066, -0.057; p=4.9×10^-145 per effect allele). However, the variant did not show significant association with risk of CHD (OR 1.01 (95%CI 0.99, 1.03; p=0.41).

In summary we replicated the association of Cystatin C with CHD risk and show a strong association of rs911119 with circulating Cystatin C. However, we find no evidence for a causal role of Cystatin C in the development of CHD.
2162M

Association of oxidative DNA damage with Folic acid metabolizing genes in Children with congenital septal defects. S. Syed1, K. Koneiti2, K. Kola3, D. Dadaala4, G. Gundimeda5, A. Akka6, M. Mufururu7, 1) Institute of Genetics, Osmania University, Hyderabad, India; 2) Care Hospital, Banjara Hills, Hyderabad; 3) National Institute of Nutrition, Tarnaka, Hyderabad; 4) Indo american cancer hospital hyderabad.

Association of oxidative DNA damage with Folic acid metabolizing genes in Children with congenital septal defects Sunayana Begum Syed*, Nageswara Rao Koneiti*, Srujana Kola*, Sujatha Dadaala**, Sandhya Devi Gundimeda**, Jyothy Akka*, Hema Prasad Mundurulu* 1 Institute of Genetics & Hospital, Osmania University, Hyderabad, India; 2 Care Hospital, Banjara Hills, Hyderabad. **National Institute of Nutrition, Tarnaka, Hyderabad *** Indo american cancer hospital, Hyderabad Background: Congenital Septal Defect (CSD) is the most common developmental anomaly occurring in 1 in 1000 live birth with unknown disease etiology. The present study was designed to understand the possible etiological factors using FISH and comet assay for chromosomal and oxidative DNA damage.

SNPs of folic acid metabolism and antioxidant genes MTHFR, MTRR, RFC and GST were studied by PCR-RFLP for the oxidative stress induced abnormal cardiac development during embryogenesis. Methods: The present study includes 162 children with CSD and their mothers and age matched controls (n=174). Blood samples were collected from Department of Pediatric Cardiology, Care hospital, Hyderabad with prior clearance from institutional ethics committee and written consent from parents. Lymphocytes were isolated for comet assay and FISH analysis. RBCs were used for the estimation of folic acid and Vitamin 12 levels. Inflammatory marker like C-reactive protein (CRP) was analyzed by ELISA for correlating with DNA damage. Genetic scores were calculated for PCR-based RFLP analysis of SNPs 677C>T, A69G, A80G, GSTM1 and GSTT1 genes. Results: Low RBC folate and high CRP levels were found in patients than the control group, demonstrating the severity of oxidative DNA damage. Fasting vitamin B12 levels in CSD and their mothers were significantly decreased than the control group. 22g/11.2 micro deletons and significant DNA damage were observed in CSD patients. In MTHFR C677T the overall mutant T allele frequency was higher in patient. Similarly the frequencies of G’G alleles of the genotypes ATRA69G/RFC1A80G were all found to be increased in CSD patients. Among the mothers of CSD children a 3 fold increase in frequencies of mutated T’G/G’ allele was observed when compared to the controls. Conclusion: Elevated DNA damage, increased CRP levels and decreased folate, vitamin B12 indicates that oxidative stress plays an essential role in CSD.

2163S

Causal role of alcohol consumption in blood levels of lipids and hemostatic factors, and risks of coronary heart disease and ischemic stroke. K. Vu1, E. Boerwinkle1,2, A.C. Morrison1, 1) School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; 2) The Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA.

Background: Moderate alcohol use may reduce the risk of coronary heart disease (CHD) and ischemic stroke (IS) through an improved blood lipid profile and reduced hemostatic factor levels. However, the observed association may be confounded. Mendelian randomization (MR) can help to examine the causality of the association. Methods: The study includes 10,893 Japanese (49.5%) and 9533 Africans (50.5%) from the China Kadoorie Biobank and the Africa-United Kingdom Obesogeny Study (UKOS), respectively. Alcohol use was measured as grams of alcohol/day (g/day). *ADH1B*2 and *ALDH2*2 (rs671) genotypes were combined into unweighted genetic scores. Alcohol use was measured using FISH and comet assay for chromosomal and oxidative DNA damage. SNPs of folic acid metabolism and antioxidant genes MTHFR, MTRR, RFC and GST were studied by PCR-RFLP for the oxidative stress induced abnormal cardiac development during embryogenesis. Results: The present study includes 162 children with CSD and their mothers and age matched controls (n=174). Blood samples were collected from Department of Pediatric Cardiology, Care hospital, Hyderabad with prior clearance from institutional ethics committee and written consent from parents. Lymphocytes were isolated for comet assay and FISH analysis. RBCs were used for the estimation of folic acid and Vitamin 12 levels. Inflammatory marker like C-reactive protein (CRP) was analyzed by ELISA for correlating with DNA damage. Genetic scores were calculated for PCR-based RFLP analysis of SNPs 677C>T, A69G, A80G, GSTM1 and GSTT1 genes. Results: Low RBC folate and high CRP levels were found in patients than the control group, demonstrating the severity of oxidative DNA damage. Fasting vitamin B12 levels in CSD and their mothers were significantly decreased than the control group. 22g/11.2 micro deletons and significant DNA damage were observed in CSD patients. In MTHFR C677T the overall mutant T allele frequency was higher in patient. Similarly the frequencies of G’G alleles of the genotypes ATRA69G/RFC1A80G were all found to be increased in CSD patients. Among the mothers of CSD children a 3 fold increase in frequencies of mutated T’G/G’ allele was observed when compared to the controls. Conclusion: Elevated DNA damage, increased CRP levels and decreased folate, vitamin B12 indicates that oxidative stress plays an essential role in CSD.

2164M

Mendelian randomisation study of alcohol and cardio-metabolic risk factors. I.Y. Millwood1, L. Li2,3, R.G. Walters4, W. Mei1, D. Bennett1, Y. Guo5, Z. Bian6, R. Peto1, R. Collins1, S. Parish1, R. Clarke1, Z. Chen1, Y. China Kadoorie Biobank collaborative group. 1) Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford; 2) School of Public Health, Peking University Health Science Center, Beijing, China; 3) Chinese Academy of Medical Sciences, Dong Cheng District, Beijing, China.

Background: To explore the causal relationship between alcohol and cardio-metabolic risk factors in a large Chinese population, a Mendelian randomisation approach was taken using alcohol metabolising gene variants ADH1B*2 and ALDH2*2. Methods: The study includes 82,155 men and women aged 30-79 years from the China Kadoorie Biobank prospective cohort, recruited in 2004-2008 from ten regions across China, with baseline data on alcohol consumption, physical measurements, random blood glucose and ADH1B (rs1229984) and ALDH2 (rs671) genotypes. Results: 10,929 men (34%) and 1014 (2%) women were weekly drinkers, consuming mean (SE) 40.9 (0.3) and 16.7 (0.6) g alcohol/day, respectively. Allele frequency of ADH1B*2 was 0.60 and ALDH2*2 was 0.20, and both varied by region. For each acetaldehyde increasing allele, alcohol intake was lower by 3.4 (0.5) g/day (ADH1B*2) and 14.0 (0.9) g/day (ALDH2*2) in men, and by 1.8 (0.8) g/day (ADH1B*2) and 5.1 (1.6) g/day (ALDH2*2) in women, adjusted for age and region. Table 1 shows the association of reported alcohol consumption (adjusted for covariates) and alcohol predicted from a weighted genetic risk score (GRS) combining ADH1B*2 and ALDH2*2 (adjusted for age and region), with outcomes, in male weekly drinkers (*P<0.05, **P<0.0001). Results in female weekly drinkers were generally consistent, but not significant.

<table>
<thead>
<tr>
<th>B e t a ( S E )</th>
<th>95% C I</th>
<th>W h i s t l e</th>
<th>W a i s t</th>
<th>B o d y f a t</th>
<th>Glucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>0.05 (0.04)**</td>
<td>0.001 (0.000)**</td>
<td>0.01 (0.001)**</td>
<td>0.01 (0.001)**</td>
<td>0.01 (0.001)**</td>
</tr>
<tr>
<td>**</td>
<td>0.10 (0.06)**</td>
<td>0.001 (0.000)**</td>
<td>0.01 (0.001)**</td>
<td>0.01 (0.001)**</td>
<td>0.01 (0.001)**</td>
</tr>
</tbody>
</table>

Conclusion: Alcohol has a strong causal effect on blood pressure, with 10 g/day (~one drink) increasing SBP by about 2 mmHg, and a modest adverse effect on glucose and measures of overall and central adiposity. Stronger effects seen using gene variants could reflect lifelong consumption patterns, and the limitations of self-reported questionnaire data.
2165S
Heritability and linkage study on heart rate variability in an isolated Arab population, L. Moon1, M.O Hassan2, Y.S. Voruganti2, D. Jay1, A. Albanaran1, A. Aslani2, R. Bayoumi2, S. Al-Yahyaee2, A.G. Comuzcio1, B.Z. Alizadeh1, I. Nolte1, H. Snieder1. 1) Department of Epidemiology, University Medical Center Groningen, University of Groningen, The Netherlands; 2) College of Medicine and Health Sciences, Sultan Qaboos University, Muscat, Sultanate of Oman; 3) Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, USA; 4) Department of Genetics, Texas Biomedical Research Institute, Texas, USA.

Background: Heart rate variability (HRV) is a reliable, non-invasive, economical reflection of many physiological and psychological factors regulating the normal heart rhythm. It is often used as an index of cardiac parasympathetic system activity. In the general population decreased HRV is an independent risk factor for morbidity and mortality. Individual differences of HRV can be attributed to genetic factors and are found to be more pronounced in subjects during stress tests. However, most published studies on the genetic influence on HRV have focused on subjects of European and African Americans descent. In Oman Family Study (OFS) we aim to: (a) estimate and quantify the contribution of genes to the variance of HRV, (b) calculate the overlap in genetic and environmental influences on HRV at rest and under stress using uni- and bivariate analyses; and (c) perform a multipoint variance component-based linkage analysis for all HRV parameters. Methods: The OFS consists of five large, multigenerational pedigrees of Arab families living in Willayat (state) of Nizwa. The strengths of the OFS are: (1) geographically isolated which provides a more homogeneous environmental exposure, (2) similar socio-economic status and health-related habits, (3) highly consanguineous and is represented by no less than 6 meioses. This may indicate shared environmental factors contribute by 6 meioses. This may indicate shared environmental factors contribute.

2167S
Characteristics of Aortic Disease Associated with ACTA2 mutations. E.S. Regalado1, D. Guo1, S. Prakash1, T.A. Bensend1, K. Flynn1, A. Estrera2, H. Safi2, D. Liang1, J. Hyland1, A. Child2, O. Al-Jaroudi2, V.S. Voruganti1, G. Jondeau3, A. Braverman4, R. Moran5, T. Morisaki6, H. Morisaki6, R. Pyentz7, J. Coselli8, S. LeMaitre9, D.M. Milewicz9, M. Montalco Aortic Consortium. 1) Department of Internal Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA; 2) Department of Cardiothoracic and Vascular Surgery, University of Texas Health Science Center at Houston, Houston, TX, USA; 3) Department of Medicine, Stanford University Medical Center, Stanford, CA, USA; 4) Connective Tissue Gene Tests, Allentown, PA, USA; 5) Department of Cardiac and Vascular Sciences, St George’s, University of London, London, UK; 6) AP-HP, Hopital Bichat, Centre National de Reference pour le syndrome de Marfan et apparentes, Paris, France; 7) Universite Paris 7, Paris, France; AP-HP, Hopital Bichat, Laboratoire de Genetique moleculaire, Boulogne, France; INSERM, U1148, Paris, France; 8) Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO, USA; 9) Cleveland Clinic, Cleveland, OH, USA; 10) Department of Bioscience and Genetics, National Cerebral and Cardiovascular Center Research Institute, Osaka, Japan; 11) Perelman School of Medicine, the University of Pennsylvania, Philadelphia, PA, USA; 12) Division of Cardiothoracic Surgery, Michael E. DeBakey Department of Surgery, Baylor College of Medicine, the Texas Heart Institute, and Baylor St. Luke’s Medical Center, Houston, TX, USA.

ACTA2 is the most frequently mutated gene causing familial thoracic aortic aneurysms and dissections. Patients carrying ACTA2 mutations have significantly lower penetrance of aortic events (88%) at variable ages ranging from 12 to 76 years and associated with a 25% mortality rate. Type A dissections were more common than type B dissections (54% versus 21%), but the median age of onset of type A dissections (27 years) was significantly younger than type B dissections (27 vs. 36 years of age). Among individuals presenting with type A dissections, the maximum diameter ranged from 3.8-9.5 cm, and one-third of these individuals had dissections at diameters <5.0 cm. Only 12% of aortic events were repair of ascending aortic aneurysms at diameters ≥200 mg/dL & ≤190 mg/dL, HDL-C ≥190 mg/dL). The overlap in genetic and environmental influences on HRV at rest and under stress using uni- and bivariate analyses; and (c) perform a multipoint variance component-based linkage analysis for all HRV parameters. Meth-
Are short telomeres a cause or consequence of hypertension in spontaneously hypertensive mice? C.L Chiu, N.L Hearn, D. Paine, N. Steiner, J.M Lind. University of Western Sydney, School of Medicine, New South Wales, Australia.

Telomere length is widely considered as a marker of biological ageing. Clinical studies have reported associations between reduced telomere length and hypertension. However, the role of telomere length in the pathogenesis of hypertension, and whether reduced telomere length is a cause or consequence of disease, remains unknown. The aim of this study was to determine whether telomere shortening occurred prior to the onset of disease in spontaneously hypertensive mice. Spontaneously hypertensive Schlager mice (BPH/2J) and their normotensive controls (BPN/3J) were used in this study. Genomic DNA was extracted from kidney tissue of 4, 12 and 20 week old male BPH/2J and BPN/3J mice (n=10/group). Relative telomere length (T/S) was measured using quantitative PCR. Linear correlation estimates were performed to analyse telomere length over time within a strain. A general linear model with repeat measures testing was used to compare rate of telomere shortening between groups. A general linear model was used to compare relative telomere lengths and gene expression between groups. P<0.05 was considered significant. Age was inversely related to telomere length. In 4 week old pre-disease animals no difference in T/S was observed between BPH/2J and BPN/3J animals (p=0.09). The rate of telomere attrition between BPH/2J and BPN/3J was significantly different (p=0.001). At 12 and 20 weeks, established disease, BPH/2J animals had significantly shorter telomeres when compared to their age-matched controls (12 weeks p<0.001 and 20 weeks p=0.004). This is the first study to show that reduced telomere length occurs after the development of hypertension, indicating that this is not the cause of hypertension in spontaneously hypertensive mice. Further studies are needed to determine the mechanisms which lead to the development of hypertension and the shortening of telomeres in these animals.
2169S
Infusions of dexamethasone loaded erythrocytes in ataxia telangiectasia patients. L. Chessa,1 V. Leuzer,2 R. Michel1, M. Piane,1 D. D’Agnan,3 A. Molinaro1, T. Venturi,2 A. Pleban2, AR. Sorensen3, E. Fazzini2, M. Marinini3, P. Ferretri Leali,2 I. Quinti1, FM. Cavaliere1, MC. Pietrogrande2, A. Finocchi1, M. Magnani4,5. 1) Dept. Clinical Molecular Medicine, Sapienza University, Rome, Italy; 2) Dept. Pediatrics Child Neurology and Psychiatry, Sapienza University, Rome, Italy; 3) Child Neurology and Psychiatry, Sapieni Civil and University, Brescia, Italy; 4) School Reproductive and Developmental Science, University of Trieste and Brescia, Italy; 5) Dept. Clinical Experimental Sciences, Spedali Civili and University, Brescia, Italy; 6) Dept. Molecular Medicine and Surgery, University Roma, Italy; 7) Dept. Pediatrics, Milano University and Fondazione IRCCS Ca’ Granda, Milano, Italy; 8) Dept. Pediatrics, Ospedale Pediatrico Bambino Gesù and University Tor Vergata, Roma, Italy; 9) Dept. Biomedical Sciences, University Carlo Bo, Urbino, Italy; 10) Erdyel S.p.A, Urbino, Italy.

Background. Ataxia Telangiectasia (AT) is a rare devastating neurodegenerative disease presenting with early onset ataxia, oculocutaneous telangiectasias, immunodeficiency, radio sensitivity, and proneness to cancer. In a previous phase II study we showed that 6 monthly infusions of autologous erythrocytes loaded with dexamethasone (EryDex®) were effective in improving the neurological impairment in young AT patients. Here we report the results of the extension of this study for an additional 24 month-period.

Methods. After the end of first trial, 4 subjects continued to be treated with monthly EryDex infusions for further 24 months and their clinical outcome was compared with that of 7 age-matched subjects who had stopped the treatment after the first 6 infusions. The protocol included serial assessment of ataxia (by International Cooperative Ataxia Rating Scale) and adaptive behavior (by Certificate of Puyhura, Adaptive Behavior Scales), and clinical and laboratory tests revealing treatment- and steroid-dependent adverse effects, if present.

Results. Patients in the extended study experienced a continuous neurological improvement with respect to their pre-treatment status while control subjects showed a progressive neurological deterioration (according to the natural history of the disease) after the discontinuation of the treatment. The delivery system we adopted proved to be safe and well-tolerated and none of the side effects usually associated with the chronic administration of corticosteroids were observed during the whole treatment period. Preliminary promising results call for a wide-scale controlled study on protracted treatment of AT patients with dexamethasone loaded erythrocytes.

2170M
Inta-nasal DDAVP administration for the prevention of massive subcutaneous hematoma in dermatan 4-O-sulfotransferase 1 (D4ST1)-deficient Ehlers-Danlos Syndrome (DDEDS). T. Koshio1, M. Ishikawa1, E. Kise1, J. Takahashi1, S. Yuzuhara1, Y. Fukushima1, 1) Dept Med Gen, Shinshu Univ Sch Med, Matsumoto, Japan; 2) Department of Plastic and Reconstructive Surgery, Shinshu University School of Medicine, Matsumoto, Japan.

Background. EDS, which could cause significant reduction of ADL and QOL. In this report, we present experiences of three patients (7-years-old girl; 8-years-old girl; 13-years-old boy) suffering from recurrent serious hematomas and receiving intra-nasal 1-desamino-8-D-arginine vasopressin (DDAVP), the only agent described to have efficacy in the control of bleeding (DDAVP), the only agent described to have efficacy in the control of bleeding.

Methods: All five cases responded more slowly. All recovered baseline or near baseline functioning. None of the side effects usually associated with the chronic administration of corticosteroids was observed during the whole treatment period. Extensive testing to rule out other causes of motor and cognitive regression was within normal limits. Based on the presence of multiple motor disturbances (slowing, grimacing, posturing), the adolescents were diagnosed with unspecified Catatonia and treated with anti-catatonic treatments (benzodiazepines/lecetocons/convulsive therapy). One patient who had a Pupillary Light Reflex (PLR) study prior to developing Catatonia was followed with PLRs over his treatment course.

Results: All five cases responded to initial treatment with a benzodiazepine, but required electroconvulsive therapy (ECT) to achieve remission. Longer duration Catatonia responded more slowly. All recovered baseline or near baseline functioning. A DS specific Diagnostic Protocol and Catatonia Improvement Scale were developed to standardize diagnosis and monitor treatment response over time. Changes in pupil constriction coincident with Catatonia implicates autonomic involvement. A survey of adolescents and young adults who attended one DS clinic as young children suggests a tentative prevalence of catatonic regression of 3%.

Conclusions. We suspect Catatonia is a common, overlooked cause of unexplained deterioration in adolescents and young adults with DS. Being unaware of Catatonia in DS in years of incapacitating illness and a risk of malignant catatonia, which may be induced by treatment with serotonergic or dopaminergic medications. Alerting geneticists to Catatonia may be a first step in preventing long-term negative consequences of life threatening neurologic syndrome and discovering the frequency and course of this disorder. PLR evidence of autonomic dysregulation may suggest why individuals with DS are at risk for Catatonia.
Phenylbutyrate increases pyruvate dehydrogenase complex activity in cells harboring a variety of defects. R. Ferreiro, A. Bourtou, M. Briquet, D. Kerr, E. Morava, R.J. Rodenburg, L. Bonafé, M.P. Baumgartner, Y. Anikster, N.E. Braverman et al. 10, N. Brunetti-Pierri et al.

Phenylbutyrate enhances PDHC enzymatic activity in vitro and in vivo by increasing the proportion of unphosphorylated enzyme through inhibition of pyruvate dehydrogenase kinases and thus, has potential for in vivo response in patients with PDHC deficiency harboring a wide spectrum of molecular defects.

217MS


Mutations in the valosin containing protein (VCP) gene cause hereditary inclusion body myopathy (hIBM) associated with Paget disease of bone (PDB), frontotemporal dementia (FTD), more recently termed multisystem proteinopathy (MSP). Affected individuals exhibit scapular winging and die from progressive muscle weakness, and cardiac and respiratory failure, typically in their 40s to 50s. Histopathologically, patients show the presence of rimmed vacuoles and TAR DNA-binding protein 43 (TDP-43)-positive large ubiquitinated inclusion bodies in the muscles. We have generated a VCPR155H+/- knock-in mouse model which recapitulates the disease phenotype and impaired autophagy typically observed in patients with VCP deficiency. Rapamycin and chloroquine at pharmacological doses have previously shown to alter the autophagic flux and autophagic signaling intermediates. Herein, we report results of administration of rapamycin, a specific inhibitor of the mechanistic target of rapamycin (mTOR) signaling pathway, in 19-month old VCPR155H+/- mice and VCPR155H+/-; knock-in mice harboring a variety of defects. Muta-
deh los of VCP, detected in most fibroblasts harboring VCP mutations, were directed to proteasome degradation. Large deletions affecting VCP have been associated with more severe disease. In the VCPR155H+/- mouse model, rapamycin treatment led to a significant increase in the proportion of unphosphorylated mTORc1, in vitro and in vivo by increasing the proportion of unphosphorylated enzyme through inhibition of pyruvate dehydrogenase kinases and thus, has potential for in vivo response in patients with PDHC deficiency harboring a wide spectrum of molecular defects.

2174M

Development of a cell-based reporter assay suitable for small-molecule drug discovery in FGFR3-inducible HEK293 cells stably expressing Klotho. S. Diener, K. Schoppert, B. Lorenz-Depiereux, K. Hadian, T.M. Strom 1. 1 Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuhemger, Bavaria, Germany; 2 Institute of Molecular Toxicology and Pharmacology, Assay Development and Screening Platform, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuhemger, Bavaria, Germany; 3 Institute of Human Genetics, Christian Albrechts Universität, Kiel, Germany; 4 Institute of Human Genetics, Iserl der Technischen Universität Munich, Munich, Bavaria, Germany.

Fibroblast growth factor 23 (FGF23) is a key regulator of phosphate homeostasis. It is of crucial importance in hereditary and acquired hypophosphatemic and hyperphosphatemic disorders. Moreover, FGF23 has emerged as a promising biomarker for the prediction of adverse clinical outcomes in patients with chronic kidney disease (CKD), as it might be related to mortality, cardiovascular abnormalities and disease progression. FGF23 is a bone-derived endocrine factor, which inhibits renal tubular phosphate reabsorption by interacting with receptor complexes of FGF receptor (FGFR) 1c and the co-receptor Klotho. As a major signalling pathway mitogen-activated protein kinase (MAPK) pathway is employed. For the investigation of FGF23 in an in vitro model we established FGFR3-inducible HEK293 cells that stably express Klotho (HEK293-KL). The induction of HEK293-KL cells by FGF23 was shown by detecting the activation of MAPK pathway, which could be reduced by the use of two known small-molecule inhibitors of MAPK pathway: SU5402 and U0126. To identify novel small-molecule compounds that modulate FGF23-FGFR1c/Klotho signaling, we have developed a cell-based reporter assay that is suited for high-throughput screening (HTS). The assay is based on the AlphaScreen SureFire platform of Perkin Elmer to monitor the phosphorylation of endogenous extracellular signal-regulated kinase 1 (ERK1) and 2 (ERK2) cellular target of HEK293-KL transfected cells treated with FGF23 in the presence of small-molecule compounds. Since increased plasma concentrations of FGF23 are the main cause of many phosphatemic disorders, a modulation of its effect could be a potential strategy for drug development and new therapeutic approaches in disorders affecting phosphate homeostasis.
2177S
Unfolded protein response induced by mutant alpha1-antitrypsin (AAT) activates JNK-MAPK pathway that modulates AAT levels and toxicity in AAT deficiency. N. Pastore1, B. Grannese1,2, C. Mueller2, J. Teckman1, N. Brunetti-Pierri1,2-3. 1) Telethon Institute of Genetics and Medicine, Italy; 2) Department of Pediatrics, Children’s Hospital of Philadelphia, Philadelphia, PA, USA; 3) Department of Biochemistry and Molecular Genetics, Boston University, Boston, MA, USA.

Alpha-1-antitrypsin (AAT) deficiency is the most common genetic cause of liver disease in children and liver transplantation is the only available treatment in severe patients. The majority of patients carry a missense mutation (lysine for glutamate at amino acid position 342) in the AAT gene that alters protein folding. Mutant AAT is prone to polymerize and aggregate in the endoplasmic reticulum (ER) of hepatocytes, causing liver injury by a gain-of-toxic mechanism. Accumulation of these polymers results in release of ER calcium, NFkB activation, and apoptosis. The involvement of unfolded protein response (UPR) in AAT deficiency has been controversial. We observed that the PiZ mouse, a transgenic mouse model that expresses the mutant human AAT gene and recapitulates the features of human liver disease, has an age-dependent accumulation of mutant AAT that decreases with aging. UPR markers ATF6, PERK, IRE1, BIP, and GRP94 were all found to be overexpressed in younger PiZ mice with more abundant load of mutant AAT. Next, we injected PiZ mice with a recombinant serotype 8 adenovirus vector that incorporate microRNA (miRNA) sequences targeting the AAT gene and resulting in mutant AAT knockdown in the liver. These mice displayed reduced activation of UPR as shown by decreased ATF6 and BIP protein levels thus confirming that UPR activation is dependent upon mutant AAT protein dosage at a given age. Finally, we observed that UPR induced by mutant AAT results in the activation of JNK-MAPK pathway in both PiZ cells and human liver samples from affected patients. Interestingly, inhibitors of this pathway decrease mutant AAT levels in a cell model of the disease, suggesting that these components might be new targets for mitigating AAT deficiency. Collectively, these data reveal new therapeutic entry points for treatment of AAT deficiency.

2178M
Case report: Efficacy of L-Citrulline supplementation in a patient with Mitochondrial Encephalopathy, Lactic Acidosis and Stroke-like episodes (MELAS) with behavioral disturbances and failure to thrive. E. Alkhuwaizi1, A. Shuen1, M. Lecointre2, M. Larose3, C. Saint-Martin4, W. Al-Hertani1. 1) Department of Medical Genetics, McGill University Health Centre, Montreal, QC, Canada; 2) Department of Nutrition, McGill University Health Centre, Montreal, QC, Canada; 3) Department of Psychiatry, McGill University Health Centre, Montreal, QC, Canada; 4) Department of Diagnostic Imaging, McGill University Health Centre, Montreal, QC, Canada.

MELAS (MELAS [MIM 540000]) is a mitochondrial multisystem condition that often manifests in childhood with seizures associated with stroke-like episodes and transient hemiparesis or cortical blindness and lactic acidosis. Therapeutic options for MELAS are limited at this time, and under trial. We report on an 8 year-old female of Sri Lankan origin, who presented with status epilepticus and found to have severe lactic acidosis (5-8.2 mmol/L). She subsequently developed mild to moderate bilateral sensorineural hearing loss. Her past medical history was significant for a seizure disorder, Wolff-Parkinson-White syndrome, failure to thrive and behavioral disturbances. Magnetic resonance imaging with spectroscopy (MRI/MRS) showed progressive volume loss of T2 and FLAIR signal in multiple cortical and subcortical regions and presence of lactate peaks. She was later confirmed to have the most common MELAS mutation (m.3243A>G) in the MT-TL1 (MIM 590050) gene. After her delayed diagnosis at 9 years of age, we started her on arginine supplementation (0.4g/kg/day) for approximately one year. Her weight remained unchanged on the 0.1-3rd percentile and height on the 3rd percentile in that time period, and she continued to deteriorate in her behavior and mood as noted on neuropsychiatric assessment. One year later, as a trial we switched her to citrulline supplementation (0.3mg/kg/day) and noted a remarkable improvement in her neuropsychiatric function, especially mood and behavior, was noted by her parents. Furthermore, admissions for seizures and stroke-like episodes were less frequent. Finally, her growth parameters increased up to the 15th-50th percentile for weight and 15th percentile for height, over a period of seven months, after initiation of citrulline. It is known that citrulline increases nitric oxide (NO) production more than arginine, and as published by El Hattab et. al., citrulline may potentially have a better therapeutic effect in MELAS, compared to arginine, with large trials currently underway. This case report illustrates impressive improvement in clinical outcome and quality of life with citrulline supplementation in our MELAS patient.

2179S
Combination Therapy To Enhance Antisense Mediated Exon Skipping for Duchenne Muscular Dystrophy. D.W. Wang1, E.I. Mokhonova1,2, L. Martinez1,2, D. Becerra1,2, M.J. Spencer1,2, S.F. Nelson1,2,5,3, M.C. Miceli1,2,3, D. Miceli1,2,3. 1) Microbiology Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA., USA; 2) Neurology, David Geffen School of Medicine, University of California, Los Angeles, California, USA; 3) Center for Duchenne Muscular Dystrophy, University of California Los Angeles, Los Angeles, California, USA; 4) Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, USA; 5) Molecular Biology Institute, University of California Los Angeles, California, USA; 6) Department of Pathology and Laboratory Medicine, University of California Los Angeles, Los Angeles, California, USA.

Duchenne muscular dystrophy (DMD) is a lethal X-linked recessive disorder primarily caused by frameshifting deletions in the DMD gene that result in a loss of dystrophin protein expression. Antisense oligonucleotide-mediated exon skipping shows promise as a therapeutic approach by restoring the DMD reading frame and correcting dystrophin protein localization to the DGC. Using a small-molecule screen, our lab has previously found that dantrolene, an FDA approved compound used to treat malignant hyperthermia, increases antisense oligonucleotide-mediated exon skipping. Previous work has evaluated dantrolene synergy with antisense oligonucleotide (AO) in the context of twice daily intraperitoneal administered dantrolene. We tested if dantrolene administered via oral gavage or incorporated into mouse chow had a similar efficacy enhancement of Dmd exon 23 antisense oligonucleotide (PMO23) skipping and dystrophin protein rescue. RNA was analyzed via qRT-PCR for amounts of skipped and full-length Dmd mRNA. Quantitative immunohistochemistry was performed to assess both rescued dystrophin localization to the DGC as well as proper localization at the sarcolemma. We have thus assessed a series of alternate and more viable methods for dantrolene delivery with a large number of replicates 3 week experiments, and determined that dantrolene mixed in chow delivers maximally effective and effective dosing. We further demonstrate that dantrolene administered with one half of the maximally effective dose of AO (greater than 100mg/kg) induces an equivalent amount of dystrophin as 100mg/kg per week AO or 300mg/kg per week AO. Thus, we have now determined new methods of dantrolene delivery that is compatible with long-term administration, and have established dosing of AO to assess both potential for enhancement of maximal effect and potential for dose saving effect in the upcoming 6 month treatment experiment. In addition to the animal studies, we have established that dantrolene enhancement of exon skipping of human exon 51 occurs in multiple independent human induced myotubes. For this work, we have established protocols for skin punch fibroblast growth and transdifferentiation into myotubes. Seven skin punch biopsies have been obtained from patients with exon 51 skippable mutations, and 5 have been reprogrammed into iDRMs. We are assessing AVI4657 mediated skipping with and without dantrolene in myotubes generated from these iDRM cell lines.
Inhibition of AKT signaling in Proteus and PROS cells: A simple model for cancer therapeutics targeting the AKT/PI3K pathway. L.G. Bie secker2,3, M.J. Lindhurst1, M.R. Younck1, D.T. Dransfield1. 1) NHGR, NIH, Bethesda, MD; 2) ArQule, Inc., Woburn, MA.

The AKT/PI3K (phosphatidylinositol-3-kinase) signaling pathway is critical for cellular growth, survival, protein synthesis, and glucose metabolism. Mutations in several genes that result in active AKT signaling have been implicated in disorders characterized by overgrowth and/or hypoglycemia and have been identified in numerous tumor samples. Identification of inhibitors that ameliorate the affects of these mutations is an important first step in the development of protocols to treat not only existing, but also newly emerging disorders such as Proteus syndrome (PS) and PROS (PIK3CA-related overgrowth-spectrum). We tested a small molecule imidopyridine pan-AKT inhibitor in skin fibroblasts harboring the AKT7 p.Glu17Lys mutation isolated from several patients with PS. Cells treated for 24 hours with 310-500 nM with this compound grown in both serum-free and serum-containing medium had reduced levels of AKT phosphorylation, with levels at the higher three doses approaching those of quiescent wild-type cells. The inactivation of AKT, as measured by AKT phosphorylation, occurred rapidly as the levels were reduced 20-50 fold within two hours of addition of 125 nM of this inhibitor to the media. Similar results were found using fibroblasts containing PIK3CA p.His147Arg or p.His1047Leu mutations. AKT1 mutation-positive cells treated with 425 nM of the inhibitor for 72 hours had a 25-40% reduction in cell viability compared to untreated cells, whereas mutation-negative cells only had a 0-20% decrease. These results indicate that basal AKT signaling in cells containing activating mutations in AKT1 and PIK3CA can be inhibited in PS and PROS cells using an imidopyridine inhibitor. Furthermore, our results suggest that AKT1 or PIK3CA mutations may not be causal in activating signals due to increased expression or disruption of cellular processes. This work has implications beyond PS and PROS as these mutations have been shown to be driver mutations in tumors. Because these cells harbor only a single mutation, as compared to the hundreds or thousands of mutations in tumors, they can be considered as single gene models of tumor regulation. Thus, the evaluation of such cells from patients with only a single mutation will simplify studies and allow investigators to assess the efficacy of a compound on a specific pathway without confounding effects resulting from additional driver and passenger mutations typically found in tumor cells.

Cycloheximide enhances skipping of mutated DMD exons synergistically with TG003. A. Nishida1, Y. Tsuchimasa1, T. Lee1, T. Takakada1, M. Matsuo1. 1) Medical Rehabilitation, Kobekagun University, Kobe, Japan; 2) Department of Pediatrics, Hyogo Medical College, Nishinomiya, Japan; 3) Department of Pediatrics, Kobe University Graduate School of Medicine, Kobe, Japan; 4) Department of Clinical Pharmacology, Kobe Pharmaceutica University, Kobe, Japan.

Duchenne muscular dystrophy (DMD) is a progressive and fatal muscle wasting disease. DMD is characterized by muscle dystrophin deficiency caused by mutations in the DMD gene. Endogenous mRNA encoding the full frame functional DMD mRNA is the most plausible way to express dystrophin in DMD therapy. In our previous report, we showed that TG003, a chemical Cik1 inhibitor, enhances skipping of mutated DMD exons and the subsequent expression of dystrophin (Nishida et al. Nat Commun 2011). It has been questioned whether TG003-mediated exon skipping is potentiated. Here, we searched for chemicals that enhance skipping of 14 mutated DMD exons. Hybrid-minigenes inserting one of mutated exons were transfected into HeLa cells and Lenti-CMV-Hyg containing mCherry were used. We found that cycloheximide (CHX), a protein synthesis inhibitor, enhanced skipping of DMD exon 31 encoding a nonsense mutation (c.4303G>T). Other protein synthesis inhibitors, anisomycin and puromycin did not enhance skipping, but metein did, indicating inhibitor-specific enhancement. As observed in TG003 treatment, CHX enhanced the skipping of mutated exons 27 and 39. CHX showed no effect on splicing of further 11 mutated or 14 wild-type DMD exons. CHX-mediated exon skipping was observed in the limited mutated exons that have shown a splicing response to TG003 for exon skipping by increasing the CHX concentrations and the CHX enhancement of the mutated exons through common mechanism. Unexpectedly, the level of exon 31-skipping with TG003 was increased higher by combined treatment with CHX than the sum of increased levels with single chemical treatment. CHX may enhance the skipping of mutated exons synergistically with TG003. This suggested that TG003 and CHX enhanced mutated exon 31 skipping in their own pathway. TG003 has been shown to increase full-length mRNA of the Cik1 gene which product phosphorylates SR protein. Accordingly, TG003 and CHX may upregulate the Cik1 mRNA level. Remarkably, combination treatment of TG003 and CHX shifted the splicing completely to produce full-length mRNA only. This indicated that CHX modulates splicing of the Cik1 gene. CHX was shown to enhance skipping of mutated DMD exon 31 in a different cell line, and this was designated as a leading compound to potentiate TG003-mediated exon skipping.
2184M
Towards treatment of Cantu syndrome. G. van Haastert, Department of Medical Genetics, University Medical Center Utrecht, Utrecht, Utrecht, Netherlands.

The CantuTreat project aims to develop a novel therapeutic approach in treatment of Cantú syndrome. Cantú syndrome is a rare genetic disorder, affecting a small number of patients (currently 100 cases are known worldwide) who suffer from multiple symptoms including hypertrophic lymphedema, distinctive facial features and cardiac abnormalities. Cantú syndrome is caused by dominant gain-of-function mutations in the ATP-dependent potassium channel subunits ABCB9 and KCNJ8. This recent discovery offers a promising opportunity for therapeutic intervention. The IKATP potassium channel is a known pharmaceutical target: pharmaceutical correction of these channels by sulfonylurea drugs results in a nearly complete cure in canine mucopolysaccharidosis I. We suggest that small quantities of IV ERT may improve in brain lysosomal storage in MPS I dogs receiving early IV ERT. We suggest that small quantities of IV ERT may reverse, but CD21 low cells were normalized, indicating improved B cell maturation. An increase in dendritic cells was also observed. Although the majority of patients with ASM deficiency (ASMD). This study was funded by Genzyme, a Sanofi company. Repeat-dose safety and exploratory efficacy of rhASM support its continued development for the treatment of the non-neurological manifestations of ASMD. This study was funded by Genzyme, a Sanofi company.

2186M
The impact of enzyme replacement therapy on immunity in Gaucher disease. R. Lingala1,2, C. Loano1, E. Kombriddi-Pastzor1, M. Plassmeyer2, M. Ryter2, M. Broman2, A. Lipinski1, A. Hebert1, O. Alpan2, Q. Goker-Alpan1. 1) Lysosomal Disorders Treatment Unit, O&O Alpan LLC, Fairfax, VA 22030; 2) Amerimmune, O&O Alpan, LLC, Fairfax, VA 22030.

Introduction: Gaucher disease (GD) is an autosomal recessive inherited disorder of metabolism caused by mutations in the GBA gene that results in the deficiency of the lysosomal enzyme glucocerebrosidase results in accumulation of glucosylcerolipids in macrophages. Patients with GD often present with an abnormal immune response that may be the result of cellular and/or humoral immune dysregulations. One of the most accepted therapies to treat GD is Enzyme replacement therapy (ERT). Here, we are examining the impact of ERT on immunity of GD patients. Cases and Methods: In an ongoing study (NCT01358188), the impact of ERT on the immunity was assessed in twenty-seven GD patients (19F/8M, mean age 40.2 yrs), five of whom were treatment-naive at the start of the study. Flow cytometry based immunophenotyping was performed from peripheral blood samples obtained before and after ERT administration. Lymphocyte subsets, memory, NK, B and dendritic cell populations were assayed along with chemokine receptors and activation markers. Results: In a treatment naive patient at baseline, prior to commencing ERT, CD4/CD8 ratio was less than 1 indicating immunosuppression, and transitional B-cells (characterized by CD21low expression) were markedly elevated suggesting a B-cell maturation defect. After establishment of stable-dose therapy, not only was the CD4/CD8 ratio reversed, but CD21 low cells were normalized, indicating improved B cell maturation. An increase in dendritic cells was also observed. Although the majority of subjects on long-term ERT exhibited normal ranges of lymphocyte subsets, in less than 20%, all of whom were diagnosed before the advent of ERT, the CD8 T-cell fraction was either elevated or expressed increased chemokine receptors like CXCR3, CCR6. Conclusions: Patients with GD can present with multiple immune abnormalities beyond the scope of activated macrophages resulting from accumulated glucosylcerolipids. Our data suggest ERT or its effects may improve immunological parameters indicating the role of ERT on long-term outcome may extend beyond the known effects of decreasing organ size and improvement of hematological parameters.

2188S
Early intravenous enzyme replacement therapy improves white matter myelination in canine mucopolysaccharidosis I. P. Dickson1, J.M. Prov- enzano1, S. Chen4, I. Nestrasil2, J. Yee1, S.H. Kan3, S.O. Le1, J. Jens3, E. Snell3, M.A. Guzman3, C. Vie4, E. Shapiro4, N.M. Ellinwood2. 1) LA BioMed/Harbor-UCLA, Torrance, CA; 2) Duke University, Durham, NC; 3) Emory University, Atlanta, GA; 4) University of Minnesota, Minneapolis, MN; 5) Iowa State University, Ames, IA; 6) St. Louis University, St. Louis, MO; 7) University of Pennsylvania, Philadelphia, PA.

Diffusion tensor imaging studies in people with mucopolysaccharidosis I (MPS I) show reduced volume and reduced fractional anisotropy in the corpus callosum, which correlates with inattention on neurobehavioral tests. We have found decreased myelination in MPS I dogs as early as 6 weeks of age, and the fractional anisotropy of the corpus callosum varies with the amount of myelin basic protein in myelin extracts. We therefore studied the impact of enzyme replacement therapy (ERT) with recombinant human alpha-L-iduronidase on myelination. MPS I dogs received intrathecal (IT) ERT 0.05 mg/kg every 3 months and/or intravenous (IV) ERT 0.58-2.0 mg/kg weekly. Untreated MPS I and normal carriers were used as controls. The expression of several myelin-related genes was reduced in the corpus callosum of MPS I dogs compared to normal carriers. Expression of myelin-related genes improved in MPS I dogs treated with IT and/or IV ERT beginning at ≤30 days of age: the expression of myelin-related genes was normal in dogs treated with 0.58 mg/kg weekly IV ERT + IT ERT, and was greatly improved in dogs receiving 1.57 or 2.0 mg/kg weekly IV ERT without IT ERT. There was no improvement in myelin gene expression in dogs treated with IT ERT beginning at age 4m, although there was some improvement in fractional anisotropy, myelin basic protein levels, and myelin lipids. We have previously published improvement in brain lysosomal storage in MPS I dogs receiving early IV ERT. We suggest that small quantities of IV ERT cross the blood-brain barrier and that is sufficient to prevent some of the neuropathology of MPS I.

2187S
An open-label, multicenter, ascending dose study of the tolerability and safety of recombinant human acid sphingomyelinase (rhASM) in patients with ASM deficiency (ASMD). M.P. Wasserstein1, S.A. Jones2, H. Soran3, G. Diaz4, B. Thurberg5, K. Culm-Merdek1, E. Komlodi-Pasztor6, T. Singh1, A.C. Puga1. 1) Icahn School of Medicine at Mount Sinai, New York, NY, USA; 2) Manchester Centre for Genomic Medicine, St. Mary's Hospital, CMFT, University of Manchester, Manchester, UK; 3) Genzyme, a Sanofi company, Cambridge, MA, USA.

Background: Enzyme replacement therapy (ERT) with recombinant human acid sphingomyelinase (rhASM) is in clinical development for the treatment of the non-neurological manifestations of acid sphingomyelinase deficiency (ASMD). In a phase-1 single-dose study in adult patients, the maximum tolerated starting dose of rhASM was determined to be 0.6 mg/kg. Objectives: The primary objective of this phase 1b study was to determine the safety and tolerability of within-patient dose escalation of rhASM in five adult patients with Niemann-Pick Disease type B (NPD B), the non-neuropathic form of ASMD. Each patient was to receive a starting dose of intravenous rhASM at 0.1 mg/kg and advance every two weeks according to a predetermined schedule to 3.0 mg/kg, or their maximum tolerated dose. Secondary objectives: to study the pharmacokinetics, pharmacodynamics, and exploratory efficacy of rhASM administered intravenously every two weeks for 26 weeks. Methods: Study assessments included continuous adverse event (AE) reporting and periodic evaluations of safety, pharmacoki-netics, pharmacodynamics, and exploratory efficacy. Safety biomarkers (e.g., C-reactive protein, bilirubin, IL-6, IL-8) were evaluated; plasma ceramide was used as a biomarker for the breakdown of sphingomyelin. Sphingomyelin content in liver was used as a pharmacodynamic endpoint. Results: The dose escalation regimen was well tolerated, with all patients reaching the maximum dose of 3.0 mg/kg. No serious or severe adverse events or deaths were reported. Related AEs consisted predominantly of infusion-associated reactions, the majority of which were mild and resolved without sequelae. A positive response to treatment with rhASM was observed in liver sphingomyelin content and several exploratory efficacy parameters, including spleen and liver volumes, pulmonary function testing, lung imaging, lipid profile, and quality of life assessments. Results will be presented. Conclusions: Within-patient dose escalation of rhASM was well tolerated. Repeat-dose safety and efficacy of rhASM will be the continued development for the treatment of the non-neurological manifestations of ASMD. This study was funded by Genzyme, a Sanofi company.
2188M

Pompe disease is an inherited lysosomal storage disease that results from deficiency in acid α-glucosidase (GAA) activity. It is characterized by progressive accumulation of lysosomal glycogen in cardiac and skeletal muscles. The only treatment currently available is enzyme replacement therapy (ERT) using a recombinant human α-glucosidase. The efficacy and safety of Fabagal® (agalsidase beta) that generates changes of plasma lysoGL-3, urine lysoGL-3 (remained within normal Survey, and Subject symptom diary at 24 weeks of treatment compared to score of Short Form McGill Pain Questionnaire & SF-36 Health Status (Cr), plasma lysoGL-3 and urine lysoGL-3, echocardiographic findings, the secondary endpoint variables were changes of urine GL-3, serum creatinine administration (1.0 mg/kg every 2 weeks) for 24 weeks. The primary endpoint study in Fabry disease patients for 24 weeks. A total of ten patients (7 male, 3 female) were included in this study. The safety and efficacy of Fabagal® were evaluated in patients with Fabry disease. We evaluated the efficacy and safety of enzyme replacement therapy (ERT) with Fabagal® in patients with Fabry disease. Methods: The study was a multicenter, open-label, and phase II study in Fabry disease patients for 24 weeks. A total of ten patients (7 male and 3 symptomatic female patients including 1 naive patient) were enrolled (n = 10). The objective was to evaluate safety and efficacy during Fabagal® administration (1.0 mg/kg every 2 weeks) for 24 weeks. The primary endpoint of efficacy was to maintain the level of plasma GL-3 within normal range. The secondary endpoint variables were changes of urine GL-3, serum creatinine (Cr), plasma lysoGL-3 and urine lysoGL-3, echocardiographic findings, the score of Short Form McGill Pain Questionnaire & SF-36 Health Status Survey, and Subject symptom diary at 24 weeks of treatment compared to the baseline. Results: Nine of 10 enrolled patients completed the study (n = 9). Treatment with Fabagal® proved efficacious since plasma GL-3 level maintained within the normal range. Also urine GL-3 value was normalized after ERT with Fabagal® compared to the baseline. There were no significant changes of plasma lysoGL-3, urine lysoGL-3 (remained within normal range), echocardiographic findings, and serum Cr level during the study period. As a result of the adverse event (AE) analysis, there was no serious adverse event or any drop out or death due to AEs. All AEs were resolved without any complications and none were related to Fabagal®. There was no significant elevation of aminotransferases. This study proves the beneficial effect and safety of Fabagal® (agalsidase beta) that generates clinically significant decrease of GL-3 with no significant adverse drug reaction in Fabry disease patients. However, further study is needed in larger number of naive patients for long-term period.

2189S

The choroid plexuses are highly vascularized structures that project into the cerebrospinal fluid (CSF) of the four cerebral ventricles. The specialized polarized epithelia of choroid plexuses produce CSF by transporting water and ions into the ventricles, and turn over at an extremely slow rate. We hypothesized that remodeling these epithelia to secrete a missing lysosomal enzyme by one-time administration of a recombinant AAV (rAAV) gene therapy vector into the cerebrospinal fluid could be an attractive and efficacious approach for long-term treatment of lysosomal storage diseases (LSD). Lysosomes function as the primary digestive units within cells and specific enzymes within lysosomes normally break down nutrients. Patients with LSDs cannot metabolize certain nutrients, resulting in diminished lifespans and reduced quality of life. There are no ideal therapeutic options presently available for the neurological manifestations of LSDs. Brain-directed recombinant enzyme replacement therapy has shown promise for several LSDs but requires repeated instillations due to short enzyme half-lives. In contrast, rAAV-mediated gene transfer to the choroid plexus would enable continuous synthesis and secretion of missing lysosomal enzymes into the CSF and steady penetration to the cerebral cortex and cerebellum. To evaluate this hypothesis, we obtained a mouse model of alpha-mannosidosis in which a targeted disruption of the lysosomal α-mannosidase (LAMAN) gene is present. Using mice from our developing colony, we devised a genotyping assay that distinguishes wild type, heterozygous and homozygous animals. We cloned the human (hs) LAMAN cDNA into an adeno-associated virus (AAV5) shuttle plasmid and documented robust LAMAN expression in transfected HK293T cells. We next generated high titer rAAV5-huLAMAN expressing huLAMAN and administered 1×10e10 or 1×10e11 viral particles to homozygous affected mice by lateral ventricle injection on day that the higher MHFeificient, dose-dependent viral transduction in the initial group of treated animals. Biochemical analyses and neurobehavioral testing in treated animals is in progress. We also established a NIH clinical protocol (14-CH-0106) to evaluate biomarkers in CSF of human subjects following delivery of the gene therapy vector into the choroid plexus. If the choroid plexus viral gene therapy approach were successful, the largest current barriers to health for patients with certain LSDs would be circumvented.

2191S
Glucose transporter 1 deficiency syndrome (Glut-1 DS) is an autosomal dominant disorder characterized by the SLC2A1 gene. The Glut-1 is expressed in the blood-brain barrier and responsible for hexose transport into the brain. The classic phenotypes with Glut-1 DS are infantile drug resistant seizures, intellectual disability (ID) and cerebellar ataxia. Seizures of Glut-1 DS can be treated with a ketogenic diet, but other symptoms persist. We detected a new mutation on SLC2A1 in a classical type patient on whom ketogenic diet was not effective. To aim more effective treatments of Glut-1 DS, we have been investigating the potentiality of gene therapy using AAV vector. As the first step, we established a functional system that monitor the ability of introduced SLC2A1 to transport glucose into cells. (Materials and methods) The patient was a 16-year old boy. He had frequent convulsion from early infancy and had severe ID. Ketogenic diet started from 6 years was not effective. After obtaining informed consent, we isolate DNA from a patient and sequenced the SLC2A1. Total RNA was also extracted and reverse transcribed to cDNA. For transfection study, we used SLC2A1 expression vector by Origene technologies (RC222696). We created plasmids with SLC2A1 mutation by ligated oligonucleotide corresponding to R333W, A405D and c.906_907insG. To observe glucose transport, we transfected wild type and mutated SLC2A1 vectors into HEK293 cells, and 2-deoxyglucose (2DG) uptake was analyzed 48 hours after gene transfection. (Results) We identified a novel frame shift mutation, c.906_907insG, and 2DG uptake was increased by wild type SLC2A1 vector or control vector transfection. 2DG uptake by mutated SLC2A1 vectors transfection to HEK293 cells decreased compared with wild type SLC2A1 vector or control vector transfection. (Discussion) Glut-1 DS is caused by haplosufficiency of the SLC2A1 and expected to benefit from gene therapy without highly gene expression level. We could achieve the increase of 2DG uptake by wild type SLC2A1 gene transfection to HEK293 cells, which was not observed by the mutated SLC2A1. This system allows us to assess the 2DG uptake by vectors SLC2A1.

2193S
Online DuchenneConnect self-report data indicates Exon 44 skippliable DMD patients have prolonged wheelchair-free survival. J.W. Ulm, T. Yamagata, 1 Department of Pediatrics, Jichi Medical University, Shimosuke-shi, Tochigi, Japan; 2 Department of Neurology, Jichi Medical University, Shimosuke-shi, Tochigi, Japan.

Duchenne muscular dystrophy (DMD) patients have significant heterogeneity in the age at which they lose the capacity for ambulation. This reflects that the disease is not a single entity but a spectrum of disorders with variable clinical severity and prognosis. The DuchenneConnect registry is the largest prospective, longitudinal registry of DMD patients and their families in the world. In this analysis, we investigated the age at which DuchenneConnect patients report the first use of a wheelchair (WCase) using self-report data on DuchenneConnect registry. At present, there is limited data available on which “skippable” exons would most likely give rise to an ameliorated phenotype, knowledge that would be of significant value in testing possible exon-skip protocols and in formulating eventual treatment regimens. In the work described here, we used online self-reported data on DuchenneConnect registry along with self-report data on DuchenneConnect registry and the DuchenneConnect, BioMarin Duchenne Muscular Dystrophy (DMD) Patient Registry and the International Duchenne Registry (IDR) along with self-report data on DuchenneConnect registry and the DuchenneConnect, BioMarin Duchenne Muscular Dystrophy (DMD) Patient Registry and the International Duchenne Registry (IDR) dataset. The dataset includes data from patients, focusing especially on their ages at loss of ambulation. We found that Exon 44 skippliable patients demonstrated significant mitigations in severity compared to Exon 45, 50, 51, and 53 skippliable patients and those with non-arrhenoblock mutations. These preliminary results suggest that online self-report data can be of tangible value in guiding further research and treatment protocols.
**2196M**
Glycogen storage disease type Ia mice receiving gene therapy are protected against age-induced obesity and insulin resistance. G. Y. Kim1, J. H. Cho1, Y. M. Lee1, C. J. Pan1, H. S. Jun1, B. C. Mansfield1, 2, J. Y. Chou1 and D. B. Bindman1, 2. 1) Section on Cellular Differentiation, Program on Developmental Endocrinology and Genetics, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD; 2) The David Novikoff Laboratory for Pediatric Research and Therapy, Division of Neurology, National Institutes of Health, Bethesda, MD, USA.

Glycogen storage disease type Ia (GSD-Ia), characterized by impaired blood glucose homeostasis, fasting hypoglycemia, and hepatocellular adenoma (HCA), is caused by a deficiency in glucose-6-phosphatase-α (G6Pase-α). In a long-term, dose-ranging study, we showed that systemic administration of rAAV-G6PC, a recombinant AAV2/8 vector expressing G6Pase-α directed by the G6PC promoter/enhancer, delivered the G6Pase-α transgene to the liver of GSD-Ia mice. The old (70-90-week-old) AAV-G6PC-treated GSD-Ia mice expressing 3-63% of normal hepatic G6Pase-α activity (AAV-mice) maintained glucose homeostasis, tolerated prolonged fasting, and showed no evidence of HCA. Interestingly, the AAV-mice exhibit better metabolic controls than their control littermates with fasting blood insulin levels closer to the normal values and a leaner phenotype. Mice overexpressing the transcription factor carbohydrate response element binding protein (ChREBP) in the liver exhibit improved glucose tolerance and are protected against insulin resistance, and ChREBP signaling is activated by G6P. The elevated levels of hepatic G6P in the AAV-mice suggest that ChREBP signaling would be activated. The AAV-mice also produced reduced levels of hepatic glucose, averaging 61-68% of the control littermates, suggesting that these AAV-mice lived under a chronic calorie restriction (CR). A hallmark of aging is mitochondrial dysfunction, and several studies have now shown that CR pathways can modulate mitochondrial function. These include: 1) the NADH shuttle systems that reoxidize cytosolic NADH produced by glycolysis, interconnecting glycolysis, lipogenesis, and mitochondrial oxidative phosphorylation; 2) intracellular levels of NAD+ that are decreased in obesity and the aged; and 3) the AMP-activated protein kinase (AMPK)/peroxisome proliferator-activated receptor-γ coactivator 1α (PGC-1α) pathway, the major regulator of mitochondrial biogenesis in response to energy depletion. We now show that during aging, AAV-mice are protected against age-induced mitochondrial dysfunction and obesity and that this correlates with: activation of ChREBP signaling; increases in glycerol-3-phosphate shuttle system expression; increases in NAD+ concentrations; and activation of the AMPK/PGC-1α pathway in the liver of the AAV-mice. Thus, long-term, systemic gene therapy is beneficial to longevity and a lean phenotype.

**2198M**
Cell Reprogramming Technologies for the Treatment of Genetic Disorders of Myelin. A. Lager1, Z. Nevin1, J. Heaney2, P. Tesar2, 3, 1) Genetics and Genome Sciences, Case Western Reserve University, Cleveland, OH; 2) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 3) New York Stem Cell Foundation, New York, NY.

Cellular therapies are a promising treatment modality for a variety of human myelin diseases. In previous work we describe reprogramming methods for generating functional oligodendrocyte progenitor cells (OPCs) from autologous sources (Najm & Lager et al Nat Biotech 2013). However, in cases of dysmyelinating diseases with genetic etiology (e.g. Pelizaeus-Merzbacher and Spastic Paraplegia Type 2), autologous cells harbor genetic abnormalities that preclude oligodendrocyte function. Here we present technology for generating functional OPCs from fibroblasts isolated from shiverer mice. Shiverer mice harbor a 20kb deletion in a locus containing myelin basic protein (MBP) and consequently lack compact myelin and exhibit severe motor deficits. We outline methods for correction of the MBP-shiverer allele in induced pluripotent stem cells (iPSCs). These methods utilize the CRISPR/Cas9 genome editing system for induction of homology-directed repair based on an exogenously supplied bacterial artificial chromosome that contains the MBP-wt allele. Gene corrected iPSCs will be differentiated into OPCs, and we hypothesize that introduction of gene-corrected shiverer OPCs back into the central nervous system of shiverer mice will allow myelination of hypomyelinating axons and behavioral improvements. This work will provide proof-of-concept for gene correction of large deletions by homologous recombination, with important clinical implications in treatment of microdeletion syndromes.

**2199S**
Rescue of lethal hypophosphatasia model mice by adeno-associated virus mediated muscle specific expression of bone targeted alkaline phosphatase. A. Nakamura1, O. Iijima1, K. Miyake1, A. Watanabe1, 2, Y. Hira1, H. Kinoshita1, T. Noguchi1, S. Abe1, T. Okada1, T. Shimada1, 2, 1) Department of Biochemistry and Molecular Biology, Division of Gene Therapy Research Center for Advanced Medical Technology, Nippon Medical School, Tokyo, Japan; 2) Division of Clinical Genetics, Nippon Medical School Hospital, Tokyo, Japan; 3) Department of Anatomy, Tokyo Dental College, Tokyo, Japan.

**[BACKGROUND]** Hypophosphatasia (HPP) is a systemic skeletal disease, caused by the deficiency of tissue-nonspecific alkaline phosphatase (TNSALP). The TNSALP knockout (AKP-/-) mice phenotypically mimic the severe infantile form of HPP and usually die by 3 weeks of age suffering epileptic seizures and apnea. Enzyme replacement therapy (ERT) is a potential approach, but repeated injection of large amounts of recombinant TNSALP is required for clinical benefit. As an alternative approach, we have recently shown that AKP-/- mice are protected by systemic delivery of the MCK promoter driven by the AAV vector containing the muscle-specific muscle creatine kinase (MCK) promoter and encoding the feasibility of muscle directed gene therapy for HPP. **[METHODS]** Self-complementary AAV vectors containing the MCK promoter were delivered after birth into AKP-/- mouse newborns using a retrograde vector to target the developing muscle. The plasma ALP activity was rapidly elevated and maintained after injection of AAV into the neonates has not yet been established. In this study, we exploit safe and clinically applicable protocol, we designed a new self-complementary type 8 AAV vector containing the muscle-specific muscle creatine kinase (MCK) promoter and examining the feasibility of muscle directed gene therapy for HPP. **[RESULTS AND DISCUSSION]** Stingent muscle specificity of the MCK promoter was demonstrated after intramuscular (IM) injection of scAAV8-MCK-EGFP (5x10¹¹ vector genome/mouse) into WT mice (n=3). Neonatal AKP-/- mice (n=10) were treated with a single IM injection of scAAV8-MCK-TNSALP-D10 (2.5x10¹² vector genome/mouse). The plasma ALP activity was rapidly elevated and maintained at a therapeutic level (>1.0 unit/ml) for at least 3 months. The treated AKP-/- mice grew well and survived over 3 months with healthy appearance (9/10), while untreated AKP-/- mice died within 3 weeks (P<0.001). Improved mineralization of the knee joints was demonstrated on X-ray images and micro CT analysis. Ectopic calcification was not detected in treated mice. These results suggest that muscle directed gene therapy using scAAV8-TNSALP-D10 would be safe and effective to cure the severe infantile form of HPP.
Functional genomics analysis of the interplay between the retrograde transport machinery and HIV-1 replication. D. Dykshorn1,2, S. Liu1. 1) John P. Hussman Institute for Human Genomics, Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL; 2) Department of Microbiology and Immunology University of Miami Miller School of Medicine.

Human immunodeficiency virus (HIV) remains a global health concern affecting millions of individuals worldwide. By attacking the host’s immune system, HIV leaves the infected individual susceptible to infection by a wide range of pathogens. Like all viruses, HIV-1 is dependent on host cell factors for successful infection, replication and release of progeny virus. Previously, a large-scale functional genomic screen was performed which identified over 230 novel factors, termed HIV-1 dependency factors (HDFs) whose silencing inhibited viral replication. Detailed analysis of these HDFs showed enrichment for factors involved in the trans-Golgi network (TGN), a pathway which had not been previously implicated in HIV replication. This included components of large multiprotein complexes that facilitate the recognition, docking and fusion of different vesicles as they are trafficked through the cell. The targeted silencing of these HDFs, including components of the conserved oligomeric golgi (COG) complex and the Golgi-associated retrograde protein/vesicular protein sorting fitying three (GRASP/VFP) complexOur results show that the targeted silencing of components of COG complex each impaired HIV-1 replication in HIV-1-susceptible cell lines and primary monocyte-derived macrophages and inhibited replication against multiple strains of the virus. More detailed analysis of the different forms of the HIV-1 genome showed that the defect in HIV-1 replication occurred prior to late RT product formation, suggesting that these HDFs are required for very early steps in the HIV-1 life cycle. These results highlight a novel pathway required for HIV-1 replication and infection which may hold promise for the development of novel approaches to prevent HIV-1 transmission.

Using Antisense Technology.

Protein misfolding diseases, also known as proteopathies, are a class of diseases that are caused by changes in protein conformation which often lead to gain of toxic function. Antisense technology utilizes RNase H mechanism to degrade RNA and subsequently remove unwanted protein products in the cell, potentially providing an optimal platform for the treatment of these diseases. We have evaluated potent antisense oligonucleotides (ASOs) for three genetically-based protein misfolding diseases: 1) Transthyretin amyloidosis, 2) Alpha-1 antitrypsin liver disease, and 3) Autosomal dominant retinitis pigmentosa due to a mutation (P23H) in the rhodopsin gene. ISIS-TTR-1A targets all mutant forms as well as wild type TTR mRNA for the treatment of ATTR. In a Phase 1 normal volunteer study, ISIS-TTR-1A was well tolerated and achieved mean reductions in plasma TTR of ~75% at the 300mg dose level. This compound is currently being evaluated in a multicenter, double-blind, placebo-controlled Phase 3 trial in ATTR familial amyloid polyneuropathy patients. In a mouse model of AAT liver disease (PiZZ mice), treatment with an antisense oligonucleotide targeting the alpha-1 antitrypsin mRNA (AAT-ASO) prevented liver disease progression after short-term treatment; reversed liver disease after long-term treatment, and prevented liver disease in young animals. Furthermore, AOS treatment markedly decreased liver fibrosis in these mice. In addition to the reduction of peripheral targets like TTR and AAT which are mainly produced by liver, ASOs can be delivered directly into the eye via intravitreal injection and distribute to all the cell layers in the eye. P23H rhodopsin transgenic rats express both mutant mouse and normal rat rhodopsin, representing an excellent model of retinitis pigmentosa. We developed ASOs targeting the diseased allele (mouse mutant Rho), but not the normal endogenous allele in this model. Using allele-specific ASO, we demonstrated a slower progression of photoreceptor degeneration and improved electroretinography (ERG) measurements 30 days after a single intravitreal injection.

Eyes injected with P23H ASO had a 128 ± 22% improved amplitude response. This study demonstrates that antisense technology provides promising treatments for genetically-based protein misfolding diseases.
2204M
Feasibility of Exon-Skipping Therapy for Juvenile Neuronal Cereoid Lipofuscinosis. M. Velinov1,2, N. Dolzhanskaya1. 1) Human Genetics, NYS Institute for Basic Research in Developmental Disabilities, Staten island, NY; 2) Albert Einstein College of Medicine, Bronx, NY.

Rationale: Juvenile Neuronal Cereoid Lipofuscinosis (JNCL) is a progressive, lethal, neurodegenerative disorder and is the most common type in the group of Neuronal Cereoid Lipofuscinoses. 95% of the JNCL patients carry a common 1 kb deletion in one or both alleles of gene CLN3. The deletion includes exons 7 and 8, and leads to reading frame shift and formation of truncated protein. No effective therapy for JNCL is available. Splicing out (skipping) of exon 9 may be induced in cells with the common CLN3 deletion using antisense oligonucleotides (AO). Such exon skipping would restore the correct reading frame and is expected to produce a modified protein with correct aminoacid sequence downstream of exon 9. Induction of exon skipping (ES) may lead to the increase of the CLN3 transcript in mutant cells by decreasing the nonsense-mediated decay. ES may also help restore some of the CLN3 protein function since the correct aminoacid sequence downstream of exon 9 would be restored. Such approach is currently used in a clinical trial for patients with Duchenne Muscular Dystrophy. Exon skipping: Exon skipping enhancer sequences within exon 9 were identified using the program ESEFinder. Four modified oligonucleotides with 2'O-methyl modified bases on a phosphothiolate backbone, complementary to splicing enhancer sequences were synthesized. Fibroblast cell lines from patient, homozygous for the 1 kb deletion in CLN3 and phenotype of JNCL were plated 24 h prior transfection and were transfected with all four AO using Lipofectamine 2000 reagent. The cells were harvested 18 h after transfection. Total RNA was isolated and used for RT-PCR analysis. The AO transfection resulted in complete exon 9 skipping in the detectable mutant CLN3 transcript. These results were confirmed with sequencing of the RT-PCR product. Increase of the transcript level: Total RNA extracted after AO treatment was used for real time quantitative RT-PCR. The total mutant transcript level increased with 30% after AO treatment compared to the non-treated mutant transcript. Intranasal administration: Two mouse-specific AO were labeled with immunofluorescent marker, and were applied intranasally in a mouse model of JNCL. The intranasally applied AO penetrated the nasal mucosa, reached the brain, and were observed in the mouse neuronal cells. Conclusions: This prove of concept study showed that AO based ES therapy for JNCL may be feasible using intranasal application.

2205S
Improving the communication skills of Down Syndrome patients through speech therapy. A. Umrigar1, A. Musso2, K. Foley1, M. Banaje2, F. Tsien1. 1) Department of Genetics, LSU Health Sciences Center, New Orleans, LA; 2) Department of Communication Disorders, LSU Health Sciences Center, New Orleans, LA.

Two major obstacles that Down Syndrome patients encounter in the process of communicating effectively are the craniofacial abnormalities and cognitive deficits associated with the syndrome. Physiological features, such as a small oral cavity, large tongue, high arched palate, and facial hypotonia impede the clear articulation of speech. In addition to the physical abnormalities, cognitive deficits further interfere with speech and comprehension. Early intervention in the forms of speech therapy and hearing loss treatments improves the communication skills of Down Syndrome patients. This study focuses on detailed communication evaluations of Down Syndrome patients enrolled in the Louisiana State University Health Sciences Center (LSUHSC) Speech Language Therapy program from 6 to more than 15 years. The patients' treatment plans and progress were examined weekly in order to determine the effect of speech therapy on the patient's comprehension and articulation of both verbal and written forms of communication. Improvement was monitored using the Goldman-Fristoe test of articulation, repetitive one word vocabulary test, expressive one word vocabulary test, and oral-motor/range of motion evaluation. In addition, these patients received audiological assessments. Concrete progress goals and objectives were constructed specific to each patient. All assessment methods demonstrated a significant improvement in effective communication skills, including progression from one-word to full sentences, vastly improved articulation, and communication in a group setting compared to those who did not receive any form of communication or speech therapy. It is therefore important for geneticists to recognize the improved quality of life benefits of communication intervention programs and refer Down Syndrome patients to speech pathologists.

2206M
The outcome of N-carbamylglutamate (Carbaglu®) therapy in a Korean patient with N-acetylglutamate synthase deficiency. K. Wool1, J. Choi2, B. Lee1, H. Yoo1, 2) Genomics, Asan Medical Center, Seoul, South Korea; 1) Department of Pediatrics, Asan Medical Center, Univ. Ulsan College of Medicine, Seoul, South Korea.

N-acetylglutamate synthase (NAGS) deficiency is a rare autosomal recessive disorder of the urea cycle. Affected individuals present with poor oral intake, vomiting, seizure, lethargy, and coma within a first few days of life. N-carbamylglutamate (Carbaglu®), a synthetic derivative of glutamic acid as well as a structural analogue of human N-acetylglutamate, is a treatment choice for NAGS deficiency. We report a 8-day-old boy with NAGS deficiency who was treated with N-carbamylglutamate (Carbaglu®). The patient was born from non-consanguineous Korean parents with a birth weight of 3.4 kg after 40 weeks of gestation. At age 8 days, he showed poor oral intake and lethargy with hyperammonemia (800 µM), which was managed by peritoneal dialysis before referral to our institution. Plasma amino acid profiles revealed elevated glutamate (1117 µmol/L), normal arginine (51 µmol/L), ornitine (49 µmol/L), and low citrulline (7 µmol/L) levels. Urine orotic acid was normal (1.77 µmol/mmolCr; range 0.2-6 µmol/mmolCr). He has been placed on sodium benzoate, carnitine and citrulline therapy with protein restriction diet with presumptive diagnosis of carbamoyl phosphate synthase deficiency. He was referred to our center at age 2 years. Genetic testing for CPS1 was negative. At age 7 years, the NAGS genes analysis using cDNA extracted from liver tissues identified compound heterozygote of p.V310A and p.H488Qfs*2, which were inherited from both parents. He showed mild mental retardation (IQ 51), epilepsy, and attention-deficit/ hyperactivity disorder. Despite protein restriction and administration of nitrogen scavenger medicines, he presented recurrent hyperammonemia, stunted growth velocity, and poor weight gain. N-carbamylglutamate (Carbaglu®) replaced since age of 8.1 years, and he discontinued nitrogen scavenger drugs. After 4 months, he discontinued special formula and has been maintained on liberal diet without protein restriction. He has maintained plasma ammonia within normal range and showed normal prepubertal growth velocity. In conclusion, N-carbamylglutamate (Carbaglu®) has been effective for NAGS deficiency to maintain stable metabolic state and growth without any complications.

2207S
Modeling Autosomal Dominant Optic Atrophy in Induced Pluripotent Stem Cells (iPSCs) and Identifying the Potential Therapeutic Targets. J. Chen1, H. Riazifar2, M. Guan1, T. Huang1. 1) Division of Human Genetics, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, USA; 2) Department of Pediatrics, Division of Human Genetics, University of California, Irvine, CA, USA.

Purpose: Many retinal degenerative diseases are caused by loss of retinal ganglion cells (RGCs). Autosomal dominant optic atrophy (DOA) is the most common hereditary optic atrophy characterized by central vision loss and degeneration of RGCs. DOA is largely caused by mutations of a nuclear gene (OPA1), which encodes an inner mitochondrial membrane protein. Currently, there is no effective treatment for these types of diseases. However, stem cell therapy holds great potential by replacing patient lost RGCs. Compared with embryonic stem cells, induced pluripotent stem cells (iPSCs) are derived from adult somatic cells, associated with fewer ethical concerns, and are less prone to immune rejection. In addition, these patient-derived iPSCs may serve as a cellular model, allowing us to closely study pathogenesis and therefore to identify potential therapeutic agents. Methods: In this study, we were able to generate specific iPSCs from the patient carrying an OPA1 mutation and diagnosed with optic atrophy. We differentiated iPSCs into RGCs by using our established protocol. Results: We found that the mutation of OPA1 in patients with optic atrophy led to significantly increased apoptosis and OPA1-iPSCs were not able to differentiate into RGCs. However, adding neuron induction medium or beta-estrogen into differentiation medium efficiently promoted OPA1-iPSC differentiation into RGCs. Conclusions: Our results suggest that apoptosis mediated by OPA1 mutations plays a very important role in DOA pathogenesis and both noggin and beta-estradiol may be the potential therapeutic agents for OPA1 mutations related optic atrophy.
2208M
1) Centre for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC, Canada; 2) Neurogenetics Unit and Rare Disease Centre, IRCCS Neuromed Pozzilli (IS), Italy.
BACKGROUND: The dominant CAG repeat expansion causing Huntington disease (HD [MIM 143100]) occurs on specific haplotypes and is the HTT gene. In humans, at least 2300 polymorphisms occur across HTT and offer potential targets for allele-specific silencing of the mutant transcript encoding the CAG expansion. The organization of all HTT polymorphisms into gene-spanning regions and the routine polymerase chain reaction testable by selective silencing of disease-associated haplotypes, is unclear and needed to guide allele-specific HTT silencing strategies.
AIMS: Our study sought to identify all genetic variants occurring on the most common HD-associated haplotypes, A1 and A2, and determine the frequency of these haplotypes among distinct HD patient populations of European ancestry. We additionally sought to demonstrate selective HTT silencing using novel HD-associated polymorphisms identified by this work.
RESULTS: Using whole-genome sequencing from the 1000 Genomes Project, and direct genotyping of diverse HD patient cohorts from Canada, Sweden, France, and Italy, we assemble complete haplotypes for all known common variants (>5%) in HTT. The most common HD-associated haplotype, A1, is uniquely defined by three transcribed polymorphisms; the second most common HD haplotype, A2, is uniquely defined by five intragenic polymorphisms. A1 is most heterozygous in Northern European patient populations, where HD is most common, whereas A2 is most heterozygous in Southern European patients. Using antisense oligonucleotides (ASOs) targeting ASO-matched poly-A tails, using polymerase chain reaction testing that selectively amplify mRNA and protein in patient-derived cells, maintaining wild-type HTT expression at untreated levels.
CONCLUSIONS: Across populations of European ancestry, silencing the A1 haplotype may offer allele-specific therapy in approximately half of HD patients. Our results suggest that targeting the A2 haplotype may allow selective HTT silencing in an additional quarter of patients across all four populations, and over half of HD patients in Italy. In combination, antisense reagents silencing A1 and A2 may offer selective silencing of mutant HTT in >70% of HD patients.

2209S
IPS-cell derived basal keratinocytes and melanocytes to study severe monogenic and polygenic diseases in patient-specific 3D tissue systems. K.M. Eckl1, D.M. de Lima Cunha2, R. Plank1, R. Casper1, H. Traupe1, M. Rauch1, M.K. Gupta1, T. Saric2, H.C. Hennes1,2,3,4,5.
1) Ctr. for Dermatogenetics, Div. of Human Genetics, Innsbruck Medical University, Innsbruck, Austria; 2) Ctr. for Dermatogenetics, Cologne Ctr. for Genomics, Univ. of Cologne; 3) Dept. of Dermatology, University Hospital Münster, Germany; 4) Inst. for Neurophysiology, Medical Center, Univ. of Cologne, Germany.
Induced pluripotent stem (IPS) cells from a somatic source like dermal fibroblasts from a patient with a known severe, often fatal, but always rare genodermatosis are a unique and excellent tool to study the pathomechanism during neuro-ectodermal differentiation, melanocyte development and epidermal stratification. Induced pluripotent stem cells (IPS cells) can be easily performed and allows stable gene correction and fast in-vitro gene-therapy approaches. Up to date, for most genodermatoses, like nevus comedonicus, ICDP (autosomal dominant ichthyosis, retinal dysplasia and polydactyly), and Sjogren-Larsson syndrome, cell lines have been generated and can be used to perform various experiments. The development of patient-specific 3D tissue models is a future goal in the potential of this exciting new therapeutic approach, warranting further studies.

2210M
A Fine Balance of Dietary Lipids Improves a Murine Model of VCP-associated Disease. K.J. Llewellyn1, A. Nalbandian1, N. Walker1, J. Tang1, A. Gomez1, V. Krimonis. Pediatric, 2501C Hewitt Hall, Irvine, CA.
Despite intense investigations, the discovery of effective novel advancements/therapies and the disease mechanisms underlying Valosin containing protein (VCP)-associated myopathies and neurodegenerative disorders remain elusive. VCP diseases, caused by mutations in the VCP gene, are clinically and genetically heterogeneous groups of disorders with manifestations varying from hereditary inclusion body myopathy, Paget’s disease of bone, frontotemporal dementia (FTD), amyotrophic lateral sclerosis (ALS), and other neurodegenerative changes. Affected individuals exhibit scapular winging and progressive muscle weakness and die from cardiac and respiratory failure. Histologically, patients display rimmed vacuoles and TAR DNA Binding Protein-43 (TDP-43)-positive ubiquitinated inclusion bodies in muscles. Currently, there are no effective treatments for patients with VCP-related myopathies. VCP mouse models carrying the common R155H mutation include several of the clinical features typical of human diseases. Here, we examined the effects of varying dietary lipid percentages on VCP155H/+ and wild type (WT) mice from birth till 5 months of age. Disease progression was monitored and analyzed using survival curves, pathological and immunohistochemical methods. Mice on the 9% lipid-enriched diet demonstrated an improvement in muscle strength measurements, histology, and autophagy signaling pathways when compared to mice on normal diet. In this report, we demonstrate VCP mice fed increasing lipid diets of 12%, 15%, 30%, and 48% showed no improvement in muscle pathology or the autophagy cascade, suggesting that significantly increased lipid diets only depict a detrimental effect in mice. Thus, a balanced lipid supplementation suggests a promising translational therapeutic strategy for patients with VCP-associated neurodegenerative diseases.

2211S
Development of autologous myogenic stem cell therapy for carriers of a heterolorphic mtDNA mutation: a proof of principle study in m.3243A>G mutation carriers. F. van Tienen1,2, E. Timmer1,2, M. Quattrocelli3, M. Sampaolesi4, I. De Coo4, H. Smeets2,3,5,6,7,8,9,10, 1) Clinical Genetics, MaastrichtUMC, Maastricht, Netherlands; 2) Research School GROW, MaastrichtUMC, Maastricht, Netherlands; 3) Interdepartmental Stem Cell Research School, KU Leuven, Leuven, Belgium; 4) Dept Neurology, Erasmus UMC, Rotterdam, Netherlands; 5) Research School CARIM, MaastrichtUMC, Maastricht, Netherlands.
Mitochondrial diseases due to heterolorphic mitochondrial DNA (mtDNA) mutations are severe multisystem disorders (frequency 1/5,000), caused by defects in oxidative phosphorylation (OXPHOS). Severe progressive myopathy and exercise intolerance occur in >50% of the mtDNA disorders with a variety of other clinical manifestations and seriously affect the quality of life and well-being. Effective therapy is currently unavailable. We hypothesize that inducing muscle regeneration by transplantation of mtDNA-mutation free autologous myogenic stem cells (pericytes/mesoangioblasts and vascular progenitors+mesenchymal stem cells) or myogenic progenitors for the OXPHOS defects may reverse myopathy and exercise intolerance, resulting in increased mobility and quality of life for heterolorphic mtDNA mutation carriers. MABs are currently the only myogenic precursors that fulfill all requirements for being used in this therapeutic strategy and a phase II/II clinical trial using allogeneic MABs is currently ongoing in DMD patients. The key characteristic of our project is transplantation of the patient’s own MABs that have been freed of the mtDNA mutation. This approach allows targeting all proximal muscles safely without adverse reactions to the transplant and the use of immunosuppressive agents. To this end, we have succeeded in generating mutation-free MABs from m.3243A>G carriers (with adult onset myopathy and a mutation load in skeletal muscle ranging from 15% to 93%) in sufficient amounts for initial clinical treatment. In the in vitro growth rate and myogenic differentiation capacity (ranging between 20%-57%) of these mutation-free MABs were comparable with MABs derived from healthy controls and their in vivo myogenic capacity was subsequently verified in immunodeficient mice. The next step is to perform a proof of principle clinical study to demonstrate the local effect of transplantation of mutation-free autologous MABs, injected in the leg of m.3243A>G carriers. Muscle regeneration, mutation load and OXPHOS capacity will be determined. From this study, we expect to get a clear insight in the potential of this exciting new therapeutic approach, warranting further steps into a systemic clinical trial.
Online Self-Report Data for Duchenne Muscular Dystrophy: observations of current natural history and explorations of therapeutic responses. A. Eskin, R. Wang, J. Ulm, C. Silverstein, I. Jankovic, A. Li, V. Miller, R. Cantor, N. Li, R. Elashoff, A. Martin, H. Peay, S. Nelson. 1) Human Genetics, UCLA, Los Angeles, CA; 2) Department of Internal Medicine, College of Osteopathic Medicine of the Pacific, Western University of Health Sciences, Pomona California and the Western Diabetes Institute, Western University of Health Sciences, Pomona, CA; 3) Patient-Crossroads, San Mateo, California; 4) Department of Biomathematics, David Geffen School of Medicine, University of California, Los Angeles; 5) Parent Project Muscular Dystrophy, Hackensack, NJ; 6) Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles.

To assess the utility of online patient self-report outcomes in a rare disease, we attempted to observe the well established therapeutic benefit of corticosteroids for Duchenne muscular dystrophy. First, we assessed how comparable an online self-report registry for Duchenne, called DuchenneConnect, compared to prior natural history data in regard to age at diagnosis, mutation spectrum, and age at loss of ambulation. Second, we assessed the correlation of reported corticosteroids usage with reported age at fulltime wheelchair use in Duchenne muscular dystrophy (DMD) only using data from 1,057 males from DuchenneConnect. Because registrants reported differences in steroid and other medication usage, as well as age and ambulation status, we could explore these data for correlations with age at loss of ambulation. Using multivariate analysis, current steroid usage was the most significant and largest independent predictor of improved wheelchair-free survival. Thus, these online data were sufficient for the retrospective observation that current steroid usage by patients with DMD is associated with a delay in loss of ambulation (a key functional endpoint difficult to assess in traditional study designs). Interestingly, reported use of Vitamin D, CoEnzyme Q10, insurance status, and age at diagnosis after 4 years were also significant, but smaller, independent predictors of longer wheelchair-free survival. This study demonstrates the utility of DuchenneConnect data to observe therapeutic differences, and highlights needs for improvement in quality and quantity of patient-report data, which may allow exploration of drug/therapeutic practice combinations impractical to study in clinical trial settings. Further, with the low barrier to participation, we anticipate substantial growth in the dataset in the coming years.


Spinal muscular atrophy, the leading genetic cause of infant mortality, is a neuromuscular disease characterized by progressive loss of alpha-motor neurons in the anterior horn of the spinal cord. It is caused by disruption of the SMN1 gene and insufficient compensation of its function by the neighboring, nearly identical paralogous gene SMN2, which harbors a splicing silencer element in intron 7 that suppresses inclusion of exon 7 critical for production of full-length SMN protein. Antisense oligonucleotide (ASO)-mediated blockade of the splicing silencer was previously shown to promote inclusion of SMN2 exon 7 in multiple mouse models of SMA and to mediate phenotypic rescue. To date, however, the molecular signature of this rescue has not been defined. Here we characterize the transcriptional changes that occur in an induced model of type III SMA and the substantial prevention or reversal of those changes in CNS tissue upon intracerebroventricular administration of an ASO promoting inclusion of exon 7. We further show that the timing of ASO administration is correlated with prevention or reversal of specific subsets of transcriptional changes. Collectively, these changes represent potential biomarkers of SMN depletion and therapeutic response to SMN repletion.
A hemizygous GYG2 mutation in Japanese siblings showing Leigh syndrome without marked elevation of lactate and pyruvate. E. Imaeda, T. Nakashima, M. Shimizu, E. Takahashi, H. Sugie, N. Makashima, Y. Tsurusaki, H. Saiatsu, K. Ogata, N. Matsumoto, N. Miyake, 1) Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 2) Division of Neurology, Clinical Research Institute, Kanagawa Children’s Medical Center, Yokohama, Japan; 3) Department of Molecular Biology, Yokohama City University School of Medicine, Yokohama, Japan; 4) Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 5) Faculty of Medicine, Immunological and Cellular Research Institute, Kanagawa Children’s Medical Center, Yokohama, Japan; 6) Department of Pediatrics, Jichi Medical University, Tochigi, Japan.

Leigh syndrome (LS; MIM#256000) is a progressive neurodegenerative disorder characterized by unique, bilateral neuropathological findings in brainstem, basal ganglia, cerebellum and spinal cord. The estimated incidence of LS is 1:40,000 live births. LS is genetically heterogeneous and the majority of the causative genes affecting mitochondrial malfunction. However, the genetic causes in the majority of patients still remain unsolved. Here, we present male sibling affected with LS showing developmental delay, convulsion, athetoid movements, nystagmus, hypotonia, spasticity, increased deep tendon reflex, ketonemia, and bilateral lesion of basal ganglia (caudate nuclei, globus pallidus, and putamen), but no marked elevation of lactate and pyruvate. To identify their genetic cause, we performed genetic analysis using whole exome sequencing. Based on the hypothesis of autosomal recessive and X-linked recessive models, we identified only one hemizygous missense mutation (c.665G>C, p.Trp222Ser) in glycogenin-2 (GYG2, isozyme a: NM_001079855) gene in both affected siblings. The heterozygous male patient was not available to undergo the self-glucosylation observed in wild-type GYG2. We proposed a pathomechanism of the GYG2 impairment in this family based on the canonical pathway of glycogen metabolism. LS is known as one of metabolic disorders affecting the majority of the causative genes affecting mitochondrial malfunction. Therefore, a comprehensive genetic evaluation of LS patients is crucial in order to identify the underlying gene defects. Here, we achieved successful genetic diagnosis of LS patient showing the brainstem involvement by only limited biochemical findings.

INTRODUCTION: There is growing evidence of urea cycle disorder (UCD) recognized in adult. Encephalopathy after gastric bypass has been reported requiring intubation and ventilatory support. Her post surgical course is notable for severe malnutrition over the preceding 4 months. Admission labs revealed a markedly elevated NHI 193μg/dL peaking at 276 [normal range 40-50] and multiple micronutrient deficiencies. Head MRI is normal. Therapy was initiated with an ammonia scavenger, ammonul and TPN nutrition support. Metabolic panel showed urine orotic acid 2.8 mmol/mmolCr (0.4-1.2), alanine 180μmol/L (240-350), arginine 30μmol/L (40-160), citrulline 5μmol/L (10-60), glutamine 366μmol/L (410-700), isoleucine 0μmol/L (30-130), methionine 11μmol/L (17-53), serine 36μmol/L (60-200), tyrosine 9μmol/L (30-120), valine 44μmol/L (140-350). Ornithine, carnitine and acylcarnitine levels were normal. With the exception of low glutamine and alanine, the metabolite findings were consistent with the well-recognized malnutrition–amenia–encephalopathy–cyanosis syndrome (OTC) deficiency. The patient recovered after aggressive treatment. Further molecular genetic analysis did not reveal a detectable mutation in the OTC gene. However, a heterozygous G209S (GGC>AGC) in exon 6 of the acyl-CoA dehydrogenase (ACBD3) gene was identified. DISCUSSION: Mutations in the ACAD gene are associated with short-chain acyl-CoA dehydrogenase (SCAD) deficiency, an autosomal recessive disorder. Only those patients homozygous or compound heterozygous for the G209S mutation have been identified. The patient was found to have a homozygous G209S short chain acyl-CoA dehydrogenase gene carrier.

Role of ACBD3 protein in the mitochondrial energy metabolism. T. Tesarova, A. Vondrackova, J. Spacilova, M. Hulova, H. Hansikova, V. Patel, 1) Department of Surgery, Metropolitan Support Service, Brigham and Women’s Hospital, 75 Francis Street, Boston, MA 02115; 2) Department of Medical Genetics, Mayo Clinic, 200 First Street SW, Rochester, MN 55905; 3) Department of Surgery, Thomas Jefferson University Hospital, Philadelphia, PA 19107; 4) Department of Pediatric Endocrinology, Kanagawa Children’s Hospital, Yokohama City University, Yokohama, Japan; 5) Department of Nuclear Medicine, Kanagawa Children’s Hospital, Yokohama, Japan; 6) Institute of Inherited Metabolic Disorders, National Institute of Health, Kanagawa, Japan; 7) Department of Pediatrics and Adolescent Medicine, National Hospital Organization, Kanto Children’s Medical Center, Saitama; 8) Department of Pediatrics, Jichi Medical University, Tochigi, Japan.

N. Hauser, Medical Genetics Dept, Childrens Hospital Central California, Madera, CA.

INTRODUCTION: Urea cycle disorders can affect single or multiple organs and can present with varied symptoms. Here, a patient was eventually diagnosed with NARP who previously presented with ataxia and retinitis pigmentosa (NARP). CONCLUSION Mitochondrial enzyme deficiencies are becoming more commonly identified in these disorders can affect single or multiple organs and can present with varied symptoms. Here, a patient was eventually diagnosed with NARP who previously presented with ataxia and retinitis pigmentosa (NARP). CONCLUSION Mitochondrial enzyme deficiencies are becoming more commonly identified in these disorders can present with varied symptoms. Here, a patient was eventually diagnosed with NARP who previously presented with ataxia and retinitis pigmentosa (NARP). CONCLUSION Mitochondrial enzyme deficiencies are becoming more commonly identified in these disorders can present with varied symptoms. Here, a patient was eventually diagnosed with NARP who previously presented with ataxia and retinitis pigmentosa (NARP). CONCLUSION Mitochondrial enzyme deficiencies are becoming more commonly identified in these disorders can present with varied symptoms. Here, a patient was eventually diagnosed with NARP who previously presented with ataxia and retinitis pigmentosa (NARP).
2218M
Growing evidence suggests that organic acidemias are accompanied by secondary mitochondrial dysfunction, decreased energy production, and oxidative stress. We previously reported that reduced and oxidized glutathione (GSH and GSSG), as measured in whole blood by LC-MS/MS, are useful biomarkers of mitochondrial dysfunction, and that patients with primary mitochondrial disorders show significant redox imbalance which further worsens during periods of metabolic crisis. Here we extend these findings to the organic acidemias, and provide direct evidence that these disorders are associated with secondary mitochondrial dysfunction. Whole blood GSH and GSSG levels were evaluated in 105 samples from 57 organic acidemia patients (29 methylmalonic acidemia, 5 propionic acidemia, 3 isovaleric acidemia). All patients were evaluated at least once during outpatient clinic visits (i.e., clinically stable), and 7 were also tested while hospitalized for metabolic decompensation. Redox potential was calculated using the Nermet equation, and compared to the mean redox potential from 59 healthy controls (-260 mV ± 6.4). Statistical comparisons were conducted using Student’s t test, with significance set at p<0.05. As a group, organic acidemia patients showed significant redox imbalance (redox potential -250 mV ± 6.8, p=0.001) with an increased level of oxidation of ~10 mV compared to controls. This was seen in all three organic acidemias studied, with no significant differences noted between the groups. Patients experiencing a metabolic crisis had significantly more oxidized redox potential (-245 mV ± 3.5) than those who were clinically stable (-250 mV ± 6.8; p=0.036). These results support the idea that measurements of blood GSH and GSSG may give insights into the contribution of oxidative stress and mitochondrial dysfunction to the pathophysiology of organic acidemias.

2219T
A newborn with persistent mild elevations of succinylacetone in bloodspot, plasma and urine, with an identified homozygous variant in the GSTZ1 gene. W. Al-Hentati1, L. You1, K. Phommamn1, S. Yak1, S. El Tarazi1, M-T. Berthier2, Y. Giguère2, T. Gagnon2, P.J. Waters2. 1) Department of Medical Genetics, McGill University Health Centre, Montreal, QC, Canada; 2) Biochemical Genetics Laboratory, Medical Genetics Service, Centre hospitalier universitaire de Sherbrooke (CHUS), Sherbrooke, QC, Canada; 3) Department of Pediatrics, McGill University Health Centre, Montreal, QC, Canada; 4) Quebec Newborn Bloodspot Screening Program, Service de Biochimie, CHU de Quebec, Quebec City, QC, Canada.
Case report: The proband is a Native Canadian female, born to non-consanguineous parents. At age 18 days, she was referred as screening positive for tyrosinemia, following slightly elevated succinylacetone (SUAC) in bloodspot (age 40 hours); hydrazine method, 1.16 micromol/1, cutoff 0.67; enzymatic method 3.53, cutoff 2.50. Meanwhile, a urine organic acid profile at age 6 days (test requested independently, for unrelated reasons) revealed mild SUAC elevation. SUAC results obtained using stable isotope dilution GC-MS on samples collected between age 4 days and 4 weeks were: in plasma, 128, 688, 617 nanomol/L (reference <24); in urine, 776, 766, 604 micromol/L (reference <34). While unremarked variants, considering the clinical course together with biochemical and molecular findings, a diagnosis of tyrosinemia type 1 is highly unlikely. Despite novel human phenotype in the literature, we considered the possibility that reduced activity of maleylacetoacetate isomerase, due to hypomorphic variants in the GSTZ1/MAAI gene (J Clin Pharmacol 2012;52:S37-84), is the most likely potential etiology for the persistent elevation of SUAC in this infant. Interestingly, molecular analysis in our patient, did reveal a homozygous variant of uncertain significance in exon 7 of the GSTZ1 gene; c.494C>T (p.Ala165Val), predicted to be pathogenic by SIFT and Mutation Taster. Parental analysis is currently under way.

2220M
Lysosomal acid lipase (LAL) is a lysosomal enzyme that is involved in intracellular lipid metabolism. Deficiency of the LAL enzyme results in accumulation of cholesterol esters and triglycerides in most tissues of the body, and causes Wolman disease and cholestereryl ester storage disease (CESD). Wolman disease is fatal within the first year of life due to severe hepatomegaly, persistent diarrhea and failure to thrive, while CESD is a milder disease that is characterized by hyperlipidemia and hepatomegaly that can be observed in childhood or develop in adulthood. LAL deficiency is an inherited genetic condition, caused by mutation in the lipase A (LIPA) gene that encodes the LAL enzyme. Spanning about 40 kb with 10 exons and coding 400 amino acids, the LIPA gene is located on the long arm of chromosome 10 between positions 23.2 and 23.3. So far more than 40 mutations in the LIPA gene have been reported, including missense, nonsense, splice site, small deletion/insertions and large deletions. We have examined twenty specimens worldwide with LAL deficiency for mutation analysis of the LIPA gene, performing bi-directional sequence analysis of coding exons and corresponding intron/exon boundaries on genomic DNA. Among mutations we identified, the common mutation c.834 G>A is about 20% of alleles; and four novel mutations have never been reported previously. These new mutations include: (1) missense mutation c.67G→A (p.G23R), (2) missense mutation c.286T→C (p.L96P), (3) small deletion TGCCAAACGCAAGCTT in c.297-311, and (4) missense mutation c.421G→C (p.Arg141Pro). Interestingly, the mutation c.421G→C has only been detected in specimens from persons of Asian ethnicity.
2222M

Introduction: Pompe disease is caused by pathogenic sequence variations in the Acid α-Glucosidase gene (GAA) causing lysosome glycogen accumulation. SNPs c.1726G>A (p.G576S) and c.2065G>A (p.E689K) were predicted to cause pseudodeficiency of GAA. Substitution p.E689K reduces GAA activity by 50%; at most. Substitution p.G576S reduces GAA activity to clinical spectrum ranges. Objective: to assess the effects of p.[G576S; E689K] and their association with mutation p.W746C found in 2 Colombian patients. Methods: Three-dimensional model of human GAA based on the N-terminal Subunit of Human Maltase-Glucoamylase complex with Acarbose (MGA PDB 2QMJ) template. Energy minimization was performed. Physicochemical properties were validated with ProtParam. Quality/Reliability of stereo chemical parameters of the homology model was confirmed with Qmean6, GROMOS and Ramachandran plot. Alignments were calculated with ClustalW and Muscle. Conserved amino acids were obtained as a consensus of MultiDisp, Consurf and ConSeq. Solvent accessible surface area value (ASA) was determined with PYMOL. The mutant model was superimposed on the wild type structure to determine the influence of amino acid replacement in the 3D structure on the basis of the total root-mean-square distance (RMSD). Results: Amino acid identity between GAA and the template was 44%. Both proteins belong to the glycosyl hydrolases family 31 in which the structure is conserved and shares similar binding site, overcoming the low sequence identity. GAA and GMA are hydrophilic with similar composition of negative and positive residues. Stereo-chemical quality of the predicted model showed 85.7%; of residues in favored regions, 8.4%; in allowed regions and 5.9%; in outlier regions. The conservation analysis recognized a highly conserved amino acid position G576, in a lesser extent E689. The preliminary ASA measurement indicated that both polymorphisms and W746C mutations are buried. Wild type predicted model value, significantly differs from the mutant model (27125.766 Å² and 30854.953 Å², respectively). The Superposition of the two models reveals that G576 is located inside the barrel structure and change the side chains of the Aα4-helix. E689 is located at N-terminus of Aα8-helix where the side chain is fully exposed. The superposition RMSD value was 0.84 Å.

2223T
Familial hypermangenesemia among Egyptian families: further delineation of the phenotype and management. M.S. Zak1, MS. Abdel-hamid2, H. Hossni2, S. Idris2, M.Y. Issa1, A. El-Safty4, A. Oraby5. 1) Clinical Genetics Department, National Research Centre, Cairo, Egypt, Dokki, Cairo, Egypt; 2) Human Molecular Genetics Department, National Research Centre, Cairo, Egypt; 3) National Institute of Neuromotor System, Cairo, Egypt; 4) Toxicology Department, Cairo University, Cairo, Egypt; 5) Pediatrics Department, Cairo University, Cairo, Egypt.

Familial hypermanganesemia or inherited manganesism is a rare metabolic error caused by recessive mutations in SLC30A10 gene. This gene is functioning to maintain homeostatic control of intestinal absorption and biliary excretion of manganese through responsible for manganese transporter. Manganese accumulates in the liver, muscle, bloodstream, and brain, specifically the basal ganglia resulting in extrapyramidal manifestations, polycythemia and liver cirrhosis. We describe 7 patients from 3 unrelated consanguineous families with cardinal manifestations of manganesism. They were five males and 2 females and their aged ranged from 14/12y to 18 years old. Six patients (Family 1(4 sibs), 2 (2 sibs)) presented with slowly regression of walking around the age of 2/12, manifested as recurrent failing due to limb dystonia. Interestingly, one patient (Family 3) had sudden acute dystonia of lower limbs with inability to walk at the age of 1 year and 5 months with suspicious of inflammatory etiology but the weak response to intravenous immunoglobulin together with the striking neuroimaging pointed to this metabolic error. All affecteds had polycythemia with hemoglobin concentration ranged from 17 -21 g/dl and blood manganese level was above 2000 nmol/L. Elevated liver enzymes with no sign of cirrhosis were in the older 2 sibs in family 1 (17 and 15 ys old), however normal enzymes were in the older affected in family 2 in spite of reaching 18 years old and was still ambulant. Brain MRI showed the typical picture of manganese deposition in basal ganglia classically on T1-weighted images, with affection of midbrain and the tegmentum of the pons as well as the middle cerebellar peduncle with no corresponding abnormality on T2-weighted scans. Molecular analysis for these 3 families revealed novel mutations in SLC30A10 gene in 2 of them. We designated oral treatment with 2,3 dimercaptosuccinic acid (DMSA), iron supplementation and levodopa that showed satisfactory preliminary results in chelating manganese with remarkable improvement of the symptoms in all and alleviation of clinical manifestations in the younger probands. This report represents further delineation of familial hypermanganesemia. It emphasizes the importance of considering this treatable metabolic error in any patient presenting with acute or insidious dystonia and polycythemia.
2224M
Abnormal phospholipid metabolism due to variants in PCYT1A causes spondylometaphyseal dysplasia with cone-rod dystrophy, J. Jurgens1, N. Sobrero1, J. Hooven-Foong2, P. Modaffi3, G. Yamamoto2, W. Barateia3, D. Bertola2, F. Collins1, J. Christodoulou1, M.B. Bober3, R. Pauli1, D. Valle4.
1) Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2) Department of Pediatrics, University of Wisconsin-Madison, Madison, WI; 3) Unidade de Genética, Instituto da Ciência, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil; 4) Western Sydney Genomics Program, Children’s Hospital at Westmead, Sydney and Disciplines of Pediatrics & Child Health and Genetic Medicine, Sydney Medical School, University of Sydney, New South Wales, Australia; 5) Division of Medical Genetics, Department of Pediatrics, A.I. duPont Hospital for Children, Wilmington, DE.

Spondylometaphyseal dysplasia with cone-rod dystrophy (SMD-CRD) is a rare, autosomal recessive condition characterized by short stature, bowing and metaphyseal flaring of the long bones, rhizomelic shortening, platyspondyly, scoliosis, and early-onset progressive degeneration of cone and rod photoreceptors. By whole exome sequencing and Sanger sequencing, Hoofer-Fong et al. (2017) identified novel homozygous or compound heterozygous variants in PCYT1A in 6 unrelated probands. Independently, Yamamoto et al. (2014) described two novel homozygous PCYT1A variants, p.Glu129Lys and p.Ser323Argfs*38, in two unrelated SMD-CRD probands. Here, we used Sanger sequencing to identify homozygous or compound heterozygous variants in PCYT1A in four additional unrelated individuals with SMD-CRD. We detected a single homozygous variant (p.Ala99Val) in 2 unrelated probands and found a separate homozygous variant (p.Tyr240His) in a third unrelated proband. In the fourth proband, we detected the compound heterozygous variants p.Ser323Argfs*38/p.S114T. PCYT1A encodes CTP-phosphocholine cytidydyltransferase α (CTCα), an enzyme which catalyzes the rate-limiting step in the de novo phosphatidylcholine biosynthesis by the Kennedy pathway. Phosphatidylcholine is the predominate membrane phospholipid in mammalian cells. Western blot analysis of CTCα expression in cultured skin fibroblasts of 3 unrelated individuals with SMD-CRD shows that the protein levels vary among SMD-CRD patients. When compared to controls, p.Ala99Val homozygotes show a 30% reduction in protein levels; p.Glu129Lys homozygotes show a 70% reduction; and p.Ala99Val homozygotes showed a 30% reduction when compared to controls. When compared to controls, p.Ala99Val homozygotes show a 30% reduction in protein levels; p.Glu129Lys homozygotes show a 70% reduction; and p.Ala99Val homozygotes showed a 30% reduction when compared to controls. When compared to controls, p.Ala99Val homozygotes show a 30% reduction in protein levels; p.Glu129Lys homozygotes show a 70% reduction; and p.Ala99Val homozygotes showed a 30% reduction when compared to controls.

2225T
Mutation spectrum of six genes in Chinese phenylketonuria patients obtained through next-generation sequencing, Y. Gu1, K. Lu2, G. Yang3, Z. Chen4, Y. Li5, L. Lin5, J. Hao5, Z. Yang3, J. Peng3, S. Cui6, J. Huang7.
1) Lianyungang Maternal and Child Health Hospital, Lianyungang, China; 2) Chinese National Human Genome Center at Shanghai Shanghai, CHINA.

Background: The identification of gene variants plays an important role in the diagnosis of genetic diseases. Methodology/Principal Findings: To develop a rapid method for the diagnosis of phenylketonuria (PKU) and tetrahydrobiopterin (BH4) deficiency, we designed a multiplex, PCR-based primer panel to amplify all the exons and flanking regions (50 bp average) of six PKU-associated genes (PAH, PTS, GCH1, QDPR, PCBD1 and GFRP). The Ion Torrent Personal Genome Machine (PGM) System was used to detect mutations in all the exons of these six genes. We tested 92 DNA samples from blood specimens from 35 patients and their parents (32 families) and 26 normal adult individuals. Using strict bioinformatic criteria, this sequencing data provided, on average, 99.14% coverage of the 108 exons at more than 70-fold depth. We found 23 novel and 32 deleterious variants in the PAH gene and six novel mutations in the PAH and PTS genes. A detailed analysis of the mutation spectrum of these patients is described in this study. Conclusions/Significance: These results were confirmed by Sanger sequencing. In conclusion, benchtop next-generation sequencing technology can be used to detect mutations in monogenic diseases and can detect both point mutations and indels with high sensitivity, fidelity and throughput at a lower cost than conventional methods in clinical applications.

2226M
Functional Characterization of RYR1 Sequence Variants Associated with Malignant Hyperthermia Susceptibility Identified through Exome Sequencing, S.G. Gonzales1,1, C.E. Kasper3, S. Perry4, S.M. Muldoon5, L.G. Biesecker1.
1) Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA; 2) Uniformed Services University of the Health Sciences, Graduate School of Medicine, Bethesda, Maryland, USA; 3) Uniformed Services University of the Health Sciences, School of Medicine, Department of Anesthesiology, Bethesda, Maryland, USA.

Malignant hyperthermia susceptibility (MHS) is a life-threatening, inherited disorder of muscle calcium metabolism, triggered by anesthesia and depolarizing muscle relaxants. Exome sequence (ES) data from ClinSeq® were screened to identify putative pathogenic MHS variants in the RYR1 and CACNA1S genes. To study the effects of RYR1 variants on channel function, we selected benign and pathogenic variants for comparison—using an in-vitro assay measuring calcium release from participant-derived Epstein-Barr virus immortalized B-lymphocytes—to determine if incidental MHS-associated variants identified from ES showed abnormal calcium release from lymphoblasts. Annotation of 870 exomes for RYR1 and CACNA1S variants used an algorithm that filtered results based on genotype quality, allele frequency, mutation type, and information from mutation databases. Sixty-three RYR1 and 41 CACNA1S variants passed the quality and frequency metrics and were scored on a four-point pathogenicity scale (1-2). Three RYR1 class 5 (p.Arg6141C, p.S1728F, p.D3986E) and one class 3 (variant of uncertain significance) p.R1667C, were selected for functional testing. Pharmacological sensitivity of the RyR receptor type-1 in EBV-immortalized B-lymphocytes derived from four ClinSeq® participants were tested and compared to MH-negative controls. RYR1 p.R614C (with published functional data) was selected as a positive control. B-lymphoblastoid cell lines were used to study the transient peak calcium release induced by the ryanodine receptor agonist 4-chloro-m-cresol (4-CMC). Calcium release (area under curve) was averaged from 12 replicates. Differences in calcium release between cell lines were analyzed by t-test. EBV-lymphocytes from class 5 variants showed increased sensitivity of calcium release to 4-CMC compared to normal controls. Elevated calcium release was observed in RYR1 p.R614C and p.D3986E. In addition, the RYR1 p.Glu129Lys and p.Ser323Argfs*38 variants were found to be pathogenic using an in-vitro assay measuring calcium release from EBV-lymphocytes to determine if incidental MHS-associated variants identified from ES showed abnormal calcium release from lymphoblasts. Annotation of 870 exomes for RYR1 and CACNA1S variants used an algorithm that filtered results based on genotype quality, allele frequency, mutation type, and information from mutation databases. Sixty-three RYR1 and 41 CACNA1S variants passed the quality and frequency metrics and were scored on a four-point pathogenicity scale (1-2). Three RYR1 class 5 (p.Arg6141C, p.S1728F, p.D3986E) and one class 3 (variant of uncertain significance) p.R1667C, were selected for functional testing.

2227T
Screening for Lysosomal Storage Disorders using an integrated enzymatic and molecular approach, O. Bodamer1, G. Ghaffari2, P. Nonega3, B. Knecht1, B. Illagan4, A. Essa5, B. Johnson1.
1) Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 2) Uniformed Services University of the Health Sciences, School of Medicine, Bethesda, Maryland, USA; 3) University of Miami, Miami, FL; 4) Genzyme-Sanoﬁ Inc, Mexico City, Mexico.

Background: Lysosomal storage disorders (LSDs) comprise a heterogeneous group of genetic conditions that affect lysosomal metabolism. The clinical phenotype typically follows a clinical continuum of progressive multi-organ involvement. Effective therapies including enzyme replacement therapy (ERT) are available for a subset of LSDs including Fabry, Pompe, Gaucher diseases, MPS I and MPS II. Additional therapies are under development. Timely diagnosis and initiation of ERT is vital to reduce and/or prevent disease morbidity. Methods: We have validated an integrated diagnostic approach utilizing enzyme testing and second tier molecular testing on dried blood spots (DBS) for 6 LSDs including Fabry, Pompe, Gaucher, Niemann-Pick type A/B, Krabbe diseases and MPS I. Samples were collected from at-risk patients based on clinical phenotype suspicious for LSD from large tertiary referral centers in Mexico. Enzyme analysis in DBS was done by tandem mass spectrometry followed by direct sequencing in samples with low enzyme activities. Results: 3206 samples were received for enzyme testing and 820 samples for direct sequencing of the GLA gene in females. 2223 samples (7%) were reflexed to molecular testing based on low enzyme activities. 3/22 (14%) of samples were diagnostic for Pompe disease, 4/19 (21%) were diagnostic for Gaucher disease, 6/11 (55%) were positive for MPS I, 9/71 (5%) were diagnostic for Fabry disease and 12/820 (1.4%) of females tested for Fabry disease carried a heterozygous GLA mutation. Conclusion: The use of an integrated diagnostic approach for the above LSDs in DBS allows timely and inexpensive diagnosis of at-risk patients. The refinement of the clinic phenotype that prompts testing for LSDs will lead to further reduction of unnecessary testing.
2228M
Lysosomal acid lipase activity in dried blood spots from patients initially suspected of Gaucher disease. V.G. Pereira1, C.F. Chaves1, J.U.S. Yamamoto1, A.M. Martins2, V. D’Almeida1. 1) Department of Psychobiology, Universidade Federal de São Paulo, São Paulo, Brazil; 2) Department of Pediatrics, Universidade Federal de São Paulo, São Paulo, Brazil.
Lysosomal acid lipase (LAL) deficiency is an autosomal recessive condition, which leads to the accumulation of cholesteryl esters and triglycerides in most tissues, causing two different lysosomal storage disorders: Wolman disease (WD) and Cholesterol ester storage disease (CESD). WD is the most severe form of the disease, leading to death before one year of age, while CESD presents a more attenuated clinical course, with highly variable phenotypes. Hepatic alterations are a common symptom between lysosomal storage disorders, mainly in CESD, Gaucher and Niemann-Pick diseases. Therefore, differential diagnosis is essential for proper treatment. In this sense, patients initially suspected of Gaucher disease, whose biochemical diagnosis had not been confirmed, may represent a high-risk population for LAL deficiency (LALD). The aim of this study was to determine LAL activity in dried blood spots (DBS) referred to Laboratório de Erros Inatos do Metabolismo/UNIFESP - Brazil with an initial suspicion of Gaucher disease, which had beta-glucosidase activity within the normal range (Gaucher-negative), as a screening for LALD patients. LAL activity was determined by a fluorimetric assay. Until now, 110 Gaucher-negative samples were analyzed and mean LAL activity observed was 45.28 pmol/punch/h (standard error = 4.17), median = 33.99 pmol/punch/h, ranging from 0 to 299 pmol/punch/h. From these samples, 38 (~35%) of them presented LAL activity below the normal range, which had been previously determined in our laboratory (~24 pmol/punch/h). These samples were then analyzed separately; in this group, all LAL activity was 12.82 pmol/punch/h (standard error = 1.23); median = 12.97 pmol/punch/h, ranging from 0 to 23.91 pmol/punch/h. All the physicians responsible for patients in the group with low LAL activity were contacted and a new DBS was requested to confirm the diagnosis, to exclude the possible interference of reduced LAL activity due to long storage periods (all DBS samples used in this study were kept at 4°C through a maximum period of one year). Although we still have not received these new DBS to confirm LALD, the high rate of LAL activity below the normal range found in our study suggests that Gaucher-negative patients indeed represent a high-risk population to screen for LALD patients. In addition, Gaucher-negative patients should also have LAL activity determined as differential diagnosis. Funding: FAPESP, CAPES, IGEIM and AFIP.

2229T
Novel strategy for the diagnosis of late onset Pompe disease using next generation sequencing technologies. S. Levesque1, E. Gravel2, S.L. Austin3, S. Gravel3, J. Keutzer3, C. Auray-Blais3, P.S. Kishnani2, 1) Department of Pediatrics, Division of Medical Genetics, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada; 2) Departments of Pediatrics, Division of Medical Genetics, Duke University Medical Center, Durham, NC, USA; 3) Genzyme Corporation, a Sanofi company, Cambridge, MA, USA.
Background: Pompe disease is a rare lysosomal storage disorder caused by acid alpha-glucosidase enzyme deficiency, and is characterized by progressive lysosomal glycogen accumulation especially in muscles. Studies have shown that early diagnosis followed with prompt initiation of enzyme replacement therapy may maximize clinical benefits. However, early diagnosis remains challenging, especially in the actual adult cohort owing to the lack of universal screening. Despite published guidelines to educate clinicians, testing for Pompe is often overlooked owing to its low frequency. Next generation sequencing (NGS) might solve this issue given that it can analyse at once a large number of genes causing diseases with overlapping phenotypes. Method: We designed a NGS assay in order to analyze the coding sequences and splice site junctions of GAA mutated in Pompe disease, along with 78 genes causing neuromuscular disorders with overlapping phenotypes. Analytical validity was then determined using DNA from 20 Pompe patients with known variations and the reference cell line NA12878. Enrichment of target sequences was performed using a custom in-solution oligonucleotide probes library (SeqCapEz®, Nimblegen), followed by NGS on the illumina MiSeq® platform. Variations were detected with our bioinformatics pipeline based on BWA and GATK. Result: We obtained a median coverage of ~90X (sequences per base) per sample, with only 0.3% of all exons showing less than 20X. All GAA exons were successfully covered with >20X. The NGS assay showed a sensitivity of 100% (95% C.I. 98-100%) for the GAA gene, based on the analysis of 197 known single nucleotide variations and indels in Pompe patients (42 different variations). In addition, recurrent deletion of exon 18 was correctly detected by read depth analysis in 5/5 Pompe patients. Sensitivity of the assay across all selected genes was similar with 100% (95% C.I. 99%-100%), as determined by comparison of observed variant with >20X coverage against known variations in NA12878. Specificity was determined to 98% (95% C.I. 95-99%), for 208 variations found in exons and splice site junctions of selected genes in the NA12878 cell line. Conclusion: We were successful in designing a highly sensitive and specific NGS assay for Pompe disease genetic disorders. NGS of patients presenting with a neuromuscular disorder of unknown etioloqy from neuromuscular clinics is in progress and results will be presented.


Molecular confirmation in our case was indispensable for the correct diagnosis with molecular results. The same mutation. The third son did not have the mutation and consequently lead to the frameshift p.Ile556Metfs*11). Mother is hemizygous for a mutation in ATP7A (c.1668-1680del, predicted to cause a frameshift in an intron). Two children have died. The first child (son) died at 2mo21d with clinical diagnosis of MD. Laboratory exams showed low serum ceruloplasmin <2mg/dL (reference range from 20-60 mg/dL) and copper <8 μg/mL (reference range from 70-160 μg/dL) confirming clinical diagnosis of MD. Treatment with copper histidine was indicated immediately but it was instituted only at 2mo27d. The reason for delay in treatment was lack of time demanding for the patient and family. The patient was admitted and treatment was instituted since 3d and he presented no neurological symptoms. Molecular analysis was performed the proband was hemizygous for a mutation in ATP7A (c.1668-1680del, predicted to lead to the frameshift p.Leu556delMet*11). Mother is heterozygous for the same mutation. The third son did not have the mutation and consequently treatment was discontinued. We support the current relevance of molecular confirmation for diagnosis and genetic counseling, once clinical findings in the neonatal period are nonspecific and early treatment must be indicated. Molecular confirmation in our case was indispensable for the correct diagnosis of MD, once biochemical markers are nonspecific of copper deficiency.

Mitochondrial Heteroplasmy and Clinical Variability in a MELAS Family. K. McCombe1,2, K. Weisigger3, T. Huang4, J. Youngblom5, L. Lee6, N. Barasa7, C.A. Valencia8, S. Packman1. 1) California State University, Stanislaus; 2) University of California, San Francisco; 3) University of Cincinnati.

Clinical variability in mitochondrial disease is challenging due to highly variable expression and penetrance. One source of variability is the existence of heteroplasmy; the co-occurrence of wild-type and mutant mitochondrial DNA (mtDNA) within a cell. Heteroplasmic levels can differ both between individuals and between different tissue types in the same individual. Manifestation of pathology in a tissue is thought to depend on that tissue's threshold for mtDNA mutation load. While a relationship between mutation load and disease expression has been shown in vitro studies, current literature is conflicting on the utilization of heteroplasmy for predictive and prognostic counseling. In the present work, we asked whether there was a correlation between mutation load, as measured in urine sediment cells, and clinical manifestations in an extended Hispanic family with the ma3243G mutations of MELAS [mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes] mutation. This study examined 23 maternal relatives of a patient diagnosed with MELAS for the presence of clinical symptoms, physical findings and degree of heteroplasmic expression in urine sediment cells. For the molecular analyses and assessment of percent heteroplasmic, mitochondrial DNA was isolated from cells in urine samples and the ma3243G mtDNA mutation was detected by real time PCR. Presence and number of clinical symptoms [patient reports] and physical findings [examination by SP and CL] were compared to percentage of mtDNA A3243G mutation detected in urine sediment cells for each individual. This study found that 65% of family members self-reported MELAS-associated symptoms, and 52% had physical findings, both neurologic and other systemic. A significant correlation was found between age and number of MELAS-associated symptoms, as well as age and number of physical findings. Interestingly, no significant correlation was found between levels of mutation load measured in urine and the number of symptoms or physical findings. We conclude that urine sediment cell heteroplasmy analysis is not a uniformly effective test to predict clinical phenotype in individuals, or small sample sizes that carry the ma3243G mutation.

Morquio A is caused by deficient activity of N-acetylgalactosamine-6-sulfatase (GALNS) resulting from mutations in GALNS. GALNS mutations are numerous and heterogeneous. To aid detection and interpretation of mutations, we summarize published mutations from 541 Morquio A patients, together with 81 published mutations not described as genotypes from individual Morquio A patients, and report a new public-access GALNS locus-specific database. 277 unique GALNS mutations were identified from 1081 alleles. Most alleles reported from patients (79%) are missense. Even the most frequent mutations are uncommon. The three most common alleles (R386C, 1113F, G301C) together only represent 14% of alleles. Significant geographical and/or ethnic origin-based allele frequency variability exists. 11% of alleles are insertions/deletions, 6% intronic, 4% nonsense. Most Morquio A-associated mutations have been reported 1-2 times. 48% of patients are homozygous for a GALNS mutation, 39% heterozygous, and 13% have only one mutation detected. Mutation detection and genotype-phenotype correlations are in part due to the heterogeneity of GALNS mutations and lack of multiple families with the same mutations. Parental testing is encouraged. Reporting new alleles facilitates distinguishing pathogenic from benign mutations. The standard for Morquio A diagnosis is still deficient GALNS enzyme activity measured in leukocytes or fibroblasts, together with normal control enzyme activities.

Expanding the toolbox for the diagnosis of mitochondrial disorders. L. Krenn1, L. Pichler2, T. Schirmayer1, T. Haack1, T. Wetland1, T. Stroim, M. Mann1, T. Meitinger1, H. Frohlich1. 1) Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany; 2) Institute of Human Genetics, Technische Universität München, Munich, Germany; 3) Max Planck Institute of Biochemistry, Martinsried, Germany.

Despite good progress in the field, many disease causing mutations in patients with mitochondrial disorders still have to be identified. We recently applied whole exome sequencing (WES) in 400 unrelated individuals with juvenile-onset mitochondrial disorder. In 88% of patients, we were unable to identify the disease causing mutation and hence additional tools are needed to promote the identification of the disease gene for the remaining 52% of cases. Sequencing of trios will expand the screening for de novo mutations, while whole genome sequencing (WGS) approaches will ease the discovery of additional DNA variation not seen by WES. However, the interpretation of the consequences of rare variants remains a challenge. To this end, we established a collection of more than 200 fibroblast cell lines from patients with mitochondrial disorders, WES and respiratory chain complex I deficiencies. We are using the coding variants from the exome data to predict the disease causing mutations in leukocytes or fibroblasts, together with normal control enzyme activities.

Two children with Phenylketonuria with normal tetrahydrobiopterin biochemical testing and normal Phenylalanine Hydroxylase gene testing. M. Oron1, S. Mooney1, H.L. Levy4. 1) Division of Genetics, Phoenix Children's Hospital, Phoenix, AZ; 2) Division of Genetics and Metabolism, Boston Children's Hospital, Boston, MA.

We report two girls with biochemically confirmed hyperphenylalaninemia who do not have detectable mutations of the Phenylalanine Hydroxylase (PAH) gene. The first girl was seen at age 16 months for developmental delay, hyperactivity, microcephaly; her plasma phenylalanine (Phe) level was 1620µmol/L (27mg/dL). Urine amino acids were normal except for increased phenylalanine. PAH gene sequencing performed on two different occasions showed no detectable mutations. Typically, PAH gene sequencing should be performed in a multiplex PCR assay including the coding and intronic sequences. The next was bidirectional sequencing of the ampliﬁed exons. Investigation of the mentioned genes in all families are ongoing. In a second patient, both parents were heterozygous carrier of the disease. The spectrum of the mutations are heterogeneous. To aid detection and interpretation of mutations in the PAH gene, we report a new mutation in a twin which is the outcome of a consanguineous marriage. The patient is an 8 years old boy with developmental delay and severe mental retardation. Homozygosity mapping showed the probable mutation in SDHD gene. Direct sequencing of the SDHD gene revealed a mutation of the gene in the patient and in the TD in C234 of the mentioned gene. Our results show that the parents are heterozygous carrier of the disease. The spectrum of the mutations underlying MSUD would expand by these previously unpublished mutations in the new PAH gene. The clinical diagnosis by genetic testing, provided the availability of Prenatal Diagnosis and pre-genetic implantation Diagnosis for demanding families. Also it helps to identify heterozygous carriers to reduce the burden of MSUD in Iran.
2239T Molecular Characterisation of known & novel mutations in Congenital Adrenal Hyperplasia patients (CYP21A2 gene): Genetic & diagnostic evaluations. R. Khajuria1, R. Prasad, R. Wali2, K. Bhansali3, S. Dhillon1, P.J. Garg1,2,3,4,5, K. Godbole1, N. Makadaria1.1) Department of Genetic Medicine, Deenanath Mangeshkar Hospital & Research Center, Pune, India; 2) Department of Medical Genetics, King Edward Memorial Hospital, Mumbai, India; 3) Department of Pediatrics, Seth GS Medical College & KEM Hospital, Mumbai, India; 4) Genetics Division,PGIMER,Chandigarh, Chandigarh, India; 5) Endocrinology,PGIMER,Chandigarh India.

Congenital Adrenal Hyperplasia (CAH) is an autosomal recessive disease, with a wide range of clinical manifestations from severe classical form to late onset form. Most cases of CAH, the inherited inability to synthesize cortisol, aldosterone with subsequent overproduction of androgens, are caused by mutations in steroid 21 Hydroxylase genetic analysis. As a complement to hormonal measurements, mutation analysis of CYP21A2 gene is potentially important tool in diagnosis of steroid 21 Hydroxylationdeficiency as well as genetic counselling. Our aim was to determine the frequency of common CYP21A2 gene mutations and identify novel mutation. Clinical and hormonal evaluations were used to categorize the patients in Salt Wasting (SW), Simple Virilizing (SV) and Non Classical (NC) forms. About 95 % of mutant alleles have apparently been transferred from linked pseudogene (CYP21P1) to active gene (CYP21A2). Molecular analysis of CYP21A2 was performed in 50 patients for detection of common mutations viz gene deletion.Q318X,R356W,V281L,E6 cluster,Ins 1 in exon 7 or L307 frameshift mutation, P267L by PCR, ACRS, restriction method and finally confirmed by sequencing . Polymorphisms viz D183E & S268T were identified in 50 patients and 50 control by DNA RFLP and sequencing. Novel mutations were identified by SSCP technique and subsequently sequencing of amplified product. Disease causing mutations were identified in patients comprising SW(n=14), SV(n=22) and NC(n=14). Single gene deletion was found with frequency of 16.4 % in SW, SV patients whereas homozygosity was found with 7.9% in these cases. Frequency of other known mutations were to be present at R356W mutation ( 20 %), Q318X (11 %), V281L (25 %), i2g (36 %), L307 (10 %) and P267L (4 %). D183E and S268T polymorphisms are also present in our population . The genotypic frequency of SW (81 %), SV (17 %) and NC (2 %) was present at R356W mutation. H365N, F306V, P357P, D234D are novel mutations in CYP21A2 gene. Each novel mutation was present at frequency of 2 %. This is a comprehensive study showing compound heterozygosity in majority of our CAH population.
2241T
DUP 24BP IN CHIT1 IN SIX MEXICAN AMERICAN POPULATIONS. T.D. Da Silva Jose, 1 J.A. Juarez, 1 A. Porras, 1 K.J. Jaurez, 1 M.T. Magara, 1 L. Sanders, 1 A. Valladares, 1 M. Cruz, 1 M. Gonzalez, 1 A. Soto, 1 J.E. Garcia. 1) Genetics, Universidad de Guadalajara, Guadalajara, Jalisco, México; 2) Laboratorio de diagnóstico de enfermedades lisosomales, División de Genética, Centro de Investigación Biomédica de Occidente, CINOMS, Guadalajara, Jalisco, México; 3) Unidad de Investigación Médica en Bioquímica Centro Médico Nacional, “Siglo XXI” IMSS. Av Cuauhtémoc 330, Ciudad de México, México.

Introduction: In chitinotriase (CHIT1), a biomarker used in Gaucher disease, 24bp duplication is the most frequent polymorphism at CHIT1 and results in deficient enzymatic activity compromising its use as biomarker. In this study we wanted to determine the allelic and genotypic frequency of dup 24bp in six Mexican American populations. Material and methods: 692 samples were analyzed: Purepechas (49), Tarahumaras (97), Huicholes (97), Mayan (139), Tenek (97) and Nahuaus (213). DNA extraction and PCR was done as published elsewhere. Statistical analysis was done by direct counting of genotype and allele frequencies and HWE were used using chi-square and Markov chain by Arlequin 3.5. Results: We found the dup 24bp (CHIT1) in distributions observed in Table 1. All groups were in HWE. The allele frequency of dup 24bp was higher than expected in all analyzed groups compared with Mexican mestizo. Table 1, dup 24 bp in Mexican Amerindian populations PopulationsNumberGenotype Frequencies (%)Allele Frequencies (%)Hardy-Weinberg wt/wt wt/Dup Dup/Dup value Purepechas4924 (49)18 (74)30.30.30.40.5 Purepechas9724 (97)10 (73)79 (27)20 0.50 0.50 0.50 Tarahumaras9750 (97)30 (58)20 (42)18 0.50 0.50 0.50 Huicholes9743 (45)12 (25)38 (65)10 0.50 0.50 0.50 Mayan13944 (32)38 (27)58 (41)14 0.50 0.50 0.50 Tenek13944 (32)38 (27)58 (41)14 0.50 0.50 0.50 Nahuaus13829 (21)69 (79)0 (11)2 0.50 0.50 0.50 Results: We found the dup 24bp (CHIT1) in distributions observed in Table 1. All groups were in HWE. The allele frequency of dup 24bp was higher than expected in all analyzed groups compared with Mexican mestizo. Table 1, dup 24 bp in Mexican Amerindian populations PopulationsNumberGenotype Frequencies (%)Allele Frequencies (%)Hardy-Weinberg wt/wt wt/Dup Dup/Dup value Purepechas4924 (49)18 (74)30.30.30.40.5 Purepechas9724 (97)10 (73)79 (27)20 0.50 0.50 0.50 Tarahumaras9750 (97)30 (58)20 (42)18 0.50 0.50 0.50 Huicholes9743 (45)12 (25)38 (65)10 0.50 0.50 0.50 Mayan13944 (32)38 (27)58 (41)14 0.50 0.50 0.50 Tenek13944 (32)38 (27)58 (41)14 0.50 0.50 0.50 Nahuaus13829 (21)69 (79)0 (11)2 0.50 0.50 0.50 Tenek13944 (32)38 (27)58 (41)14 0.50 0.50 0.50

2242M

Mucopolysaccharidosis type VI (MPS VI) is a rare, autosomal recessive lysosomal storage disorder caused by deficiency of the enzyme N-Acetyl Galactosaminase 4-sulfatase resulting from mutations in the Aryl Sulfatase B (ARSB) gene. In the present study, molecular and functional characterization of the ARSB gene was done for twenty three patients and their families who were enzymatically confirmed to have Mucopolysaccharidosis type VI. A total of fifteen mutations were identified, of which eleven were novel mutations: p.W353X, p.H393R, p.S403Yfs, p.P445L, p.W450L, and p.W450C and three known mutations (p.D54N, p.A237D and p.S320R). All the novel mutations were confirmed not to be polymorphic variants by performing sequencing in 50 normal individuals from the same ethnic population.

2243T
Antioxidant enzymes gene expression in Fabry Disease along the circadian rhythm. A.C. Barris-Oliveira, V. D’Almeida. Department of Psychobiology, Universidade Federal de São Paulo, São Paulo, São Paulo, Brazil.

Fabry disease (FD) is caused by a mutation in the GLA gene, located in the X chromosome, which leads to a deficiency in the α-galactosidase A and thus, globotriaosylceramide storage in lysosomes. The symptoms of the disease are not completely explained by the lipid storage and pathophysiological changes observed in this disorder have been correlated with the activation of secondary pathways, as oxidative stress. Our group has already investigated the erythrocytes levels of antioxidant enzymes in FD patients and has observed higher catalase (CAT) activity levels and also, higher levels of total glutathione. The generation of reactive oxygen species and their removal have already been described as processes that vary along the day, following a circadian rhythm. The objective was to investigate gene expression of the antioxidant enzymes: superoxide dismutase (SOD1), glutathione peroxidase (GPX1) and CAT in FD and control fibroblasts along a circadian rhythm. Four fibroblast cultures were used, divided equally in FD and control cultures, from female and male individuals. To induce the circadian rhythm, it was used the serum shock protocol with DMEM 50% fetal bovine serum for 2 hours. The RNAs, extracted at the following times: T1, T4, T10, T16, T22 and T28, were further converted in cDNA and used for qPCR. The genes expressions were plotted in graphics in relation to 2684, used as endogenous control gene. Comparing the cell cultures from Fabry Disease Woman (FDW) with Control Woman (CW), SOD1, GPX1 and CAT showed similar circadian patterns of expression, and there were statistical differences for SOD1 at T28, for GPX1 at T10, T16 and T22, and for CAT at T16; besides, the amplitude of expression from CW seemed higher. Comparing the cultures from Fabry Disease Man (FDM) with Control Man (CM), the circadian patterns of expression looked different for all genes, and there were statistical differences for CAT at T4, T16 and T28; moreover, FDM presented lower amplitude of expression. Comparing CW and CM for all genes, the circadian expressions followed similar patterns, but CW tended to have higher amplitude of expression, as expected. With the data obtained, we conclude that FD cell cultures show a variable gene expression of antioxidant enzymes along the circadian rhythm, and this variation differs between FDW and FDM. Comparing information with other studies, we found higher amplitudes of expression between FD male and female patients. Funding: CAPES, FAPESP, AFIP, IGEIM.
Introduction. Lysosomal storage diseases (LSD) are a group of disorders associated with the accumulation of one or more specific substrates in the body and cause damage. In many individuals with GD, especially Type I, enzyme replacement therapy (ERT) can reduce the enlargement of the liver and spleen and improve anemia and thrombocytopenia. Recently, we encountered a 35-year-old man with GD manifested gastrointestinal involvement during ERT. Case: The patient was monozygote twin and both were diagnosed with GD by bone marrow biopsy showing marked increase in foamy histiocytes and lipid laden macrophages at 9 years old. They took a splenectomy at 10 years old. ERT was begun in the subject at the age of 24. They visited our hospital 3 years ago and have been treated with ERT as dosage of 60units/kg administered every other week. Their biomarkers including angiotensin converting enzyme and acid phosphatase were decreased but osteolytic bone lesions still remained and he was suffered from bone pain at that time. Whole body magnetic resonance imaging (MRI) showed severe bone necrosis in right hip and both tibia and liver was not enlarged. The subject complained of dyspepsia and difficulty of ingestion of any pills 3 years ago. He performed esophagogastroduodenoscopy (EGD) from another hospital and the result revealed normal. He continued complaining of dyspepsia despite medication with proton pump inhibitor. A follow-up EGD showed multiple nodular yellowish lesions on duodenum by gastrofibroscopy. Gastric biopsy was done and pathologic finding showed that nodular lesions consisted of Gaucher cell infiltrations. We increased the dose of cerezyme from 60IU/kg to 80IU/kg and performed genetic testing. Direct sequencing of GBA gene showed the compound heterozygous mutations, c.259C>T (p.R48W) and c. 5118G>A (p.R257Q). The mutations (R48W and R257Q) are known as rare mutations. The R48W regarded as mild mutation was reported in non-neuronopathic GD manifesting bone disease, whereas R257Q was previously reported in both acute neuronopathic and non-neuronopathic GD. We increased the dose of cerezyme from 60IU/kg to 80IU/kg and are planning to repeat gastrofibroscopy 6 months later. Conclusion: As ERT might change the natural course of GD during ERT, close monitoring of adverse events and sharing the patient’s course during ERT would be required.
oral galactose supplementation. In summary, PGM1 l

2247T Characterization of the Beta-Galactosidase Protein Isolated from Brain of Normal Sheep and from a Unique Ovine Model of GM1-Gangliosidosis. O.A. Nevin, A.J. Ahern-Rindell. Biology, University of Portland, Portland, OR.

Background: β-galactosidase (β-gal) is a lysosomal hydrolase that catalyzes the removal of terminal β-linked galactose moieties from glycoconjugates. Human β-gal is encoded by the GLB1 gene and synthesized as an 88 kDa precursor that is processed into a mature 64 kDa enzyme within the lysosome. Patients with mutations in the GLB1 gene develop GM1-Gangliosidosis (GM1 [MIM 230500]), a Lysosomal Storage Disorder (LSD) recognized by a decrease in β-gal activity in fibroblasts. Seven α-neuraminidase (α-neur). These enzymes associate and are stabilized within the Lysosomal Multienzyme Complex (LMC), a structure responsible for facilitating step-wise substrate degradation in the lysosome. We hypothesize that a missense mutation in the GLB1 gene alters the structure of β-gal and compromises the formation of the LMC, producing a secondary deficiency in α-neur activity.

Methods: Brain tissues were homogenized in lysis buffer containing protease inhibitors, centrifuged, and supernatant collected. Western blotting was used to determine the size of β-gal in ovine brain. Protein lysates were denatured and size-separated using SDS-PAGE. Ovine β-gal was detected using a feline-stimulated antibody that cross-reacted. A glyceraldehyde-3-phosphate dehydrogenase antibody was used as a protocol control (~38 kDa band) and followed by film capture of chemiluminescence. Results: Approximately 26, 72, and 79 kDa bands were consistently detected in normal and normal brain. Conversely, the ~79 kDa band was absent or a smaller band (~51 kDa) β-gal was isolated from normal and GM1-affected sheep fibroblasts. Subsequently, we found a C714T DNA sequence variation causing a C229F missense mutation in exon 6 of the GLB1 gene. These findings support our hypothesis that the disease-causing mutation alters the structure of β-gal, but not its size. Preliminary results suggest that tissue-specific protein expression plays a pronounced role in the size of β-gal. Post-translational modifications and/or the presence of β-gal isoforms may explain this size variation. Further analysis of the unique GM1 ovine model is in progress to characterize the specific interactions of the LMC components, to discover the underlying molecular mechanism of this disorder, and to create an effective therapeutic treatment for GM1.

2248M Galactose Supplementation Improves Glycosylation in PGM1-CDG. T.E. Gadomski, K.J. Scott, G. Preston, R.O. Crandall, L.A. Pro, T. Kozicz, E. Morava. Hayward Genetics Center, Tulane University School of Medicine, New Orleans, LA.

Congenital disorders of glycosylation (CDGs) are phenotypically diverse genetic syndromes caused by impaired glycoprotein synthesis. Phosphoglucomutase 1 deficiency (PGM1-CDG, MIM612941) is a newly discovered and possibly under-diagnosed disorder characterized by decreased glycolysis, hyperglycemia, abnormal liver function, muscle involvement and endocrine dysfunction. Intercellular cell adhesion molecule 1 (ICAM-1) serves as a useful biomarker for the measurement of cell surface glycosylation. Dietary galactose supplementation is a promising treatment for PGM1-CDG. We present the cases of two female patients, Patients 1 and 2, ages 16 and 19 respectively, who presented with rhabdomyolysis, hypogonadotropic hypogonadism and hypoglycemic episodes. Each patient was diagnosed with PGM1-CDG. Fibroblasts from both patients were cultured and stained for ICAM-1 to measure glycosylation. In vitro galactose complementation assays with repeat ICAM-1 staining were then performed on fibroblasts to assess for improved glycosylation. ICAM-1 staining in fibroblasts from both patients revealed decreased glycosylation compared to a normal control. Five recombinant expression systems were used to identify a novel mutations in the PGM1 gene, leading to a deficiency of N-acetylgalactosamine-6-sulfate enzyme which is encoded by this gene, resulting in the accumulation of keratan sulfate and condroitin sulfate in certain tissues. It is expressed as generalized skeletal dysplasia with corneal clouding, resulting in short stature, pectus carinatum, platyspondyly, odontoid hypoplasia, kyphoscoliosis, and genu valgum. Cardiovascular and respiratory systems can also be affected. This phenotype has a broad variability associated to the more than 180 mutations that have been reported. Until recently there was no effective therapy for treating the disease aside from a multidisciplinary approach, and this year was approved the enzymatic replacement therapy. Objective: to report the case of two siblings identifying a new mutation. Methods: 9 and 6 six year old brothers with clinical characteristic of severe Morquio syndrome, a deficit of enzymatic activity of GALNS enzyme, GALNS gene was analyzed by PCR and exon sequencing. Results: two mutations were found, a previously reported one in exon 3 (c.280C>T,p.R94C) and a new heterozygous variant in exon 9 (c.989G>A,p.G333D), with software analysis predicting it as probably damaging. Conclusion: These brothers have received an approach by a multidisciplinary team including pediatrician, orthopedist, geneticist, endocrinologist, neurosurgeon, physiatrist and neuropsychologist and a multidisciplinary approach by a multidisciplinary team including pediatrician, orthopedist, geneticist, endocrinologist, neurosurgeon, physiatrist and neuropsychologist and we hope that this case will be useful for the best control of the disease.

2250M Identification of genetic mutations in Malaysian patients with fructose-1,6-bisphosphatase deficiency. L.H. Ng¹, A.A. Nor Azimah², H.Y. Leong¹, B.C. Chen¹, H. Muzhirah¹, M.Y. Zabedah³, H. Anasuliza³, Y.K. Chor³, Y. Yuniata³. ¹Genetics Department, Hospital Kuala Lumpur, Jalan Pahang, 50586 Kuala Lumpur, Malaysia; ²Molecular Diagnostics and ProVen, Specialised Diagnostics Centre, Institute for Medical Research, Jalan Pahang, 50586 Kuala Lumpur, Malaysia; ³Biochemistry unit, Specialised Diagnostics Centre, Institute for Medical Research, Jalan Pahang, 50586 Kuala Lumpur, Malaysia; ⁴Paediatric Department, Sarawak General Hospital, Jalan Hospital, 93586 Kuching, Malaysia.

Fructose 1,6-bisphosphatase (FBPase) deficiency (MIM #229700) is an autosomal recessively inherited disorder of gluconeogenesis. Patients present notably with recurrent episodes of hypoglycaemia, ketosis and lactic acidosis, which can be fatal if not treated appropriately. Urinary analysis in acute samples may reveal increased glycerol and glycerol phosphate. However these biochemical diagnostic markers rapidly normalized following treatment with fructose, they are easily missed. FBPase (EC3.1.3.11) is encoded by FBP1 (NM_000507.3) gene on chromosome 9q22.2-q22.3. Genetic analysis of FBP1 is useful in diagnosing FBPase deficiency, avoiding the need of invasive liver biopsy. Although more than 20 FBP1 gene mutations have been described, there has been limited investigation into the genetics of this disorder in Malaysian patients. We studied 5 patients (2 males, 3 females) from 4 unrelated families with clinical symptoms and laboratory findings consistent with FBPase deficiency. All of them have recurrent episodes of hypoglycemia. Urinary fructose and glycerol phosphate were only detected in 2 patients. Bidirectional sequencing of all seven FBP1 coding exons plus approximately 50 base pairs of flanking non-coding intronic DNA on either side of each exon is performed using genomic DNA from unaffected genetic consanguous parents. We identified a previously reported homozygous missense mutation in exon 7, c.778G>A (p.[Gly260Arg]). Three other unrelated patients from non-consanguinous parents have novel mutations discovered. One of them has a homozygous missense mutation in exon 5, c.472C>T (p.[Arg158Trp]). One patient has a homozygous deletion c.392delT [p.(Val131Glyfs*71)] in exon 4 in association with a heterozygous splice site mutation at c.568-2A>C in intron 5. Another patient has a homozygous deletion c.603delG [p.(Val201Valfs*3)] in exon 6. This mutation causes a frame-shift change after Valine-201 and creating a new reading frame that ends at a stop at position 3. All the newly identified mutations in the FBP1 gene were not found in normal controls in our population and are predicted to result in loss of FBPase activity. Following their genetic diagnosis, all patients had a multidisciplinary approach by a multidisciplinary team including pediatrician, orthopedist, geneticist, endocrinologist, neurosurgeon, physiatrist and neuropsychologist and our results showed that multidisciplinary approach is recommended for the best control of the disease.

MORQUIO SYNDROME: NEW HETEROZYGOUS MUTATION OF THE GALNS GENE IN TWO SIBLINGS FROM SOUTH-WEST COLOMBIA. M.F. Hernandez-Amarez¹, F. Ruiz-Botero¹, S. Eschler², H. Pachajos³. ¹Centro de Investigaciones en Anomalías Congénitas y Enfermedades Raras (CIACER), Universidad Icesi, Cali, Colombia; ²Centogene, Rostock, Germany; ³Fundación Clínica Valle del Lili, Cali, Colombia.

Intravenous Mucopoly saccharidosis type IV A, or Morquio syndrome, is an autosomal recessive lysosomal storage disorder that is caused by mutations on the GALNS gene, leading to a deficiency of N-acetylgalactosamine-6-sulfatase enzyme which is encoded by this gene, resulting in the accumulation of keratan sulfate and condroitin sulfate in certain tissues. It is expressed as generalized skeletal dysplasia with corneal clouding, resulting in short stature, pectus carinatum, platyspondyly, odontoid hypoplasia, kyphoscoliosis, and genu valgum. Cardiovascular and respiratory systems can also be affected. This phenotype has a broad variability associated to the more than 180 mutations that have been reported. Until recently there was no effective therapy for treating the disease aside from a multidisciplinary approach, and this year was approved the enzymatic replacement therapy. Objective: to report the case of two siblings identifying a new mutation. Methods: 9 and 6 six year old brothers with clinical characteristic of severe Morquio syndrome, a deficit of enzymatic activity of GALNS enzyme, GALNS gene was analyzed by PCR and exon sequencing. Results: two mutations were found, a previously reported one in exon 3 (c.280C>T,p.R94C) and a new heterozygous variant in exon 9 (c.989G>A,p.G333D), with software analysis predicting it as probably damaging. Conclusion: These brothers have received an approach by a multidisciplinary team including pediatrician, orthopedist, geneticist, endocrinologist, neurosurgeon, physiatrist and neuropsychologist and we hope that this case will be useful for the best control of the disease.


2251T

Hereditary hemochromatosis is a relatively common autosomal recessive disorder caused by the high absorption and deposition of iron in major organs. This leads to organ dysfunction, and eventually arthritis, cirrhosis, cardiomyopathy, diabetes mellitus, hypergonadism and hepatocellular cancers. Early diagnosis and treatment are of utmost importance in the prevention of end stage organ damage. In more than 90% of hereditary hemochromatosis patients, HFE gene mutations have been responsible for the clinical manifestations. The aim of this retrospective study is to investigate the frequencies of common mutations found in individuals referred for HFE gene molecular analysis to a single genetic center in the Aegean region of Turkey. A total of 230 individuals were tested for HFE mutations between the years 2005 and 2014.June 1. The HFE common polymorphisms (H63D, C282Y, S65C, Q283P, E168Q, E168X, W169X, P160delC, Q127H, H636, V59M, and V53M) were studied using polymerase chain reaction. Of 230 patients studied for HFE mutations, 73% were found to have no mutation and 27% had mutations either homozygously or heterozygously. Three distinct mutation (C282Y, H636, S65C) and five different genotypes, with the exception of mutation free genotype, were detected in the group studied. Although no one was homozygous for C282Y mutation, one patient was heterozygous and another compound heterozygous for this mutation. Five patients (2.1%) were homozygous for H636 mutation, with another fifty three patients (23%) showing heterozygosity for it. One patient had the genotype H636/C282Y and one patient was heterozygous for S65C mutation. Among the mutant alleles H636 was the most prevalent allele with a frequency of 94.1%. In conclusion; C282Y mutation of HFE gene, which causes severe form of hemochromatosis and is common in Northern Europe, was not to be found in the patients, whereas the frequency of H636 mutation was found to be 18.1% in the patients studied. HFE gene screening can be a useful tool for the genetic counseling of the patients and their family members.

2252M
An infant with hyperhomocysteinemia, methylmalonic aciduria, and an atypical cellular distribution of protein-bound cobalamin. M. Pupavac1, F. Petrella1, D. Watkins1, S. Fahiminiya1, J. Muenzer1, J. Majewski1,2, D. Rosenblatt1. 1) Human Genetics, McGill University, Montreal, Quebe, Canada; 2) McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada; 3) University of North Carolina, Chapel Hill, North Carolina, USA.

In mammalian cells, derivatives of vitamin B12 (cobalamin, Cbl) are required as cofactors for two enzymes. Adenosylcobalamin (AdoCbl) is required in the mitochondria as a cofactor to methylmalonyl CoA mutase (MCM), and methylcobalamin (MeCbl) is required in the cytoplasm as a cofactor to methionine synthase (MS). Inborn errors of Cbl metabolism can present with elevations of methylmalonic acid and/or homocysteine in the blood and urine. A sixteen-month-old boy of Hispanic ethnicity was investigated because of elevated serum methylmalonic acid and homocysteine levels. Additional clinical findings included cardiac defects (ventral septal defect, long QT syndrome), cleft palate, hypospadias, hyperbilirubinemia, and one patient was heterozygous for S65C mutation. Among the mutant alleles H636 was the most prevalent allele with a frequency of 94.1%. In conclusion; C282Y mutation of HFE gene, which causes severe form of hemochromatosis and is common in Northern Europe, was not to be found in the patients, whereas the frequency of H636 mutation was found to be 18.1% in the patients studied. HFE gene screening can be a useful tool for the genetic counseling of the patients and their family members.

2253T
Insightful investigation of mtDNA integrity in affected tissues of patients with mitochondrial disorders. J. Wang, J. Lin, X. Tian, V.W. Zhang, E.S. Schmitt, LJ. Wong, Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Introduction: Mitochondrial biogenesis and maintenance defects are due to mutations in a group of nuclear genes, including TK2, SUCCL2, RRM2B, SUCLG1, DGUNG, MPV17, TOMP, POLG, POLG2, and OPA1. Affected tissues usually showed mitochondrial DNA (mtDNA) depletion, multiple deletions, or point mutations. MtDNA multiple deletions are known to be associated with mutations in nuclear genes involved in mtDNA biosynthesis. Investigation of mtDNA alterations in affected tissues can reveal impacts of nuclear gene defects on mtDNA integrity and their correlation with clinical phenotype. Methods: The mitochondrial genome is amplified by long range PCR followed by massively parallel sequencing (MPS) to assess mtDNA point mutations, large deletion and multiple deletions. MtDNA content was examined by Real Time-qPCR. Nuclear gene was analyzed by target gene capture/MS or Sanger sequencing. Results: In a total of 34 patients with nuclear gene mutations, mtDNA deletion or multiple deletions in muscle or liver were detected in 25 patients (73.5%). Ten patients with mutations in MPV17, DGUGK, TK2, SUCGL1 or POLG had mtDNA deletion in the affected tissues, and presented with encephalohapathopathy or encephalomyopathy at young age (1 month to 4 years). Multiple deletions were detected in 11 patients with mutations in RRM2B, POLG, COL10c2 or OPA1. These patients presented with myopathy at adulthood (17 to 78 years old, mean 46). Four patients with encephalopathy (4 to 66 years old) and mutations in TK2, RRM2B or POLG had mtDNA deletion and multiple deletions. The number of low heteroplasmic variants is significantly increased compared with tissues without mutations. Conclusions: MtDNA multiple deletions and depletion are secondary defects to defects in nuclear genes responsible for mtDNA biosynthesis and maintenance. Mutations in the POLG, TK2 or RRM2B gene can cause both mtDNA depletion and multiple deletions. Mutations in the MPV17, DGUGK or SUCGL1 gene are mostly associated with mtDNA depletion in patients with early onset of encephalopathy or hepatopathy, whereas autosomal dominantly inherited mutations in COL10C2 or OPA1 were associated with multiple deletions without mtDNA depletion. This study provides insights into the mtDNA integrity and its correlation with nuclear gene mutations and phenotypes. MtDNA depletion and/or deletions in affected tissues may guide molecular diagnosis.

2254M
X-chromosome inactivation in females heterozygotes for Fabry disease. L. ECHEVARRIA1, K. BENISTAN1, A. TOUSSAINT1, A. BOUROUG3, A.A. HAGEG2, D. ELADARI1, F. JABBOUR1, I. DETRAIT1, C. BELDJ1, P. DE MAZANCOURT1, P. GERMAIN1. 1) Medical Genetics, University of Versailles, Garches, France; 2) Molecular biology, CHU Cochin, Paris, France; 3) Cardiology, CHU Ambroise Paré, Boulogne-Billancourt, France; 4) Cardiology, HEGP, Paris, France; 5) Physiological, HEGP, Paris, France; 6) Biophysics, CHU Raymond Poincare, Garches, France; 7) EA 2493, UFR des Sciences de la Santé, Montigny, France.

Background: Fabry disease (FD) is an X-linked genetic disorder caused by the deficient activity of lysosomal α-galactosidase A. Inheritance is X-linked, so that while males are usually severely affected, female patients range from asymptomatic to clinical disease equal to that of males. The underlying reasons for this phenotypic variability in female patients are still poorly understood. Aims: To evaluate the existence of X-inactivation in females with FD, its variability and concordance between tissues, and its potential contribution to clinical phenotype. Patients and methods: 56 female patients were enrolled in the study. Extensive clinical work-up was carried out and two validated scores assessed global clinical severity. Kidney function was evaluated using measured GFR, while magnetic resonance imaging (MRI) evaluated left ventricular mass index, left posterior wall thickness and interventricular septum thickness. Alpha-galactosidase A residual activity was measured in leukocytes. X-inactivation status was analysed using DNA methylation studies at the Hpa II sites in the highly polymorphic CAG repeat in the first exon of the androgen receptor gene in four different tissues (leukocytes, mouth smears, skin biopsies and urine sediment). Results: Disease severity increased with age. Skewed or highly skewed X-inactivation was found in 29 percent of the study population, and 46.5% had skewed X-inactivation and residual α-galactosidase A activities, global severity scores of FD, progression of cardiomyopathy and deterioration of kidney function was evidenced. Discussion: This study shows that X chromosome inactivation variability in females with FD, its concordance between tissues and its potential role in the natural history of the disease and as a decision making criteria for enzyme replacement therapy in female patients.
2255T

Multiple gene variants cause a phenocopy for Lysinuric Protein Intolerance: A case report. S. Lipski1, A. Hebert1, O. Alpan2, S. Pattison2, A. Tolun3, S. Huguenin1, E. Haverfield3, J. Neidich1, O. Goker-Alpan1, 1) Medical Genetics Center, Asan Medical Center, Seoul, South Korea; 2) Dept. Pediatrics, Asan Medical Center, Seoul, South Korea.

Glycogen storage disease (GSD) ia is caused by mutations in the G6PC gene. This condition is characterized by hepatomegaly, doll-like face, hypoglycemia, lactic acidosis, dyslipidemia and hyperuricemia. The aim of the study was to characterize clinical and molecular features and evaluate late complications in Korean patients with GSD ia. Thirty Korean patients (18 males and 12 females) from 26 unrelated families were diagnosed based on clinical data and biochemical analysis. The mean age of diagnosis was 8.8 ± 9.9 years (range, 8 month to 42 years) and the follow-up period was 11.0 ± 7.0 years (2.4-43 years). Most patients (60%) presented with hypoglycemia and frequent epistaxis (37%) were common findings as well. Serum lactate and uric acid levels at diagnosis were 27.44 ± 32.33 mmol/L and 7.91±2.4 µmol/L, respectively. Serum cholesterol and triglyceride levels were high (mean level: 222.8 ± 56.1 and 639.8 ± 507.1 mg/dL, respectively) in 55% of cases. In genetic analysis, c.648G>T was the most common (45/52 alleles, 86.5%), followed by p.G122D (2 alleles), p.G222R (1 allele), p.S326P (1 allele), p.R83H (1 allele), p.F51S (1 allele) and p.Y128* (1 allele). Allopurinol and uncoupled carnitine were given to all patients. Fourteen patients (46.7%) received fibrates or HMG-CoA reductase inhibitors for hyperlipidemia. Biochemically, GSD ia is characterized by circulating deficiency of arginine and ornithine, suggestive of P5CSD. The presented case has motor and speech delay. The diagnosis was confirmed by measuring plasma levels of proline, ornithine and arginine. Its deficiency (P5CSD) is characterized by neurological and connective tissue abnormalities.

2257T

Two new unrelated cases of Pyrroline-5-carboxylate synthase deficiency: a Canadian perspective. Y. Trakadis, M. Berry, S. Fox, A. Khan, C. St Martin, D. Morel, D. Buhas, McGill U., Montreal, Canada.

Objectives: Pyrroline-5-carboxylate synthase catalyzes the biosynthesis of proline, ornithine and arginine. Its deficiency (P5CSD) is characterized by neurological and connective tissue abnormalities. 11 cases have been reported, with abnormal amino acids profile in only two families. Two new cases will be described here. Methods: A literature review is performed. Results: Both patients (B.A. and B.M.) are of Inuit origin and carry the same novel mutation in ALDH18A1 (homozygous c.544A>G). The two families are non-consanguineous and unrelated. B.M. has had a history of irritability, tremor and hypotonia since birth. He developed bilateral cataracts at 6 months. Plasma amino acids showed low Proline, Arginine, Ornithine and Histidine, suggestive of P5CSD deficiency. ALDH18A1 sequencing confirmed the diagnosis. At 2 years old he still had jitting episodes, mild developmental delay, short stature, and feeding problems. His brain MRI demonstrated an unusual T2 hyperintensity in the pons and bilateral widening of the temporal / insular CSF space. B.A. presented at 10 months old with bilateral subcapsular cataracts and recurrent vomiting. She had mild dysmorphic features and stretchy skin. The low level of Citrulline and Arginine, as well as the low/normal Proline and Ornithine levels were evoking a defect in pyrroline-5-carboxylate synthase. Conclusions: Although rare, P5CSD should be in the differential diagnosis of metabolic causes for early cataracts. A founder effect may exist in the Inuit population of Quebec.
2265T NORRBOTTINIAN VARIANT OF GAUCHER DISEASE IN SOUTHERN ITALY: LONG TERM FOLLOW-UP. M. Grisolla 1, S. Simon 1, F. Ceravolo 1, E. Pascale 1, M. Filocamo 1, P. Strisciuglio 2, D. Conticello 1 1) Pediatrics, University “Magna Graecia”, Catanzaro, Italy; 2) Centro di diagnostica genetica e biochimica delle malattie metaboliche, Istituto G. Gaslini, Genoa, Italy; 3) Department of Pediatrics, University Federico II, Naples, Italy.

The Norrbottian type of Gaucher disease has been described many years ago, as due to a unique mutation which may have happened in or before the 16th century in northern Sweden and is a well defined nosological entity with a characteristic course of clinical manifestations. We report the results of a long-term follow-up of four patients affected by this Gaucher type III. The patients, originating from South of Italy, present with clinical features and progression of disease, in particular for the skeletal involvement, comparable with the “Norrbottian” Swedish phenotype. The patients (3M, 1F), median age at diagnosis was 1.8 years (range 0.5-6.2 years) belonging to three different families, had been diagnosed at a median age of 3.4 years (range 1.1-8 years). The clinical manifestations at diagnosis were hepatosplenomegaly, thrombocytopenia, anemia and growth retardation. The [L444P]→[L444P] genotype was confirmed in all of them. Bone involvement consisting, primarily, of bone pain, bone crisis, osteopenia and bone abnormalities (such as Erlenmeyer flask deformity) detected by X-ray, was present in all four patients. With age a progressive kyphoscoliosis, due to wedging of vertebral bodies, occurred in all patients together with a restrictive ventilatory defect. All of the four patients had cognitive delay and other neurological manifestations began to appear at a median age of six years (range 4-10 years), as horizontal gaze, intentional tremor, seizures for one patient and depression for two different patients. The patients started Enzyme Replacement Therapy (ERT) at a median age of 11 years (range 5-22 years) and were followed during the years of treatment with regular clinical observations, biochemical tests and psychometric testing. The duration of ERT was associated with significant improvements in platelet count, haemoglobin, liver and spleen volumes, while neurological involvement seemed no to benefit of ERT so as skeletal manifestations.

2266M A Novel mutation in the Methylenetetrahydrofolate Reductase (MTHFR) gene in a Turkish child: A Case Report. M. ARSLAN 1, S. VURUCU 1, H.I AYDIN 1, B. UNAY 1, R. AKIN 1 1) GULHANE MILITARY MEDICAL SCHOOL, DEPARTMENT OF CHILD NEUROLOGY, ANKARA, Turkey; 2) TURGUT ÖZAL UNIVERSITY, DEPARTMENT OF PEDIATRIC METABOLISM DISEASES, ANKARA, Turkey.

Severe methylenetetrahydrofolate reductase deficiency is an autosomal recessive metabolic disorder of folate metabolism causing elevated levels of homocysteine and decreased levels of methionine. A 12-year-old boy with nonspecific developmental delay was admitted to our clinic with a 2-week of history of inability to walk, muscle weakness, speech disorder, confusion, and visual hallucinations. The patient was found to have high serum homocysteine and low-normal serum methionine, cerebral and cerebellar atrophy. Molecular genetic analysis identified homozgyous p. F564V (c.1690T>G) MTHFR mutation in the patient.


2261T The role of inflammation in vascular disease in the MPS I canine model. M. Vera 1, S. Le 1, S. Kan 1, P. Dickson 1, R. Wang 1 1) Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA; 2) CHOC Children’s, Orange, CA.

Inflammation has come to be appreciated as an important pathophysiologic mechanism in the mucopolysaccharidoses (MPS). Inflammatory markers and patterns of acute inflammation have been identified in the joints and cardiovascular systems in several MPS animal models. Undoubtedly, inflammation causes a significant portion of the morbidity and progressive decline suffered by MPS patients. Arterial disease has been well described in mucopolysaccharidosis type I (MPS I) and consists of intimal-medial thickening with proliferation of vascular smooth muscle cells and extracellular matrix remodeling. Evidence of inflammatory pathways involving Toll-like receptor 4 (TLR4) and transforming growth factor beta (TGFβ) signaling have been observed in these arterial lesions. Though the inciting events stimulating these pathways have not been identified. In this report we have studied the role of macrophages as effector cells in the initiation of the vascular disease phenotype. Our model hypothesizes that macrophages are recruited to sites of vascular endothelial damage or dysfunction caused by lysosomal storage where they secrete cytokines that stimulate an inflammatory cascade.
2264M
Autoimmune thrombocytopenia in a patient with Hunter Syndrome: a rare association. A. Rufus, J. Flores, B. Croke, K. Fernandez, P. Sandford, R. Antony. Department of Pediatrics, University of Illinois College of Medicine, Peoria, IL.

Background: Individuals with Hunter syndrome develop coarse facial features, hepatosplenomegaly, cardiovascular disease, skeletal dysplasia, pulmonary disease, and immunologic deficits and deafness from about 2 years with phenotypic variation. Hematological complications of Hunter syndrome are typically mild thrombocytopenia and neutropenia. We describe our experience managing a patient with Hunter syndrome with severe autoimmune thrombocytopenia and neutropenia.

Case Report: A 13-year-old boy with Hunter syndrome presented with fever, fatigue, poor oral intake and decreased urine output. Initial hematological indices: WBC 4600/microitre, Hemoglobin 5.3g/dl, Platelet count 7000/microitre, and ANC 1940/microitre. Parvovirus B19, varicella, EBV, and CMV studies were negative. Blood and bone marrow examinations (pancytopenia, dyserythropoiesis, no abnormal cell lines and no hemophagocytosis) ruled out hematological malignancy and HLH. Pericardial effusion was assumed to be viral or secondary to anemia. Despite multiple platelet transfusions his Hemoglobin (<6 g/dl) and platelet count (<100000/microitre) failed to increase. The patient then developed acute intracranial hemorrhage and pulmonary hemorrhage necessitating oscillatory ventilation. Steroid dose was increased to IV methylprednisolone (MP) 12mg/kg/day and daily plasmapheresis was commenced (continued for 11 days). Platelet infusions, Keppra, Labetalol and Clozapine were used for supportive care. His PICU course/platelet count were as follows: Day 3: Platelets = 94000 (MP 2 mg/kg/day), Day 13: Platelets = 195000 and Hemoglobin 10.2 (MP 1mg/kg/day). Day 17: EVD/ Ventilation discontinued. Day 45: MP stopped. On stopping the MP his platelet count slowly dropped to 147000 so MP restarted (30 mg Q8h) and IVIG 400mg/kg/day was administered for 4 days. This permitted us to wean and discontinue the MP by day 82. (Platelet =253000).

Weekly immune replacement therapy was then commenced and patient's platelet counts have been consistently above 150000. Conclusion: With this report of the first case of Hunter syndrome associated with autoimmune thrombocytopenia we hope to add to the body of knowledge regarding the management of hematological complications of Hunter syndrome.

2266M
Optic nerve atrophy in methylmalonic acidemia (MMA): natural history, pathological findings and experience with anti-oxidant therapy. I. Manoli1, W.M. Zein2, J.L. Sloan1, E. Harrington1, Y. Wang3, J. Zhang2, B.P. Brooks2, A. Hamosh2, C.C. Chan3, C.P. Venditti1. 1) Genetics and Molecular Biology Branch, NHGRI, NIH, Bethesda, MD; 2) Ophthalmic Genetics & Visual Function Branch, NEI, NIH, Bethesda, MD; 3) Immunopathology Section, NEI, NIH, Bethesda, MD; 4) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

Optic nerve atrophy (ONA) is increasingly recognized as a complication of isolated MMA, caused by a deficiency of the mitochondrial enzyme methylmalonyl-CoA mutase (MUT). Multisystem manifestations include growth failure, pancreatitis and renal disease, but patients can also suffer acute neurological events such as bilateral globus pallidi strokes and ONA. The natural history and disease pathophysiology remain unknown. 26 patients with isolated MMA (17 mut, 4 cblA and 5 cblB, including 3 with organ transplant) underwent ophthalmologic evaluations, including serial optical coherence tomography (OCT), through a dedicated natural history protocol. A total of 56 OCT studies were performed, with 12 patients contributing 2-15 serial measures. Of 77 patients with isolated MMA evaluated, 10 mut patients (6M, 4F) presented with symptoms of acute or subacute visual loss. Age of presentation ranged from 7 to 27 years. Visual loss was bilateral and symmetric and 8 patients were legally blind. Thinning of the retinal nerve fiber layer (RNFL) preceded symptomatic presentation in 2 patients. There were no apparent environmental triggers shared between patients, and MR imaging variably revealed thin optic nerves. Ophthalmic pathology, assessed by light and electron microscopy using autopsy samples from a 14.5y old mut patient with normal visual acuity, showed thinning of the macular outer retinal layers, with minimal thinning of intraretinal ultrastructure of enlargement, edema, and degeneration in the photoreceptor inner segment, the retinal ganglion cell and vascular endothelium, as well as corneal epithelium and keratoocyte but not the retinal pigment epithelial (RPE) layer. Coenzyme Q10, vitamin E, idebenone and intravenous N-acetylcysteine were employed with variable success in 5 of the patients. Thinning of the RNFL progressed with near complete loss of vision in all over the ensuing years. Our experience suggests that 1) ONA is a complication of isolated MMA; 2) the RNFL can deteriorate over a long period of time, even before the onset of OCT; 3) pre-symptomatic diagnosis; 2) cell-intrinsic mitochondrial morphological changes in the retina, similar to what has been observed in the proximal tubule epithelial cells in the kidney, suggests that cellular autonomy may exist in MMA in isolated MMA, and possibly organ transplantation may not prevent progression of disease, but can ameliorate the rate of loss of visual function.
2267T
Ophthalmic manifestations of Cobalamin C disease occur independent of metabolic control and prenatal treatment. B. Brooks\(^1\), A.H. Thompson\(^2\), N. Carrillo-Carrasco\(^3\), J. Sloann\(^4\), I. Manoli\(^5\), W.M. Zein\(^6\), C.P. Vendi\(^7\)
\(^1\) National Eye Institute, NIH, Bethesda, MD; \(^2\) National Human Genome Research Institute, NIH, Bethesda, MD; \(^3\) National Center for Advancing Translational Sciences, NIH, Bethesda, MD.

Cobalamin C (cblC) disease, an inborn error of intracellular vitamin B12 metabolism, is caused by mutations in the MMACHC gene and leads to impaired intracellular synthesis of adenosyl- and methylcobalamin, which produces both methylmalonic acidemia (MMA) and hyperhomocysteinemia. To examine the clinical course of ocular manifestations of cblC in relation to genotype and metabolic parameters over time we conducted a retrospective, observational case series study at the NIH Clinical Center. The 25 cblC patients ranged in age from 2 to 27 years at last ophthalmic visit, and follow-up ranged from 0 months to 83 months. There were a total of 69 visits: 14 patients were seen more than once, for a median follow-up duration of 36 months. Nystagmus was present in 64% of patients and strabismus 36%. Despite significant elevation of plasma homocysteine, lens dislocation was never observed. The median visual acuity at last ophthalmic visit of the 18 patients with measurable acuity was 1.15 LogMAR (20/280). Retinopathy with a prominent macular component was present in 72% of patients, with MMACHC c.271dupA homoygotes (14) showing the most extensive degeneration. Retinopathy was accompanied by optic nerve pallor and vascular changes in 68% and 44% of patients, respectively. Empirc treatment with prenatal vitamin B12 in two cases resulted in children with less cognitive dysfunction and better visual function than their affected older siblings. The retinopathy, while delayed in its progression compared to that of the older sibling, nonetheless progressed. These data suggest that maternal prenatal vitamin B12 therapy may be palliative, but does not prevent ocular symptoms or eventual ocular progression. Our longitudinal study reports the ocular status of the largest group of patients with cblC systematically examined at a single facility over an extended period of time. There are general differences in the progression and severity of macular degeneration, optic nerve pallor, and vascular attenuation between homozygous c.271dupA patients and compound heterozygote genotypes. Tempo and chronicity of the ophthalmic manifestations may be helpful in distinguishing cblC disease displays a developmental as well as a degenerative phenotype.

2268M
GENOTYPE-PHENOTYPE CORRELATION IN FABRY PATIENTS DETECTED BY LYSOSOMAL NEWBORN SCREENING, J. Navarrete\(^1\), D. Cervantes\(^1\), A. Limon\(^2\), R. Del Valle\(^3\), R. Delgado\(^4\), Dept Gen, Hosp Sur PEMEX, Mexico City, Mexico; \(^2\) Pediatrics Division, Hosp Sur PEMEX, Mexico City, Mexico; \(^3\) Gerencia Medicina Preventiva, Subdireccion Servicios de Salud PEMEX; \(^4\) Pediatrics Division, Hosp Reg Villahermosa, Tab, Mexico; \(^5\) Faculty of Medicine, Universidad Anahuac, Mexico City, Mexico.

The goal of newborn screening is an early detection of inborn errors of metabolism diseases. In Mexico we began newborn screening since 1977 with very few inborn errors of metabolism such as phenylketonuria, galactosemia, congenital hypothyroidism, sickle cell anemia and cystic fibrosis. Petroleos Mexicanos is a big governmental institution with approximately ten thousand workers and their families. Since 2005 a larger screening has been done to all newborns in this institution through all the country. We test for most aminoacidopathies including acidurias, hemoglobinopathies, G6PD deficiency, adrenal hyperplasia, cystic fibrosis, and biotinidase deficiency; since August 2021 we included primary immunodeficiencies, Gaucher disease, Niemann-Pick (A/B) disease, Pompe disease, Krabbe disease, and MPS 1. We analyzed our results from August 2012 to May 2014. We have found 4 newborns with Fabry disease confirmed with enzyme activity and molecular analysis; and 4 patients with Pompe disease, three were pseudodeficiencies and one was late onset presentation. Since the most lysosomal prevalent disease in our cohort was Fabry disease, we describe our findings and compared them with other populations, and we make a close follow up of all newborns with pathological mutations and study and analyze their families. We also look for biomarkers of the disease so we can know the timing and speed of the mutation and the course of the disease as soon as possible.

2269T
Biochemical, Molecular and Clinical Heterogeneity in Very-Long-Chain Acyl-CoA Dehydrogenase Deficiency. The Atlantic Canadian Experience. J. Gillis\(^1\), S. Dyack\(^2\), D. Skidmore\(^2\), H. McDonald\(^1\), J. Farrell\(^1\), M. Chapman\(^1\), N. Kureshi\(^2\).
\(^1\) Dept Medical Genetics, IWK Health Centre, Halifax, NS, Canada; \(^2\) Dalhousie University, Halifax, NS, Canada.

Very long-chain acyl-CoA dehydrogenase deficiency (VLADD, MIM 202147) is a recognized rare genetic disorder of fatty acid metabolism caused by mutations in the gene ACADVL. The disorder is classified into three forms depending on the time of onset and severity of illness. VLADD demonstrates autosomal recessive disease inheritance. Accurate diagnosis and early detection are crucial for favourable clinical outcomes for these patients and for this reason VLADD is now part of most expanded newborn screening (NBS) programs including the Atlantic Canadian provinces (Nova Scotia, New Brunswick and Prince Edward Island). Diagnosis is established through; acylcarnitine analysis by MS-MS of plasma or a dried blood spot specimen, molecular genetic analysis and measurement of residual enzyme activity in lymphocytes. However, very little remains known of the clinical course of this highly variable condition and there are no established guidelines for management. Since screening began in the Maritimes a decade ago, we have ascertained more true positive cases than expected, demonstrating a greater prevalence for the condition. We have identified novel and deleterious mutations in those ascertained via NBS, as well as in those born post-NBS for VLADD and clinically symptomatic; notably we have identified a novel splice site variant in patients of Acadian ancestry. Our primary objective is to characterize the biochemical, molecular and clinical phenotype of patients with a diagnosis of VLADD, including those identified via NBS in the Atlantic Maritime provinces. We have identified 17 patients known to our clinic ranging in age from birth to 25 years. Despite their asymptomatic presentation those ascertained on NBS were found to have two deleterious mutations and low level of residual enzyme activity (REA) on follow up analysis. Review of clinical data showed that ≤10% REA correlates with development of symptoms in our patients including; significant hypoglycemia, myopathy, cardiomyopathy and death. Moreover our data also suggest that Individuals with REA within 10-15% are also at risk to develop symptoms. Our study indicates that our current NBS method is sensitive and adequate to identify infants with VLADD and that acylcarnitine testing on NBS profile on NBS may possibly predict biochemical phenotype correlated to REA and allow for earlier risk assessment for cardiac, hepatic and metabolic complications associated with this condition.

2270M
GAL-1-P levels and GAL Gene mutations in infants following abnor- mal newborn screening for galactosemia in South Florida. S. A. Hos- seinl, C. Hung, B. J. Illagan, G. Ghaffar, B. Johnson, O. Bodamer, Division of Genomic Medicine, Department of Human Genetics, University of Miami, Miller School of Medicine, Miami, FL 33136.

Background: Galactose-1-phosphate uridylytransferase (GALT) deficiency (OMIM #228400) is a lysosomal recessive inherited condition with an incidence of 1 in 30,000 to 60,000 newborn infants. Galactosemia is one of the 29 core conditions for newborn screening in the United States. No data are available for confirmatory testing in South Florida. Methods: We received blood samples from 42 infants following an abnormal newborn screening for low GALT activity from August 2012 until May 2014. All infants were born in the State of Florida and screened through the State Newborn Screening Laboratory in Jacksonville. Confirmatory testing samples were analyzed at the Clinical Biochemical and Molecular Diagnostic Laboratories, Department of Human Genetics at the University of Miami for erythrocyte galactose-1-phosphate (GALT-P) levels and for GALT gene sequencing, respectively. Standard laboratory protocols were employed. The laboratories are accredited by the College of American Pathologists (CAP), and AP certified and participate in proficiency testing. Results: 16 infants had the D/G genotype (GALT-P: mean +/- SD: 21.8 mg/dl +/- 16.6), 18 infants were carriers for either D or G variants (GALT-P: 7.1mg/dl +/- 10.4), and 8 were without detectable mutation (GALT-1-P levels ranged 1 to 1.4 mg/dl). GALT-P levels were significantly different between infants with mutation, carriers, and D/G infants (p<0.005). The Duarte variant c.940A>G (p.Asn314Asp) in cis with the promoter GTCA deletion (c.-116,-119delGTCA) is present on 25% of alleles. 22% of alleles carried the pathogenic mutations c.404C>T (p.Ser135Leu) or c.583A>G (p.Glu195Arg). Two carriers had the Duarte variant without, GTCA promoter deletion. We did not observe any infants with classic galactosemia. Discussion and conclusions: Confirmatory testing following abnormal newborn screening for galactosemia is straightforward provided GAL-1-P can be analyzed within a few hours followed by timely molecular testing.
Newborn screening and the incidences of inherited metabolic and newborn defects of metabolic disorders by genomic sequencing technology.

By typing rare variation in these genes, including that flagged as pathogenic by either population incidence of these diseases, we infer that the vast majority of (20 variants) and 23% of compound heterozygous protein-altering mutations with MAF ≤ 0.01 are present, and 76 protein-altering variants show homozygous (20 variants) or compound heterozygous (56 variants) state in at least one person. Several genes have higher homozygous frequency than expected from Hardy-Weinberg equilibrium. We found one homozygous or compound heterozygous protein-altering mutation with MAF ≤ 0.01 in a FAOD, AAD or OAD related gene in 17, 28, and 19 people, respectively, and measure compound heterozygous protein-altering variants (13 variants) are pathogenic. Moreover, based on the annotated in the MENA region. There is significant benefit in expanding national NBS for multiple disorders. In most of the region, screening is limited to a few endocrine and metabolic disorders in 10 of 17 MENA countries. Rates of annual births in the region. Incidence rates had no universal screening program, which together comprised 87.8% of annual births in the region. Incidence rates were identified for 66 inherited metabolic and endocrine disorders in MENA. Rates of disease were found to be higher in MENA countries than in Caucasian neonates of Middle Eastern descent. Published data on MENA NBS programs are insufficient—only a small number of MENA countries, accounting for 12.2% of regional births, have widespread neonatal genetic screening. The primary outcome was the mean DCS score on the validated Decisional Conflict Scale (DCS) (1-100, higher scores indicating higher decisional conflict). Secondary analyses explored associations with respondent characteristics, their understanding of key messages, and cognitive complexity. The mean (95% CI) DCS score was 27 (25-29). We observed statistically significant differences in DCS score between recipients and non-recipients of test messages. The heel-prick pain message (p<0.01) and the storage/secondary use of bloodspots message (p<0.05) were associated with lower mean DCS scores. Specific knowledge questions were more likely to be answered correctly if the participant had received the corresponding message. The mean DCS score declined significantly with increasing number of messages received (test for linear trend, p<0.01). The study has high internal validity, but needs to be replicated in more diverse populations, using a broader set of outcomes. The factorial survey design is a useful approach in this context.

Treatment of hyperammonemia in the neonate is at best problematic. Many factors complicate treatment. Discovering the hyperammonemia requires a good degree of suspicion. Current treatment requires removal of ammonia by Renal Replacement Therapy (RRT), either by hemodialysis or CVVH (continuous veno-venous hemofiltration). Confounding the treatment is the size of the patient (neonate) who usually weigh between 2.5 and 3 kg, and the placement of catheters suitable for dialysis. “The optimal RRT prescription for neonatal hyperammonemia remains unknown,” as was concluded by Spinale, et al, in a recent article in Pediatric Nephrology (2013). We decided to poll nephrologists as to their approach to RRT for hyperammonemia in a neonate. An informal telephone poll was conducted. Nephrologists were asked their approach to treating a 3 kg neonate with hyperammonemia. Findings: 1. The start of therapy was cited as the important factor for successful outcome 2. Catheter placement whether by ICU staff, interventional radiology or Pediatric surgery was not as much of a concern as was the speed with which the line could be placed so RRT could begin. 3. Catheter size was generally agreed to be a 7 French for this size infant. 4. Area of placement was optimally felt to be the internal jugular in most cases, but again speed of access would override the area of placement. 5. Hemodialysis vs CVVH: The division here seemed to depend on the experience of the nephrologists. Those more comfortable with hemodialysis preferred that mode of ammonia removal, while other felt that higher flows with CVVH could achieve similar results and was more easily accomplished with CVVH. Standardization of treatment is difficult because it depends on the availability and experience of the treating nephrologist and factors related to each institution.
Induction of immune tolerance in MPS I patients initiating enzyme replacement therapy with Aldurazyme. G.F. Cox¹, R. Giugliani¹, P.V. Novakov², S. Richards¹, Y. Xue¹. 1) Genzyme, a Sanofi company, Cambridge, MA; 2) Medical Genetics Service/HCPA, Department of Genetics UFRGS and INAGEMP, Porto Alegre, RS, Brazil; 3) Department of Clinical Genetics, Moscow Research Institute for Pediatrics and Children Surgery, Moscow, Russia.

Objectives: Mucopolysaccharidosis I (MPS I) is a rare autosomal genetic disorder caused by deficiency of the lysosomal enzyme α-L-iduronidase and the subsequent accumulation of its substrates, the glucosaminoglycans (GAGs) dermatan and heparan sulfates. This results in progressive and debilitating multi-organ disease. Enzyme replacement therapy with laronidase (Aldurazyme®) is approved for the treatment of patients with MPS I. Studies in MPS I dogs have shown that antibodies to laronidase affect its biodistribution and ability to clear glucosaminoglycans (GAGs). In patients with MPS I, high antibody titers correlate with less urinary GAG reduction, but not with clinical response. This study was designed to determine whether an immunosuppressive regimen that induced immune tolerance in dogs could be used in patients with MPS I. Methods: This open-label clinical trial (NCT 00741338) in patients with severe MPS I (5 years of age) included a Tolerance Induction Period (TIP) followed by an Immune Challenge Period (ICP). The immunosuppressive regimen consisted of cyclosporine A (CsA), azathioprine, and low-dose weekly Aldurazyme (0.058 mg/kg). Following gradual discontinuation of immunosuppressants, patients received full-dose Aldurazyme (0.58 mg/kg) for a total of 24 weeks. Immune tolerance was defined as an anti-laronidase IgG antibody titer ≤1:3200 at the end of the ICP. A sequential adaptive design allowed for modification of the immune tolerance regimen in a second patient cohort if the initial regimen was unsuccessful (<20% of 3 patients immune tolerant). Safety evaluations focused on signs of infection or organ toxicity, blood pressure elevation, and bone marrow suppression. Results: None of the 3 patients in Cohort 1 and 1 of 3 patients in Cohort 2 achieved immune tolerance. A second patient in Cohort 2 maintained a low IgG titer after 17 weeks of full-dose Aldurazyme, but discontinued the study due to infusion-associated reactions and was considered a treatment failure. CsA levels were variable and required frequent modifications. There were no deaths or other serious adverse events. Conclusions: The results of the study are inconclusive. The risk, benefit, and feasibility of immune tolerance induction should be reconsidered in patients with MPS I. This study was supported by the Genzyme/BioMarin Joint Venture.

Clinical response to eliglustat in treatment-naïve patients with Gaucher disease type 1: Post-hoc comparison to imiglucerase in a real-world setting. R. Manksö, J.S. Taylor, J. Angeli, M.J. Peterschmidt. Genzyme, a Sanofi company, Cambridge, MA.

Introduction: Gaucher disease type 1 is a multi-systemic lysosomal storage disorder resulting from acid β-glucosidase deficiency. The standard of care for more than two decades has been enzyme replacement therapy. Eliglustat is a novel oral substrate reduction therapy for adults with Gaucher disease type 1.

Objective: To compare long-term treatment response to eliglustat and imiglucerase in treatment-naïve patients with Gaucher disease type 1.

Methods: Four-year data from eliglustat-treated patients in an open-label study (NCT00358150, N=26) and 18-month data from a randomized, double-blind, placebo-controlled study (ENGAGE: NCT00891202, n=20 in eliglustat arm) were compared to 75 matched imiglucerase-treated patients enrolled in the ICGG Gaucher Registry who had received at least 15 U/kg/2 weeks.

Results: At baseline, hematologic parameters were similar in the two groups but eliglustat patients had slightly larger spleens and livers. Time course and degree of improvement were similar for eliglustat- and imiglucerase-treated patients for most parameters. After 4 years, mean spleen volume decreased by 63% and 48%, mean liver volume decreased by 27% and 30%, mean platelet count increased by 95% and 99%, and mean hemoglobin level (g/dL) increased by 2.27 and 0.71 in eliglustat and imiglucerase patients, respectively. Improvements in lumbar spine and total femur Z-scores were consistently higher in the eliglustat group at all time points; however, bone data were limited from the imiglucerase-treated patients. The Z-score increases observed with eliglustat were higher than those observed by Wenstrup et al. (2007, J Bone Min Res) during low- to high-dose treatment with imiglucerase (0.06-0.13 Z-score/year) in patients who had similar mean baseline bone mineral density.

Conclusion: Although not a head-to-head trial, this post hoc analysis suggests that eliglustat, in treatment-naïve patients with Gaucher disease type 1, results in improvements in organ volumes and hematologic parameters that are comparable to those observed with imiglucerase in a real-world setting.
2282M ENCORE: A randomized, controlled, open-label non-inferiority study comparing eliglustat to imiglucerase in Gaucher disease type 1 patients stabilized on enzyme replacement therapy: 24-month results. T.A. Burrow1, T.M. Cox2, G. Drellichman3, R. Cravo4, M. Balwani5, A.M. Martins6, E. Lukina7, B. Rosenbloom8, L. Ross9, J. Angell10, A.C. Puga11. 1) Cincinnati Children’s Hospital Medical Center, Cincinnati, OH; 2) University of Cambridge, Addenbrooke’s Hospital, Cambridge, UK; 3) Hospital de Niños Ricardo Güirreñes, Buenos Aires, Argentina; 4) HEMORIO, Rio de Janeiro, RJ, Brazil; 5) Mount Sinai Hospital, New York, NY, USA; 6) Universidade Federal de São Paulo, São Paulo, SP, Brazil; 7) Hematology Research Center, Moscow, Russia; 8) Cedars-Sinai/Tower Hematology Oncology, Beverly Hills, CA, USA; 9) Genzyme, a Sanofi company, Cambridge, MA, USA.

Introduction: Gaucher disease is an autosomal recessive disorder caused by deficiency of the lysosomal enzyme acid β-glucocerebrosidase (β-glucocerebrosidase or glucosylceramidase), resulting in progressive substrate accumulation and a spectrum of debilitating visceral, hematologic, and skeletal manifestations. Eliglustat is a novel oral substrate reduction therapy in development for Gaucher disease type 1. This open-label Phase-3 trial (ENCORE, NCT00943111, Genzyme, a Sanofi company) evaluated eliglustat and imiglucerase in patients who had reached pre-specified therapeutic goals after ≥3 years of enzyme replacement therapy. We report efficacy data from the 12-month primary analysis period (PAP) and the first 12 months of the extension period during which all patients received eliglustat.

Methods: Patients were randomized 2:1 eliglustat: imiglucerase. The primary efficacy endpoint was percent of patients remaining stable on a composite of spleen, liver, hemoglobin, and platelet parameters. As this was a non-inferiority trial, efficacy analyses were performed on the per-protocol population (99 eliglustat, 47 imiglucerase patients).

Results: Eliglustat was non-inferior to imiglucerase: after 12 months, 85% of eliglustat and 94% of imiglucerase patients maintained all four goals (lower bound of 95% CI of difference -1.6% within the pre-specified [-25% non-inferiority margin]). One hundred and forty five (91%) of the 159 patients treated in this study completed 24 months of treatment. Preliminary 12-month extension data demonstrate continued stability in spleen volume, liver volume, platelet count and hemoglobin level in most of the 99/106 patients who continued on eliglustat and most of the 46/53 patients who received imiglucerase in the PAP and then eliglustat in the trial extension. Most adverse events were mild or moderate in severity. In general, the overall number, seriousness and type of adverse events were similar for all patients during day 1 to month 24 is similar to the profile for eliglustat randomized patients in the PAP as well as similar to the safety profile in patients who switched from imiglucerase to eliglustat after 12 months.

Conclusions: In the Phase 3 ENCORE study, most patients maintained clinical stability while on eliglustat for 12 or 24 months.


Introduction: Gaucher disease type 1 (GD1) is a multi-systemic lysosomal storage disorder resulting from acid β-glucocerebrosidase deficiency. Eliglustat, an experimental oral substrate reduction therapy for adults with GD1, is metabolized mainly by CYP2D6, a polymorphic cytochrome P450 isozyme. Clinical studies used a dose-escalation scheme (range from 300 to 150 mg BID) to ensure plasma eliglustat steady-state pre-dose concentrations (C\text{\textsubscript{\text{trough}}}) above 5 ng/mL; however, this method has limitations in clinical practice. An alternative dosing regimen of 100 mg BID was evaluated for CYP2D6 genotypes 100 mg BID dose (-31.9%). In the Phase 3 randomized placebo-controlled study of GD1 patients switching from enzyme replacement therapy to eliglustat (ENCORE, NCT00943111), 55 of 88 IM/E/PM patients (62.5%) were dosed at either 50 or 150 mg BID (maximum dose allowed) based on C\text{\textsubscript{\text{trough}}}. Modeling confirmed that the percent change in spleen volume (primary endpoint) was similar whether patients were dosed at 50 or 150 mg BID (-27.8%, observed) or if all IM/E/PM patients had received a 100 mg BID dose (-31.9%). In the Phase 3 randomized placebo-controlled study of GD1 patients switching from enzyme replacement therapy to eliglustat (ENCORE, NCT00943111), 55 of 88 IM/E/PM patients (63.2%) were dosed at either 50 or 150 mg BID (maximum dose allowed) based on C\text{\textsubscript{\text{trough}}}. Modeling predicted that the change in spleen volume was similar whether patients were dosed at 50 or 150 mg BID (-27.8% observed) or if all IM/E/PM patients had received a 100 mg BID dose (-6.3%) and that this dosing regimen would not result in any additional treatment failures. For all patients, projected eliglustat exposure at the 100 mg BID dose was within the range observed in the clinical studies (i.e., would not be expected to affect safety).

Conclusion: A simplified dosing regimen for eliglustat based on CYP2D6-genotype-predicted phenotype (100 mg BID for IM/EM patients) would achieve exposure, efficacy, and safety results similar to dosing based on plasma eliglustat concentrations.


Objectives: Pompe disease (PD) is a lysosomal storage disorder caused by deficiency of alpha-1-4-glucosidase which leads to an accumulation of glycogen in lysosomes. The aim of this study is to evaluate the profile of patients with PD who are treated at the Reference Center of Inborn Errors of the Universidade Federal de São Paulo - Brazil (CREIM). Methods: Medical records of patients with confirmed diagnosis of PD were analyzed.

Results: Six patients, one male (16.5%) and five female (83.5%) were analyzed. One patient (16.5%) died at seventeen years due to respiratory complications. Patients were divided into two groups: onset of symptoms in childhood (CG) and initial symptoms in adulthood (AG). There were no patients censored. One patient (16.5%) died at seventeen years due to respiratory complications. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option. ERT was used as an attempt to prevent storage during the elapsed time from birth to transplant/engraftment. Whether transplantation or other treatments were used was an “optional” option.
Inhibition of Hepatic Mitochondrial Metabolism during Systemic Immune Activation. P.J. McGurie1, S. Matsumoto2, K. Saito2, J. Senac1, S. Colognesi1, T.N. Tarasenko1, 1) National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 2) National Cancer Institute, National Institutes of Health, Bethesda, MD; 3) National Institute of Child and Health Development, National Institutes of Health, Bethesda, MD.

The current study demonstrated that hepatic mitochondrial dysfunction is a common feature of systemic immune activation, and may contribute to the development of metabolic complications.

POSTERS: GENETIC DISORDERS

Precise targeted gene correction of arginase-1 deficiency using single-stranded oligodeoxynucleotides with TALENs or the CRISPR/Cas9 system. Y. Y. Sin, M. Morales, C.M. McCracken, C.D. Funk. Department of Biomedical and Molecular Sciences, Queen’s University, Kingston, Ontario K7L 3N6 Canada.

Arginase-1 (ARG1) deficiency, a rare autosomal recessive disorder, leads to hyperammonemia with progressive neurological impairment, growth retardation and infrequent episodes of hyperammonemia. This disorder is caused by mutations throughout the eight exons and several splice sites of the ARG1 gene located on chromosome 6q23. Currently, there is no cure and pharmacological treatment is limited. Our study aims to correct ARG1 mutation alleles through nuclease-induced homologous recombination (HR) in the presence of homologous donor repair templates. Using human cells [HEK293, fibroblasts, induced pluripotent stem cells (iPSCs)], we employed engineered transcriptional activator-like effector nucleases (TALENs) and the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9) system, designed de novo, to target and modify specific DNA sequences at the ARG1 gene locus. Here, we show that these genome editing tools efficiently cleave chromosomal DNA in a precise, predictable and robust manner, with mutation frequencies (indels) of up to 50%. Moreover, we designed short single-stranded oligodeoxynucleotides (ssODNs) as donor repair templates, which carry a unique novel restriction enzyme site and two flanking homology arms on both sides of the target region.

Sequence-independent targeting, these ssODNs introduce desired DNA sequence edits in exons 2, 3 and 7 in conjunction with the delivery of TALENs or CRISPR/Cas9 and a guide RNA (gRNA). Correctly targeted clones are being isolated via limiting dilution, PCR screening and/or RE digestion. Overall, our results demonstrate the versatility of using both TALEN and CRISPR/Cas9 systems to provide proof-of-concept for gene manipulation in human cells and sequence alterations within a defined locus can be achieved simultaneously by introducing ssODNs alongside. iPSCs have been generated from dermal fibroblasts obtained from individuals with ARG1 deficiency. It is anticipated that HR-mediated direct gene corrections in ARG1 deficiency patient-derived iPSC lines may prove useful for autologous cell-based therapies.

Genotype and Phenotype of Vietnamese patients with ornithine transcarbamylase (OTC) deficiency, K.Ngoc. Nguyen1, D.Chi Vu1, M. Chi Nguyen1, H.Viet Dau1, T.Anh Ta2, G.H Kim3, H.W Yoo2. 1) Department of Endocrinology, Metabolism, Genetics, National Hospital of Pediatrics, Hanoi, Viet Nam; 2) Medical Genetics Center, Aasn Medica Center, Korea.

Ornithine transcarbamylase (OTC) deficiency, a partially dominant X-linked disorder, is the most common inherited defect of the urea cycle. The disease presents variable severity affecting both males and females. Molecular diagnosis of OTC deficiency was the first time in Vietnam.

Methods: Three patients were diagnosed OTC with the criteria of hyperammonemia, hypocitrullinemia and orotic aciduria. 10 exons and their respective exon-intronic boundaries of the OTC gene were sequenced using genomic DNA. Results: There were two boys with newborn - onset form (the age of onset was 2 days and 8 days of age) and 1 girl with late - onset form (the age of onset was 18 months of age) in our study. One of two cases with newborn - onset form had two older brothers died at 5 and 7 days of age due to unknown coma. The rest were the 1st child. The initial symptoms of two newborn - onset cases were poor feeding, then vomiting, convulsion, comma and apnea. The initial symptoms of the late - onset form (the age of onset was 2 days and 8 days of age) and 1 girl was vomiting, convulsion, comma and right hemiplegia. Management of hyperammonemic crisis of two newborn - onset cases included glucose infusion, L arginine, L carnitine, hemofiltration, restricted protein diet therapy (metabolic formula). Management of one late - onset case included glucose infusion, L arginine, L carnitine, restricted protein diet therapy (metabolic formula). One case of newborn - onset and one case of late on - set crisis survived (1 year old and 6 years old). The rest died at 8 days old. Analysis of OTC gene detected 3 different mutations in 3 cases: c.77G>A (p.Arg26Gln), c.289+5G>C (IVS3+5G>C), c.422G>A (p.Arg141Gln) heterozygous. Conclusion: Newborn - onset OTC deficiency occur in Vietnamese male patients (n=1); newborn - onset OTC deficiency occur in Vietnamese female patients. The outcome depends on the diagnosis.
MEDNIK syndrome: clinical and biochemical delineation of the copper metabolism phenotype and response to zinc therapy. D. Martineilli,1,2 M. Ben Tartousi,3 C.A. Drouin,4 C. Dionisi-Vici,4 N. Braverman1,3, S.G. Kaler1. 1) Section on Translational Neuroscience, Molecular Medicine Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, Bethesda, MD, USA; 2) Division of Metabolism, Bambino Gesù Children’s Hospital, IRCCS, P.S.S., Rome, Italy; 3) Genetic Department, Montreal Children’s Hospital, QC, Canada; 4) Department of Dermatology, Centre Hospitalier Regional Grand-Portage, Riviere-des-Loups, QC, Canada; 5) McGill University-Montreal Children’s Hospital Research Institute, QC, Canada.

MEDNIK syndrome (MIM 609313) - an acronym for mental retardation, enteropathy, deafness, neuropathy, ichthyosis, keratoderma - is caused by mutations in the AP1S1 gene, encoding the sigma 1A subunit of adaptor protein complex-1 (AP-1). AP-1 plays a crucial role in the intracellular trafficking of transmembrane proteins, including the copper-transporting ATPases.

ATP7A is the genetic cause of Menkes disease, whereas ATP7B is responsible for Wilson disease. In addition to identifying rational treatment approaches required for normal brain growth and maturation, as illustrated in patients with Menkes disease, were normal in the one subject studied to date. This patient's CSF copper level was increased, corroborating the diagnosis of Menkes disease, whereas ATP7B gene was normal. All patients presented with hypocupremia, increased serum free copper level, increased urinary copper excretion, and liver copper accumulation. The hepatic component of the illness has been shown to be due to copper deposition in the liver. The disease is characterized by skeletal deformities and mental retardation, the latter often preceded by death. More severe clinical presentation of the mutation is the MELAS syndrome (Mitochondrial Myopathy, Encephalopathy, Lactic Acidosis and Stroke-like episodes). The aim of our study was to search for m.3243A>G mutation among patients sent for MIDD or MODY testing, who lack mutations in the most common MODY genes (i.e. GCK, HNF1A or HNF4A) and those with diabetes without a high proportion of deafness.

In summary, we report the first description of a new syndrome, MEDNIK syndrome, characterized by clinical and biochemical features consistent with Menkes disease. The new symptoms and signs in our patients were growth retardation, especially short stature, and severe mental retardation. The disease is characterized by skeletal deformities and mental retardation, the latter often preceded by death. More severe clinical presentation of the mutation is the MELAS syndrome (Mitochondrial Myopathy, Encephalopathy, Lactic Acidosis and Stroke-like episodes). The aim of our study was to search for m.3243A>G mutation among patients sent for MIDD or MODY testing, who lack mutations in the most common MODY genes (i.e. GCK, HNF1A or HNF4A) and those with diabetes without a high proportion of deafness.

In summary, we report the first description of a new syndrome, MEDNIK syndrome, characterized by clinical and biochemical features consistent with Menkes disease. The new symptoms and signs in our patients were growth retardation, especially short stature, and severe mental retardation.

Background and aims. Diabetes and deafness - the leading clinical features of Menkes disease - increase the likelihood of identifying a heteroplasmic M3243A>G mutation in mitochondrial DNA. Higher levels of the diabetes phenotype are observed in those with a high proportion of mutant coexisting with the wild-type sequence in heteroplasmy. We have performed serial clinical and neurological evaluations of three subjects with MEDNIK syndrome under zinc therapy (approved for the treatment of Wilson disease because oral zinc limits copper absorption at the intestinal level). All patients had mental retardation, increased serum free copper level, increased urinary copper excretion, and liver copper accumulation.

The hepatic component of the illness has been shown to be due to copper deposition in the liver. The disease is characterized by skeletal deformities and mental retardation, the latter often preceded by death. In summary, we report the first description of a new syndrome, MEDNIK syndrome, characterized by clinical and biochemical features consistent with Menkes disease. The new symptoms and signs in our patients were growth retardation, especially short stature, and severe mental retardation.

Background and aims. Diabetes and deafness - the leading clinical features of Menkes disease - increase the likelihood of identifying a heteroplasmic M3243A>G mutation in mitochondrial DNA. Higher levels of the diabetes phenotype are observed in those with a high proportion of mutant coexisting with the wild-type sequence in heteroplasmy. We have performed serial clinical and neurological evaluations of three subjects with MEDNIK syndrome under zinc therapy (approved for the treatment of Wilson disease because oral zinc limits copper absorption at the intestinal level). All patients had mental retardation, increased serum free copper level, increased urinary copper excretion, and liver copper accumulation.

The hepatic component of the illness has been shown to be due to copper deposition in the liver. The disease is characterized by skeletal deformities and mental retardation, the latter often preceded by death. In summary, we report the first description of a new syndrome, MEDNIK syndrome, characterized by clinical and biochemical features consistent with Menkes disease. The new symptoms and signs in our patients were growth retardation, especially short stature, and severe mental retardation.

In summary, we report the first description of a new syndrome, MEDNIK syndrome, characterized by clinical and biochemical features consistent with Menkes disease. The new symptoms and signs in our patients were growth retardation, especially short stature, and severe mental retardation.
2294M

Objectives: Laronidase (Aludrazyme®) is an enzyme replacement therapy approved for the treatment of patients with mucopolysaccharidosis I (MPS I), a genetic disorder caused by deficiency of α-L-iduronidase and subsequent accumulation of glycosaminoglycans (GAG) throughout the body. Most MPS I patients treated with laronidase in clinical trials have developed anti-laronidase antibodies. Theoretically, antibodies might alter the biological effects of laronidase, inhibit its enzymatic activity, limit its cellular uptake into target tissues, and/or causing an allergic reaction. The objective of this meta-analysis was to determine whether there are any differences in anti-laronidase antibodies and clinical outcomes, hypersensitivity reactions, and urinary GAG reduction in MPS I patients receiving laronidase. Methods: Data from 73 patients enrolled in 4 clinical trials of laronidase were used for this meta-analysis. Patients received treatment with the labeled dose of laronidase (0.58 mg/kg/wk) for 26 to 208 weeks. Efficacy parameters included 6-MWT, % predicted FVC, liver volume, range of motion, left ventricular mass, investigator global assessments, and uGAG levels, as appropriate, in individual studies. Safety and immunogenicity parameters included infusion-associated reactions (IARs), anti-laronidase IgG antibodies (seroconversion time to seroconversion, peak titer, overall exposure over time), enzyme activity inhibition and enzyme cellular uptake inhibition assays, and hypersensitivity testing (i.e. IgE complement activation, and serum tryptase).

Results: Based on the analysis of selected clinical outcome measures, there was no apparent correlation between anti-laronidase IgG antibodies and efficacy of laronidase. Likewise, there was no apparent relationship between antibodies and the occurrence of potential allergic reactions or IARs. There was, however, a statistically significant inverse correlation between uGAG reduction and antibody titer (p=0.0001), consistent with individual study results. Conclusions: This meta-analysis describes the most comprehensive systematic review of immunogenicity data from clinical studies of Aludrazyme (including long-term follow-up), and involves patients with all MPS I phenotypes. Although the presence of antibodies correlated with impaired urinary GAG clearance, no impact on the clinical efficacy or safety of laronidase was observed. This study was funded by the Genzyme/BioMarin Joint Venture.

2296M
Antisense U1-snRNAs-based correction of g.9273C>T and g.9331G>A GLA deep intronic mutations which cause Fabry disease. L. Ferri1,2, A. Cacciotti2, G. Covello3, D. Perrone4, R. Guerrini1,2, M.A. Dentì2, A. Morroni1,2, 1) Department of Neurosciences, Psychology, Pharmacology and Child Health (NEUROFARBA), University of Florence, Florence, Italy; 2) Paediatric Neurology Unit and Laboratories, Neuroscience Department, Meyer Children’s Hospital, Florence, Italy; 3) Centre for Integrative Biology, University of Trento, Trento, Italy; 4) Department of Biology and Evolution., University of Ferrara, Ferrara, Italy.

Mutations in the GLA gene cause enzyme deficiency of alpha-galactosidase A (α-GAL A) which leads to the X-linked Fabry disease, a lysosomal storage disorder of the glycosphinoglipids metabolism. GLA produces two transcripts, the major mRNA which encodes α-GAL A and a minor mRNA whose function is unknown. The minor mRNA isoform arises by an alternative splicing and carries a cryptic exon between exons 4 and 5 which contains a premature stop codon. Two GLA deep intrinsic mutations, the g.9273C>T and g.9331G>A, have been described being the cause of a pathological unbalanced expression of these two GLA transcripts, leading to α-GAL A deficiency and Fabry disease. Antisense-induced exon skipping represents a promising approach to correct such kind of mutations and can be obtained by the use of antisense oligonucleotides (AOEs) or antisense modified U1snRNAs (U1asRNAs). Hence, our research focuses on the setting up of an Exon Skipping based approach to correct the deep intronic mutations of the GLA gene. We designed and produced a set of AOs and of U1asRNAs specific to silence both g.9273C>T and g.9331G>A mutations. We also designed three GLA minigene to express the two studied mutations and the wild-type. We tested our antisense molecules against the minigines by Real-Time PCR and by Western Blot analysis. U1asRNAs that we designed promoted the recovery of the α-GAL A activity up to the wild-type level for both g.9273C>T and g.9331G>A GLA mutations. Western blot showed an increment of α-GAL A protein produced as a consequence of the treatment with antisense molecules. The analysis of transcripts showed the antisense-induced exon skipping of the cryptic exon with an increment of the major GLA mRNA. Our study suggests that antisense molecules are able to develop an approach that can be potentially used in the correction of splicing defects in Fabry disease. Authors wish to express their gratitude to: Fondazione Meyer ONLUS, Florence, Italy (Young Researcher grant to Lorenzo Ferri); Fondazione Meyer ONLUS (Regione Toscana, Italy grants CRO PSE 2007-2013); Italian Ministry of Health and Regione Toscana (Ricerca Finalizzata2011-02347694); AMMeC Foundation, Italy. 

2298T
Best Policy for helping LSD patients and Families: Collaboration between Charity Foundation, MOH, and pharmaceutical companies. F. Hadipour1,2, A. Hadipour2,3, A. Torabi2,3, Z. Hadipour1,2, Z. Hadipour1,2
1) Medical Genetics, Sarem Cell Research Center & Hospital, Tehran, Tehran, Iran; 2) YASNA Charity Foundation, Tehran, Iran; 3) Golden Daric CO, Tehran, Iran.

The lysosomal storage disorders (LSDs) are due to deficiencies of lysosomal enzymes caused by mutations of genes that encode the enzyme, proteins involved in transportation, and/or cofactors. Lysosomal enzymes degrade most biomolecules. This process is crucial for cells. LSDs result in accumulation of products of normal metabolic degradation. More than 45 LSDs are known, and occur in about 1 in 5,000-7,700 live births. LSDs have recessive inheritance. The presence of “founder” mutations in Middle East with a high degree of consanguinity is expected to lead to high prevalence of LSDs. Treatment varies across the LSDs. The ERT of LSDs is expensive comparing with the minimum wage of Iranian; so most of patients could not afford it._charities play a vital role in facilitating treatment of rare disease especially LSD. Charities exist to help people, serve in areas where there are little or no resources available, and provide aid by educating, setting up programs, providing necessary medical supplies. YASNA is a non-profit, non-political, and non-governmental charity supported by people. YASNA strive to care for patients with rare disease. Main goal of YASNA include: Diagnosis and treatment, prevention, through improvements in public health, and education. When Registration and subsidizing of ERT for each kind of LSD is done in MOH; ERT would be provided for patients in special centers. Iranian MOH subsidized 90% of the price of ERT for LSD. But remaining 10 % should be paid by patients. According to The minimum wage of Iranian which is 248 $ per month; most of patients could not afford to pay it, and they discontinued their treatment. During 3.5 years YASNA charity Foundation could covered 50% of price of ERT for 60 MPS type 1 patients; and 70% of it for 45 Pompe patients whose are living in territory of IRAN(for Iranian or resident in Iran). YASNA’s missions are providing access to medical care, and afforded for them regular counseling and training meeting. Genetic counseling, Genetic testing and sometimes prenatal diagnosis. For example genetic testing is done for 12 / 160 patients of Iranian Cystinosis. For countries with situation like IRAN, high rate of consanguinity marriage leads to high rate of rare diseases; and The minimum wage is less than price of ERT per month. The best solution for helping Patients and families is collaboration between pharmaceutical companies, MOH, and charity foundation. MAHAK foundation & YASNA foundation are 2 good examples of such collaboration in Iran.
**2297T**

Nutrient Intake and Growth Patterns among Adult and Pediatric Subjects with Urea Cycle Disorders (UCDs) Participating in Glycerol Phenylbutyrate (GPB) Clinical Trials. D. Hook1, G.A. Diaz2, B. Lee3, J. Bartley4, N. Longo5, W. Berquist6, C. Le Mons7, I. Rudolph-Angelich, M. Porter8, B.F. Scharschmidt9, M. Mokhtarani10. 1) Miller Children’s Hospital , Long Beach, CA; 2) Icahn School of Medicine at Mt Sinai, New York, NY; 3) Baylor College of Medicine, Houston, TX; 4) University of Utah, Salt Lake City, UT; 5) Stanford University, Palo Alto, CA; 6) National Urea Cycle Disorders Foundation, Pasadena, CA; 7) Hyperion Therapeutics, Brisbane, CA.

**Objective:** To analyze protein and caloric intake among adult and pediatric UCD patients as well as growth in pediatric patients participating in the GPB (HPN-100, RAVICTI®) clinical trials. Methods: Forty-four adult patients completed a 28 day, randomized, double-blind, placebo-controlled crossover comparison of 24-hr ammonia exposure on GPB vs. sodium phenylbutyrate (NaPBA). Weekly 3-day diet histories were collected. Prescribed and actual protein and caloric intakes were compared to those for healthy individuals (HI) derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC). Forty-nine pediatric patients ages 2-mo to 17-yrs received GPB for 12 months at a dose equivalent to their previously prescribed NaPBA dose. Prescribed protein, calorie, height, weight and BMI were assessed at enrollment (baseline) and throughout the 12 month open-label study. Results: In adult patients, mean protein intake was 53-56% of that for HI, greater than UCD-REC, and accounted for 10% of caloric intake vs. 15% for HI. Total caloric intake was ~70% of that for HI and 20-25% less than UCD-REC. There was a positive correlation between blood urea nitrogen (BUN) and total protein intake (r=0.65, p=0.0003). In pediatric patients, mean prescribed protein and caloric intake were 40% to 64% and 92% to 100% of HI and 54% to 97% and 68% to 87% of UCD-REC, respectively. Mean height, weight and BMI z-scores were within 2 standard deviations of normal population at baseline and did not change during the study. BMI correlated negatively with prescribed protein or calories expressed as g/kg/day or kcal/kg/day, respectively, in both adults (protein: r=-0.641; p=0.007; calories: r=-0.798 p<0.001) and pediatrics (protein: -0.329; p=0.022; calories: r=-0.362 p=0.024), but positively and weakly when expressed as total daily protein (g/day) or calories (kcal/day) in both populations. After 12 months of GPB dosing, the percent-of-tight ammonia control (<0.5 ULN) deserves further study as an approach to improving long-term outcome in UCD patients.

**Conclusions:** Both pediatric and adult UCD patients derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC). Forty-nine adults (HI) derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC) participated in the GPB (HPN-100, RAVICTI®) clinical trials. Methods: Forty-four adult patients completed a 28 day, randomized, double-blind, placebo-controlled crossover comparison of 24-hr ammonia exposure on GPB vs. sodium phenylbutyrate (NaPBA). Weekly 3-day diet histories were collected. Prescribed and actual protein and caloric intakes were compared to those for healthy individuals (HI) derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC). Forty-nine pediatric patients ages 2-mo to 17-yrs received GPB for 12 months at a dose equivalent to their previously prescribed NaPBA dose. Prescribed protein, calorie, height, weight and BMI were assessed at enrollment (baseline) and throughout the 12 month open-label study. Results: In adult patients, mean protein intake was 53-56% of that for HI, greater than UCD-REC, and accounted for 10% of caloric intake vs. 15% for HI. Total caloric intake was ~70% of that for HI and 20-25% less than UCD-REC. There was a positive correlation between blood urea nitrogen (BUN) and total protein intake (r=0.65, p=0.0003). In pediatric patients, mean prescribed protein and caloric intake were 40% to 64% and 92% to 100% of HI and 54% to 97% and 68% to 87% of UCD-REC, respectively. Mean height, weight and BMI z-scores were within 2 standard deviations of normal population at baseline and did not change during the study. BMI correlated negatively with prescribed protein or calories expressed as g/kg/day or kcal/kg/day, respectively, in both adults (protein: r=-0.641; p=0.007; calories: r=-0.798 p<0.001) and pediatrics (protein: -0.329; p=0.022; calories: r=-0.362 p=0.024), but positively and weakly when expressed as total daily protein (g/day) or calories (kcal/day) in both populations. After 12 months of GPB dosing, the percent-of-tight ammonia control (<0.5 ULN) deserves further study as an approach to improving long-term outcome in UCD patients.

**Conclusions:** Both pediatric and adult UCD patients derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC). Forty-nine adults (HI) derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC) participated in the GPB (HPN-100, RAVICTI®) clinical trials. Methods: Forty-four adult patients completed a 28 day, randomized, double-blind, placebo-controlled crossover comparison of 24-hr ammonia exposure on GPB vs. sodium phenylbutyrate (NaPBA). Weekly 3-day diet histories were collected. Prescribed and actual protein and caloric intakes were compared to those for healthy individuals (HI) derived from the National Health and Nutrition Examination Survey and current recommendations for UCD patients (UCD-REC). Forty-nine pediatric patients ages 2-mo to 17-yrs received GPB for 12 months at a dose equivalent to their previously prescribed NaPBA dose. Prescribed protein, calorie, height, weight and BMI were assessed at enrollment (baseline) and throughout the 12 month open-label study. Results: In adult patients, mean protein intake was 53-56% of that for HI, greater than UCD-REC, and accounted for 10% of caloric intake vs. 15% for HI. Total caloric intake was ~70% of that for HI and 20-25% less than UCD-REC. There was a positive correlation between blood urea nitrogen (BUN) and total protein intake (r=0.65, p=0.0003). In pediatric patients, mean prescribed protein and caloric intake were 40% to 64% and 92% to 100% of HI and 54% to 97% and 68% to 87% of UCD-REC, respectively. Mean height, weight and BMI z-scores were within 2 standard deviations of normal population at baseline and did not change during the study. BMI correlated negatively with prescribed protein or calories expressed as g/kg/day or kcal/kg/day, respectively, in both adults (protein: r=-0.641; p=0.007; calories: r=-0.798 p<0.001) and pediatrics (protein: -0.329; p=0.022; calories: r=-0.362 p=0.024), but positively and weakly when expressed as total daily protein (g/day) or calories (kcal/day) in both populations. After 12 months of GPB dosing, the percent-of-tight ammonia control (<0.5 ULN) deserves further study as an approach to improving long-term outcome in UCD patients.
2299T
Gene trap of NDUFS4 in mouse is a viable model of mitochondrial complex I deficiency.

B. Graham1, T. Dongii2, Y. Li3, M. Ge2, S. Ather2, R. Lucas1, R. Pautert1, J. Noebels4, X. Weihrans4, W. Craigert5
1) Dept Molecular & Human Genetics, Baylor College of Medicine, Houston, TX;
2) Dept Pediatrics, Baylor College of Medicine, Houston, TX;
3) Dept Molecular Physiology & Biophysics, Baylor College of Medicine, Houston, TX; 4) Dept Neurology, Baylor College of Medicine, Houston, TX;
5) Dept Neurology, University of Maryland School of Medicine, Baltimore, MD

Mitochondria are essential for many fundamental cellular processes, including energy production, ROS metabolism, Ca2+ homeostasis, and apoptosis. Dysfunctional mitochondria frequently result in multigland disease that established in neurological and cardiac defects. Mutant mice exhibit complex I deficiency and Leigh-like neurological phenotype in humans that typically result in death within 3-16 months of birth. Heart dysfunction is also a common phenotype in primary mitochondrial disease, typically presenting in the context of multi-systemic disease and can manifest as a cardiomyopathy and/or as cardiac arrhythmias. From the International Gene Trap Consortium, mouse embryonic stem (ES) cells that contain a gene trap insertion in Ndufs4 were obtained and used to generate transgenic animals. Homozygous mice showed reduced levels of NDUFS4 protein in various tissues by Western blot. While Ndufs4 mutants exhibit partial embryonic lethality, surviving adult mutants are able to survive for at least eighteen months. Ndufs4 mutant animals exhibit partial complex I deficiency, increased oxidative stress in brain and heart, and defects in cellular and mitochondrial respiration. They also exhibit neuronal dysfunction manifested by increased oxidative stress in brain and heart, and defects in cellular and mitochondrial respiration.

2300M
A Molecular Genetic Study of Indian Patients with Glycogen Storage Disorders reveals Six Novel Mutations

M. Bhat1,2, R.K. Sanjeev3, S. Jagdeo2
1) Centre for Human Genetics, Bangalore, India; 2) Indira Gandhi Institute of Child Health, Bangalore, India; 3) Fetal Care Research Foundation, Chennai, India; 4) Base Hospital, Delhi, India

Glycogen Storage Disorders (GSDs) are a set of autosomal recessive conditions caused by mutations in genes that govern different aspects of the glycogen metabolic pathway. Depending on the enzyme that is deficient, the glycogen metabolic pathway is altered, resulting in abnormal glycogen metabolism and accumulation in various tissues. GSDs are typically classified into three categories based on the glycogen metabolic pathway: Type I (GSD-I), Type II (GSD-II), and Type III (GSD-III). While Type I and Type III GSDs are relatively well understood, Type II GSDs are less well characterized.

GSD-I (type I) is characterized by the enzyme deficiency of G6PC, which catalyzes the conversion of glucose-6-phosphate to glucose-6-phosphatase. Patients with GSD-I often suffer from hepatic and muscular symptoms, such as hypoglycemia and myoglobinuria.

GSD-II (type II) is caused by mutations in the gene encoding the enzyme phosphorylase b kinase (PBPK). Patients with GSD-II may present with hypoglycemia, muscle weakness, and fatigue.

GSD-III (type III) is caused by mutations in the gene encoding the enzyme UDP-glucose 6-dehydrogenase (UGDPH). Patients with GSD-III may present with hypoglycemia, muscle weakness, and fatigue.

Our study aimed to identify novel mutations in the GSD-I, GSD-II, and GSD-III genes in Indian patients with GSDs. A total of 66 patients were recruited from five different hospitals in India. The patients were screened for mutations in the GSD-I, GSD-II, and GSD-III genes using Sanger sequencing. We identified six novel mutations in the GSD genes, including two novel missense mutations in G6PC, one novel frameshift mutation in PBPK, and three novel missense mutations in UGDPH.

These findings contribute to the understanding of GSD genetics and may aid in the diagnosis and management of GSD patients in India.
2303S
The UDP Self Study Short Course for Genome Wide Analysis. T.C. Markello1,2, C.F. Boerkoel1,2, W.A. Gahl2, E. Valkanis1,2,3, E.D. Flynn2, W.P. Bone2, A.E. Links2, C.J. Markello2, P.J. Pemberton2, D.R. Adams1,2
NIH Intramural Sequencing Center, NIH Genomic core, and UDP Clinical Team. 1) Med Gen Branch, NIH/NHGRI, Bethesda, MD; 2) NIH Undiagnosed Diseases Program, Office of the Director Common Fund and NHGRI, NIH, Bethesda, MD; 3) Clinical Director, NHGRI/NIH, Bethesda, MD.
Since 2008 the NIH Undiagnosed Diseases Program has found many types of Mendelian and non-Mendelian genetic and molecular mechanisms for human diseases in patients referred to the program. These include: Homozygous recessive cases of identity by descent with parental consanguinity, compound heterozygous recessive alleles including double point mutation pairs, point mutation/deletion pairs and double deletion null allele pairs, unique de novo dominant new mutations, de novo dominant with transmission to the third generation mutations, mosaicism including both fixed ratio and progressive telomere shortening, duplications, and a form of variant discovery we call extreme novel findings. These discoveries all occurred in single nuclear families with one exception and typically with only one or two probands. Many are CLIA certified and some are published, with the remainder in various stages of cell biological validation. To teach the techniques that systematically find these types of variants in single families, we have taken 20 such cases and generated a set of self-study exercises, along with a set of 4 narrated power point presentations, and a total of 150 pages of exercise and reference manuals. These cases include annotated VarSifter (and/or .vcf) files, bam files, take up approximately 800GB of storage, but to run these files currently requires a Windows based 64bit machine with a minimum of 64GB of RAM. We are currently anonymizing the cases and will make them available to the community. As the data is collected and presented, we will set up educational modules to help us define the most informative and comprehensive way of communicating about this subject. The module consists of: 1. refreshment of genetics through polling 2. case studies 3. genetic gift card 4. practical (bitter tasting and genetic variants) 5. future perspectives. A total 210 GPs participated in the 195-minute genetic training session and afterwards an evaluation form was completed by each participant. There was a difference in appreciating the educational module between the ‘old’ and the new generation GPs; the overall rating was 7.9 (scale from 1-10). The majority of the participants was especially interested in Pharmacogenomics (drugs/genes) and pedigree based clinical decision making. And, of great importance, the training low-versed the threshold to control the genetic disability and paved the way to a new generation of geneticists.

2305S
“Hey Doc, here is my DNA sequence, what do you think?” M. Kriek1,2, H. Smia1, C.E. van der Meer1, I. Arnold2, E. Aten1, M. Koopmans1, A. van Haeringen1, M.H. Braeunling1, T. Vrijenhoek2, J.T. den Dunnen1,2,3,4,5,6,7,8,9
1) Clinical Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 3) Workgroup Promotion of expertise General Practitioners Rijndijk, Netherlands; 4) Center for Genome Diagnostics, Department of Medical Genetics, University Medical Center Utrecht.
When my General Practitioner (GP) fell from his chair when I opened my laptop to check whether the drug prescribed matched with my genome, I realized that there was a problem with my DNA-based medicine. Although DNA-based medicine is at the door-step, GPs are not trained in this direction and therefore they lack the knowledge to correctly interpret complex clinical data. To bridge this gap, we developed an unique educational project that prepares physicians for their role in the near future: they were trained to translate genome-wide analysis data to the patient’s specific clinical situation. The educational module was set up in close collaboration with members of the target group (e.g. GPs) to help us define the most informative and comprehensive way of communicating about this subject. The module consisted of: 1. refreshment of genetics through polling 2. case studies 3. genetic gift card 4. practical (bitter tasting and genetic variants) 5. future perspectives. A total 210 GPs participated in the 195-minute genetic training session and afterwards an evaluation form was completed by each participant. There was a difference in appreciation of the educational module between the ‘old’ and the new generation GPs; the overall rating was 7.9 (scale from 1-10). The majority of the participants was especially interested in Pharmacogenomics (drugs/genes) and pedigree based clinical decision making. And, of great importance, the training low-versed the threshold to control the genetic disability and paved the way to a new generation of geneticists.

2303S
Effectiveness of Personalized Genetic Education Modules. A. Jenks, S. Darabi, J. Eggert. Clemson University, Clemson, SC., USA.
Effectiveness of Personalized Genetic Education Modules Applied genetic therapies and personalized medicine urge the necessity of genetic education for clinicians. Specifically there is a need for education about how to utilize available genetic technologies for practice and how to translate available genetic information into specialty practice. In 2012 Powell et al. reported that only %15 of clinicians feel knowledgeable enough to utilize genetic medicine in their clinic. If genetic medicine is to converge with personalized medicine, then doctors in specialty practice need to be equipped to better utilize and interpret genetic technologies. The purpose of this study was to determine the effectiveness of specialty department personalized genetic education modules for clinicians in hospital settings using pre and post-test measures. For the project, two researchers provided independent presentations on specially focused genetics in the clinic of two different departments a large community hospital in the Southeast region of the United States. Before the presentation, both researchers met with department staff to discuss need, observed the assigned departments and developed the presentation content based on identified needs. Pre and post-tests were assigned based on the presentations and given to the staff in each department to determine the effectiveness of the presentation to increase clinicians’ understanding of genetics and attitudes toward genetic medicine. Overall the researchers found little increase (%9,61) in genetic knowledge. This research is a baseline study to determine the effectiveness of specialty department based personalized education modules for clinicians. Results from this project indicate that healthcare professionals are already being taught the theory of genetic medicine but that application based modules are necessary for clinicians to be open to the importance of personalized genetic medicine to their specialty practice. Reference Powell, K. P., Cogsett, E.D. Flynn, E. Valkanis, E. Aten, M. Koopmans, A. van Haeringen, M.H. Braeunling, T. Vrijenhoek, J.T. den Dunnen, A. van Haeringen, M.H. Braeunling, T. Vrijenhoek, J.T. den Dunnen, M. Kriek, H. Smia, C.E. van der Meer, I. Arnold, E. Aten, M. Koopmans, A. van Haeringen, M.H. Braeunling, T. Vrijenhoek, J.T. den Dunnen. (2012) Clinical Genetics, Leiden University Medical Center, Leiden, Netherlands; 2) Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 3) Workgroup Promotion of expertise General Practitioners Rijndijk, Netherlands; 4) Center for Genome Diagnostics, Department of Medical Genetics, University Medical Center Utrecht.

2306S
Teaching Genomic Medicine to Physicians - this is our responsibility as medical geneticists! I. Maya1,2, L. Basel-Vanagaite1,3,4,5, E. Taub6, A. Kollman7,8,9, D.M. Behar2, R. Tomashov-Matar1, R. Sukenik-Halevi2, A. Reches2, M. Hubshman Weiss1,3,4, N. Orenstein5,6,7,8, D. Marom10, M. Sho- hat1,2,3,4,5,6,7,8,9,10
1) Rabin Medical Center, Petah Tikva, Israel; 2) The Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; 3) Pediatric Genetics, Schneider Children’s Medical Center of Israel, Petah Tikva, Israel; 4) The Raphael Recanati Genetic Institute and Felsenstein Medical Research Center, Rabin Medical Center, Petah Tikva, Israel; 5) Institute of Human Genetics, Soroka University Medical Center, Beerseheba, Israel; 6) Ben-Gurion University of the Negev, Beersheba, Israel; 7) Rambam Medical Center, Haifa, Israel; 8) Meir Medical Center, Kfar Saba, Israel; 9) Souraski Medical Center, Tel Aviv, Israel; 10) Pediatrics A, Schneider Children s Medical Center of Israel, Petah Tikva, Israel.
Background: Due to the transition of genetic knowledge from research laboratories into clinical practice, the new field of “Genomic Medicine” has emerged. Primary care practitioners are facing this era with inadequate knowledge and skills in medical genetics and many are unaware of the technical, ethical, legal and psychosocial implications of genetic testing. The challenge is to develop strategies that facilitate the translation of genomic medicine into clinical practice that require novel approaches to educate existing and future clinical physicians from all medical fields. Methods: We initiated a “genomic education” program for the purpose of providing physicians from different medical fields and phases of their medical carrier an advanced knowledge in genomic medicine. In the last 2 years since we initiated the program we are organizing 4 courses a year, both in Rabin Medical Center and in Tel Aviv University. We emphasized the main take-home messages for physicians, defined as: 1) need and capability to evaluate various genetic diseases, recognition of the mode of inheritance from pedigrees, guidelines for decision-making on which technique to use, interpretation of test results and their clinical implications. Results: To date, 111 physicians have participated in 7 courses of our “genomic education” program, which included lectures, workshops and guided tours in genetic laboratories. The physicians that participated in our program came from 14 different hospitals and 2 HMO in Israel. Two thirds where senior physicians defined as specialists for more than 10 years, and 1/3 were junior physicians (both residents, fellows and young specialists). In the “post-course” examination the average score was 56%, the median was 55% (range 20%-80%), whereas in the “post-course” examination it was 78%, Median 80% (range 40%-100%). The average improvement in score as a result of the course, both for junior and senior physicians was 20% (p-value 0.001). The physician in our program reported a very high level of satisfaction from the theoretical and practical knowledge they acquired as well as the concept of a “one-week update course.” Conclusion: A one-week “genomic education” program is an effective strategy to update physicians regarding the advances in Genomic Medicine in order to improve their care of patients.
Medical Student Confidence and Knowledge about Hereditary Retinoblastoma. J. Govindavani, A. Nath, W. Blazey, D. McMahon, T. Chan, D. Tegay, B. Krishnamachari. NYIT College of Osteopathic Medicine, Old Westbury, NY.

Background: In the pediatric population, hereditary cancer predispositions are associated with ~5-10% of all cancer diagnoses. Early identification and treatment of hereditary cancer syndromes often significantly improves likelihood of survival. The level of physician knowledge regarding pediatric hereditary cancer syndromes is currently unknown. Assessing pediatric hereditary cancer knowledge amongst physicians and trainees, including medical students, is crucial. The goal of this study was to assess perceived confidence in knowledge, and actual knowledge, of pediatric hereditary cancer syndromes in medical students, with a focus on retinoblastoma (RB).

Methods: Medical students answered questions regarding confidence in knowledge about hereditary cancer and knowledge-based question derived from national clinical guidelines and recently published literature.

239 students completed the survey. Results: When analyzed by year in school, the four years differed in confidence of their knowledge of hereditary cancer (p<0.02), with the highest percentage of students feeling confident in year 2 (42.0%). However, there was no statistically significant difference between class years in terms of answering knowledge questions correctly. Students who felt confident in their knowledge of hereditary cancer were more than 2 times more likely to know that the risk for extra-ocular primary neoplasms is increased in individuals with hereditary RB [OR 2.30 (1.08, 4.91)] and were more than 3 times more likely to know “that genetic screening of the RB1 gene should be considered in any child with bilateral RB” [OR 3.42(1.11, 10.54)], after adjusting for age and years in school. Conclusion: Confidence was associated with knowledge, but knowledge was not associated with increased years of schooling. Students felt most confident in their second year, and less so in their third and fourth years. The increased confidence during year 2 may have due to students preparing to take their initial board exam. More research is needed to understand the cause of the differences in confidence and knowledge between the four years of medical school classes with extension to assessment of knowledge and confidence in medical residents and attending physicians.

Preparation of future physicians to practice genomic medicine: Lessons from the first two years of the University of Miami Master of Science in Genomic Medicine program. W.K. Scott1,2, S.E. Hahn1,2, S.C. Sacharow1,2, M. Tekin1,2, E. Rampersaud1,2, B. Johnson1, G. Ghaffari1, E.M. Bendik1,2, J.M. Vance1,2, 1) Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL; 2) John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL.

Needs Assessment: One barrier to the implementation of genomic medicine in the clinic is the concern among physicians that they do not have the knowledge needed to use genomic information in patient care. Lowering the barrier requires training at three levels: medical school, residency, and continuing education. While medical schools should be the leaders in this educational effort, curricula are saturated with topics required for subsequent board exams.

Rationale: To address this barrier in medical school, we have created a 4-year Master of Science program in Genomic Medicine (MSGM) that is completed concurrently with the medical curriculum. Students apply to the program in the first semester of medical school and instruction begins in the second semester.

Course Content: MSGM topics were selected by faculty with expertise in genomic medicine, genetic counseling, clinical human genetics, and laboratory science. Topics include: ethical, legal, and social issues in genomic medicine; computational methods; bioinformatics; clinical laboratory methods and technology; pharmacogenetics, and research ethics. The first 3 semesters also contain courses in clinical applications of genomic medicine that examine cases from physician systems studied in the medical curriculum.

Teaching Strategies: A flipped classroom model is used the first 2 years. Didactic lectures, small group discussions, and interactive modules are provided online for self-directed learning. Lessons are reinforced by weekly small-group discussions of core topics and examples from medical literature. The final 2 years include a 4-week clinical genetics clerkship and a practicum consisting of clinical experiences performed with a faculty mentor.

Outcomes: In total 17 students (5 in the first class, 12 in the second) have enrolled in the MD/MSGM program. All have continued to be successful in the medical curriculum. All reported satisfaction with the current structure and curriculum and not as a path to research careers. We believe this is a viable approach to prepare future physicians to incorporate genomic medicine into their daily clinical practice.

As an increasing number of personalized medicine programs have begun to pre-emptively genotype pharmaco genetic variants, the utilization of clinical pharmaco genetic testing is escalating. The recent publication of clinical practice guidelines has facilitated this process, but concern over the lack of expertise with these tests in the curriculum and in training programs is increasing. We surveyed all third-year residents at Mount Sinai, New York City, to determine the extent of interest in pre-emptive genotyping, the number of students who would support such a program, and the number of students who would support it if the testing could be performed at a reduced cost. The survey was conducted online in March 2014. The final survey was sent to 773 Mount Sinai residents, and 257 residents responded. The survey collected information about personal and family history, as well as demographic information. The results of the survey were analyzed using descriptive statistics.

The results of the survey showed that the majority of residents would support the implementation of a pre-emptive genotyping program at Mount Sinai. However, the majority of residents would not support a program if the testing cost was not reduced.

...)
2314S
Attitudes and Response to Genetic Risk Information Among Adolescents. C. Bloss1, B. Darst1, B. Daley2. 1) Scripps Genomic Medicine, Scripps Translational Science Institute, La Jolla, CA; 2) High Tech High, San Diego, CA.

Genome sequencing with pediatric populations is currently being performed, however there is little data on the impact of genetic risk information on younger individuals. Furthermore, most work in this area is focused on specific clinical populations, and the voices of minors themselves are often overlooked with most studies focused on parents. The current, school-based study assessed attitudes and response to genetic risk information in a group of tenth graders before and after an educational seminar on genomics and personalized medicine. Students recruited from a local group of San Diego charter schools attended an educational seminar on genomics and genomics and then received information about a test of genetic risk for complex disease and drug response traits within the context of an analogue study design. Specifically, each participant was randomly assigned to one of four conditions, which varied by the number of diseases or drug response traits that showed a high risk result. Students completed demographic, attitudinal, and behavioral assessments before and after the seminar and exposure to the analogue vignette. 140 students across three biology classrooms were offered participation. Of those, 80 students enrolled and completed both the baseline and follow-up assessments. All participants were in the 10th grade and ranged from 15 to 17 years of age. There were no significant differences in emotional response between students in the four genetic risk conditions. Analyses of follow-up data also showed significantly lower anxiety towards genome testing than the already low levels reported at baseline before the seminar. Further, after attending the genomics seminar, students' knowledge of genomics increased significantly and a large majority of students said that they would consider undergoing genetic testing if it seemed warranted. Our findings suggest a lack of adverse emotional responses among adolescents exposed to genetic risk information. While analogue methods cannot exactly replicate reality, previous studies suggest that responses to risk information observed in such studies are similar to those seen in actual clinical situations. Although studies with larger samples are needed, these results also highlight the utility of education programs for minors to increase genomics literacy. The findings contribute to understanding the relationship between sequence, structure, and function.

2315S
Teaching the relationship between genotype and phenotype with public data and Molecule World™ on the iPad. TM. Smith, SG. Porter. Digital World Biology LLC, Seattle, WA.

It is impossible to grasp fundamental concepts in biology without understanding the relationship between sequence, structure, and function. Modern data collection technologies are creating enormous data resources that can be used to help students understand these relationships, but currently are underutilized. This is due to the fact that easy-to-use tools that meet teachers' needs for clear instruction are not yet commonplace. Filling the gap between the embarrassment of data riches and practical classroom use requires three things: user-friendly tools, content that demonstrates specific applications with interesting stories, and packages that combine instruction, assessment, and inquiry-based investigations.

Digital World Biology is addressing this need with its on-line courses and mobile apps. The on-line courses increase students' computer literacy while using standard tools like Cn3D, Blast, ORF finder, and multiple databases, in directed and exploratory ways, and help students better understand biology as well gain a better appreciation for the value of the data and the field of bioinformatics. In response to nearly two hundred interviews with K-12 and college teachers and students, we created Molecule World and the Molecule World DNA Binding Lab™ iPad apps to display 3D-data from multiple structure databases (MMDB, PDB, and PubChem) using a novel rendering engine that allows us to uniquely highlight chemical properties and sequence orientation. The ability to display and highlight sequences within molecular complexes enables exploration into the relationships between structure, function, and evolution. This activity is used in new genetics courses, integrated biolabs, in professional development workshops, and many demonstrations, supports the hypothesis the being able to view and simultaneously interact with data improves teaching capabilities and student engagement. This work has been supported by NSF grant IIP 1315428.

2316S
Teaching the genome generation, a laboratory intensive high school genomics and ethics course. C. Wray1, M. McKennan1, D. Waring2. 1) Genomic Education, The Jackson Laboratory, Bar Harbor, ME; 2) Personal Genetics Education Project, Harvard Medical School, Boston MA.

To increase genetic and genomic literacy and train the next generation of scientists, we have developed, tested, and assessed an intensive Genomic Biology course, entitled Infinite Variations, Personalized Medicine and Genomics. The course is multi-faceted and includes bench molecular biology experiments, bioinformatics exercises and focused bio-ethics discussions. The goals of the course are four-fold. First the course introduces the complexity of human genetic and genomic variation and stresses the importance of individual variants. Rather than discussing mutant and wild-type alleles, the course highlights the fact that every genome is rife with single nucleotide polymorphisms (SNP), insertion-deletions (indels), and structural or copy number variants (CNV). Second, at the hands-on skill level, the course instructs students in the use of molecular assays used to detect genetic variation. Four experimental modules and a panel of de-identified human genomic DNA samples are used to demonstrate and assay for SNPs, indels, and CNVs; simple PCR-restriction digest assays are taught and compared to higher throughput molecular beacon PCR techniques and traditional sequencing. The third goal of the course is to introduce students to Bioinformatics and the use of public databases to learn about genetic diseases and genomic variation in humans. The fourth, final and vital goal of the course is challenge students with relevant ethical dilemmas that anchor genomic biology within real-world situations. Students are engaged in discussions covering genetic privacy, genetic pre-determinism, prenatal genetic testing, and biomedical research ethics. Over the most recent three year period a pre-course and post-course assessment strategy has been used to measure learning outcomes. In all three years the student participants have shown significant gains in genomic and genetic literacy and knowledge of bio-ethics; on average, 8 correct answers on the post-course assessment increase by 24%. The course is currently being re-designed as a teacher professional development exercise in order to increase dissemination across New England high schools.

2317S

Nowadays, genetic research, and the application of these studies to everyday life, is more attractive to the public, especially as it can be used to find cures for diseases. Also as scientists, it is important to increase the interest for science in new generations to ensure future advancements of the field. The Spanish School of Washington DC promotes the study and practice of the Spanish language for children born in the United States, but with parents of Spanish origin. This education takes place through extracurricular classes and cultural activities, conducted in Spanish, outside of their English-speaking schools. One such activity was a visit to “Genome: unlocking life’s code,” an exhibit at the National Museum of Natural History. Children were divided into three groups of 12, each group ranging in ages from 7 to 15 years old. The explanation in Spanish started with the basic concepts of DNA, which focused on human genetics; from how migrations are marked in our mitochondrial DNA, to how the advances in genetics let us diagnose diseases and look for new drugs and treatments. Although some of the concepts were difficult to explain to some of the younger students, they appreciated the opportunity to take examples from real life and discussing their genotype with our phenotype. To illustrate the concept of DNA, at the end of the lesson the students were given the opportunity to work as scientists. We performed an experiment that consisted in extracting DNA from strawberries and bananas through a protocol that used common household items. Both procedures were used in this activity and at the end of many of the experiments the students were successful in the extraction of DNA. The students and parents reacted positively to the lesson and especially the experiment: they kept the protocol and many expressed their desire to try it again at home. In summary we can say that this two year old project has helped to promote science in new generations, and make young students familiar with scientific concepts in Spanish, increasing their knowledge of this language.
Accelerating public awareness of personal genetics. D. Waring1, M. Gelbart2. 1) Depart of Genetics, Harvard Medical School; Personal Genetics Education Project Boston, MA; 2) Marnie Gelbart Harvard Medical School Department of Genetics Personal Genetics Education Project Boston, MA.

The Personal Genetics Education Project (pgEd.org) is using a unique and expanding platform of strategies to address the growing gap between what researchers are learning at the frontiers of genetics and what the public understands. Here, we focus on our most recent accomplishments in the domains of policy, entertainment, gaming, and education. First, in 2014, pgEd conducted a Congressional briefing to draw policymakers’ attention to scientific advances in personal genetics and the importance of raising public awareness. Second, pgEd is expanding its emphasis on advancing awareness through television and film through partnerships with Hollywood, Health & Society and the Science & Entertainment Exchange. Third, pgEd is continuing to develop the on-line game, Map-Ed (map-ed.org) for spreading awareness. Our 2014 expansion of Map-Ed’s topics include the microbiome and genetic discrimination. To date, Map-Ed has 5,400 players on all 7 continents. Fourth, pgEd is building our up-to-date and freely available online curriculum that addresses the benefits and implications of personal genetics. Finally, pgEd’s efforts training teachers to engage young people on the interdisciplinary topic of personal genetics are extending beyond the biology classroom and into health, social studies, and literature. pgEd is continuing to set new goals, including engaging faith-based communities, and is seeking partners and collaborators to expand our vision and reach.

Celebrity disclosures and information seeking: The case of Angelina Jolie. R. Juthe1, A. Zaharchuk2, C. Wang2. 1) Office of Communications and Education, National Cancer Institute, Rockville, MD; 2) iDoxSolutions, Inc., Bethesda, MD; 3) Department of Community Health Sciences, Boston University School of Public Health, Boston, MA.

Background: On May 14, 2013, actress Angelina Jolie disclosed that she had a deleterious BRCA1 mutation and underwent a preventive bilateral mastectomy to reduce her inherited cancer risk. Her disclosure generated media frenzy and much discussion about her potential “effect.” We sought to document the impact of her disclosure on information-seeking behavior, specifically regarding online cancer genetics and risk reduction resources available from the National Cancer Institute (NCI). Methods: Using Adobe Analytics, daily page views for 11 online resources, including resources written for the public and health professionals, were tracked over a 9-week period from April 23, 2013 through June 23, 2013. Online usage data were also obtained for 4 resources over a 2-year period (2012-2013). Source of referral by which viewers located a specific resource was also examined. Results: Jolie’s disclosure led to a dramatic and immediate increase in traffic to NCI’s online resources. The Preventive Mastectomy fact sheet written for the public saw the largest increase with 69,225 page views on May 14, representing a 795-fold increase compared with the previous Tuesday. The Cancer Genetics Services Directory for the public saw a 31-fold increase in page views from May 7 to May 14. A fivefold increase in page views was observed for the health professional Physician Data Query (PDQ) Genetics of Breast and Ovarian Cancer summary in the same timeframe. Health professional PDQ summaries on the genetics of other inherited cancers, including skin cancer and colorectal cancer, received increases of up to 65% from May 7 to May 14. Two-year data demonstrated the magnitude and longevity of Jolie’s effect. A substantial increase from 0% to 49% was seen in referrals from news outlets to NCI’s breast cancer resources from May 7 to May 14. Conclusions: Celebrity disclosures, such as Jolie’s disclosure of her BRCA1 mutation status and her risk management decision, can dramatically influence online information-seeking behaviors. Efforts to capitalize on these disclosures to ensure easy access to accurate information are warranted.

REU Site: Research Experiences for Underrepresented Minority Undergraduates in Basic Science and Genetic Research at Louisiana State University Health Sciences Center (LSUHS). K. Foley, A. Umrigar, A. Musto, H. Farris, F. Tsien. LSUHS, New Orleans, LA.

The Louisiana State University Health Sciences Center (LSUHS) in New Orleans provides research training for a period of ten weeks for up to ten Louisiana underrepresented minority (URM) undergraduates interested in exploring graduate school and biological research careers, particularly in genetics. In the New Orleans area, Blacks account for 60.1% of the population, Hispanics 5.0%, and Native Americans 0.3%, compared to 12.6%, 16.3%, and 0.9%, respectively in the US. The field of genetics is constantly evolving and it is therefore imperative for students who are pursuing biological science careers to be proficient in hypothesis-driven scientific problem solving, current methodology, data analysis, and communication skills. The objectives of the proposed REU program in basic biology research include: 1) Scientific problem solving 2) Learning state-of-the-art laboratory techniques 3) Developing scientific communication skills including teaching at a genetics workshop to high school students 4) Acquiring resume and CV writing and career education resources 5) Participate in a number of seminars and interactive workshops. Students present posters at an end of the summer symposium which are judged by local scientists. Additional seminars regarding graduate school professional development are presented by currently enrolled genetics graduate students and invited faculty. Students complete anonymous pre- and post-internship surveys designed to measure their self-assessed research interest, demographics, and achievement. The students will be followed up for a period of at least 3 years following the completion of the program to gauge the degree to which their experiences have had a lasting influence on their respective career paths. Students will also use an on-line REU assessment tool. We will stay in close contact with the students via a yearly electronic survey. Following the REU program, interns also have the opportunity to teach genetic concepts to New Orleans area schools during the academic year, an important skill that will enhance their future careers as scientists.
Gene hunting with IMG-ACT: Integrating genomics research into the undergraduate biology curriculum. K. Moitra. Dept of Biology, Trinity Washington University, Washington DC, DC.

Genome annotation is the process of linking biological information to gene sequences. It involves annotating predicted genes and connecting these genes to protein functions. The integration of genomics research into the undergraduate biology curriculum provides students with the opportunity to become familiar with state-of-the-art bioinformatics tools and answer original research questions. Our purpose with this research project was to upscale the research experience so that large numbers of students could have access to research. One way to achieve this is through integration of research with the classroom experience. Students in an introductory genomics class annotated predicted ABC (ATP Binding Cassette) genes of Methanothermobacter thermotofusus using the Integrated Microbial Genomics-Annotation Collaboration Toolkit (IMG-ACT) developed by the Joint Genome Institute. This research project is linked to the much larger scientific question: What can the sequence of an organism’s genome tell us about its overall biology? A variety of pre-course and post-course tests and surveys were conducted to assess if the project engaged students and generated interest in scientific research. Preliminary results suggest that it did. This session will discuss: initiation and overview of the project, integration of research into the biology curriculum, up-scaling research projects for involvement of larger numbers of students, student learning outcomes and data collected from student surveys/assessments, including how future career choices were impacted by this research project. The session will also highlight building and scaffolding research skills in a research-intensive genetics course and the integration of research into the curriculum.

Development of a molecular test and human pedigree analysis on Alpha-1 Antitrypsin Deficiency. A. Olson, D. Caporable. Biology, University of Wisconsin-Stevens Point, Stevens Point, WI.

Alpha-1 Antitrypsin Deficiency (Alpha-1) is a monogenic recessive disorder affecting roughly 1 in 5,000, making it 1 of 3 of the most common genetic diseases to cause death among adult Caucasians in the United States. The deficiency is caused by a mutation of the SERPINA1 gene, located on chromosome 14, which codes for the serpin peptidase inhibitor alpha-1 antitrypsin. Inheriting two mutated variants can result in Emphysema, COPD, and/or liver disease between ages 40-60. In one such case, an individual was recently diagnosed with this disorder and was found to be homozygous for the affecting variant “Z”, rendering each parent a carrier. This raised several concerns for other family members, who then wished to seek genetic testing. Current genetic tests include restriction digestion of amplicons, which can be less reliable or repeatable due to partial digestion issues. The objective of my undergraduate research was to design a novel genetic test for the SERPINA1 “Z” single-nucleotide polymorphism (SNP) based on DNA sequence comparisons - that is reliable, repeatable, noninvasive and relatively easy to perform - to facilitate the detection and diagnosis of Alpha-1 patients. Since the “Z” variant is a result of a transition missense mutation, whereas a guanine (found in the wild type form denoted as “M”) is replaced by an adenine, our methods included designing a forward and reverse primer pair that would successfully amplify this SNP region and generate DNA sequences representing an individual’s genotype that is easily interpreted. Eighteen members of the Alpha-1 affected family were used as controls to verify the reliability of this molecular test. Amplified products were cycle-sequenced by the fluorescent Sanger method and electrophoresed using an automated capillary system. After protocol optimization, homozygous wild types (MM), heterozygous carriers (MZ), and homozygous mutants (ZZ) became easily discernable among the respective electropherograms. Heterozygote electropherograms displayed equal representations (concentrations) of adenine and guanine at the polymorphic site. In addition, a genogram was constructed that successfully illustrated the congruence of their genotypes with their respective phenotypes of this disorder. We conclude that our molecular method for identifying M/Z genotypes associated with Alpha-1 Antitrypsin Deficiency is a reliable test that can be easily adapted for research or diagnostic purposes.
2326S
Sex and Gender Differences in Health: Educational and Collaborative Outreach to Genetics Researchers, Clinicians, and Students. M.R. Ten- nan†1, M.E. Edwards, H.F. Norton1, N. Schaefer*1. 1) Health Science Center Libraries, University of Florida, Gainesville, FL; 2) Health Science Center Libraries, University of Florida, Gainesville, FL.

Purpose: Historically, medical and scientific thinking has been based largely on male anatomy, disease presentation, and response to therapy. This focus presents ethical and scientific challenges. We embarked on an outreach project to promote equitable research sensitive to sex and gender differences in the basic and clinical sciences, including genetic research. Given this, we distributed gender-sensitized materials and information to researchers and students at our institution. Through National Library of Medicine/NIH Office of Research in Women’s Health funding, a team of librarians partnered with researchers and clinicians to promote awareness of the need for further research on sex and gender differences in health, facilitate collaboration among such practitioners and students, and make research in this area more accessible. Setting/Methods: A large land grant institution with contiguous main and health center campuses. “Collaborating with Strangers” (CoLAB) workshops were held to facilitate collaboration among participants interested in this discipline. Affiliates represented the gamut of health, life, and social sciences, as well as clinical and research realms. Instructional sessions related to the science of sex and gender differences in health were performed for a number of cohorts, including graduate and undergraduate students in genetics and genomics. In order to make sex and gender differences research more accessible within and external to the institution, an open access publication fund was created. Researcher records in the campus finding tool (VIVO) were augmented with publication and education records. Conclusions: In its first year the team hosted two CoLABs with 37 attendees; a third session has been held in Year 2. Evaluations suggest that participants connected with other researchers and gained potential collaborators. Instructional sessions for genetics students yielded positive feedback. In Fall 2021, gender differences researchers were included in the Institutional Repository (IR). In the second year of funding the team collaborated in the university’s outreach to high school students and teachers, catching students early in the pipeline when they contemplate careers in science and medicine. Results/Conclusions: In its first year the team hosted two CoLABs with 37 attendees; a third session has been held in Year 2. Evaluations suggest that participants connected with other researchers and gained potential collaborators. Instructional sessions for genetics students yielded positive feedback. In Fall 2021, gender differences researchers were included in the Institutional Repository (IR). In the second year of funding the team collaborated in the university’s outreach to high school students and teachers, catching students early in the pipeline when they contemplate careers in science and medicine.

2326S
Molecular Genetic Studies in Indian Patients with Ectodermal Dysplasia. s. kushmakar. pediatrics, All India Institute of Medical Science, New Delhi, India.

Introduction: Ectodermal Dysplasia (ED) is a group of about 170 heritable disorders that affect the ectoderm, the outer layer of tissue. Each of the roughly 170 ED syndromes represents a different combination of abnormalities. Inheritance pattern are variable include autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive. Here we report our experience with X-linked hypohidrotic or anhidrotic ectodermal dysplasia. TheEDA gene which causes this disease is present on Xq12 - q13 and comprises of 12 exons. Eight of which encode the transmembrane domain. Mutations have been reported from different groups, including missense, splicing, small deletions and insertions. Objectives: Our objectives were to identify the spectrum of mutations in Ectodermal Dysplasia patients and to plan a stratagy for molecular diagnosis so that optimal management, genetic counsel- ing, and prenatal diagnosis can be offered. Material and methods: Thirty (25male and 5 Female) patients with clinical features suggestive of ED were included in the study. In two families mothers also had mild features suggestive of ED. Four of the families were consanguineous. DNA was extracted from 30 patients and their parents. PCR amplification was carried out using 8 pairs of primers and sequencing was done . All the patients who were found negative for EDA gene sequence variation were subjected to screening of EDAR gene. Results: 14 patients out of 30 were reported with an EDAR variant and out of 14. 8 were reported and 6 patients were found to be a reported mutations, c.2 T>A (exon 1) , c.463 C>T (exon 3), c.466 C>T (exon 3), c.1045 G>A (exon 9) and c.820 T>A (exon 8) were found in one patient each. Five novel mutations, c.502+1 G>T, c.524+2 G>T, c.655+1 G>T, c.631+1 G>T, c.900 C>A were identified in six patient each. Two patients were found to have a splice site mutation. Protein prediction software, polyphen2/SIFT predicted that none of the novel changes has a pathogenic effect. In EDAR gene 10 patients were reported with splice site mutation, 4.8,12 and IVS11. Conclusion: We detected variation of EDA and EDAR genes with variable phenotypic of Ectodermal Dysplasia. This study should be useful for genetic counseling and prenatal diagnosis for the families enrolled in study and in expanding the database on EDA and EDAR gene.
2330M Patient Perspectives on the Use of Electronic Health Records for Research: The MI-GENES Study. R.A. Haddad1, H. Joun1, I.N. Isser1, J.B. McCormick2, C.G. Chute3, T.J. Kullo1. 1) Department of Internal Medicine, Mayo Clinic, Rochester, MN; 2) Division of Biomedical Ethics, Departments of Internal Medicine and Health Sciences Research, Mayo Clinic, Rochester, MN; 3) Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN.

Background: The use of electronic health records (EHRs) is increasing due to federal mandates. The myocardial infarction genes (MI-GENES) study, an eMERGE genomic medicine implementation pilot, is exploring the integration of genomic results into the EHR. The views of participants in genomic studies on the use of EHRs for research are not known. Methods: The MI-GENES study randomized adults aged 40-65 years, at intermediate risk for coronary heart disease (CHD) and not on statins, to receive either the 10-y risk of CHD based on conventional risk factors versus the 10-y risk based on a genetic risk score and conventional risk factors. We assessed participants' perspectives regarding the use of EHRs for clinical care and research using the health information national trends survey (HINTS). This survey was completed by all study participants at the initial study visit.

Results: The survey was completed by 212 study participants recruited from the community (mean age 58.8±5.1 y, 47% male). The majority (n=208, 96.1%) felt physicians should be able to share medical information in EHRs and 204 (96.2%) felt confident in the mechanisms in place to maintain confidentiality of their EHRs. Most participants (n=139, 90.1%) felt it was important to be able to opt out of any EHR data sharing or reporting but only 119 (56.1%) used the internet to communicate with their medical providers. A significant proportion (n=154, 72.6%) was aware about the availability of direct-to-consumer genetic testing. Conclusion: Among participants in a genomic medicine implementation study, the majority supported the use of EHRs for research and felt confident in the mechanisms in place to maintain confidentiality of their EHRs.

2331T A systematic review of individuals’ perspectives on broad consent and data sharing in the United States. N.A. Garrison1,2, M. McPheeters1,2, N.A. Sathé3, R.R. Walden2, E.W. Clayton1,2, 1) Center for Biomedical Ethics and Society, Vanderbilt University, Nashville, TN; 2) Department of Pediatrics, Vanderbilt University, Nashville, TN; 3) Vanderbilt Evidence-Based Practice Center, Institute for Medicine and Public Health, Vanderbilt University, Nashville, TN.

Objective: The purpose of this systematic literature review was to synthesize and evaluate evidence related to individuals’ willingness to provide broad consent for biobank research and data sharing. Study Design: We searched bibliographic databases (MEDLINE, via the PubMed interface, Web of Science, National Reference Center for Bioethics Literature databases (EthxWeb, GenETHX), and Dissertation Abstracts International) for original research conducted in the United States and published since 1990. Two reviewers independently screened studies using a tool for evaluation criteria using DistillerSR™, an online systematic review tool. Next, the two reviewers independently assessed the quality of studies using a predefined instrument. Any disagreements that arose between reviewers were resolved through discussion. Results: The final sample consisted of 46 articles. Most studies involved surveys (n=26), followed by focus groups (n=8), mixed methods (n=9), analyses of consent forms (n=2), and interviews (n=1). Overall, women were more often included in studies, and the quality of studies ranged from good (n=18), to fair (n=26), and poor (n=2). This presentation highlights our major findings. First, individuals were more willing to provide broad consent if the samples are de-identified. Respondents’ preference for broad consent over other types of consent (such as tiered or study-specific) was stronger when there was awareness of demographic and costs associated with maintaining large biobanks and when the privacy concerns were adequately addressed. Second, most respondents would allow data derived from their samples to be shared with many researchers. They are interested in the research that will provide researchers with the resources to carry out large-scale genomic research studies that would not be otherwise possible. As genomic testing becomes increasingly incorporated into clinical settings, biobanks now have the potential to serve a greater role in supporting clinical care and biobank researchers as they design ways to effectively use the left over samples for research. Introduction: Biobanks are repositories of biological specimens and health information that provide researchers with the resources to carry out large-scale genomic research studies that would not be otherwise possible. As genomic testing becomes increasingly incorporated into clinical settings, biobanks now have the potential to serve a greater role in supporting clinical care and biobank researchers as they design ways to effectively use the left over samples for research. The objective of this study was to examine the use and sharing of samples for research. We aimed to determine the extent to which respondents supported the use and sharing of samples for research. The aim of this study was to determine the extent to which respondents supported the use and sharing of samples for research. The aim of this study was to determine the extent to which respondents supported the use and sharing of samples for research.Methods: This study was conducted in the United States and published since 1990. The survey was completed by 212 study participants recruited from the community (mean age 58.8±5.1 y, 47% male). The majority (n=208, 96.1%) felt physicians should be able to share medical information in EHRs and 204 (96.2%) felt confident in the mechanisms in place to maintain confidentiality of their EHRs. Most participants (n=139, 90.1%) felt it was important to be able to opt out of any EHR data sharing or reporting but only 119 (56.1%) used the internet to communicate with their medical providers. A significant proportion (n=154, 72.6%) was aware about the availability of direct-to-consumer genetic testing. Conclusion: Among participants in a genomic medicine implementation study, the majority supported the use of EHRs for research and felt confident in the mechanisms in place to maintain confidentiality of their EHRs.

2332M Attitudes and concerns related to placing genomic information in the electronic medical record: a survey of biobank participants. A. Fiksdal1, J. Olson2, K. Maschka3, J.B. McCormick4. 1) Biomedical Ethics Program, Mayo Clinic, Rochester, MN; 2) Health Sciences Research, Mayo Clinic, Rochester, MN; 3) The Hastings Center, Garrison, NY.

Introduction: Biobanks are repositories of biological specimens and health information that provide researchers with the resources to carry out large-scale genomic research studies that would not be otherwise possible. As genomic testing becomes increasingly incorporated into clinical settings, biobanks now have the potential to serve a greater role in supporting clinical care and biobank researchers as they design ways to effectively use the left over samples for research. Introduction: Biobanks are repositories of biological specimens and health information that provide researchers with the resources to carry out large-scale genomic research studies that would not be otherwise possible. As genomic testing becomes increasingly incorporated into clinical settings, biobanks now have the potential to serve a greater role in supporting clinical care and biobank researchers as they design ways to effectively use the left over samples for research. Methods: Methods: This study was conducted in the United States and published since 1990. The survey was completed by 212 study participants recruited from the community (mean age 58.8±5.1 y, 47% male). The majority (n=208, 96.1%) felt physicians should be able to share medical information in EHRs and 204 (96.2%) felt confident in the mechanisms in place to maintain confidentiality of their EHRs. Most participants (n=139, 90.1%) felt it was important to be able to opt out of any EHR data sharing or reporting but only 119 (56.1%) used the internet to communicate with their medical providers. A significant proportion (n=154, 72.6%) was aware about the availability of direct-to-consumer genetic testing. Conclusion: Among participants in a genomic medicine implementation study, the majority supported the use of EHRs for research and felt confident in the mechanisms in place to maintain confidentiality of their EHRs.

2333T Research Use of Stored samples (RUST)-Community Perspectives from a Developing Country. S. Ramalingam. Molecular Medicine, PGSInstitute of Medical Sciences and Research, Coimbatore, Tamilnadu, India.

Introduction: Human biomedical samples including blood, tissue and other body fluids are very important for biomedical research. Rapid development in the newer technologies increases the scope of bio banking in developing countries. In countries like India with a large number of diseases there is a great need to store samples of left over samples and tissues which are being used for biomedical research. However there are lot of ethical and legal implications in the reuse of these samples for research. Bio banks are often sought for international collaborative research to generate newer findings that are generalizable across populations in the world. Lacks of clear guidelines compounded with lack of awareness of biomedical research especially with stored samples are important problems in countries such as India. Data on this in a populated country such as India with a huge potential for bio banking makes this study an important one. Aims: To study the perspectives and attitudes of the community towards the use of stored sample for biomedical research. Methodology: A cross sectional survey from patients and general population at a tertiary care hospital in South India and general population from field practice area of a communitry medicine was carried out. Data from this was analyzed using SPSS 19 software. Results: There were 225 respondents with 52.4% females and 47.6% males. 62% of the population did not want to store their samples for research. 88% opined that they will consent if their treating doctor informs them about the use of samples for research and explains about the scope of research. There was a significant association between age and consenting for research. The older the age, more likely they are consenting for use of their left over samples. (p<0.05) but this difference was not observed between the sexes. Though they agreed for the ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any ethical and research implications of such practices must be addressed as the process unfolds. Incorporation of population in the research study is a must, if any
2334M
Genes for Good: An Online Community Study of Genes, the Environment, Health and Disease. J.R. Forster1, W. Liu2, M. Zawistowski3, J. Wu4, K. Briggs5, S.T. Vrieze6, G.R. Abecasis1. 1) Department of Biostatistics, Center for Statistical Genetics, University of Michigan, School of Public Health, Ann Arbor, MI; 2) School of Information, University of Michigan, Ann Arbor, MI; 3) Medical School, University of Michigan, Ann Arbor, MI.

Genetic studies with traditional recruitment strategies utilizing in-person assessments, tissue sampling, and medical evaluations are precise, but also very expensive and exclusive. Not everyone can join these traditional studies because of limitations on available time, geography, among others. On a research team at the University of Michigan and designed to break down these research participation barriers, we have developed a database driven dynamic survey and passive monitoring platform for large scale genetic association studies that uses the internet and social media to recruit participants and provide online surveys that can be completed by participants when convenient for them. The software is currently available as a Facebook App, allowing us to engage thousands of participants quickly and cost-effectively through a popular and widely-used social media platform.

Participants complete a minimum number of surveys about their health and behavior are eligible for DNA analysis. We send participants split kits to collect saliva in a simple and non-invasive manner. Extracted DNA is genotyped with a reference set of GUCS and coding variants. To motivate participants to contribute and promote literacy about genetic information, we have developed web-based software applications that provide participants information about their genetic ancestry and their survey results. Users can download their survey results in full and receive their own raw genetic data. At the time of data sharing, participants are asked to sign a data use agreement, and provide a certificate of confidentiality, and expect to present data on participation rates, attitudinal measurement accuracy, and cost efficiency of the research design.

Participants will be snowball sampled, in a process that will be started by the investigators using twitter and posts on their Facebook wall. Our software platform will then rely on existing Facebook social networking to encourage active participants to invite their Facebook friends to participate and to advertise the study by posting to their Facebook wall. We ultimately want to open our platform to any researcher with a relevant scientific question. They will be able to use this research tool, including the genetic and phenotypic data we have collected, to answer a wide range of questions about the links between genetic information and health.

2335T
Practical Solutions for Protecting Individual Genomic Privacy. J. Fellay1, J.L. Raisaro2, Z. Huang3, M. Humbert4, E. Ayday3, P.J. McLaren2,1, A. Telenti1, J.P. Hubaux2. 1) School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Switzerland; 2) Swiss Institute of Bioinformatics, Lausanne, Switzerland; 3) School of Computer and Communications Sciences, Ecole Polytechnique Fédérale de Lausanne, Switzerland; 4) Institute of Microbiology, University Hospital and University of Lausanne, Lausanne, Switzerland.

The increasing availability of genomic data has major implications for personal privacy. The issues raised by genomic privacy reside at the crossroads of medicine, computer science, legislation and public policy. We here describe the development and deployment of privacy enhancing technologies that aim to find the optimal balance between usability and privacy of genomic data in clinical care and in biomedical research. First, we propose a privacy-preserving algorithm for genetic risk testing in clinical care that uses homomorphic encryption and distributed key generation. After genomic data (e.g. sets of variants from whole genome sequencing) are generated by a certified institution, they are encrypted and stored at a centralized “storage and processing unit” (SPU). Our architecture, while preserving data privacy, enables a medical unit to retrieve the encrypted information (without biometric authentication) and to use it for individualized care. We deployed this solution in a pilot pharmacogenomics study of 180 patients participating in the Swiss HIV Cohort Study. Retrieval and processing of encrypted genotypes, for a test using 50 markers (SNPs), take less than 1 second on commodity hardware. An interim analysis showed this to be acceptable to clinicians as both usability and privacy of genomic data are preserved. Second, we developed a system to protect the privacy of mapped short reads (e.g. bam files). Millions of sequencing reads from individual genomes are stored at a SPU in encrypted form. We developed an automated data provenance system that allows us to determine which subset of the reads without revealing the nature of the request to the SPU. In addition, the SPU can mask partial parts of the retrieved reads if they are not in the requested range or not consented by the patient (e.g., regions revealing a known disease). Finally, we have developed a novel methodology, based on model checking and formal model validation, to estimate the erosion of genomic privacy of an individual when genomic data of some of his/her relatives are publicly available. We showed that a target genome can be reconstructed by relying on Mendelian inheritance. As a result of this inference attack, we proposed different possible definitions of genomic privacy metrics.
Whole genome sequencing of children: Consent, parental choice, and the hunt for secondary variants. N. Monfared1, C. Shuman1, J.A. Anderson2, R. Hayeems1, M. Szego3,6,7, R. Zlotnik Shaul6,7,8, M.S. Meyn5,6,7, S. Bowdin4,5,9. 1) The Centre for Genetic Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; 3) Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; 4) Centre of Clinical and Metabolic Genetics, Dept of Paediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 5) Department of Paediatrics, University of Toronto, Toronto, ON, Canada; 6) Joint Centre for Bioethics, University of Toronto, Toronto, ON, Canada; 7) McLaughlin Centre, University of Toronto, Toronto, ON, Canada; 8) Centre for Clinical Ethics; Providence Healthcare, St. Joseph's Health Centre and St. Michael's Hospital; Toronto, ON, Canada; 9) Department of Bioethics, The Hospital for Sick Children, Toronto, ON, Canada.

Background: Traditionally, consent processes preceding single or multigene genetic tests have focused on preparing individuals for test outcomes related to clinical presentation and/or family history. WES and WGS can also reveal medically important variants unrelated to the primary diagnosis (secondary variants) thereby challenging existing genetic counseling models. As part of The Hospital for Sick Children’s Genome Clinic, a research project designed to model routine use of WGS in paediatric genomic medicine, we are studying the process of informed consent for WGS of children and the choices parents make regarding return of secondary variants (SVs) and carrier variants. Methods: We offer WGS to children undergoing clinical molecular testing for a suspected genetic disorder. Prospective participants (child and parent trios) are counseled regarding the nature of WGS and the opportunity, to learn about the child’s pharmacogenomics as well as childhood and adult-onset medically actionable variants (MAVs) and carrier status variants in over 2800 disease genes listed in NHGRI’s Clinical Genomic Database. Parents are also offered the option to undergo targeted testing for adult onset MAVs and carrier status variants identified in their child. Results: Of the 321 families approached to date, 54% agreed to participate; 42% declined. The initial consent session averaged 45-60 minutes with parents and children participating in the decision making process. The mean lag between the first counseling appointment and enrollment is 9 days, with most families requiring more than one appointment in this interval. There was no consensus among families regarding return of SVs. 58% of those enrolled chose to learn about all secondary adult-onset MAVs and 63% chose to learn about gene carrier variants, while 17% elected not to learn of any SVs. The most common concerns raised by families declining to learn about SVs included the psychological burden of identifying SVs in the child and themselves, potential impact on other family members, and/or fear of life insurance discrimination. Among parents who decided to learn of their child’s adult-onset MAVs and carrier variants, 68% chose to learn their own status for the same risk variants. Conclusion: Our findings highlight the need for genetic counseling models, decision aids and resources to evolve in order to support the complex decision making processes necessary for clinical use of WES/WGS.
2340M
Attitude toward genetic research on children and informed assent. Z. Yamagata1, I. Ishiyama2, K. Muto3, J. Minari4, G. Yoshizawa5, K. Kato6, 1 Department of Biomedical Ethics, Neuhaus, Germany; 2 Department of Health Sciences, School of Medicine, Univ Yamanashi, Yamanashi, Japan; 3 Teikyo-Gakuen Junior College, Yamanashi, Japan; 4 Department of Public Policy, Inst Medical Science, Univ Tokyo, Tokyo, Japan; 5 Department of Biomedical Ethics and Public Policy, Graduate School of Medicine, Osaka Univ, Osaka, Japan.

[Background] The aim of this study is to determine the attitudes of the general Japanese public toward genetic research on children and to clarify factors related to such attitudes based on nationwide surveys conducted in 2014.

[Methods] From the general population, 2,400 people (age = 20-69) were selected using a stratified two-phase sampling method. In a mail survey administered in February 2014, the participants were surveyed regarding the following topics: (1) their attitudes toward genetic testing, the genetic testing of children, and obtaining blood donations from children for research; (2) their perspective regarding informed consent; (3) their level of scientific literacy regarding genomics; and (4) their demographic information and socio-economic status. [Results] The response rate was 56.4% (1,354/2,400). We calculated genomic literacy scores by considering the participants' (1) knowledge of genomic terminology, (2) contextual understanding of genomic terminology, and (3) awareness of the benefits and risks of genomic studies. Conducting genetic testing on children for disease susceptibility was favored by 58.4% of participants. Regarding obtaining blood donations from children, 53.2% approved, 11.3% disapproved, and 34.8% were undecided. A higher proportion of participants with high genomic literacy levels approved of obtaining blood donations from children. The multiple logistic analysis odds ratio regarding genomic literacy was 1.28 (95% confidence interval: 1.05-1.50). Regarding whether seeking consent from children is acceptable if the child understands the details of the research and 42.5% answered that seeking consent from a child is acceptable if the child can judge the pros and cons of participating in the research. These responses were associated with genomic literacy. [Discussion] Many birth cohort studies have been conducted worldwide. The results of this study suggest that people's genomic literacy is essentially related to people's perspective on genomic research on children.

2341T
EuroGentest Guidelines for Diagnostic Next Generation Sequencing. P. Bauer1, G. Matthies2, M. Alders3, A. Corveleyn4, S. Eck5, I. Feenstra6, V. Race7, H. Yntema8, Participants of the EuroGentest Workshop on Diagnostic NGS Guidelines. 1 Institute of Medical Genetics and Applied Genomics, Univ Tuebingen, Tuebingen, Germany; 2 Genetikzentrum, Kiel, Germany; 3 Department of Clinical Genetics, Academic Medical Centre (AMC), Amsterdam, Netherlands; 4 Center for Human Genetics, Leuven, Belgium; 5 Department of Clinical Genetics, Academic Medical Centre (AMC), Amsterdam, Netherlands; 6 Department of Human Genetics, Radboud University Medical Center (RUMC), Nijmegen, Netherlands; 7 Department of Clinical Genetics, Academic Medical Centre (AMC), Amsterdam, Netherlands.

The use of Next generation sequencing (NGS) for clinical diagnostics is increasingly recognized. While the exploitation for research purposes has made tremendous advancements in human genetics, the diagnostic use brings challenges at different levels including data production and storage, and interpretation of results. Although several diagnostic NGS guidelines have been issued by the American, Australian, Dutch and British genetic professional societies, a couple of relevant topics have not yet been addressed. The participants of a EuroGentest Workshop on Diagnostic NGS Guidelines, were also working on compiling, integrating and completing these guidelines. Among other statements, the guideline delivers three major definitions: “diagnostic utility”, “quality scoring for NGS tests”, and “reportable range”. We believe that defining the ‘diagnostic utility’ of the NGS test is the laboratory’s first duty when preparing to offer diagnostic NGS. Our scoring system for the different NGS assays depends on quality and comprehensiveness. With this system, referring physicians, patients, and stakeholders in the health system will be enabled to compare different tests offered at the market. This scoring system is new, although not feature in any other guideline. As far as ‘reportable range’ is concerned, we propose the use of 3 specific percentages depending on the reference (technical target, coverage of transcript in a gene panel, coverage with reference to the genome) which will again allow to compare individual tests in an easy way by using runs, between tests and between laboratories. The guidelines propose a generic template for reporting NGS results as well. While dealing with informed consent, unclassified variants and unsolicited findings, again from the laboratory standpoint, is already addressed in aforementioned guidelines, the distinction of diagnostic and research is a very relevant topic when it comes to the “duty to recontact”. The latter is elaborated with a practical solution.

2342M

[Background] Direct-to-consumer genetic testing (DTCGT) in the US and China has been suspended due to alerts by the authorities in these countries. However, DTCGT can still be purchased in Japan. The Ministry of Health, Labour and Welfare (MHLW) in Japan has ignored these international alerts for a long time. The Ministry of Trade and Industry (METI) have conducted market research three times since 2007, to create a new market and to educate consumers in this field. The METI considered the regulations governing DTCGT and attached importance to quality control, scientific evidence, and informed consent. There is at present no ban on genetic discrimination or legal protection of personal genomic information in Japan. Several surveys have been conducted to clarify questions regarding public or consumer attitudes to DTCGT; however, there is still little data available in East Asia. [Purpose] The aim of this study is to clarify public attitudes towards DTCGT and the unique regulations in Japan. Method: We conducted a web-based questionnaire survey to investigate general perceptions in 2014. In total, 7,390 Japanese citizens completed the questionnaire (RR = 39.5%). We analyzed this data and compared it with past datasets. [Results] Of the respondents, 21.8% knew about DTCGT, and just 1% had purchased it before. Regarding actionable diseases, 50% of the respondents reported that they would like to receive their results via their physicians. Only 10% wanted to receive the results by post or on a website. We asked respondents to report their willingness to undergo 4 types of genetic testing based on scientific evidence. About 50% reported that they would like to undergo genetic testing for single gene Mendelian diseases with certain scientific evidence. However, 25-26.4% responded that they would undergo such testing even if there was no scientific evidence. More than 70% of the respondents reported that they needed ban on DTCGT. The perceived need for regulation on genetic discrimination in employment and in insurance was 59.7% and 52.8% respectively. [Discussion] A limitation of our study was that the percentage of respondents who reported “I can’t decide” was relatively high, compared with past results in other countries. Japanese respondents showed less interest in DTCGT. Before promoting DTCGT in Japan, legislation for protection of personal genomic information is needed.

2343T
Development of the Clinical NGS Industry in a Shifting Policy Climate. M.A. Curnutte1, K.L. Frumovitz2, J.M. Bollinger1, A.L. McGuire3, D.J. Kuffman4, 1) Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX; 2) Genetics and Public Policy Center at Johns Hopkins University, Washington, DC.

There is active debate about appropriate regulation of the rapidly evolving clinical next generation sequencing (NGS) industry. Interviews with industry leaders (n=19) and a web-based search of NGS companies (n=94) suggest the industry is developing services along the NGS pipeline partly in response to this regulatory uncertainty. Currently, few companies offer the full range of technology, software, and services needed to perform clinical testing. Rather, the industry tends to sell laboratories Research Use Only (RUO) products that do not stand alone as a clinical test. Clinical laboratories with expertise in laboratory standards, reimbursement, and clinical reporting then assemble these components into medical NGS assays that are classified as laboratory-developed tests (LDTs), subject to little NGS-specific regulation. Manufacturers of components of the LDTs have few clear standards to meet. This is occurring against a backdrop of evolving policy, technology, and insurance coverage. A better understanding of how the NGS business is developing in response to the current policy environment should inform regulatory efforts.
2344M
The protection of genetic privacy in the European Union and the proposed data protection reforms. A.C. de Paor, Centre Disability Law & Policy, Natl Univ Ireland, Galway, Galway, Galway, Ireland.

With developments in genetic science and technology, genetic information is becoming more widely accessible. However, the increasing availability of genetic information raises many ethical and legal issues that may threaten advancing scientific research in the absence of sufficient regulation. One such issue is the violation of genetic privacy. Genetic information is sensitive personal information that can reveal intimate details about an individual and an individual's family. In light of the sensitive nature of genetic information, the potential abuse is clear, as is the desire to protect such information from access and disclosure. Therefore, there are compelling reasons to maintain the privacy and confidentiality of genetic information. In the European Union (EU), the Data Protection Directive (introduced in 1995) provides a framework that protects the privacy of personal information. However, there is currently no substantive provision for genetic data and no specific reference to this category of information, indicating a gap in the privacy protection for genetic data in the EU. Recent developments indicate an intention to include genetic data within the scope of EU data protection laws, with the proposal of a new Regulation. With this new Regulation, the European Commission aims to develop an updated data protection framework. It is committed to reform data protection legislation, in line with the realities of today's society, and changing norms. The draft Regulation identifies 'genetic data' as a category of personal data designated for special protection. 'Genetic data' is defined broadly to include 'all data, of whatever type, concerning the characteristics of an individual that are inherited or acquired during early prenatal development', thereby presumably incorporating all genetic data including family medical history. Although this draft Regulation has encountered delay and a lack of consensus, it will enhance the protection of personal information (including genetic information) in the EU, when it is introduced. As regards protection of genetic privacy in third party contexts, this draft Regulation is welcomed as expressly recognising genetic information as a category that deserves protection. This paper analyses the issue of genetic privacy in the EU. It examines the current data protection framework, as well as the proposed reform of the EU data protection framework and the impact that these reforms will have on the protection of genetic data.

2345T
Towards an ethics “safe harbor” for global genomic research. E.S. Dove1, M.H. Zawawi2, E. Lévesque3, J. Simard4, B.M. Knoppers5. 1) Human Genetics, McGill University, Montreal, Quebec, Canada; 2) Molecular Medicine, Laval University, Quebec City, Quebec, Canada.

Genomic researchers are becoming increasingly globally connected and collaborative. Though global data-driven research holds great promise for disease discoveries, the underlying research ethics review systems in many of the world challenges improvements in human health and paradoxically may not improve respect for persons who participate in research. Case reports illustrate that the current system is costly, fragmented, inefficient, inadequate, and inconsistent. There is an urgent need for criteria to determine whether there is “substantial equivalency” between the principles and procedures, ethics review and governance to be able to share data across jurisdictions. Building on the international privacy “safe harbor” model that was developed following the adoption of the European Data Protection Directive in 1995, we propose a federated “Safe Harbor Framework for International Ethics Review Equivalency” that could facilitate the harmonization of ethics review of specific types of data-driven international genomic and/or genetic research, where there is global transboundary research ethics norms and principles. The Safe Harbor Framework would consist in part of a newly constituted organization (provisionally called the International Federation for Ethics Review, or IFER), formed by a voluntary agreement among countries, granting agencies, philanthropes, institutions and healthcare, patient advocacy, and research organizations. IFER would be both a central ethics review body and also a forum for review and follow-up of policies concerning ethics norms for international research projects. It would be built on five principle elements: 1) registration, 2) compliance review, 3) monitoring and enforcement of ethics review, 4) education and training, and 5) public participation. A Safe Harbor Framework would create many benefits for researchers, patients, and the general public. Research participants would enjoy uniform adequate protection, while researchers would enjoy enhanced ethics review efficiency. One of the ethical concerns that the Safe Harbor Framework seeks to address is the administrative hassle, and redundant regulatory hurdles. Most importantly, society would enjoy the maximization of the potential benefits of genomic and disease research.

2346M
Analysis of ethical and social issues of large scale genome cohort/biobanking projects in Japan. J. Minari1, Y. Hanabusa1, S. Toda2, F. Nagamori2, K. Kato1,2,3, 1) Graduate School of Medicine, Osaka University, Suita, Osaka, Japan; 2) Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan; 3) Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Japan.

In the last decade or so, many large scale genome cohort/biobanking projects have been developed all over the world. Their aim is to establish infrastructures for disease and health research. Many of the projects collect a large number of blood samples and associated lifestyle information from healthy people, while some projects collect patient data with a variety of diseases.

In Japan, several large scale projects have been established starting from early or middle 2000s. One well-known example is the Biobank Japan project which has a collection of DNA and blood serum samples of 200,000 patients. Other projects started relatively recently. The Tohoku Medical Megabank has been set up in the Tohoku area which has been hit by the big earthquake in March 2011. It has begun collecting samples in 2013 with the aim of establishing two cohorts of 150,000 people. It has also completed whole genome sequencing of 1000 individuals in the fall 2013.

For the ethical, legal and social issues that have to be addressed, projects in Japan tended to work on them relatively independently. Many projects have been initiated, but each project struggles to establish their own policy without sufficiently learning from existing projects. In order to maximize the resources and efforts used for the projects and to overcome the shortcomings aforementioned, it has become imperative to exchange experiences beyond individual projects. To this end, we have conducted a cross project comparison of ethical and social issues as well as governance mechanisms of several large scale projects.

Our analysis has shown several features of the current Japanese projects. One is that many of them are considering data sharing as a necessary activity, while sharing materials is still controversial especially with foreign researchers. Another interesting feature is that healthy participants for cohort studies are often recruited in conjunction with the system of health monitoring in the local community, which is feasible based on the existing trust between participants or community and researchers. Our findings are valuable for creating a common framework of research governance in the global community.
2348M
Concerns of researchers and physicians regarding the chromosomally integrated HHV-6. V. Noël1, R. Drouin2, L. Flamand3, C. Bouffard1. 1) Université de Sherbrooke, Sherbrooke, Canada; 2) CHR du Grand-Portage, Rivière-du-Loup, Canada; 3) Université Laval, Québec, Canada.

INTRODUCTION: The human herpesvirus 6 (HHV-6) has the unique capacity, among HHVs, to integrate to the telomeres of the chromosomes (ciHHV-6) of nucleated cells after infection; ciHHV-6 can therefore be found in gametes and be transmitted both from one generation to the next, and through organ and tissue donation. This worrisome situation affects more than 70 million people around the world and may be associated with more than 50 illnesses (pneumonitis, encephalitis, infertility, cancers, etc.), of which some are fatal. Despite its consequences on the health of individuals and populations, this phenomenon has attracted little interest in scientific and public health milieus. To highlight the importance of the problem, several researchers and physicians publish warnings in purely scientific journals. Thus, our objectives are to understand the positions of these researchers and the solutions they propose, in order to identify their clinical and socio-ethical concerns regarding the effects of ciHHV-6.

METHODOLOGY: 1) Comprehensive narrative synthesis of the literature: a) research by key words in databases; b) thematic inductive analysis; c) interpretative synthesis.

RESULTS: Of the 90 articles analysed, 62 contained opinions, recommendations or lines of question bearing on: a) population health, b) grafts and transplants, c) diagnostic tests, d) clinical and therapeutic care and e) pathogenicity. Certain authors also raise the idea of a screening test, as a solution for the negative effects of ciHHV-6 on population health. Finally, the problems identified by researchers and the solutions they propose invite us to revisit the notion of reproductive autonomy, as well as the principle of beneficence. We will conduct interviews based on our results with researchers and physicians who are members of the HHV-6 Foundation.

CONCLUSION: This research demonstrates the importance of possessing knowledge making it possible to reflect on: a) the medical and socio-ethical consequences of ciHHV-6, b) the relevance of offering screenings test and c) patient care. It could contribute to the establishment of measures that would act early on the consequences of ciHHV6, to support the work of researchers and physicians, to improve and harmonize practices for the wellbeing of affected individuals and populations.

2349T
Regulating Gamete Donation in the US: Ethical, Legal and Social Challenges. M. Sabatello1,2,3. 1) Psychiatry, Columbia University, New York, NY, USA; 2) Center for Global Affairs, NYU, New York, NY, USA.

The practice of gamete donation has received growing attention in the past two decades. As an increasing number of individuals and couples resort to this practice, and the children born as a result (donor-conceived-children) are reaching maturity, many are grappling with the question how to address the dilemmas that arise. Do recipient-parents have a duty to tell their donor-conceived child about his/her genetic origins? Should the identity of the donor be disclosed or remain anonymous? Does the child have a right to know her conception story and to receive information, including identifying, medical, and genetic information, about the donor? And if a donor-conceived-child has a right to know, who has the duty to tell her about it? Focusing on the US context, this paper explores these questions having in mind the larger question of what do we as a society owe children born as a result of assisted reproduction, especially gamete donation. It underscores the ethical, legal and social dilemmas that arise, comparing and contrasting with international developments in this area. It highlights the medical justifications for regulating this field, explores the emerging so-called right of the conceived child about his/ her genetic origins ("the right to know"), and considers the challenges such a right evokes to existing principles of medical ethics in the US as well as other broader societal implications of such a right.

2350M
Illuminating the changing landscape from newborn screening to newborn sequencing: Ethical, psychological, and societal implications for research and policy-making in the genomics era. L. Bush1, K. Rothenberg2, 1) Pediatric Clinical Genetics, Columbia University Medical Center, New York, NY; 2) Genomics and Society, National Human Genome Research Institute, Bethesda, MD.

As the genomics and bioethical communities engage in newly-funded research investigating the application of whole exome/genome sequencing both on healthy newborns (NBSeq) within the context of a population-based newborn screening (NBS) initiative as well as with ill newborns as a diagnostic test, it is critical that the complexity of emerging issues are illuminated to enhance understanding and exploration of the challenges ahead – the ethical, legal, psychosocial, and policy implications on our multi-cultural society. Historically, both NBS and genetic testing of children have been fraught with controversy - such as informed consent autonomy issues, psychological impact of false positives, wide variability in access, what constitutes the best interest of the child - and concerns raised will accelerate in parallel with the mounting information generated by genomic sequencing.

Foresight and thoughtful planning are especially needed in the research domain when infants are involved. Ethical deliberation with diverse voices is essential to set policy, particularly since the potential benefits of genomic findings on a newborn co-exist with a multitude of potential risks. By expanding identification of conditions exponentially through the implementation of new comprehensive genomic technologies, NBS as a social justice lever may be jeopardized with NBSeq as some states lack resources to provide necessary follow-up, especially as many variants of uncertain significance are discovered. While tolerance to uncertainty varies, the potential for heighten anxiety must be seriously considered. Moreover, the ability to protect the interests of vulnerable newborns by traditionally mandating NBS for some conditions may be eroded with NBSeq due to informed consent and return of results blurring boundaries among a public health screening measure, research, and clinical test paradigms. From an ethical and societal perspective, this is really tough terrain. To facilitate reflection and ethical inquiry among ASHG attendees and enrich the policy process, contextual nuances surrounding this changing genomic landscape are illuminated.
Incidental findings, will be each time more common after the implantation of whole-exome sequencing: terminology, desire for information, types/phenotypes are sufficiently important that they should be sought out and characterized the potential extent of this issue, we evaluated how many genes are associated with additional phenotypes across the entire Online Mendelian Inheritance in Man (OMIM) catalog, a publicly accessible database of human genes and genetic disorders (OMIM.org). Methods: We recorded how many and what kind of MIM phenotypes were noted as associated with each gene listed in OMIM. Initially, only phenotypes with an assigned MIM number were counted. Results: As of 5/14/2014, 879 of the 4555 genes in the catalog (19%) listed with multiple phenotypes (2-16) were associated with more than ten phenotypes. When not restricted to phenotypes that had been assigned MIM numbers, 1075 (24%) of these genes listed multiple phenotypic relationships (range 2-16). Conclusion: A substantial fraction of the genes in the OMIM catalog are associated with multiple phenotypes, which could be conveyed inadvertently when returning genetic results. Pleiotropy appears to be a pervasive issue that has not been addressed in current incidental findings guidelines, which position in clinical practice to deal with these extraneous yet easily-discoverable relationships. Future recommendations will need to consider the potential impact of such pleiotropic relationships when determining how best to return individual genetic results.

Debate continues over whether, how and when people should be allowed to learn about incidental findings, will be each time more common after the implantation of whole-exome sequencing: terminology, desire for information, types/phenotypes are sufficiently important that they should be sought out and characterized the potential extent of this issue, we evaluated how many genes are associated with additional phenotypes across the entire Online Mendelian Inheritance in Man (OMIM) catalog, a publicly accessible database of human genes and genetic disorders (OMIM.org). Methods: We recorded how many and what kind of MIM phenotypes were noted as associated with each gene listed in OMIM. Initially, only phenotypes with an assigned MIM number were counted. Results: As of 5/14/2014, 879 of the 4555 genes in the catalog (19%) listed with multiple phenotypes (2-16) were associated with more than ten phenotypes. When not restricted to phenotypes that had been assigned MIM numbers, 1075 (24%) of these genes listed multiple phenotypic relationships (range 2-16). Conclusion: A substantial fraction of the genes in the OMIM catalog are associated with multiple phenotypes, which could be conveyed inadvertently when returning genetic results. Pleiotropy appears to be a pervasive issue that has not been addressed in current incidental findings guidelines, which position in clinical practice to deal with these extraneous yet easily-discoverable relationships. Future recommendations will need to consider the potential impact of such pleiotropic relationships when determining how best to return individual genetic results.

As the technical and diagnostic merits for whole-exome sequencing are assessed, parallel research is underway to elucidate and meet the challenges it may present for clinicians and patients. NCGENES is a NIH-CSER study aiming to address these challenges in >500 genetic patients. One of the project goals is to explore patient perspectives and educational needs to inform results return of both diagnostic and incidental findings. To aid this study, six focus groups were held in November and December 2013 with a total of 40 participants from the UNC genetics clinic: 31 adult patients (23 hereditary cancer / 8 general genetic) and 9 parents of minor or intellectually disabled patients. The group discussions utilized a vignette to explore several areas related to incidental findings resulting from whole-exome sequencing: terminology, desire for information, types/phenotypes are sufficiently important that they should be sought out and characterized the potential extent of this issue, we evaluated how many genes are associated with additional phenotypes across the entire Online Mendelian Inheritance in Man (OMIM) catalog, a publicly accessible database of human genes and genetic disorders (OMIM.org). Methods: We recorded how many and what kind of MIM phenotypes were noted as associated with each gene listed in OMIM. Initially, only phenotypes with an assigned MIM number were counted. Results: As of 5/14/2014, 879 of the 4555 genes in the catalog (19%) listed with multiple phenotypes (2-16) were associated with more than ten phenotypes. When not restricted to phenotypes that had been assigned MIM numbers, 1075 (24%) of these genes listed multiple phenotypic relationships (range 2-16). Conclusion: A substantial fraction of the genes in the OMIM catalog are associated with multiple phenotypes, which could be conveyed inadvertently when returning genetic results. Pleiotropy appears to be a pervasive issue that has not been addressed in current incidental findings guidelines, which position in clinical practice to deal with these extraneous yet easily-discoverable relationships. Future recommendations will need to consider the potential impact of such pleiotropic relationships when determining how best to return individual genetic results.
2356M
Compare and Contrast: A cross-national study across UK, USA and Greek Experts toward Return of Incidental Findings from Clinical Sequencing. E.G. Gourna, N. Armstrong, S.E. Wallace. Heath Sciences, University of Leicester, Leicester, United Kingdom.

Guidance for the return of incidental findings (IFs) from clinical sequencing is being drafted at national and international levels. However, no studies have as yet been published comparing attitudes of experts across different countries to determine where similarities and differences of opinion and practice lie. We investigated attitudes toward returning IFs from clinical sequencing across UK, USA and Greek experts. Thirty in-depth interviews were conducted with experts including clinical geneticists, lab-based geneticists and experts with legal background. The majority of experts expressed numerous concerns and these were consistent regardless of their country. Major differences were only observed between experts according to their professional background. Clinical geneticists were mainly driven to return only medically actionable findings while the other experts also favored returning IFs that had personal utility for patients. Clinical geneticists were more inclined to order targeted tests instead of Next-Generation Sequencing (NGS) to avoid finding IFs, while lab-geneticists considered NGS as a good source of useful information for the patient’s health. All experts saw the need for extensive pre-test counseling where patients could be helped to make informed choices to receive IFs, but post-test decisions were seen as more complicated and genetic counseling was considered crucial. No consensus was reached regarding the dissemination of information to family members. In all three countries, the final decision to disclose or not to disclose, usually lied with the clinical geneticist, who was also responsible for communicating the findings and was occasionally supported by a multidisciplinary team. Our study showed more similarities than differences across countries. Professional background seemed to play a more important role, while even experts with the same professional background relied on their previous experiences to inform their attitudes. Our experts could only agree that medically actionable results should be communicated to patients. As well, they acknowledged how time and labor intensive the decision process is, and asked for clear guidance to support them. However, the type of guidance seen as most appropriate differed according to the infrastructure and resources of their health system. We concluded that while over-arching principles can be created at the international level, guidance for specific practice needs to be country-specific.

2357T
Participant satisfaction with a preference-setting tool for the return of individual research results in pediatric genomic research. I. Holm1,2,3, B. Lee4, S. Ziniel5,6, P. Bacon7, S. Savage1, K. Christiansen8, E. Weitzman2,3,9, R. Green10, N. Huntington11,11. 1) Division of Genetics and Genomics, Boston Children’s Hospital, Boston, MA; 2) Department of Pediatrics, Harvard Medical School; 3) Manton Center for Orphan Disease Research, Boston Children’s Hospital; 4) McLean Hospital; 5) Division of Adolescent/Young Adult Medicine, Boston Children’s Hospital; 6) Center for Patient Safety and Quality Research, Program for Patient Safety and Quality, Boston Children’s Hospital; 7) Johns Hopkins University School of Medicine; 8) Division of Genetics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School; 9) Children’s Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, Boston Children’s Hospital; 10) Partners Personalized Medicine; 11) Division of Developmental Medicine, Boston Children’s Hospital.

Although return of individual research results (IRR) to participants in genomic biobank research may be desirable, the feasibility of IRR return and how to incorporate participant preferences are unclear. There is skepticism that participants truly understand the implications of their choice of IRR to receive suggesting that incorporating preferences may be unrealistic. The role of preferences is particularly complex in pediatrics where a parent receives IRR on their child. We developed an online preference-setting tool for the return of IRR based on the preventability and severity of a condition, dimensions that resonated with parents in interviews. The tool also allows parents to opt-out of receiving results from 4 categories: mental illness, developmental disorders, childhood-onset degenerative conditions, and adult-onset conditions not treatable during childhood. We tested this tool in an online survey emailed to parents of patients <18 years of age at Boston Children’s Hospital. Parents were randomized to the hypothetical scenario that their child was enrolled in 1 of 4 biobanks with different policies for return of IRR on their child: 1) receive no IRR (None); 2) receive all IRR (All); 3) given a choice to receive all or no IRR (Binary); and 4) use the preference setting tool to choose categories of IRR to receive (Granular). Parents were given a hypothetical research result report for their child.

In this analysis we report on participant satisfaction with the process, biobank, and hypothetical results received. There were 11,391 emails delivered with a response rate of 24%. The granular group was the most satisfied with the process, biobank, and hypothetical results received while the None group was least satisfied with each (p<0.001). The None group was least likely to agree that the biobank they were randomized to was beneficial (p<0.001) and most likely to agree it was a “bad thing” (p<0.001). The Binary and All groups rated satisfaction nearly as high as the Granular group. No significant difference was found between the groups regarding agreement/disagreement with the statement that the biobank was harmful. Recognizing that the results returned were hypothetical, our data suggest that return of IRR may increase satisfaction for biobank participation. Our results also suggest that compared to limited or no choice on results to receive, providing participants with the ability to designate their preferences leads to the greatest satisfaction.
2358M
Patients’ Perceptions of Whole Genome Sequencing Results and Plans to Use Non-Actionable Findings. L. Jamali1,2, J.O. Robinson1, P. Lupo1, J. Blumenthal-Barby1, L. Feuerman1, J. Vassy1, K.D. Christensen1, M.J. Stashinski1, J. Wycliff1, R.C. Green4, A.L. McGuire1 for the MedSeq Project. 1) Baylor College of Medicine, Houston, TX, United States; 2) Johns Hopkins Berman Institute of Bioethics, Baltimore, MD, United States; 3) Section of General Internal Medicine, VA Boston Healthcare System, Boston, MA, United States; 4) Department of Medicine, Brigham and Women’s Hospital, Boston, MA, United States.

Background: There is much debate over best practices for returning results from clinical whole-genome sequencing (WGS) to patients. An emerging approach is to categorize WGS results into ‘bins’ based on their clinical utility. How patients interpret and intend to use WGS results is not well understood. Methods: The MedSeq Project is a randomized clinical trial exploring the use of WGS in cardiology and primary care settings. At baseline (before WGS), 106 patient-participants were surveyed and a subset of 28 completed in-person, semi-structured interviews. Transcripts of recorded interviews were consensus coded and analyzed thematically. Target enrollment of 200 patient-participants is expected by summer 2014. Results: Consistent with previous studies, the majority of patient-participants (87%) wanted all their WGS results. However, qualitative data show that the majority of interviewed patient-participants (n=24) had difficulty distinguishing among result types. When asked, they struggled to provide clear examples of results associated with preventable or treatable conditions as distinct from results with less clinical utility. These difficulties did not reflect poor study understanding or genetic literacy. Many patient-participants (n=17) wanted results associated with non-preventable or treatable conditions to inform their professional or financial decisions, and cardiologists and genomic health counselors did not easily distinguish among the types of results described in WGS informed consent documents. They value WGS results for both medical and non-medical reasons. These data challenge the conceptual framework underlying many emerging models of WGS results disclosure, which define the clinical utility of results and base disclosure practices on these definitions.

2359T
What do young breast cancer patients want to learn about individual results from genome sequencing? K.A. Kapthingst1, J. Ivanovich1, B. Biesecber2, J. See1, L. Dressler3, P. Goodfellow1, R. Dresser3, M. Goodman4, M. Ray5, S. Bell2, K. Walton1. 1) Division of Public Health Sciences, Washington University School of Medicine, St. Louis, MO; 2) Social and Behavioral Research Branch, National Human Genome Research Institute, Bethesda, MD; 3) Mission Health, Fullerton Cancer Center, Asheville, NC; 4) College of Medicine, The Ohio State University, Columbus, OH; 5) Washington University School of Law, St. Louis, MO.

Background: Many communication challenges arise with return of individual results from genome sequencing. While the type of genome sequencing results that should be returned has been hotly debated, few findings are available to inform what patients would like to know about a result. These data are needed to develop communication strategies to return WGS results, particularly among populations targeted for early application of genome sequencing such as patients diagnosed with breast cancer at a young age. Methods: We conducted 60 semi-structured individual interviews with young women diagnosed with breast cancer at age 40 or younger to investigate what they would like to learn about individual genome sequencing results. We stratified recruitment by family history of breast cancer and BRCA1/2 mutation status to examine differences by these factors. Interviews examined interest in return of individual results for six types of gene variants (e.g., affects risk for a preventable disease, affects treatment response, uncertain clinical significance) as well as what content participants would like to learn about each type of result. Two coders independently coded all interview transcripts; analysis was based on consensus codes and a thematically driven approach. Results are presented using NVivo. Results: Across different types of genome sequencing results, participants were most interested in learning about the health implications of the variant: “long term, what to expect...I would want to learn as much as I could so I could know what to do with it”. Of the six content types, participants most wanted to learn about the familial mutation, and cause of the variant (e.g., hereditary, environment). We observed some differences in themes across strata. For example, the theme of wanting to learn about the familial mutation and cause of the variant was most important to participants who were women who have known BRCA1/2 mutation. Some participants also highlighted the importance of clear communication in returning results: “I think I’d want to know everything that was in the scientific literature around that, but in layman’s terms...” Conclusions: Participants most wanted to learn about the health implications of individual genome sequencing results, but were also interested in specific information about variants. These findings can inform the development of strategies to communicate with patients about genome sequencing results.

2360M

The debate surrounds the return of incidental findings to participants in genomics research. Should researchers offer results to family members after the participant’s death? Prior law and policy protect individual privacy, choice, and control, preserving each person’s right not to learn the results of clinical validation were: acting on the recommendation of a medical professional, receiving a notification of a genetic test result, and learning about the familial mutation, and 1 had a variant of uncertain significance. Six of 29 (21%) participants validated results in a CLIA compliant laboratory, and the research and clinical results were concordant in all cases. Five additional participants shared their intent to confirm their research results but had not obtained the results. Four of 27 (15%) participants validated their research findings, 4/6 (67%) had disease-causing mutations. For the 8 participants who participated in in-depth interviews, reasons for pursuing or not clinical validation were: acting on the recommendation of the research team and future clinical care vs. lacking insurance coverage and perceiving limited personal benefits. The study results provide important empirical data on how research participants value result confirmation in CLIA compliant laboratories, and suggests the need to better communicate its advantages.

2361T
How Research Participants Value Result Confirmation in CLIA Compliant Laboratories. M.Y. Laurino1,2, A. Truth1, L. Tenney2, D. Fisheries, G.P. Jarvis3, P. Newcomb2,4, D. Veenstra5, S.M. Fullerton1,6. 1) Institute for Public Health Genetics, University of Washington, Seattle, WA; 2) Fred Hutchinson Cancer Research Center, Seattle, WA; 3) Division of Medical Genetics, University of Washington, Seattle, WA; 4) Pharmaceutical Outcomes Research and Policy Program, University of Washington, Seattle, WA; 5) Department of Epidemiology, University of Washington, Seattle, WA; 6) Institute of Behavioral and Human Sciences, University of Washington, Seattle, WA.

The Clinical Laboratory Improvement Amendments of 1988 (CLIA) requires that clinical laboratories meet quality standards to ensure that patient care is not compromised. While the CLIA certification process is one of the main justifications supporting return of results from genetic and genomic research, there is disagreement about the necessity of validation in a CLIA compliant laboratory prior to returning research results. Under the CLIA research exception, results of tests performed in non-CLIA compliant laboratories may be returned with the caveat that such results are used for informational purposes only and not for clinical care. At present, limited data exist regarding the extent to which research participants act on or interpret results of tests performed in non-CLIA compliant laboratories. To address this gap, this study examined whether participants at the Fred Hutchinson Cancer Research Center Colan Cancer Family Registry Site pursued the recommendation to confirm their research results in a CLIA compliant laboratory by using a mixed-methods approach. Of registry participants who received Lynch syndrome (LS) related findings, 19/26 (77%) supported offering results despite privacy protections, suggesting that the incidental findings debate must broaden to include consideration of return to family.
Ethical implications of incidental findings found by array-CGH in a routine clinical population. M. Lefebvre1,2, D. Sanlaville1, N. Marie1, C. Christel Thauvin-Robinet2, E. Gautier1, F. E. Cader2,23, M. Ajenjar2, S. Chantot-Bastarda2, P. Bottin2, B. Hero2, E. Pipara2, S. Mouton10, N. Chassaing1, L. Lepinspina12, S. Manouvrier-Hanu13, M. Marti-Dramard14, A. Goldberg15, C. Ferrec16, S. Odent17, C. Le Caignec18, B. Gilbert-Dussardier15, A. Toutain16, E. Schaefer16, N. Philip17, F. Giulian23, L. Lambert24, P. Kuentz25, S. Julia17, L. Olivier-Faivre1,2. 1) Centre de Génétique et Centre de Référence Anomalies du Développement et Syndromes Malformatifs de l’Est, FHU-TRANSLAD, CHU Dijon, France; 2) GAD “Étude Germline et Syndromes Malformatifs de l’Est, FHU-TRANSLAD, CHU Dijon, France; 3) Genetics Service, Hospices Civils de Lyon, Hôpital Femme-Mère-Enfant, and Eastern Biology and Pathology Centre, Bron Cedex, France; 4) Département de Génétique, Hôpital Necké-Enfants Malades, AP-HP, Paris, France; 5) Service de Génétique, Hôpital Pitié Salpêtrière, Paris, France; 6) APHP, Hôpital Armand Trousseau, Service de Génétique et d’Embryologie Médicales Paris, France; 7) Service de Pédiatric, Hôpital Jean Verdier, Assistance Publique - Hôpitaux de Paris, Bondy 93143, France; 8) Department of Neuropediatrics, Armand Trousseau Hospital, APHP, Paris, France; 9) Cytogentic Laboratory, Jean Verdier Hospital, Bondy, France; 10) Département of Clinical Genetics, Bordeaux Children’s Hospital, CHU de Bordeaux, Bordeaux, France; 11) Service de Génétique Médicale et de la Cytogenétique, CHU, Toulouse, Paris, France; 12) Cytogentic Laboratory, Chambery Hospital, Chambery, France; 13) Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, France; 14) Unité de Génétique clinique, Hôpital Nord, CHU Amiens, France; 15) Unité de Génétique Médicale, CHU Rouen, Rouen, France; 16) Laboratoire de génétique moléculaire, CHU, Brest, France; 17) Service de Génétique Clinique, CLAD-Ouest, Hôpital Sud Rennes, France; 18) Génétique Clinique, CLAD-Ouest, CHU de Nantes, Nantes, France; 19) Genetics, University Hospital La Milétrie, Poitiers, France; 20) Service de génétique, Centre Hospitalo-Universitaire Tours, Tours, France; 21) Service de génétique médicale, Hôpital de Hautepière, Strasbourg, France; 22) Département de Génétique Médicale, Hôpital d’Enfants de La Timone, Marseille, France; 23) Service de Génétique Médicale, Hôpital de l’Archet Notre-Dame, Nice, France; 24) Unité de Génétique Clinique, Centre Hospitalier, CHU, B. Gilbert-Dussardier, B. Heron, France; 25) Centre Hospitalier Universitaire de Besançon, Besançon CEDEX 25030, France.

Microarray-based comparative genomic hybridization (aCGH) is commonly used to assist in diagnosing patients presenting with a phenotype compatible with a microdeletion syndrome or with intellectual disability with or without malformation. Since aCGH is interrogating the whole genome of an individual for micro rearrangement, there is a risk to be confronted to incidental findings (IF), defined as a finding that has potential health importance for the propositus or his family and is not related to the indication of research scientists, bioethicists, lawyers, epidemiologists, and clinical geneticists. The preliminary outcome of the meeting 1) retrospectively, guidance should allow for returning results and to develop a method to return results. From that meeting NCHS temporarily closed the DNA bank to new proposals in 2012. NCHS commissioned the National Academy Science Committee on National Statistics to conduct a two-day workshop. “Guidelines for Returning Individual Results from Genome Research Using Population-Based Banked Specimens” was held in February 2014. The workshop included participants from the Federal government, universities, and the private sector from a variety of specialty areas including research scientists, bioethicists, lawyers, epidemiologists, and clinical geneticists. The preliminary outcome of the meeting 1) retrospectively, guidance was not to return results due to consent language and lack of CLIA certification for lab processing of the DNA specimens and 2) prospectively, consensus should allow for returning results and to develop a method to return results.

Receiving input and guidance from a panel of experts was a valuable approach to understanding and reconciling the evolving issue of returning results for genomic research particularly for historic DNA banks.

Parental decision-making for children enrolled in exome sequencing research and attitudes toward receiving variants of uncertain significance. J. Sapp1, M. Crenshaw1, D.A. Dong1, L.G. Biesecker1, B.B. Biesecker1,1 National Human Genome Res Inst, Bethesda, MD; 2) Genzyme, Boston, MA.

Genome sequencing is increasingly being employed to identify the cause of rare diseases. Yet policies and practices regarding the interpretation, return, validation, and release of genomic research data to participants are currently evolving and under debate. Many participants in genomic research are children whose parents consent and make decisions regarding the return of results on their behalf. In an NIH study to find the genetic causes of rare conditions we explored participants’ preferences regarding the return of both primary and secondary findings. We conducted semi-structured interviews with 22 parents of 13 pediatric probands with rare disorders. Parents were asked to reflect on the notion of being offered choices in learning their child’s variant results and about their attitudes toward receiving variants of uncertain clinical significance (VUCS). All parents appreciated the opportunity to make individual choices about receiving categorized variant results for their child, and seven parents stated that they would not have participated in genomic research if they had not had the opportunity to learn their child’s sequence results. Parents had mixed attitudes toward the possibility of receiving their child’s VUCS. Most parents with positive attitudes about receiving these variants discussed their hope for the information could prove useful in their child’s future and planned to retain and occasionally revisit these results over time. Nine parents held negative or neutral attitudes toward VUCS because of the ambiguity inherent in these findings and perceived them to be of less utility. Parents’ own personal experiences with uncertainty informed their attitudes toward learning more about their child’s VUCS: some parents referenced their comfort with uncertainty as part of coping with their child’s illness while others perceived receipt of these results as an undesired additional burden and worried about where they would find expertise to meaningfully interpret findings of this nature. Our findings suggest that while research scientists continue to sequenc- ing where both primary and secondary variant results are offered to parents of pediatric research participants can expect enhanced participant engagement, informed consent, and decision-making. Further understanding of how these attitudes affect longitudinal outcomes of research studies may inform discussion of best practices regarding the return of genomic data to research participants.
2365T
“Bring it on!” Preference setting for secondary results from exome sequencing using My46. H. Tabor1,2, J. Crouch1, S.M. Jamal1, A.G. Shankar3, U.H. Ye4, M.J. Bamshad2,3, 1) Treuman Katz Ctr Pediatric Bioethics, Seattle Children’s Hospital, Seattle, WA; 2) Dept of Pediatrics, University of Washington, Seattle, WA; 3) Dept of Genome Sciences, University of Washington, Seattle, WA.

Exome sequencing is widely used in clinical and research settings, yet it is unclear whether and how individuals should be allowed to set their preferences about what secondary results (SR) they want to receive. It is also unknown how different patient populations will respond to the opportunity to receive a variety of SR, which SR they will choose, and why or why not. We recruited 145 people whose exomes (58 adults) or child’s exomes (87 parents) were previously sequenced to select SR preferences using My46, a web-based results management tool. Participants selected “yes,” “no” or “I am undecided” for each of 11 SR categories (SRC): carrier status, medication response, ancestry, and disease risk (e.g., blood, cancer, etc.). The majority (79%) selected “yes” for all SRC (83% of parents, 75% of adults). A subset (n=51) were interviewed about SRC selection, including 12 who refused ≥1 SRC. Most interviewees (92%) described optimism and excitement; 31% volunteered that they were not worried or anxious. A subset (n=11) simultaneously expressed ambiguity and nervousness about possible “bad” results but said that optimism about benefits outweighed concerns. Participants used both family history and their or their child’s condition or future health prioritise their SR preferences. The average number of SRC refused was 1.8 for parents and 3.5 for adults. The SRC most frequently refused by parents was ancestry (n=7) and by adults was brain and nervous system (n=10). Two parents refused SRC to defer SR until adulthood (carrier status) or for a priori carrier status (carrier status). These findings suggest that most people are comfortable receiving all or most SR, at least when organized categorically, but a subset want to be able to refuse, defer or prioritize specific SR for reasons that may vary across populations and conditions. This conclusion is tempered by the fact that a limited number of SR examples were offered for each SRC, raising the question as to whether offering more examples with a broader range of impacts would facilitate better informed and/or more selective decision-making. Either way, it is clear that a one-size-fits-all opt-out SR policy in which all SR are packaged together will poorly accommodate many people.

2366M
Returning findings within longitudinal cohort studies: the 1958 Birth Cohort as an exemplar. S. Wallace1, N.M. Walker2, J. Elliott3, 1) Health Sciences, University of Leicester, Leicester, United Kingdom; 2) Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory Department of Medical Genetics Cambridge Institute for Medical Research University of Cambridge Cambridge, UK; 3) Centre for Longitudinal Studies Institute of Education London, UK.

Population-based, prospective longitudinal cohort studies are considering the issues surrounding returning findings to individuals as a result of genomic and other medical research studies. While guidance is being developed for clinical settings, the process is less clear for those conducting longitudinal research. Research was conducted on behalf of The UK Cohort and Longitudinal Study Enhancement Resource programme (CLOSER) to examine consent requirements, process considerations and specific examples of potential findings in the context of the 1958 British Birth cohort in order to inform future decision making regarding policy. Beyond deciding which results to return, there are questions of whether re-consent is needed and the possible impact on study, whether there is a need to introduce third parties such as genetic services to assist in the feedback process and how that will be managed, and what resources are needed to support and manage the feedback process. Recommendations are made for actions a cohort study should consider taking when making vital decisions regarding returning findings. Any decisions need to be context-specific, arrived at transparently, communicated clearly, and in the best interests of both the participants and the study.

2367T
Participant preferences regarding the return of mental health related research results from a pediatric biobank and associations with social-demographic factors, comfort and concerns with novel health information. E. Weltzman1,2,3, S. Zniel1,2, S. Savage4, K. Christensen1, N. Huntington5, P. Bacon6, C. Cacioppo7, R. Green7,8, I. Holm4,9, 1) Department of Pediatrics, Harvard Medical School, Boston, MA; 2) Division of Adolescent/Young Adult Medicine, Boston Children’s Hospital; 3) Children’s Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, Boston Children’s Hospital; 4) Center for Patient Safety and Quality Research, Program for Patient Safety and Quality, Boston Children’s Hospital; 5) Johns Hopkins University School of Medicine; 6) Division of Genetics and Genomics Boston Children’s Hospital; 7) Division of Genetics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School; 8) Partners HealthCare Center for Personalized Genetic Medicine; 9) Manton Center for Orphan Disease Research, Boston Children’s Hospital.

Preferences for return of individual research results (IRR) from a pediatric biobank may depend on features of a disease category, which can hold social, historical and medical meanings. Mental illness (MI) is one such category encompassing common problems that may complicate other conditions and carry stigma. To investigate these issues we developed a preference-setting tool for IRR based on the preventability and severity of a condition, allowing exclusion of 4 disease categories including MI. We consented and enrolled parents of Boston Children’s Hospital patients <18 years of age to an online study, randomizing them to 1 of 4 hypothetical biobank designs: 1) receive no IRR (None); 2) receive all IRR (All); 3) may choose all or no IRR (Binary); and 4) use the preference setting tool for return of IRR (Granular). Following an educational video, granular arm participants completed a survey about preferences, social and demographic factors and perceptions about new health information. We report on preferences concerning MI IRR. Of 11,394 invited parents 2,718 (24%) participated in the survey, including 1,026 randomized into the granular arm (mean age 44, 91% female, 86% white). Of these 12.8% opted-out of MI IRR. No differences were found among participants who did and did not opt-out of mental health results in sex, race, ethnicity, education, history of participating in medical research or employment in the healthcare setting. Participants who set preferences for return of severe IRR and preventable IRR were less likely to opt out of MI results than individuals who excluded these results (adjusted ORs and 95% CIs for MI IRR as a function of severe and preventable IRR: 0.42, 0.32, 0.57, p<.001; and, 0.73, 0.57, 0.93, p=.01). Participants who reported a family history of developmental delay were less likely to opt out of MI IRR (8.5%) than were participants who did not (15.3%) (adjusted OR 0.62, 95% CI 0.4, 0.96, p<.05). Participants who reported a family history of MI were less likely to opt out of MI IRR (10.6%) than participants who did not (15.2%) (p<.05) but results were not significant in multivariate models. Most participants set preferences to return MI IRR concerning their child. Doing so was associated with preferences for IRR about severe and preventable conditions. Few other factors distinguished this group. Findings suggest high willingness to be informed of a common and complex problem.
2368M  
Uptake and motivations to learn incidental genome sequencing results among cancer patients. E. Glogowski1, Y. Bombard2, S. K. Schrader1, J.G. Hamilton1, M. Salemo1, M. Cornes1, S. Patil1, M.J. Massie1, V. Joseph1, Z. Stadler1, S. Lipkin1, K. Offit1, M. Robson1, *equal contributions.  
1) Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Service, New York, NY; 2) University of Toronto, Li Ka Shing Knowledge Institute, Toronto, ON, Canada; 3) Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center, New York, NY; 4) Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY; 5) Weill Colleage, New York, NY.

Background: There are increasing efforts to disclose incidental results of whole genome or exome sequencing (WG/ES) to patients and research participants; however, limited data exist as to individuals’ interest and motivations to learn incidental results. In this study, we aimed to understand how participants would react to learn incidental actionable results in WG/ES disease risk information. Methods: We are offering the return of incidental results to patients whose genomes/exomes are sequenced as part of one of several research projects (ACMG) has recommended that pathogenic variants (PVs) in 56 genes, corresponding to 24 diseases, should be returned as incidental findings. Participants: We are offering the return of incidental results to patients whose genomes/exomes are sequenced as part of one of several research projects. Uptake and motivations to learn incidental results. This underscores the need to establish clear purpose for recommendations. Results: Of 31 eligible patients invited thus far, 8 have enrolled in the study, resulting in a 29% uptake rate (average age: 55, 8 college-educated, 4 married and 3 employed). Thirteen (42%) were lost to follow-up and 10 (32%) declined via phone call or two phone calls were needed for engagement. Motivations included: curiosity (6/8); interest in incidental disease risks for themselves (8/8) and/or relatives (6/8); and a desire to improve health (7/8) and help research (7/8). Reasons for declining study participation included: worry about the study; confidentiality; difficulty traveling or getting to the medical center to discuss results; and not about getting or keeping insurance (4/10) and/or a job (1/10); and a lack of desire to learn genomic results (4/10) and/or disease risks (4/10). Additional reasons for declining the study offered were: a recent cancer recurrence, time constraints, and a preference for targeted geriatric counseling. A majority of individuals ascertained through cancer discovery research decline to learn their incidental genomic findings due to a lack of desire, worry about learning their results, or perceived barriers to participation. Disinterest in incidental findings among cancer patients appears to arise from a preference for targeted, cancer-specific information. These preliminary results provide insights into cancer patients’ interest levels, motivations and perceived barriers, which may inform study designs aimed at returning incidental WG/ES results in the oncology setting.

2369T  
Penetration of Actionable Incidental Genomic Findings in Exomes from the Framingham Heart Study. N.B. Gold1, A.G. Bick2, H.M. McLaughlin4, C. Kraft1, E. Rehm1, H.G. Seidman1, C.E. Seidman4, R.C. Green1, *1) Division of Genetics, Department of Medicine, Brigham and Women’s, Harvard Medical School, Boston, MA; 2) Department of Genetics, Harvard Medical School, Boston, MA; 3) Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 4) Partners HealthCare Personalized Medicine, Boston, MA; 5) Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA; 6) Division of Cardiology, Department of Medicine, Brigham and Women’s Hospital and the Howard Hughes Medical Institute, Harvard Medical School, Boston, MA.

Background: The American College of Medical Genetics and Genomics (ACMG) has recommended that pathogenic variants (PVs) in 56 genes, corresponding to 24 diseases, should be returned as incidental findings. The ACMG guidelines were developed by an interdisciplinary panel of genetics professionals, representatives from the medical community, and patients. Their guidelines were based on scientific and clinical criteria to determine which disease is supported by evidence. The guidelines were intended to provide guidance about whether, and if so, how to return individual results. Several studies have examined patients’ hypothetical preferences for the return of incidental results; however, little is known about participants’ actual experiences. We report on a qualitative interview study of research participants’ perspectives on the medical and psychological impacts of incidental results. Methods: Participants were recruited from the ClinSeq® study, which enrolls mostly health participants and offers genome sequencing results to individual participants. Participants were offered a choice about receiving results unless they are urgent medical significance, and given results in person by a physician and a genetic counselor, or by phone. Twenty-four participants (67% of those eligible) agreed to participate and were interviewed. Participants received results for a wide range of conditions, the majority of which were medically actionable and pertained to cancer, cardiomyopathy or arrhythmia susceptibility. The interviews were recorded, transcribed, and coded. Half of the interviews were reviewed by an independent, second coder; 80% consensus was achieved and discrepant codes were reconciled. Most participants reported sharing their result with at least one other person and had no changes to their medical care. Participants stated that they valued the result regardless of the medical utility. One participant (17%) who had a low-risk result, was said by a participant to be “disappointed” but not “distressed.” A high proportion of participants (92%) reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions. Participants were asked whether they would tell their parents or children about their result. Participants reported sharing their result with at least one other person, while only 25% reported changing their medical care. Participants reported that their medical care was their result. Only 25% of participants reported that their results would impact their future health decisions.
2372M
Does prior comfort with new health information influence parents’ preferences for receiving genetic research results about their children? N. Hunt, C. Egreg, S. Ziniel, R. Banerji, C. D. Royal, J. L. McCauley, M. A. Quintero, R. Martynek, J. L. McCauley. Department of Pediatrics, Harvard Medical School; 3) Division of Genetics and Genomics, Boston Children’s Hospital; 4) Division of Adolescent and Young Adult Medicine, Boston Children’s Hospital; 5) Center for Patient Safety and Quality Research, Boston Children’s Hospital; 6) Johns Hopkins University School of Medicine; 7) Children’s Hospital Informatics Program, Boston Children’s Hospital; 8) Division of Hematology and Bone Marrow Transplantation, Brigham and Women’s Hospital; 9) Partners Personalized Medicine; 10) Manton Center for Orphan Disease Research, Boston Children’s Hospital.

Background: In genomic biobank research, questions remain about how to maximize benefit and minimize harm when returning individual research results (IRRs) to parents about their children. Result- and parent-specific factors can influence a parent’s reaction to receiving IRRs. We examined the impact of comfort with health information on parents’ preferences for receiving IRRs about their children.

Methods: An online survey was completed by 2,718 parents of recent Boston Children’s Hospital patients (response rate 24%). Respondents were randomized to one of three survey groups: 1) “No choice”—half receive All IRRs, half receive None; 2) “Binary choice”—parent chooses to receive All or None; 3) “Granular choice”—parent sets preferences for IRRs based on preventability and severity of conditions, plus options to opt-out of specific condition categories (mental health, developmental disorders, degenerative conditions, adult-onset conditions). Parents then viewed a hypothetical IRR report reflecting their survey condition preferences. Following this, they were asked if they would want IRRs, HIC-N least likely and HIC-A were in-between. In the Binary group, most HIC-A chose All results, in a proportion similar to the other groups. Parents were then asked to identify the most preferred and the least preferred IRRs for the Binomial and Granular conditions, respectively. Parents were also asked about the impact of the IRRs on their perceptions of their children, the potential relationship between comfort and IRRs, and finally how the results would be shared with family members.

Results: Parents were satisfied with receiving IRRs, but were divided about whether they wanted to receive them, with HIC-N least likely and HIC-A more likely to want IRRs. HIC-N least likely and HIC-A were in-between. In the Binomial group, most HIC-A chose All results, in a proportion similar to the other groups. Parents were then asked to identify the most preferred and the least preferred IRRs for the Binomial and Granular conditions, respectively. Parents were also asked about the impact of the IRRs on their perceptions of their children, the potential relationship between comfort and IRRs, and finally how the results would be shared with family members.

Conclusion: The finding that the Binomial group, most HIC-A chose All results, in a proportion similar to the other groups. Parents were then asked to identify the most preferred and the least preferred IRRs for the Binomial and Granular conditions, respectively. Parents were also asked about the impact of the IRRs on their perceptions of their children, the potential relationship between comfort and IRRs, and finally how the results would be shared with family members.

2374M
Direct-to-Consumer Ancestry Testing: Psychological and Behavioral Impacts. D.L. Boedl1,2, A.G. Lucy,3, L. Anniello1,2, O. Libiger1,2, V. Banerji1,2, E.J. Topor2,4,5, L. Kessler1, N.J. Schork1,2, C.S. Bloss1,2. 1) Scripps Translational Science Institute, La Jolla, CA; 2) Scripps Health, La Jolla, CA; 3) Scripps Clinic Medical Group, La Jolla, CA; 4) Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA; 5) Arcadia University Graduate School of Science.

Introduction: There is a growing population of consumers who are purchasing direct-to-consumer (DTC) genetic ancestry tests. Ancestry.com, a DTC ancestry testing company, recently reported that more than 400,000 individuals purchase their AncestryDNA test every year. Despite the interest, however, few studies have examined the impact of ancestry testing at an individual level. The current study aimed to evaluate attitudes towards ancestry testing and characteristics associated with response to testing in a large, diverse cohort of predominantly White Americans. A longitudinal cohort study originally designed to assess response to DTC genomic testing for complex disease risk. SGHi participants who completed a baseline health assessment were also provided with personal genetic ancestry testing performed by a team of scientists at the Scripps Translational Science Institute. Individuals who viewed their ancestry results were invited to complete an additional optional survey (http://www.stsweb.org/sghi) to assess response to the information provided. Participants also self-reported ancestry priority prior to receiving the ancestry results.

Results: 3,465 individuals were provided with personal genetic ancestry testing, 1,317 logged into the website to view their results, and a subscale of 322 (24%) completed the ancestry response survey. Although 46% of respondents reported that their ancestry results were surprising or unexpected, 39% reported a change or possible change in their perception of their cultural roots, 24% a change in the likelihood that they would travel to certain parts of the world, and 21% that the results reshaped their personal identity. A majority (77%) reported that the test results influenced their decisions about sharing results with their family members, and 12% reported the intention to share results with a healthcare provider.

Conclusion: Data from this selected sample of ancestry test recipients suggests that although genetic ancestry testing may be a significant event, it does impact a notable proportion of individuals with respect to cultural perceptions, personal identity, and propensity to share genetic information with family and healthcare providers. Our findings suggest that ancestry information may be particularly useful in promoting sharing and thus peer-peer education in genomics.

2375T
Motivations for Participation in Genomic Research in Hispanics vs. Non-Hispanics. M.L. Cuccaro1,2, J.A. Wessels1,3, C.P. Manrique1, M.A. Quintero1, R. Martynek1,2,3,4, J.L. McCauley1,2,5, D.L. Boedl1,2,3,4,5, J.L. McCauley1,2,3,4,5, A.G. Lucy1,2, O. Libiger1,2, V. Banerji1,2, E.J. Topor2,4,5, L. Kessler1, N.J. Schork1,2, C.S. Bloss1,2. 1) Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA; 5) Arcadia University Graduate School of Science.

Background: Involvement in Hispanics. Larger samples across different diseases can help elucidate differences in motivations to participate in genomic research may differ by ethnic group and that the impact of participation in genomic research to help find a cure (HI 56%, NHI 56%, p=0.16). Hispanics constituted (80%) of the sample. Among the 101 individuals surveyed 95% had agreed to participate in the MS genetics study. Individuals were asked to identify from 11 possible choices, the reasons which most influenced their decision. The most frequently cited reasons for participating in the MS genetics study were helping to find a cure (56%), having the disease (46%), and helping future generations (37%). We tested these three reasons differed by ethnic group using Pearson chi-square. Hispanics and non-Hispanics did not differ in their rates of endorsement of participating in genomic research to help find a cure (HI 56%, NHI 56%, p=0.97) or helping future generations (HI 34%, NHI 50%, p=0.26). However, the groups differed significantly on endorsement of having the disease as a reason to participate in the study. Specifically, 52% of the Hispanics endorsed having MS as reason for participating in genomic research to help find a cure (HI 56%, NHI 56%, p=0.015). We also tested whether these reasons differed by age group and found no differences. However, among the youngest individuals (18-35 years) 74% endorsed finding a cure as a reason for participating vs. 47% for the older participants.

Discussion: There are several motivations for participation in genomic research among Hispanics and non-Hispanics. These motivations may differ by ethnic group and may be the reasons that participants report for choosing to participate. Our data suggest that Hispanics are less likely to participate in genomic research than non-Hispanics. This underrepresentation has the potential to exacerbate existing health disparities. Increasing the engagement of minorities in genomic research is one strategy to reduce health disparities and to identify population specific genetic risks. This study is the first to identify cultural perceptions associated with genetic risk for multiple sclerosis (MS). We surveyed individuals who had either agreed or refused to participate about their reasons for this decision. Our dataset (N=101) was predominantly female (80%) and mainly white (95%). The most frequently cited reasons for participating in the MS genetics study were helping to find a cure (56%), having the disease (46%), and helping future generations (37%). We tested these three reasons differed by ethnic group using Pearson chi-square. Hispanics and non-Hispanics did not differ in their rates of endorsement of participating in genomic research to help find a cure (HI 56%, NHI 56%, p=0.97) or helping future generations (HI 34%, NHI 50%, p=0.26). However, the groups differed significantly on endorsement of having the disease as a reason to participate in the study. Specifically, 52% of the Hispanics endorsed having MS as reason for participating in genomic research to help find a cure (HI 56%, NHI 56%, p=0.015). We also tested whether these reasons differed by age group and found no differences. However, among the youngest individuals (18-35 years) 74% endorsed finding a cure as a reason for participating vs. 47% for the older participants.
2376M

Reasons for declining preconception carrier testing using genome sequencing: implications for research and practice. M. Gilmore1, C. Young2, M. Leo3, J. Davis2, T. Kaufmann2, E. Esterberg2, F. Lynch4, B. Wilford2, K.A.B. Goddard2. 1) Department of Medical Genetics, Kaiser Permanente Northwest, Portland, OR; 2) Center for Health Research, Kaiser Permanente Northwest, Portland, OR; 3) Truman Katz Center for Pediatric Bioethics, Seattle Children's Hospital, Seattle, WA.

Genome sequencing is an emerging technology for use in clinical practice, and its utility is not well understood for applications such as preconception carrier testing. An important facet of understanding the implementation of this new technology is to investigate reasons why potential participants decline expanded carrier testing using genome sequencing. We are investigating preconception expanded carrier testing in generally healthy couples. Potential participants are identified because their provider ordered at least one preconception carrier test, typically cystic fibrosis screening. They are contacted by phone to see if they are interested in the study. If they decline participation, they are invited to complete a survey on their reasons for declining as well as demographic information. We collect similar information for those who consent to join the study. We are comparing participants and decliners on factors such as race, education level, income and characteristics related to knowledge of genetics, such as knowing a family with a child with a genetic condition, to determine whether there may be any participation bias. In addition to the survey, we have implemented qualitative approaches including journaling by study staff and focus groups to enhance our understanding of reasons why potential participants decline the use of expanded preconception carrier testing using genome sequencing. Of the women who declined participation at any time after initial contact (N=45), seventy-two percent had completed the decliner survey. The most common reasons cited were a lack of time (fifty-five percent), not interested in participating (twenty-four percent), didn't want to know about their risk (twenty-one percent), and worried about discrimination (nine percent). Synthesis of qualitative data support the reasons captured in the survey. The high response rate for the decliner survey may suggest a high level of interest in preconception expanded carrier testing for this population. Concerns about discrimination are less common than we anticipated as a reason to decline to participate. Further analysis of decliner survey results will enhance our understanding of genomic sequencing technology implementation in the clinical setting.

2377T

Personal Genomics Online in Australia: A mixed methods study of Australian consumers’ knowledge, attitudes and experiences of direct-to-consumer personal genome testing. J. Savard1, J. Mooney-Somers1, A. Newson1, F. Kerridge1,2, 1) Centre for Values, Ethics and the Law in Medicine (VELIM), School of Public Health, The University of Sydney, New South Wales, Australia; 2) Hematology Department, Royal North Shore Hospital, New South Wales, Australia.

Health, including both one’s current and future health, is increasingly being understood in genetic terms, with personal genomic information marketed as a means by which individuals can “know” and “control” their health and reduce their likelihood of illness. Unfortunately, there is often a discrepancy between the expectations that consumers have of what personal genome tests may provide and what explanatory power they actually have. In part, this is because we still lack scientific and epidemiological knowledge needed to usefully interpret the information these tests divulge. In this paper we present findings from the Australian Genetics Online Study - a mixed methods exploration of the beliefs, knowledge, expectations and experience of Australian consumers regarding direct-to-consumer personal genome testing (DTC-PGT). The first phase of the study, which involved an online public survey, explored lay knowledge, expectations and experiences of DTC-PGT. The results of this study suggest that Australian consumers have an interest in genetic testing - both for health and ancestry purposes, and many value this information as much as they value other sources of personal information - but appear disinclined to pursue it. The second phase of the study, semi-structured interviews with potential and current consumers of a personal genome testing service, revealed a wide range of motivations for pursuing testing along with varying degrees of understanding with respect to what the test could (and did) tell the consumer about themselves, their future health and/or their ancestry. The final phase of the project, an auto-ethnography, documented the researcher’s (Jacqueline Savard’s) journey through the processes of becoming a consumer of DTC-PGT. The results of the Genetics Online Project provide the first quantitative and qualitative account of DTC-PGT in Australia. These findings will be discussed in light of ongoing debates about the impact that publicly accessible genetic testing is having on the design and delivery of health care, information about health care and about the impact of genetics on the social construction of health, illness and identity.

2378M

A novel scale to assess perceptions of uncertainty in genome sequencing information. B. Biesecker1, W. Klein2, K. Lewis3, L. Biesecker2, P. Han4. 1) Soc Behavioral Res Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892; 2) Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892; 3) Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892; 4) Maine Medical Center Research Institute, Center for Outcomes Research and Evaluation, 81 Research Dr Scarborough, ME 04074.

The scope of uncertainty in genomic sequence information is vast. Yet our understanding of how adults perceive medical uncertainty is poorly understood. Such perceptions are likely to play a key role in deciding to undergo genomic sequencing and in understanding and acting on the findings. Guided by a taxonomy of medical uncertainty developed by Han and colleagues, we aim to further understanding of perceptions of uncertainties. The NHGRI ClinSeq® cohort study represents a strategic opportunity to assess predictors of decisions to learn and use genomic sequencing information. We present analytic data on baseline perceptions of uncertainty originating from a novel scale. Methods: The Perception of Uncertainties of Genome Sequence Results (PUGSR) scale assesses perceptions of uncertainty specific to genomic sequencing. Development of items was informed by focus group findings. Ten items assessed three sub-domains of uncertainty: medical, affective and trustworthiness. Results: Five hundred seventeen ClinSeq® participants completed the scale prior to making a decision about whether to learn their sequence results. There was a normal distribution in responses with an overall mean uncertainty score of 3.5/5 (SD 0.58) with &alpha; high internal consistency (0.839). Results from the exploratory factor analysis revealed three factors: medical (four items), affective (four items) and trustworthiness (two items) uncertainty about sequence information. Perceptions of medical uncertainty were highest and contributed most to the variance (43%). These data may be used to guide informed consent for genome sequencing, to clarify the role of uncertainties in decision-making, and to anticipate uncertainties of greatest concern to recipients of genome sequencing information.

1) Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada; 2) Institute of Health Policy, Management and Evaluation, University of Toronto, ON, Canada; 3) Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, ON, Canada; 4) Division of Paediatric Medicine, Department of Paediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 5) Department of Pediatrics, Faculty of Medicine, University of Toronto, ON, Canada; 6) Department of Family and Community Medicine, Mount Sinai Hospital, University of Toronto, ON, Canada; 7) Newborn Screening Ontario, Children’s Hospital of Eastern Ontario, Ottawa, ON, Canada; 8) Department of Pediatrics, Faculty of Medicine, University of Ottawa, ON, Canada; 9) Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, ON, Canada; 10) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, ON, Canada; 11) Division of Respiratory Medicine, Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 12) Department of Psychology, The Hospital for Sick Children, Toronto, ON, Canada; 13) The Institute for Clinical Evaluative Sciences, Toronto, ON, Canada.

Objectives: The majority of infants with false positive (FP) newborn screening (NBS) results for cystic fibrosis (CF) are eventually identified as carriers of one CF mutation. Carrier detection can enable reproductive planning, provided results are appropriately shared. We examined carrier communication, carrier testing and family planning following NBS detection of carrier infants.

Methods: We surveyed mothers of infants confirmed to be FP after confirmatory testing (Time-1) and one year later (Time-2) to ascertain self-reported communication, intended and actual carrier testing, and subsequent reproductive choices. Mothers were invited for an interview. We report cross-sectional analyses for Time-1, and longitudinal analyses for carrier testing uptake by Time-2.

Results: We received completed surveys from 134 of 245 mothers (55%) at Time-1 and 96 of 214 (45%) at Time-2, and interviewed 66 mothers. 77 of 127 mothers (61%) reported one mutation in their infants (i.e., carrier-mothers); 30 (24%) reported no mutations and 20 (16%) were unsure (presumed non-carriers). Carrier testing uptake was 55% (n=23/42) among carrier-mothers. 13 intended to have testing uptake was 55% (n=23/42) among carrier-mothers. 13 intended to have carrier-mothers told relatives they may be carriers (n=70/77), as did 48% (n=24/50) of non-carrier mothers. Results were expected to influence family planning for 35% of carrier-mothers (n=26/74) and 24% of non-carrier mothers (n=11/46). Interviewees valued carrier information whether or not they intended to share it with relevant relatives.

Conclusion: While mothers describe reproductive value in carrier results, survey data are equivocal given moderate intended and actual CT uptake and limited influence on family planning.

Personalized medicine (PM) is based on the analysis of an individual’s cumulated genomic, behavioral and environmental data. It aims to evaluate, diagnose, and predict the risk of suffering from diseases and to determine the most efficient and safe preventative and therapeutic approaches. The 4P medical approach (predictive, preventive, personalized and participative) adds a participative dimension in which citizens are responsible for and in control of their own health. Prenatal diagnosis (PND) and medically assisted procreation (MAP) are also affected by the new role of individuals in questions of health. Yet, as the consequences of decisions made in the context of PND are not the same for the fetus and embryos as they are for adults and children, this situation raises particular clinical and socio-ethical concerns. However, while a vast literature exists on PM, it does not appear to take into account PND and MAP. Yet the information that genomic tests reveal about a fetus and embryo raise major clinical and socio-ethical issues that deserve to have their place in the PM movement. Based on the PM literature, our objectives were the following: 1) create a repertoire of the major PM issues, 2) determine whether PND and MAP have been discussed in terms of PM, and 3) identify the clinical and socio-ethical issues that emerge.

METHOD: Scoping review of academic and grey literature (perspectives, legal documents, government reports, etc.).

RESULTS: Of the 96 articles analyzed, 5 major themes emerged: the level of responsibility, equity, attitude to understanding, genetic counseling and the importance of the opinion of citizens. Yet, few articles discuss PND and MAP in terms of PM (10/98). When the subjects are broached, it is in the context of non-invasive prenatal testing (4/10) and the instances involve concerns such as social control of reproduction, eugenics and the risk of coercion. No study mentioned participation and the opinions of citizens.

CONCLUSION: From a clinical as well as a socio-cultural level, the participative dimension of PM implicates a societal responsibilization that will be problematic in the context of PND and MAP. It is imperative to develop knowledge based on convincing data concerning the positions of stakeholders.
2382M

Influence of the social environment on the development of breast cancer through epigenetic modifications: A comprehensive scoping review. O. Martin¹, R. Dreux¹,², C. Bouffard³, 1) Division of Genetics, Department of Pediatrics, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, Québec, Canada; 2) CHR du Grand-Portage, Department of Obstetrics-Gynecology, Rivière-du-loup, Québec, Canada.

For 20 years now, several research groups have attempted to establish a link between stressful events and susceptibility to developing breast cancer, but results have been contradictory. While some show evidence supporting the direct link between acute and chronic stress at a young age, traumatic events and breast cancer development, others suggest that stressful events cannot be considered risk factors. Nonetheless, it has been established that epigenetic modifications may play a predominant role from the onset of breast cancer and throughout its development. We can’t dismiss the idea that social environments that submit an individual to an acute and chronic stress might modify methylation patterns, thereby increasing susceptibility to certain chronic illnesses including cancer. The results presented here stem from a comprehensive and comparative analysis of academic literature bearing on the potential of the social environment, mainly acute and chronic stress, to induce epigenetic modifications linked to breast cancer among individuals with or without a known mutation in terms of a breast cancer susceptibility gene. Methodology: Comprehensive scoping review of academic literature: 1) research by keywords in: databases (Pubmed, EBSCO, etc.), reference lists of key papers, hand searching of journals not indexed on databases. Rigorousness criteria: English and French articles, from 1980-2014, epigenetics, breast cancer; 2) comparative thematic analysis. Results: Despite several publications on epigenetic modifications and the environment as factors of susceptibility to breast cancer development, the social environment remains a largely unconsidered subject. While it is ever more accepted that acute and chronic stress might be a susceptibility factor, no study has yet been able to draw a real correlation between epigenetic modifications induced by stress and the development of breast cancer. Important knowledge gaps currently exist regarding the links between social environments, stress, epigenetic modifications and breast cancer. Among other things, the effect of these relationships in a heterozygous patient for a BRCA1 or BRCA2 gene remains unknown. Conclusion: There is still a great deal to be done to obtain sufficient data to establish that social environments and human behaviour may induce epigenetic modifications that may increase the risk of developing breast cancer. However, the latest discoveries in the field tend to lean towards this possibility.

2383T

Genetic Ancestry Testing and Identity: Exploring the Relationship. C.M. Wolpert¹, J.D. Powell¹, K. Haynie¹, R. Kittles¹, C.D. Royal¹,², 1) IGSP, Duke University Medical Center, Durham, NC; 2) AAAS, Duke University, Durham, NC; 3) Dept. of Political Science, Duke University, Durham, NC; 4) University of Illinois Chicago, Illinois.

Genetic ancestry testing (GAT) uses individuals’ DNA to estimate their genealogical and geographic ancestry. Scholars have hypothesized that GAT provides information that may cause individuals to reshape their identities. Yet, there are limited empirical data to support this view. Furthermore, there is very little discussion in the research literature about the types of identity that are likely to be affected by GAT. In this study we sought to investigate whether and how GAT test takers (N = 455), the majority of whom self-identified as non-white, spontaneously referenced identity in relation to GAT. Three time points were investigated: Time 1 (pre-test), Time 2 (post-test - immediately after receiving test results), and Time 3 (8-10 year follow-up). Test takers completed self-report instruments at all three time points and a semi-structured interview at Time 3; both quantitative and qualitative data were analyzed. At Time 1, 48% (n = 218) of the 455 GAT test takers spontaneously referenced identity. Further analysis of these data revealed that different dimensions and types of identity were referenced: identity exploration, including ethnic identity exploration; personal identity; essential identity; and group identity (in decreasing order of frequency). At Times 2 and 3, individuals started referencing identity commitment. At Time 3, test takers had high levels of ethnic identity commitment as measured with the Multigroup Ethnic Identity Measure-Revised (MEIM-R) and ethnic identity achievement as measured with the Ethnic Identity Scale (EIS), both of which are associated with well-being. Hence, in this data set, identity was a salient issue for about half the participants, with identity exploration being most frequently discussed prior to testing. Subsequent to GAT, some test takers referenced identity commitment. Thus, it appears that before testing potential test takers may use GAT as a tool for exploring their identity and after GAT some individuals commit to a new or existing identity.
2384S

Psychosocial factors associated with the uptake of contralateral prophylactic mastectomy among BRCA1/2 noncarriers with newly-diagnosed breast cancer. J.G. Hamilton1, M. Salomo2, K. Amoroso2, M. Sheehan2, M. Harlan Fleischut2, E. Glogowski2, B. Siegel2, A.G. Arnold2, E.E. Salo-Mullen3, J. Hay1, K. Offit2, M.E. Robson2. 1) Behavioral Sciences Service, Department of Psychiatry & Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY; 2) Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY.

Women who are newly diagnosed with breast cancer can consider whether to have a contralateral prophylactic mastectomy (PM) to reduce their future risk of cancer in their unaffected breast. Pre-surgical BRCA1/2 genetic testing may provide valuable risk information to guide this choice, given that BRCA1/2 mutation carriers have a 27-37% ten-year risk of developing a new cancer in their unaffected breast. Despite BRCA1/2 mutation noncarriers’ much lower risk of contralateral breast cancer (which can be as low as 3-10%), some still choose to undergo a PM; the factors that motivate BRCA1/2 noncarriers to select PM are not well understood and warrant investigation. Thus, as part of a larger study of the impact of pre-surgical BRCA1/2 testing, we examined the association between various sociodemographic and psychosocial factors and BRCA1/2 noncarriers’ subsequent decision to undergo PM. Self-report questionnaire data from 75 BRCA1/2 noncarriers (mean age=44.1 years, range=29-59) were analyzed. A sizeable minority of the BRCA1/2 noncarriers (29.3%) elected to undergo a PM after learning their mutation status (as compared to 88% of the 8 BRCA1/2 carriers in the sample). Bivariate and multivariable analyses indicated that perceiving that one’s physician had recommended PM (OR=7.79, p=0.02), perceiving greater risk for breast cancer (OR=1.86, p=0.04), and perceiving greater pros of PM (OR=1.34, p=0.006) were all significantly associated with noncarriers’ decision to undergo PM. However, factors including age, Ashkenazi Jewish ethnicity, breast cancer-related distress, perceived cons of PM, and decisional conflict regarding PM were not significantly related to the decision to undergo PM (all ps>0.05). Results demonstrate that although noncarriers’ decision-making regarding PM was unrelated to sociodemographic and emotional factors, their cognitive perceptions of contralateral disease risk, surgical benefits, and physician recommendations were particularly important. Future studies should examine the content of physician communication regarding PM and mediating risk in greater detail, and explore how these conversations shape and interact with women’s past experiences, emotions, and beliefs to influence their cancer risk management decision-making.

2385S

The Effect of Disclosing Coronary Heart Disease Genetic Risk on Shared-Decision Making. T.M. Kruiselbrink1,2, H. Jouni1, R.A. Haddad3, I.N. Issel1, V.M. Montori1, J.J. Kullo1, 1) Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, MN; 2) Center for Individualized Medicine, Mayo Clinic, Rochester, MN; 3) Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, MN.

Background: Whether disclosure of genetic risk influences shared-decision making for reducing coronary heart disease (CHD) risk is not known. We investigated the effect of disclosing genetic risk for CHD on shared decision making during a patient-physician encounter. Methods: The Myocardial Infarction Genes (MI-GENES) study randomized participants 40-65 y old, at intermediate CHD risk, and not on statins, to receive the 10-y risk of CHD based either on conventional risk factors alone (CRF) or CRF plus a 27 SNP genetic risk score (GRS) (CRF*GRS). CHD risk was disclosed by a genetic counselor during a 30 min scripted session. Each participant then met with a physician to engage in shared-decision making regarding the need for statin therapy. This encounter was facilitated by the Stanford Choice decision aid (Weimiller et al, Arch Intern Med, 2007) modified to display both conventional and genetic CHD risk, with and without statin therapy. Afterwards, study participants were asked to complete a validated 11-question shared decision-making questionnaire (SDM-Q). Responses to each question were transformed to agree (1) and disagree (0) yielding a shared decision making score ranging between 0-11. Results: To date, 122 participants have completed the second study visit (mean age 59.1±5y, 48.3% male). CHD risk (risk =8.4±3.4%) estimated from CRF was communicated to 62 study participants (59±5.3y, 50% male) and 60 participants (59.1±4.7y, 53.3% male) received CHD risk estimated from CRF*GRS (risk = 9.2±7.4%). There was no difference in shared-decision making perception in the two groups (10.7±1 vs. 10.4±1.9, P=0.34). Specifically, there was no difference in participants’ perception of adequate consultation time between the two groups: CRF vs. CRF*GRS (62 vs. 58.7% vs. 62 (48.3%), P=0.23, respectively). Incorporating the GRS in the decision-making process did not affect participants’ perception of sufficient discussion of different treatment options including use of statins vs. other medications or lifestyle modifications, selection of the treatment, advantages of each treatment, and feeling of inclusion in the treatment decision (P<0.05 for all). Conclusion: We demonstrate that disclosure of genetic risk for CHD does not affect participants’ perception regarding shared-decision making. Participants in both study groups perceived that they had sufficient consultation time to discuss their CHD risk, different treatment options, and selection of treatment.
2386S
Adopting Genomes - Motivations of Adopted Persons when seeking Personal Genomic Services. N.M. Baptista1,2, D.A. Carere2, J.R. Dugan1, T.A. Moreno2, J.L. Mountain1, S.A. Bradley2, J.S. Roberts2, R.C. Green1, the PGen Study Group. 1) Division of Genetics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 02115; 2) School of Medicine, Gold Coast Campus, Griffith University, QLD 4222, Australia; 3) Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115; 4) Pathway Genomics Corporation, 4045 Sorrento Valley Blvd., San Diego, CA 92121; 5) 23andMe, Inc. 1390 Shorebird Way, Mountain View, CA 94043; 6) Department of Health Behavior and Health Education, University of Michigan School of Public Health, 1415 Washington Heights, Ann Arbor, MI 48109. Background: The use of personal genomic testing (PGT) by adopted individuals is largely unstudied. Using data from the Impact of Personal Genomics (PGen) Study, a longitudinal study of new PGT customers of 23andMe and Pathway Genomics, we conducted an exploratory analysis to characterize adoptees in the PGen population, and examined their motivations for seeking testing and their satisfaction with PGT results. Methods: PGen Study participants were surveyed before receiving results (baseline), and 2 weeks and 6 months after receipt of results. Of the 1,464 baseline survey respondents eligible for follow up, 1,042 completed the 6 month survey. At baseline, participants rated the importance, on a 5-point scale (“very important,” “somewhat important,” “not at all important”), of 12 possible motivations for seeking PGT. At 6 months, satisfaction with PGT results was measured by endorsement of 6 statements on a 5-point scale from “strongly disagree” to “strongly agree.” Associations between adoption status and each motivation or satisfaction statement were analyzed using chi-square tests and logistic regression, adjusting for age, gender, race, education, company, and prior PGT. In all analyses, responses were dichotomized to “very important” and “somewhat/not at all important” for motivations, and to “strongly/somewhat agree” and “neither agree nor disagree/strongly somewhat disagree” for satisfaction statements. Results: Of 1,042 6 month survey respondents, 57 (5%) were adoptees. In comparison to non-adoptees, adoptees were younger (42±13 years vs. 47±16 years, p<01), had lower education (63% vs. 81% college degree or higher, p<.01). No significant difference was found between adoptees and non-adoptees in gender, race, income, or prior PGT. Adoptees were more strongly motivated to seek PGT because of their limited family health history (86% vs. 32% considered it “very important,” p<.0001), and learned their personal risk of specific diseases (OR=2.2, 95%CI 1.1-4.4, p<.03). Adoptees were no more satisfied than non-adoptees with the health information they gained, but were more likely to believe that seeking PGT was a wise decision (OR=4.7, 95%CI 1.1-19.9, p<.03). Conclusions: Adoptees were younger, less highly educated, and more motivated than non-adoptees to seek genetic testing because of limited family health history and an interest in their disease risk. Belief that PGT was a wise decision may reflect a knowledge gap that adoptees wanted to fill.

2387S
Expert knowledge shapes decision-making for couples receiving uncertain prenatal chromosomal microarray testing results. M.A. Rubel, K. Kellem, F.K. Barg, B.A. Bernhardt. University of Pennsylvania, Philadelphia, PA. Chromosomal microarrays (CMA) are increasingly utilized in prenatal settings to detect copy-number variants (CNVs) of probable or clinical significance not detectable by conventional cytogenetics. Up to 20% of microarray testing results in some kind of positive finding, including both pathogenic and uncertain results, which can complicate decision-making about pregnancy. When faced with ambiguous results, patients and their partners often turn to “expert forms” of “expert knowledge,” through conversations with clinicians, internet resources, medical literature, and personal support networks. This study aims to investigate how couples utilize expert knowledge in the face of uncertain test results to navigate the decision-making process to terminate or continue a pregnancy. This qualitative study documents experiences of couples receiving uncertain prenatal CMA results from diagnosis centers in the US via phone interviews. Men and women separately discuss their experiences with CMA testing, understanding of and emotional response to the results, factors affecting decision making, and needs throughout the testing process. Interview transcripts were coded using NVivo10. Of the 18 male partner/female patient transcripts analyzed to date, most participants receiving uncertain results sought information from multiple sources. Female patients show a propensity to emphasize the impact of information in making their decision, cite various information sources, and seek expert advice from networks of family and friends. After receiving results from clinicians, women are predominantly information-seekers, with many reporting independently locating relevant clinical literature through internet searches. While men and women both reported finding expert knowledge helpful, some noted a fine-line between feeling informed by versus anxious from expert knowledge. In cases where uncertain results were not bolstered by additional understanding, expert knowledge could become “toxic knowledge,” mediating greater anxiety and emotional hardship. Male partners refer to the results and decisions as uncertain with greater frequency and tend to primarily rely on clinicians and their female partner to find and disseminate results, often seeing their role to be the primary emotional support. This study highlights both partners’ varying utilization of expert knowledge as critical factors in the decision-making in a pregnancy with uncertain outcomes.

2388S
A genomic decision aid linked to the electronic health record to disclose coronary heart disease risk and enable shared decision-making. K. Shamerer1, H. Jouini2, R. Chaudhry2, A.K. Dalenberg3, V.M. Montori3, I.J. Kullo1. 1) Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, MN; 2) Division of Primary Care Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN; 3) Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, MN. Background: The myocardial infarction genes (MI-GENES) study is a randomized controlled trial of disclosing genetic risk of coronary heart disease (CHD) leveraging the electronic health record (EHR). We created a genomic decision aid integrated within the EHR to disclose CHD risk to study participants and enable shared decision-making regarding statin therapy to reduce risk. Methods: The Generic Disease Management System (GDMS), a web-based clinical decision support system used at the point of care and integrated within the EHR, was configured to estimate the 10-y CHD risk based on conventional risk factors extracted from the EHR. A link within GDMS provides access to the MI-GENES decision aid which has been proven effective for CHD risk communication (Weymiller et al, Arch Intern Med, 2007). This tool aids with initiation of statin therapy depicting not only the 10-y risk for CHD but also the potential benefit of using statins and their related costs/side effects. A feature was added to the tool enabling the genetic counselor, physician, and patient to visualize the change in 10-y CHD risk after incorporating the genetic risk score (GRS) based on 27 genetic susceptibility loci. CHD risk was disclosed by a genetic counselor in a 30 min scripted session using a pictogram that displays how many patients with the same risk profile will suffer a myocardial infarction in the ensuing 10 years. Afterwards, the participant meets with a physician for shared decision-making regarding the need for statin therapy. The tool is also equipped with a report generating function and a “frequently asked questions” page for additional asks. Results: The tool is being used for disclosure of genetic risk of CHD in the MI-GENES study. So far, we have randomized 212 participants (47% males, mean age 58.8±5.1y) to receive the 10-y probability of CHD estimated by conventional risk factors vs that estimated from conventional risk factors plus a GRS. Patients are surveyed for satisfaction with genetic counseling and physician encounters, understanding of genetic risk, and shared decision-making. Conclusion: In summary, we have developed and successfully implemented a genomic decision aid integrated within the EHR to disclose CHD risk and enable shared decision-making in the clinical setting. The results of the study will facilitate development of best practices for incorporating probabilistic genetic risk for common diseases in the EHR.
2389S

Background: Women with an estimated >20% lifetime risk of breast cancer are candidates for more aggressive clinical management, including screening at younger ages, at more frequent intervals, and with more sensitive technologies, i.e., breast MRI. Family history is a key component of most models for estimating breast cancer risk, but family history analysis may be insufficient to identify at-risk individuals carrying moderately penetrant pathogenic mutations due to limited sibship sizes in contemporary families. We utilized pedigree simulation to estimate the probability that a female proband, who is a carrier of a pathogenic mutation conveying a moderate increase in breast cancer risk, will be identified as having at least a 20% lifetime breast cancer risk as determined by the Claus model. Methods: The SIMLA and SLINK pedigree programs were used to simulate 200 three-generation pedigrees each for 2, 3, 4, or 5-member sibships. The proband was assumed to be a 40-year old female carrying one copy of an autosomal dominant pathogenic mutation conveying a ~25% risk of breast cancer to age 80. Simulated pedigrees were one-sided and limited to either the maternal or paternal side segregating the disease allele. Phenotypes were simulated according to age-dependent liability classes modeled from the Surveillance, Epidemiology, and Ends Results (SEER) breast cancer incidence data. Resulting pedigrees were assessed by the Claus model to determine the proband’s eligibility for modified medical management. Results and Discussion: Analysis of simulated pedigrees indicate that <9% of female probands, carrying a pathogenic mutation conveying a ~25% risk of breast cancer, would receive modified clinical risk management based on the Claus model risk assessment. Thus, genetic testing may be critical for identifying individuals carrying pathogenic mutations in moderate penetration breast cancer susceptibility genes who would benefit from increased surveillance, as outlined in current professional society guidelines. Although population screening for moderately penetrant gene mutations may not yet be economically feasible, clinicians may wish to consider broader pan-cancer testing using a panel composed of both high and moderate penetrance genes when screening the patient for other cancer risks, such as colon, endometrial, and ovarian cancers.

2390S
Colorectal cancer screening for people with a family history: should recommendations vary by age? I. Lansdorp-Vogelaar1, S.K. Naber2, K.M. Kuntz2, N.B. Henriksen3, M.S. Williams4, N. Calonge4, K.A.B. Goddard5, D.T. Zallen6, E.P. Whitlock7, T.G. Ganiats8, C.M. Rutter9, E.M. Webber2, A.C.J.W. Janssens10, M. van Ballegooijen1, A.G. Zauber11, 1) Department of Public Health, Erasmus Medical Center Rotterdam, Rotterdam, Netherlands; 2) Department of Health Policy & Management, University of Minnesota, Minneapolis, MN, United States; 3) Group Health Research Institute, Seattle, WA, United States; 4) Intermountain Healthcare, Salt Lake City, UT, United States; 5) Genomic Medicine Institute, Geisinger Health System, Danville, PA, United States; 6) The Colorado Trust, Denver, CO, United States; 7) Center for Health Research, Kaiser Permanente Northwest, Portland, OR, United States; 8) Department of Science and Technology in Society, VirginiaTech, Blacksburg, VA, United States; 9) Department of Family and Preventive Medicine, University of California, San Diego, CA, United States; 10) Department of Epidemiology, Emory University, Atlanta, GA, United States; 11) Memorial Sloan-Kettering Cancer Center, New York, NY, United States.

Objective. People with first degree relatives (FDRs) diagnosed with colorectal cancer (CRC) are at an increased risk for developing CRC. It is therefore recommended to screen them more frequently than the general population. Although the relative risk for colorectal cancer decreases with the age of the person at risk, none of the screening recommendations take this age into account. The aim of this study is to determine the potential benefits of age-specific screening guidelines for people with a family history of CRC. Methods. For the age groups 30-44, 45-49, 50-54, 55-59, 60-64, 65-69 and 70+, we based the relative risk of having 1, 2, 3 and 4 or more FDRs diagnosed with CRC as compared to the general population on a literature review. These relative risks were incorporated into the MISCAN model to estimate costs and effects of colonoscopy screening strategies varying in age range and interval for people with 1, 2, 3, or ≥4 FDRs diagnosed with CRC. For each age and level of FDR, we determined the most cost-effective screening schedule with an incremental cost-effectiveness ratio (ICER) comparable to that of screening in the general population. Because past screening could influence the cost-effectiveness of future screening, we assessed the influence of screening history on the optimal screening interval in a sensitivity analysis. Results. While effectiveness of colonoscopy screening increased with the number of FDRs people have, total costs of care decreased. Colonoscopy screening was even cost-saving in people with a family history. Consequently, screening should start earlier at age 40 (35 for those with ≥4 FDRs) in people with a family history. The optimal screening interval ranged from 5 years for people with 1 FDR, to 1-3 years for people having 4 FDRs or more. Despite the decreasing relative risk with age, the optimal screening interval was similar for all age groups.

Conclusions. Although the relative risk of people with FDRs diagnosed with CRC decreases with age, optimal screening intervals do not vary with age but do vary with number of affected FDRs. People with a family history should be offered screening from age 40 (35 if ≥4 FDRs) with an interval of 5 years in case of 1 FDR, 3 years in case of 2-3 FDRs and 2 years in case of ≥4 FDRs.
2391S

Improving pedigree capture: Development and validation of an interview-optimized iPad app to eliminate the need for paper. J.M. Miller1, M.J Italia2, Ogasawara3, T. Dechene1, A. Wilkens2, C.J. Gaynor2, J. Krantz2, P.S White6,5-1. 1) Center for Biomedical Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA; 3) Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA; 4) Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA; 5) Department of Pediatrics, University of Pennsylvania, Philadelphia, PA; 6) Department of Pediatrics, Cincinnati Children's Hospital, Cincinnati, OH; 7) Department of Biomedical Informatics, University of Cincinnati, Cincinnati, OH.

Large gene panel, exome and whole genome sequencing-based clinical tests produce data at a scale that is time intensive and difficult to interpret. Trio and inheritance analyses are increasingly utilized as a means to reduce the volume of variants that must be reviewed. In addition to the sequence data, these analysis methods gain benefit from well-documented pedigree and family history data. Electronic capture of pedigree and family member data provides a way to use this data for subsequent analysis, review, and use. However, while a number of electronic solutions have been developed, these systems typically do not perform well in a busy clinical setting. Consequently, these systems have not generally supplanted the use of paper-based manual abstraction, especially for pedigree representation, resulting in decreased efficiency and increased risk of transcription error. We hypothesized that a multi-touch iPad application could compete with paper on clinical usability, producing both suitable clinical documentation and structured data useful for inheritance analysis in the laboratory. As part of the NHGRI's Clinical Sequencing and Exploratory Research (CSER) Program, we developed Proband, an iPad app for capturing pedigrees, available in the Apple App Store. The app enables the user to create complex family pedigrees by implementing nomenclature outlined by the Pedigree Standardization Work Group (PSWG). The primary method of user input is a series of intuitive touch gestures. To streamline data entry, the app relies on the user's context to make the appropriate functionality available at the right time. Proband runs on a local data store and generates pedigree results as PDF or XML. We performed a controlled user evaluation study on genetic counselors. Each counselor was asked to recreate a pedigree after viewing a training video. The counselors then completed the System Usability Scale (SUS), a standard ten-question survey used to measure user satisfaction. The mean score was 90, well above the average SUS score. Qualitative responses indicated a functional convenience equivalent to, and for some tasks, exceeding that of paper. Initial user feedback at our hospital as well as other institutions has been very positive. In one instance, a counselor in a busy clinic completely replaced paper with Proband, creating over 100 pedigrees to date. We continue to evaluate and improve Proband for use in clinical diagnostic environments.

2393S

Diagnostic challenges and behavior problems in Rwandan patients with disorders of sex development: Ethical issues in African context. L. Matesa1,2, N. Suzuki1,2, M. Kamijima1,2, J. Caberg1,6, D. Dechene1, A. Wilkens2, G. Gaynor3, A. Krantz2, F. Rutagarama3, O. Karangwa4, A. Gasana5, J.H. Caberg2. 1) Medical Genetics, National University of Rwanda, KIGALI, Kigali, Rwanda; 2) Center for Human Genetics, University of Liege, Belgium; 3) Department of Pediatrics, Rwanda Military Hospital, Rwanda; 4) Department of Urology, Rwanda Military Hospital, Rwanda.

Background Disorders of sex development (DSD) comprise a variety of congenital diseases with anomalies of the sex chromosome, the gonads, the reproductive ducts and genitalia. The most common DSDs result from disruption of androgen levels and activity that affect later embryonal development, such as congenital adrenal hyperplasia and androgen insensitivity syndrome (AIS). DSDs are always challenging and very difficult to manage. Socio-economic and cultural aspects have a great impact on decision making regarding the management of these conditions. The situation is more complicated in resource-poor settings like in Africa, where access to education and medical care is limited in both quantity and quality of infrastructure, diagnostic tools and medical professionals. In addition, traditional values and beliefs are also very strong in various cultures and sexual issues are taboo subjects. Methods The present study is a 5-year prospective descriptive cohort of patients with suspicion of DSDs referred to our genetic clinic between January 2009 and January 2014 for genetic investigations and counseling. All patients underwent abdominal ultrasound or MRI and hormonal analysis before genetic testing including karyotype and molecular tests. Results In total, 49 patients aged between 1 and 39 years were clinically evaluated in the PRG. Of these, 45 females met the criteria for inclusion. Of these, 37 cases came from male-to-female gender assignment. The most common DSDs associated with behavior problems in the majority of cases. The majority were diagnosed either with sexual ambiguity and hypospadias, or microenorrhea, or primary amenorrhea, or poor development of secondary sexual characteristics, or primary infertility. In most cases the primary method of sex assignment was based on the ultrasound and MRI revealed absence of uterus and ovaries. The FSH, LH or testosterone hormones revealed major abnormal values in more than 60% of patients. The AIS and Rokitsansky syndrome were observed in the majority of these patients. The choice of gender identity after karyotyping raised several psychological and ethical issues in the majority of adults patients. The outcome of surgery was successful for the social integration of some of these patients. Conclusion The present study showed that patients with DSDs have major behavior problems in African context. They suffer from extreme anxiety and very high psychological behavior disorders related to their gender identity. Our data suggest that gender assignment has to be avoided before expert and multidisciplinary evaluation especially in young patients.
2394S
Making sense of diagnostic uncertainty after newborn screening for cystic fibrosis. C.J. Barg1, R.Z. Hayeems1,2, F.A. Miller1, Y. Bombard1,2, P. Durie1,2, J.C. Carroll3, P. Chakraborty1,3, B.K. Potter4, K. Tam1,5, L. Taylor1, E. Kerr1,2, D. Davies1, J. Milburn2, K. Keenan4, A. Price1,5, F. Ratjen4,5,11, A. Guttmann1,4,5,14
1) Institute of Health Policy, Management & Evaluation, University of Toronto, Toronto, ON, Canada; 2) Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, ON, Canada; 3) Li Ka Shing Knowledge Institute, St. Michael’s Hospital, Toronto, ON, Canada; 4) Division of Paediatric Medicine, Department of Paediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 5) Department of Pediatrics, Faculty of Medicine, University of Toronto, ON, Canada; 6) Department of Family & Community Medicine, Mount Sinai Hospital, University of Toronto, ON, Canada; 7) Newborn Screening Ontario, Children’s Hospital of Eastern Ontario, Ottawa, ON, Canada; 8) Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada; 9) Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, ON, Canada; 10) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, ON, Canada; 11) Division of Respiratory Medicine, Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 12) Department of Psychology, The Hospital for Sick Children, Toronto, ON, Canada; 13) Department of Pediatrics, London Health Sciences Centre, London, ON, Canada; 14) The Institute for Clinical Evaluative Sciences, Toronto, ON, Canada.

Objectives: Diagnostic uncertainty related to cystic fibrosis (CF) has long been challenging to understand and manage. Newborn screening (NBS) for CF extends these challenges to the early days of life. We explored the parental experience of diagnostic uncertainty arising from NBS. Methods: Drawing from a mixed methods prospective cohort study of screen positive and screen negative control infants, we report on qualitative interviews with parents of children who received inconclusive results for CF after NBS and follow-up testing. Through the tertiary care centre that manages the majority of screen positive care in Ontario, we recruited parents of children identified since the inception of CF NBS in Ontario (2008). We used qualitative descriptive analysis to transcribe data. Results: We conducted 19 interviews with 14 parents of infants with uncertain diagnoses, ranging from 3 months - 4 years in age. Five parents completed interviews at two time-points, separated by one year. We learned that parents gain support through research involvement, but struggle to understand the meaning of an uncertain diagnosis in the face of an apparently healthy newborn, worry about their infant’s health-related vulnerability, and fear labeling and over-medicalization from continued medical surveillance. Time appears to mitigate concerns. Conclusion: The experience of diagnostic uncertainty is deeply challenging for some families, particularly in the early newborn period. These results should inform decisions by NBS programs and clinical teams about protocols for testing, clinical follow-up, and support for families of those who screen positive, but warrant triangulation with survey data and longitudinal follow-up.
2396S
Numeracy, Genetic Knowledge, and Perceived Risk for Coronary Heart Disease in the MI-GENES Study. H. Joung1, I.N. Isseh1, R.A. Haddad1, K.D. Christensen2, R.R. Sharp3, R.C. Green2, I.J. Kullo1. 1) Division of Cardiovascular Diseases, Department of Medicine, Mayo Clinic, Rochester, MN; 2) Division of Genetics, Department of Medicine, Brigham & Women’s Hospital & Harvard Medical School, Boston, MA; 3) Biomedical Ethics Program, Mayo Clinic, Rochester, MN.

Background: Genetic knowledge, numeracy, and risk perception likely influence how patients interpret and respond to genetic test results. We assessed these characteristics in participants of the Myocardial Infarction Genes (MI-GENES) study that is exploring integration of genomic results into the electronic health records. Methods: Participants were residents of Olmsted County, aged 40-65 y, at intermediate risk for coronary heart disease (CHD), and not on statins. They were randomized to receive the 10-y risk of CHD based either on conventional risk factors alone or conventional risk factors plus a 27 SNP genetic risk score. We assessed numeracy, genetic knowledge, and self-perceived risk for CHD at baseline before disclosing CHD risk. Numeracy and genetic knowledge were assessed using previously validated surveys. We adapted the Health Information National Trends Survey (HINTS) to assess CHD risk perception.

Results: The mean age of 212 participants was 58.8±5y, 47% were male, all were non-Hispanic white, 21% had completed high school and 62% had graduated from college or graduate school. At baseline, participants scored highly on numeracy; 42.9% answered all of the 8 questions correctly and 31.1% answered 7 questions correctly. For genetic knowledge, 22% answered ≥14 of 16 questions correctly while most participants (62%) answered between 9-13 questions correctly. The mean perceived 10-y risk was significantly higher than the estimated 10-y risk for CHD (25.6%±20.3% vs 8.5%±4.1%, respectively, P<0.01). Although patients overestimated their risk for CHD, the majority (80%) described their risk as the same or lower than that of healthy peers. Participants with family history of CHD were more likely to perceive increased genetic susceptibility to CHD than participants without such history (63% vs. 26%, respectively, P<0.01). Higher numeracy was associated with lower self-perceived risk for CHD (r =0.2, P=0.004). Genetic knowledge scores were not associated with perceived risk regardless of family history (P=NS). Conclusion: Early adopters of genetic testing for CHD in this study were facile with numbers and knowledgeable about genetics, but greatly overestimated their risk for CHD. Interestingly, numeracy but not genetic knowledge, was inversely correlated with perceived CHD risk. Our results indicate that education and counseling protocols for genetic testing should also address numeracy misperceptions about disease risk.

2397S
Genomic Sequencing in the Infant Population: Exploring Parental Motivations, Expectations and Utilization of Sequencing Results in the Tell Me More Study. F. Facio1, B. Solomon1, S. Haga2, E. Klein1, K. Huddleston3, A. Khromykh1, S. Hull3, B. Berkman2, E. Sutton4, K. Hurley5, J. Evans2, J. Vockley6, J. Niederhuber7. 1) Inova Translational Medicine Institute, Falls Church, VA; 2) Institute for Genome Sciences & Policy, Duke University, Durham, NC; 3) Bioethics Core, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD; 4) Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI; 5) Teachers College, Columbia University, New York, NY; 6) Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC.

Advances in genome sequencing (GS) technologies have increased their use in research and clinical practice. However, many barriers exist to translating GS to routine clinical care, including a lack of data documenting motivations, expectations, and behavioral outcomes as well as parental perspectives and behaviors. We describe findings from a large, prospective trio-based whole GS study within a multi-hospital health system (~20,000 annual deliveries, half of which are at the main study hospital). Approximately 10% of these deliveries enroll in the longitudinal GS study, largely limited by full trio availability. We have enrolled ~1500 trios to date (we accrue ~80-100 trios/month). Reasons for participation include potential benefits related to finding medically relevant information, altruism, lack of cost, enthusiasm regarding science and genomics, and the fact that the study is conducted directly through the same health care system. Reasons for study decline include unwillingness to undergo phlebotomy (especially fathers due to the need for an additional blood draw), privacy concerns and questions related to government access to genetic/genomic and other health information, concerns about potential effects on insurance coverage, and stigma associated with research, especially related to genetics/genomics. Focus groups of study participants have shown an overwhelming desire to know both specific and general genetic/genomic information. As an initial step, we have started returning CLIA-validated pharmacogenetic results within the focus groups to determine the impact of return of these results. To explore the sociobehavioral aspects of GS further, we have also designed the Tell Me More study, which aims to investigate parental motivations for enrolling or declining GS, their expectations about GS results, and the use and impact of such information. This nested study is recruiting a total of 40 participants - 20 parents who opted to pursue GS for their newborn (intervention group) and 20 who declined (control group), and uses phenomenological methods to explore parental experiences and decision-making processes. Parents in both groups complete interviews post informed consent and parents in the intervention group complete interviews post disclosure of results. Ultimately, this study will generate results that can be used as building blocks for larger projects and inform best practices in clinical applications of GS in the infant population.
2398S
Supporting the International Rare Diseases Research Consortium: achievements and challenges. P. Lasko1,2, B. Cagniard3, S. Höhn1, L. Lau1, S. Peixoto4, S. Ayme2, IRDiRC Consortium. 1) International Rare Diseases Research Consortium (IRDiRC) brings together members that have agreed to common goals and principles and to work in a coordinated and collaborative manner within a multinational consortium to advance research in this critical area. Its over 35 members are public and private research funding organizations and companies that have each dedicated over 10 million US$ to research into rare diseases. Present members are from Europe, North America, Asia and Australia. IRDiRC’s two main objectives are to deliver 200 new therapies for rare diseases and the means to diagnose most rare diseases by the year 2020. To accomplish this, IRDiRC engages in a range of initiatives. The first one is to support financially strategic research projects. The funding agencies remain independent in their calls for proposal but discuss jointly their strategy and agree on main areas to be supported. The second instrument is to identify and highlight strategic infrastructures which could be used by the researchers to speed up their projects. The third is to identify and promote standards to be used to make data interoperable and as easy to access and share as possible. Many working groups have been put in place to advise the consortium in these and other areas, and their suggestions are reviewed by three scientific committees who directly advise the executive board. A label “Recommended by IRDiRC” has recently been created to identify preferred tools, protocols, and guidelines that directly contribute to IRDiRC goals. Consortium achievements are posted in a timely manner on its website (www.irdirc.org), enabling all the activities of its many committees and working groups to be as transparent as possible. As part of its dissemination and communication plan, IRDiRC publishes feature articles on strategic research projects and infrastructures and offers on its website a space for its members to communicate. It organized an International conference in Dublin in 2013 which was the real public launch of this initiative. The next IRDiRC conference will be focused on fostering international collaboration with Asia to be held in Shenzhen (China) on 7-9 November 2014 thanks largely to support from BGI.

2399S
Assessment of factors that should be addressed in prenatal counseling for non-invasive prenatal test. L. Wang1, Y. Gu1, Q. Meng1, X. Tang1, H. Wang1, S. Yang1, H. Mao1, F. Liu1, J. Zhang2, H. Wu1, Q. Shi1, N. Zhong2. 1) Department of Prenatal Diagnosis, Lianyungang Maternal and Child Health Hospital, Lianyungang, China; 2) Department of Human Genetics, New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY.

Objectives: To investigate the potential factors affecting the women’s attitude toward using a non-invasive prenatal test (NIPT) as well as knowledge of prenatal diagnosis in pregnant women, and the clinical efficiency of identifying fetal chromosomal aneuploidies in pregnant women using an NIPT.

Method: A survey of the potential factors (age, education, and residential area) affecting pregnant women’s choice of prenatal diagnosis and their understanding of prenatal diagnosis that was conducted in our hospital. Moreover, 917 women with high-risk pregnancies were invited to participate in an NIPT trial. Complete karyotyping of the amniotic fluid was performed to validate each abnormal case identified in NIPT. Results: Women older than 35 years, or 25 years, were statistically significantly inclined to choose NIPT as their prenatal diagnosis procedure (P < 0.05). The percentages of women knowing of the purpose, advantage, limitation and costs of NIPT were less than that of amniocentesis. Moreover, an average of 24.46% women understood the indications for NIPT. The NIPT trial with validation by full karyotyping of the amniotic fluid identified 3.60% high-risk pregnancies with fetal aneuploidies, including 2.73% trisomy 21 (Tr18), 0.33% trisomy 18 (Tr18), 0.44%Turner syndrome (45, X), and 0.11% Klinefelter’s syndrome (47, XXY). In addition, in the NIPT, one woman with maternal mosaicism (45, X and 46, XX) was found, and the other cases of false positive results found with NIPT caused by confined placental mosaicism (CPM) were also identified after karyotype validation. Conclusions: Pregnant women’s age was the primary factor affecting women’s choice for NIPT, and more education on NIPT is necessary to improve the clinical application of NIPT. Meanwhile, discordant positive results found with NIPT can be caused by confined placental mosaicism and maternal mosaicism, which underlines the need to confirm all positive NIPT results via conventional invasive testing and karyotyping at NIPT’s current stage of development.

2400S
Genetic Counseling for Psychiatric Disorders: Exploring the Psychiatric Health Professionals’ Perspective in the United Kingdom. S.E. Jenkins1,2, M. Arribas-Aylon1, 1) Wessex Clinical Genetics Service, Southampton, Hampshire, United Kingdom; 2) Institute of Medical Genetics, Cardiff University, Cardiff, Wales, United Kingdom.

Background: Historically genetic counselling is not offered for psychiatric disorders in the United Kingdom through regional clinical genetics departments provided by the NHS. The purpose of genetic counselling through these specialist centres is to provide patients and their families with information about the nature and consequences of a disorder, in addition to the probability of developing symptoms of the disorder and passing it on to future generations. The options available to the patient and their families in management and treatment are discussed, and in some cases genetic testing can then be explored. Recent genomic advances identifying a greater genetic contribution to mental illness is anticipated to increase the demand for genetic psychiatric counselling and is expected to cause significant changes in the management and treatment of these disorders.

Methodology: The study employed both quantitative and qualitative methods of research to explore the attitudes and experiences of psychiatric health professionals with respect to identifying a potential for genetic counselling services within clinical psychiatry in the UK. Results: Data analysis revealed the following emerging themes: -Demand for genetic counselling; -Responsibility for genetic counselling provision; and -Barriers for the service. Evaluation of the data led to the conclusions that although demand for psychiatric genetic counselling has not been voiced in the UK at present, psychiatric health professionals believe that such a service would be a useful and desirable. Genomic advances which identify susceptibility loci for psychiatric disorders may have significant implications for genetic counselling in clinical psychiatry if these discoveries lead to genetic testing. Psychiatric health professionals describe clinical genetics as a skilled profession capable of combining complex risk communication with much needed psychosocial support. However, the possibility of such a service is confronted with a range of barriers and challenges including, but not limited to, the complexities of uncertainty in psychiatric diagnoses, patient engagement and ethical concerns regarding limited capacity and increased suicidality.

2401S
Autism Spectrum Disorder in Taiwan: Parents’ report. J. Ye1, T.Y Huang2, S.X Zhao1, L. Xu1, E. Jung2, Y.Y Wu1, D. Tsai2, L.S. Chen3. 1) Department of Health and Kinesiology, Texas AM University, College Station, TX; 2) Department of Special Education, National HsinChu University of Education, HsinChu, Taiwan; 3) Department of Educational Psychology, Texas AM University, College Station, TX, USA; 4) Department of Psychiatry, Chang Gung Memorial Hospital- Linkou Medical Center, Chang Gung University College of Medicine, Tao-Yuan, Taiwan; 5) Department of Social Medicine, National Taiwan University College of Medicine, Taipei, Taiwan.

Background: In the United States, one in 50 school-age children are affected by Autism Spectrum Disorder (ASD). As a global challenge, limited research has been conducted on ASD outside the U.S. This study reports preliminary findings and characteristics of children with ASD at pre-schools and elementary schools in Taiwan. Methods: A total of 443 parents of children with ASD, recruited from pre-schools and elementary schools of Hsinchu city and county, Taoyuan and Miaoli counties, completed the paper-and-pencil survey. Results: The average maternal and paternal ages at the time of ASD child’s birth were 29.47 years (SD=4.6) and 33.28 years (SD=5.33), respectively. The majority of the affected children were predominantly boys (6.9:4.1). The age of confirmatory diagnosis for ASD was 4 years (SD=2.2). About one-fifth (18.5%) of the parents were unaware about the classification of ASD that affected their child. Most of the parents (73.5%) reported that they received either very low or low support from the society for their children. Conclusions: Our study indicated the average age when children were diagnosed with ASD was beyond the recommended age of two. Some parents were ignorant about the type of ASD associated with their child’s ASD and most parents reported low support from the society. The possibility of such a service is confronted with a range of barriers and challenges including, but not limited to, the complexities of uncertainty in psychiatric diagnoses, patient engagement and ethical concerns regarding limited capacity and increased suicidality.
Screening tests for carriers of recessive disorders interrogating specific mutations in selected genes using SNP-genotyping approaches have been developed. By contrast, next generation sequencing (NGS) technologies enable deep gene sequencing in a cost-efficient manner. We developed an NGS-based approach targeting 215 genes causing prevalent and severe recessive diseases for testing in vitro fertilization (IVF) couples and donors/recipients in donation programs. We measured overall sensitivity and specificity of different known genetic mutations present in a blinded training set. Bioinformatics analysis is keystone in the process, as it combines several algorithms optimized for the identification and annotation of different types of mutations (point mutations, indels, copy-number and rearrangements). Our experience with NGS gene panels showed extremely high sensitivities (>99%) for all kinds of mutations. For carrier test validation, we obtained DNA from 57 unrelated individuals: 39 patients and 18 previously genotyped controls. The validation set was composed of 49 different known mutations (67 in the patients), including 29 SNVs, 13 indels and 25 CNVs causing different diseases: cystic fibrosis, phenylketonuria, spinal muscular atrophy, hypothyroidism, thalassemia, factor V deficiency and Duchenne muscular dystrophy. All but one (48/49) different mutations were correctly scored in the blinded study and only one deletion-type mutation remained undetected. This information allowed us to finely tune the algorithm to reach maximum sensitivity. All single nucleotide changes were validated and no known recessive mutations were called in the control samples. To date, we evaluated the spectrum of mutations found in 113 patients referred for carrier testing by IVF clinics. We have identified 62 carriers (54%) and a total of 56 different mutations. We detected multiple carriers for common mutations in the European population but we also identified rare mutations and novel deleterious variants that are predicted to be pathogenic (nonsense, frameshift mutations and mutations at splice site positions), as well as collected evidences for discrepancies between disease prevalence and population frequency of apparently pathogenic mutations. A significant proportion of the identified mutations, specially the novel mutations, would have not been detected by traditional methods. NGS has higher detection rates resulting in lower residual risks.
Posters: Health Services Research

587

2403S

Inherited cancer predisposition in children: challenging issues faced by a genetic clinic in a pediatric oncology hospital. F.T. de Lima1,2, V.F.A. Melo3, C.R.P.D. Macedo2, N.S. da Silva2, A.M. Cappellano2, R.V. Gouveia2, E.M.M. Caran1. 1) Gynecology and Obstetrics Department, UNIFESP-EPM, São Paulo, São Paulo, Brazil; 2) Pediatric Oncology Department - GRAACC (Support Group for Children and Adolescents with Cancer), UNIFESP-EPM, São Paulo, São Paulo, Brazil; 3) Genetics and Morphology Department, UNIFESP-EPM, São Paulo, São Paulo, Brazil. INTERVENTION: Genetic predisposition in childhood cancer is variable, and the access of children with neoplasms to genetic services skilled in hereditary cancer and dysmorphology is limited, even in well-developed countries. OBJECTIVES: We described the structure, the patients’ clinical characterization and challenges faced by a new genetic clinic, in a public pediatric oncology hospital in Brazil. METHODS: We reviewed the initial proposal for the clinic’s structure, the protocols and charts of all patients seen from August 2012 to December 2013, as well as the patient’s database. RESULTS: The initial structure was centered in patients’ care, with standard initial clinical evaluation protocol, and follow-up protocols for specific conditions, e.g., neurofibromatosis (NF) and tuberous sclerosis (TS). Appointments occur once a week. The main reason for referral was a diagnosis with known genetic implications, followed by association of cancer and malformations. Other reasons included cancer family history and concerns about recurrence in another child. Almost half of the 172 patients (152 families) seen had NF (46/26.7%) or retinoblastoma (41/23.8%); 37 (21.5%) patients had other diagnosis. The exact diagnosis was not obtained in 48 patients (27.9%), classified as suspected malformation syndromes associated to cancer (25), or suspected familial/hereditary cancer (23). Most were seen just on in 113/65.7%). Issues faced included an appreciable number of patients needing evaluation, lack of knowledge about the goals of genetic counseling and about family history. To overcome some, a folder with orientations was created for easy referral. The need for a stronger and closer interaction with the clinical staff was also observed, and periodic grand rounds were proposed. Due to the hospital structure, focused on childhood cancer care, the identified at-risk family members were referred to external health care. In Brazil, there were no public provision of genetic testing, and as a result patients cannot afford their genetic counseling and their sample stored on a biobank, with all legal and ethical policies being observed, hoping for a change in this situation. CONCLUSIONS: A genetic cancer clinic in a pediatric oncology hospital, focused on care, research and education, helps the identification of a specific tumor predisposition or a congenital malformation syndrome in children with cancer, adding value on patient’s management and counseling of at-risk relatives.

2404S

Implementing a universal public health policy for the care of rare diseases in Brazil. D.D.G. Horovitz1, M.J.B. Aguilar1, V.E.F. Ferraz2, M.F. Galera3. 1) Centro de Genética Médica, Instituto Fernando Figueira / Fiocruz, Rio de Janeiro, RJ, Brazil; 2) Faculdade de Medicina da Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil; 3) Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil. INTRODUCTION: Genetic predisposition in childhood cancer is variable, and the access of children with neoplasms to genetic services skilled in hereditary cancer and dysmorphology is limited, even in well-developed countries. OBJECTIVES: We described the structure, the patients’ clinical characterization and challenges faced by a new genetic clinic, in a public pediatric oncology hospital in Brazil. METHODS: We reviewed the initial proposal for the clinic’s structure, the protocols and charts of all patients seen from August 2012 to December 2013, as well as the patient’s database. RESULTS: The initial structure was centered in patients’ care, with standard initial clinical evaluation protocol, and follow-up protocols for specific conditions, e.g., neurofibromatosis (NF) and tuberous sclerosis (TS). Appointments occur once a week. The main reason for referral was a diagnosis with known genetic implications, followed by association of cancer and malformations. Other reasons included cancer family history and concerns about recurrence in another child. Almost half of the 172 patients (152 families) seen had NF (46/26.7%) or retinoblastoma (41/23.8%); 37 (21.5%) patients had other diagnosis. The exact diagnosis was not obtained in 48 patients (27.9%), classified as suspected malformation syndromes associated to cancer (25), or suspected familial/hereditary cancer (23). Most were seen just on in 113/65.7%). Issues faced included an appreciable number of patients needing evaluation, lack of knowledge about the goals of genetic counseling and about family history. To overcome some, a folder with orientations was created for easy referral. The need for a stronger and closer interaction with the clinical staff was also observed, and periodic grand rounds were proposed. Due to the hospital structure, focused on childhood cancer care, the identified at-risk family members were referred to external health care. In Brazil, there were no public provision of genetic testing, and as a result patients cannot afford their genetic counseling and their sample stored on a biobank, with all legal and ethical policies being observed, hoping for a change in this situation. CONCLUSIONS: A genetic cancer clinic in a pediatric oncology hospital, focused on care, research and education, helps the identification of a specific tumor predisposition or a congenital malformation syndrome in children with cancer, adding value on patient’s management and counseling of at-risk relatives.

2405S

Effect of Co-Payment on Behavioral Response to Consumer Genetic Testing. J. Outlaw1, L. Arinelli2, W. Liu3, D. Boeldt1, N. Wineinger1, N.J. Schork4, E.J. Topol5, C.S. Bloss1. 1) Rady School of Management, UCSD, San Diego, CA; 2) Scripps Translational Science Institute, La Jolla, CA. Existing research in consumer behavior suggests that a person’s perception and usage of a product post purchase may depend on the price paid. As improvements in technology make genetic testing cheaper and more accessible to increasing numbers of consumers, it is important to consider the effect marketing practices may have on the adoption of and response to testing. In particular, it is unclear how the price paid for a genetic test may influence the consumer’s subsequent behavior. In the current study, we examined the effect of price paid for a direct-to-consumer (DTC) genetic test on consumer post purchase behavior. Participants were enrolled in the Scripps Genomic Health Initiative (SGHI), a longitudinal cohort study originally designed to determine the impact of DTC genomic testing on consumers. Co-payment amounts for the test varied from $150 to $470, representing a price increase over time to encourage early enrollment. A small proportion of participants received fully subsidized testing from their employer. Of the subjects recruited for the study, 2,037 completed behavioral and psychological assessments before and approximately 5.6 months after genomic testing. Results indicated that with respect to clinical screening test completion, consumers that paid for the genetic test were significantly more likely to both obtain follow-up genetic tests (p<0.05), compared to those whose test was fully subsidized. Prior to receipt of their genetic test results, neither group differed in their propensity to visit a physician. In addition, neither group showed any significant changes in diet, exercise, or anxiety after receiving genetic risk results. The out-of-pocket price paid for DTC genetic tests may influence how results are interpreted and acted upon by consumers. Cognitive dissonance theory may help explain the increase in screening propensity for paying consumers. Such individuals may obtain follow-up screenings as a way of confirming the value of the test and validation of their initial decision to expend personal resources to obtain it. Understanding of co-payment effect on behavioral response to genetic testing may help shape public policy in this rapidly evolving area of health care.
2408S Cree Leukoencephalopathy and Cree Encephalitis Carrier Screening Program: Evaluation of Knowledge and Satisfaction of High School Students

**Background:** Cree Leukoencephalopathy (CLE) is a neurodegenerative condition with autosomal recessive transmission, characterized by progressive neurological degeneration, including cognitive decline, motor dysfunction, and psychiatric symptoms. Despite its prevalence among the Cree population in Northern Quebec, Canada, awareness and knowledge among high school students are limited.

**Aim:** To assess knowledge and satisfaction of Cree high school students regarding CLE through a carrier screening program.

**Methods:** A carrier screening program on CLE was implemented at three Northern Quebec high schools (Northern Quebec, Canada). The program included education sessions, a pre-test, and a post-test survey to assess knowledge and satisfaction.

**Results:** Of 224 eligible students, 90 (40%) consented to the carrier screening (A). Of these, 74% answered survey A, and of those 42 answered survey B. Ages ranged from 14-20 years. Seventeen (19%) were either pregnant or already had at least one child. Eleven (12%) reported a positive family history for CLE or Cree. When comparing groups who answered both surveys, improvement was observed in knowledge and satisfaction with the CLE-CLE CSP.

**Conclusion:** The Cree CLE CSP effectively increased student knowledge and satisfaction, promoting awareness of CLE and its impact on Cree communities.

**Authors:** Y. LeClair-Blain, V. Gosselin, A. Beardskin, A. Tangen, J.A. Mitchell, B.J. Wilson, A. Richter, A.M. Labeuge, J.E. Torrie, and S. Laberge.

2407S Utilizing telemedicine to support informed decision making and an expanded access to cancer genetic services in community clinics

**Background:** Given high demand and limited workforce, many patients do not have access to genetic providers. Telemedicine (TM) has been used to expand services to low access populations and could increase the use of pre-test counseling, informed decision-making, and appropriate use of genetic services.

**Methods:** A population-based carrier screening program (CSP) started in 2006 for CLE carrier screening in the James Bay Cree communities (Northern Quebec, Canada). Surveys were distributed to Cree high school students (14 years) and women of reproductive age and their partners, in pre- and post-tests (V1 and V2) using TM.

**Results:** Of 100 patients approached, 83% consented to the carrier screening. Of those, 83% had disconnections but were completed by TM. Among 34% of patients reported technical difficulties, 94% were satisfied with the carrier screening process.

**Conclusion:** Telemedicine delivery of cancer genetic services is feasible, increases knowledge, identifies carrier mutations and family history related cancer risk, and is associated with high satisfaction, suggesting an innovative model for delivery in community practices without institutional access to genetic providers.

**Authors:** J. Le Clue-Blain, V. Gosselin, A. Beardskin, A. Tangen, J.A. Mitchell, B.J. Wilson, A. Richter, A.M. Labeuge, J.E. Torrie, and S. Laberge.
2409S

Determinants of the value of genetic testing in clinical decision-making. B. Lerner1, N. Marshall2, S. Oishi1, A. Lantor1, A.B. Hamilton2,3, E. Yano2,3, M.T. Scheuner1,2,3,4
1) Veterans Hospital Administration, Boston, MA; 2) VA Greater Los Angeles Healthcare System, Sepulveda, CA; 3) VA Greater Los Angeles Healthcare System, Los Angeles, CA; 4) David Geffen School of Medicine, University of California, Los Angeles, CA; 5) UCLA Fielding School of Public Health, Los Angeles, CA.

Objective: The value of genetic testing in clinical decision-making depends on factors within the environmental, organizational, provider, patient and encounter domains. We sought to identify determinants within those domains that are associated with favorable value ratings for the use of genetic tests for six clinical indications. Methods: We administered a key-informant Web-based survey to clinical chiefs of neurology, medical oncology, gastroenterology, cardiology and primary care at Veterans Health Administration facilities. The value of genetic tests was rated on a 5-point scale, then dichotomized into "not at all, minimally and moderately valuable" and "very to extremely valuable." Bivariate and logistic regression analyses identified associations between determinants and the value ratings for each indication. Results: The response rate was 63%; (n=353). Genetic testing was rated as very-extremely valuable for the indications: inform clinical management (56.6%), assist with reproductive options (50.1%), assist with life planning (43.9%), confirm a suspected diagnosis, (39.9%), and confirm an established diagnosis (32.3%). The most frequently identified determinants associated with favorable value vs. unfavorable ratings regardless of indication were a culture that fosters genetic testing (ranging from OR 4.8, CI 2.3-9.9, p<0.001 to OR 2.3, CI 1.2-4.6, p<0.01) and the availability of genetic testing guidelines (ranging from OR 2.8, CI 2.3-5.9, p<0.001 to OR 1.9, CI 1.1-3.2, p<0.05) to the presence of a genetic counselor (OR 2.4, CI 1.2-4.5, p<0.01) and a genetic counselor who is part of the genetic counseling team (OR 2.2, CI 1.3-3.6, p<0.001); provider expertise using primary care as the reference (ranging from OR 0.3, CI 0.2-0.5, p<0.01 to ability to OR 2.8, CI 1.1-7.2, p<0.5), gain genetic consultations (OR 2.4, CI 1.2-4.7, p<0.01), the ability to order genetic tests (OR 2.6, CI 1.3-5.2, p<0.01); and patient demand for testing (ranging from OR 1.8, CI 1.0-3.2, p<0.05 and OR 1.6, CI 0.94-2.9, p<0.1). Conclusion: We identified several determinants associated with the value of genetic tests in patient care. These findings support the need to strengthen genetic counseling efforts within healthcare organizations adopting genetic testing services. Such programs should focus on strengthening a culture that facilitates testing, promoting guidelines for genetic testing, educating providers and patients about genetic testing, and ensuring access to genetic counseling.

2410S

Lean - production management rules applied to a genomics core facility. J. Altmtller1, C. Becker2, P. Nürnberg1, 1) Cologne Center for Genomics, Cologne, Germany; 2) Institute for Human Genetics, University of Cologne, Germany.

Lean management is a production management systems approach, originally coming from the automobile company Toyota. It was reviewed to have an enormous value creation per employee, superior to anything known before. Starting 25 years ago, this success model was investigated in the western world, standardized, adopted, and expanded in many fields of production. The Next Generation Sequencing platform with its large variety of applications, fast changing protocols, and comparatively small projects intensified the need for professional workflow and capacity planning and quality control implementation. Adopting lean principles, our core facility steadily improves service orientation, data quality and cost effectiveness. With this presentation we will show how lean management tools can help both research and diagnostic labs to maximize flexibility, minimize waste, and reduce mistakes.

2411S

Newborn screening for cystic fibrosis: role of primary care providers in caring for screen positive children. J.C. Carroll1, R.Z. Hayeems1,2, F.A. Miller3, C.J. Barg4, Y. Bombard1, P. Durie2,5, P. Chakraborty2,6, B.K. Potter1, J.P. Bytautas1,10, K. Tann1, L. Taylor12, E. Kerr13, C. Davies2, J. Milbum1, F. Ratjen5,6,12, A. Guttman1,2,5,6,14
1) Family & Community Med, Mount Sinai Hospital University of Toronto, Toronto, ON, Canada; 2) Institute of Health Policy Management and Evaluation, University of Toronto, Toronto, Canada; 3) Child Health Evaluative Sciences, The Hospital for Sick Children Research Institute, Toronto, Canada; 4) Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada; 5) Division of Paediatric Medicine, Department of Paediatrics, The Hospital for Sick Children, Toronto, Canada; 6) Department of Pediatrics, Faculty of Medicine, University of Toronto, Toronto, Canada; 7) Newborn Screening Ontario, Children's Hospital of Eastern Ontario, Ottawa, Canada; 8) Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, Canada; 9) Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Canada; 10) Division of Health Care and Outcomes Research, Toronto Western Research Institute, Toronto, Canada; 11) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Canada; 12) Division of Respiratory Medicine, Department of Pediatrics, The Hospital for Sick Children, Toronto, Canada; 13) Department of Psychology, The Hospital for Sick Children, Toronto, Canada; 14) The Institute for Clinical Evaluative Sciences, Toronto, Canada.

Context: Expanded newborn screening (NBS) has increased positive screening results, prompting attention to the role of primary care providers (PCP) in providing care to children who screen positive for genetic disorders. Objective: To explore PCPs' role in caring for children who receive a positive CF NBS result in their practice in the previous 6 months. Intervention: Mailed questionnaire inviting about their role in caring for children with CF, uncertain CF diagnosis or are CF carriers. Participants: PCPs in Ontario, Canada, identified by Newborn Screening Ontario as having a positive CF NBS result in their practice in the previous 6 months. Only 50% completed surveys. Survey respondents were: 65% family physicians/nurse practitioners, 21% pediatricians, 14% midwives. Most PCPs (81%) agreed they have an important role to play in NBS. For infants confirmed to have CF, for routine well-baby care: 22% PCPs would provide total care, 69% share with a specialist, 9% refer to a specialist completely; for infants with uncertain CF diagnosis: 49% would provide total care, 46% share with a specialist, 5% refer to a specialist completely and for CF carriers: 89% would provide total care, 6% share with a specialist, 2% refer to a specialist completely. Most PCPs were extremely/very confident (54%) or moderately confident (35%) in providing reassurance about the health of CF carriers. Only 24% knew how to order CF carrier testing for parents, 66% knew how to refer for prenatal diagnosis. Conclusion: The majority of PCPs are willing to treat infants with a range of CF NBS results in some capacity. It is concerning that 11% indicated that carriers should have specialist care involvement and only 54% were very confident reassuring about carrier status. This raises issues about the possible medicalization of carrier status, prompting the need for specific education for PCPs about genetic disorders and the meaning of genetic test results.
2412S
An assessment of clinician and researcher needs for support in the era of genomic medicine. C.A. Brownstein1,2,5, S.K. Savage1, S.I. Zniev2,3,4, I.A. Holm1,2,3, J. Stoler1,2, D.M. Margulies1,2,4,7, 1) Genetics-Research Connection, Boston Children’s Hospital, Boston, MA; 2) Department of Pediatrics, Harvard Medical School; 3) Division of Adolescent/Young Adult Medicine, Boston Children’s Hospital; 4) enter for Patient Safety and Quality Research, Program for Patient Safety and Quality, Boston Children’s Hospital; 5) Manton Center for Orphan Disease Research, Boston Children’s Hospital; 6) Division of Developmental Medicine, Boston Children’s Hospital; 7) Center for Biomedical Informatics, Harvard Medical School.

Introduction: To assess and characterize the need for a centralized service to support the utilization of genomic testing at Boston Children’s Hospital (BCH), we conducted a web-based survey of clinicians and researchers regarding their past, current, and anticipated future use of next generation sequencing and their anticipated needs for support. Results: We received 283 survey responses from 36 departments and divisions (5.5% response rate). A large proportion of clinician respondents anticipate that they will use exome/genome sequencing (44.8%) and/or candidate gene panels (50% clinician) within the next year. About an equal percentage of researcher respondents anticipate the need for exome/genome sequencing (48.0%), while a smaller percentage anticipates the use of candidate gene panels (31.8%). Estimated volume data show an increase in the use of these tests in the future compared to past estimates. However, few respondents (13.6%) said that they felt “Completely Ready” or “Pretty Much Ready” to incorporate NGS into their clinical practice or research. Respondents indicated varying degrees of need for a diverse list of support services, with interpretation and clinical correlation assistance ranked as the most needed services.

Conclusion: The results of this study highlight the current and growing need for various types of education, support, and assistance (e.g., data and result interpretation) and make a convincing argument for hospitals and medical centers to invest in the development of institutional genomic services. Some of the key challenges to the success of such genomic service models will involve the need for more precise, powerful, and efficient interpretation tools; expanded data sources; and expert resource networks to inform the interpretation process, not to mention issues of cost, billing, and reimbursement around these components. Data from this study will guide BCH’s establishment of a new institutional service, Interpretive Genomic Services (IGS), which will support clinicians and researchers in their use of NGS testing and data. BCH will leverage existing institutional expertise in genetics and genomics, genetic counseling, bioinformatics, and many subspecialties to create a true multi-disciplinary service.
A rapid evidence review for the inclusion of genetic data in clinical care for a common, complex disease. J. Malinowski, E.W. Clayton, D.C. Crawford. 1) Center for Human Genetics Research, Vanderbilt University, Nashville, TN; 2) Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN.

Personalized medicine (PM), defined as differentiated clinical care based, in part, on the genetic background of an individual, has been implemented for cancer treatments and to guide pharmaceutical interventions for various clinical traits. Despite a recent history of direct-to-consumer (DTC) companies offering risk prediction for dozens of diseases and traits based on the genotyping of hundreds of thousands of single nucleotide polymorphisms (SNPs), the clinical validity and utility of these SNPs for common, complex diseases have not been adequately investigated. We performed a rapid evidence review in an academic setting to identify analytic evidence that the inclusion of genetics in routine clinical care improves health outcomes for complex disorders. We identified a common, complex disease (hypothyroidism) and a related quantitative trait (serum thyroid hormone (TSH) levels) with several known genetic associations, including FOXE1 rs7850285, PTPN22 rs24766012, SH2B3 rs3184504, VAV3 rs4915077, and HLA region rs2517532. These variants were previously used by a DTC company to report risk of developing hypothyroidism. We developed key questions to identify analytic validity, clinical validity, clinical utility, and the ethical, legal, or social issues (ELSI) pertaining to the use of these genetic data in the clinical setting. Using medical subject headings (MeSH) terminology, we performed a PubMed search for articles, completed abstract and full text reviews, and analysis of the results. All data collection and review were maintained in REDCap databases. Six hundred and thirty abstracts were reviewed, and analysis of the results. Nineteen articles were full text reviewed (54.9%). Fifteen (4.3%) of the full text reviewed articles included odds ratios or effect sizes and p-values for the genetic association. None (0%) of the articles provided evidence of clinical validity or utility for the genetic associations. This rapid review was completed in seven months over two academic semesters with ten individuals. Our study supports others that have found a general lack of published data to indicate genetic research findings have demonstrable clinical utility. Despite these negative results, our rapid review methodology served as an example of deploying an evidence review in an academic setting to systematically identify the essential data required to accurately assess the utility of including genetic data to improve health outcomes for common, complex diseases.

Genetic Test Recipient as a Source of Data: A Novel Approach to Obtaining Genotypic and Phenotypic Data for Input into Genomic Databases. B.E. Kirkpatrick, E. Riggs, M. Giovanni, R. Green, A. Janze, P. Krautschied, J. Krier, C.L. Martin, D. Mettville, D. Murray, H. Rehm, D. Riethauser, W. Rubinstein, B. Smith-Packard, C. Tan, K. Wain, W.A. Faucett on behalf of the Clinical Genome (ClinGen) Resource. 1) Kaiser Permanente National Health Research Center, Division of Genomics and Department of Medicine, Brigham and Women’s Hospital, Boston, MA; 2) GeneDx, Gaithersburg, MD; 4) ARUP Laboratories, Salt Lake City, UT; 5) Laboratory for Molecular Medicine, Partners HealthCare Personalized Medicine, Cambridge, MA; 6) Department of Laboratory Medicine and the Division of Genetics, Boston Children’s Hospital, Boston, MA; 7) Department of Pathology, Harvard Medical School, Boston, MA; 8) InVitae, San Francisco, CA; 9) National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD; 10) Department of Human Genetics, University of Chicago, Chicago, IL; 11) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN.

Large datasets of genomic variants and observed phenotypic consequences have been identified as important resources of information that will help clinicians and researchers understand the extensive variability in the human genome and its impact on disease. The Clinical Genome (ClinGen) Resource, together with the National Center for Biotechnology Information (NCBI), is developing ClinVar, an open-access database of genomic variants, associated phenotypic information, and clinical classifications with supporting evidence. Genomic databases including ClinVar are incomplete, and various avenues are being investigated to increase the amount of information available. As genotypic information in isolation has limited value, the availability of phenotypic information is crucial to variant interpretation. Previous efforts have focused on obtaining linked genotypic and phenotypic information from laboratories, researchers, and clinicians. These approaches often yield limited information due to lack of time, resources, and/or incentives. In contrast, patients themselves are valuable sources of information regarding their own medical histories, and they may be highly motivated to share this information for the benefit of advancing scientific understanding. By developing an online patient registry portal, ClinGen seeks to obtain detailed phenotypic information provided by patients. Registry participants will complete an online consent process explaining the project and implications of genomic data sharing. Each participant will be asked to complete a general, review-of-systems-type online health questionnaire; follow-up questionnaires will obtain detailed information on health issues identified on the original survey. Participants also will be asked to share a copy of their genetic test results, which will be curated by the registry coordinator. This approach will allow for the accurate representation of genotype information. Strategies for allowing the linkage of data from different sources on the same individual into ClinVar are in development. As far as the researchers are aware, this is the first attempt to obtain information directly from genetic test recipients for the purposes of building genomic databases.


Obesity and overweight for all age groups are major concern in both economically developing and developed countries. Physical and mental health is directly related to obesity and hypertension. This study aimed to compare anthropometric status, blood pressure and oxygen saturation between athletic and non-athletic females. The athletic and non-athletic female subjects aged 18-25 years were selected from Guru Nanak Dev University, Amritsar, India. Anthropometric, physiometric and socioeconomic variables along with oxygen saturation was recorded for all the study samples. The data was analyzed on SPSS v. 18.0 using correlation, ANOVA, univariate and multivariate regression analysis. In this study weight gain, BMI, SBP and DBP were found higher in athletic female group as compared to non-athletic group. Whereas, waist circumference, hip circumference, waist-hip ratio and skinfolds were higher in non-athletic female group. The percentage of saturated oxygen content has been found maximum in non-athletic group.
2417S  
The NINDS Repository: A Public Collection of Biomaterials for Disease Modeling, Gene and Biomarkers Discovery in Neurological Research.  
C. Tam1,2, G. Balaburski1, S. Heil1, K. Reeves1, J. Santana1, J. Siloyo1, M. Seif1, K. Gwinn1, M. Sutherland2, R. Corveaux3, C. Pérez1.  
1) Coriell Institute for Medical Research, Camden, NJ; 2) National Institute for Neurological Disorders and Stroke-NIH, Bethesda, MD, USA.

Neurological diseases are a major public health concern but the pathological mechanisms in neurodegenerative disorders remain largely not understood. The major challenges in disease mitigation reside on lack of reliable genetic and molecular biomarkers for causatic, diagnosis, and progression monitoring; and limited reproducible cellular models for research. The National Institute of Neurological Disorders and Stroke (NINDS) Repository, funded by NINDS, was established with the mission of providing high quality biospecimens as a strategy to facilitate and accelerate research in neurological diseases. The NINDS Repository collects biosamples and de-identified clinical data from diverse patients diagnosed with various neurological disorders as well as neurologically normal controls. In addition, the NINDS Repository features collections of patients-derived fibroblasts and induced pluripotent stem cells (iPSC) with well-defined mutations as essential research tools for understanding the pathological mechanisms and establishes cellular models for neurological diseases. Recently, the NINDS Repository has broadened its collections to include whole blood RNA, plasma, serum, cerebrospinal fluid, and urine, to facilitate biomarker research utilizing longitudinal samples from both affected and neurologically healthy individuals. Since its establishment, biomaterials from more than 44,000 individuals with cerebrovascular diseases, Parkinsonism, motor neuron diseases, epilepsy, Tourette syndrome, Dystonia, Huntington’s disease and neurologically-normal controls have been banked in the NINDS Repository. The NINDS Repository has established validated standard operating procedures and rigorous quality control assessments that span the life cycle of all biospecimens collected to provide premium samples. The NINDS Repository aims to ensure and implement standardization for collecting and processing across all samples. In addition, the NINDS Repository utilizes secure and integrated laboratory information management systems to monitor inventory, processing, storage, and distribution of biospecimens, and facilitates sample-data association by cross-referencing with other databases. By developing such a centralized collection of human biospecimens and their associated de-identified clinical data, the NINDS Repository thus provides a vital resource for research designed to discover and validate genetic and proteomic biomarkers of neurological disorders.

2418S  
Patients and caregivers as sources of innovative ideas and solutions: A multiple case study approach.  
V. Francisco1,2, R. Oliveira2, P. Oliveira2, L. Zejinovic2, S.A. Oliveira1,2.  
1) Instituto Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisboa, Portugal; 2) Instituto Gulbenkian de Ciência, 2780-156 Oeiras, Portugal.

Recent academic literature shows that patients and caregivers are a significant source of innovative solutions related to their medical condition. To date, little is known about the process by which these innovations emerge, how they diffuse, and how they impact the lives of patients and caregivers. In this work we follow a multiple-case study approach to map a set of patient innovations and adoptions of patient innovation cases, and systematically explore how and how far patients and caregivers innovate in the health care field. In addition, we propose some explanations for why patients and caregivers stop at a certain stage of progression of the innovation process. We conducted 15 extensive semi-structured interviews with patients and caregivers of the following group of diseases: spinal cord injuries, Angelman syndrome, epidermolysis bullosa, cerebral palsy, and hemiparesis. These individuals shared with us their experiences with their disease and their efforts, or the lack of them, to overcome specific problems related to their health condition. This includes 4 “holistic” case studies and 26 “embedded” case studies. With the information from these cases, we analyze patient innovation paths and present them in the fall-offs conceptual framework. Through a cross-case analysis, we find that duration of the experience with the disease, complexity and pressure of a certain situation, belonging to a group or a community, and perceived value of a solution are among the most important reasons that impact how far patients and caregivers take their innovations. As a result of our multiple-case analysis we present a set of propositions from which future research in the field is warranted.
2420S

Multiplex hereditary cancer panel testing for individualized cancer risk assessment is increasingly common, but the yield of panels is not yet known. We conducted a retrospective IRB-approved chart review of 317 cancer genetics clinic patients tested with a multiplex cancer gene panel. We analyzed the yield of panels versus the traditional single gene testing for specific hereditary syndromes. Panels were performed by Myriad (n=242), Ambry (n=64), Fulgent (n=9), University of Washington (n=1), and City of Hope (n=1). Our cohort included 142 patients from the USC Norris Comprehensive Cancer Center and 175 from the Los Angeles County + USC Medical Center. The mean age was 48.8 years (SD=12.5) and 85.6% were female; 80.1% had a history of breast cancer (n=254); race/ethnicity reported was 48.9% Hispanic, 34.4% Caucasian, 4.4% Black, 10.0% Asian, and 2.2% other. Among 317 panel test results, 17.0% contained at least one deleterious mutation (n=54) and 47.3% contained at least one variant of uncertain significance (VUS) (n=150). Of 242 Myriad MyRisk panel results, 16.9% had at least one deleterious mutation (n=41) and 48.4% had at least one VUS (n=117). Of the 64 Ambry panel results (CancerNext, Ovalt, ColoNext, ColoNext and others), 12.5% had at least one deleterious mutation (n=8) and 42.2% had at least one VUS (n=27). Among 129 patients with breast cancer, 6.2% carried a mutation in BRCA1 or BRCA2 (n=8) and 9.3% carried a mutation in BARD1, CDH1, CHEK2, MUTHY, PALB2, RAD50, or TP53 (n=12). Among 69 patients with colon cancer, 2.9% carried a mutation in CDH1, CHEK2, MUTHY, or PMS2 (n=10) and 2.9% had a mutation in RAD50 or RAD51C (n=2). Thus, multiplex panel testing increased mutation detection over a traditional approach in 9.3% of breast cancer patients and 2.9% of colon cancer patients. Single gene assays for these specific genes, therefore, should be supplemented by multiplex genetic testing to identify the supplemental yield of panels was 7.4% for ovarian cancer (2/27), 11.7% for endometrial cancer (2/17) and 22.2% for gastric cancer (2/9). In the overall cohort of 317 patients, 7.6% (n=24) had a mutation identified on a panel that would not have been detected as a standalone test. In most cases, these mutations led to enhanced surveillance recommendations for the patient and family members. We conclude that multiplex panel testing increases the yield of mutation detection and adds to the capability of providing individualized cancer risk assessment.

2420T

Introduction: As the genetic testing landscape expands in both the complexity and quantity of clinically available testing products, the task of finding, comparing, and selecting appropriate genetic tests becomes increasingly burdensome. The genetic testing landscape can be divided into two categories: single-gene tests and panels. While tests, which interrogate a single genetic target are relatively simple to group, panels, which interrogate multiple genetic targets, prove considerably more difficult. We set out to develop an organizational schema that allows for the side-by-side comparison of similar panel testing products. Methods: Leveraging our data resource of all currently available genetic tests from more than 200 CLIA-certified US testing laboratories, we tested the efficacy of four unsupervised machine learning algorithms in clustering multi-target genetic tests (panels) based solely on their diagnostic targets (genes). Using hyper-parameter optimization and domain expertise, we identified a high performance algorithm, DBSCAN, and generated a novel set of genetic test clusters. Next, we manually curated the algorithmically-generated clusters to provide greater clinical relevance. Results: Our dataset initially contained 1875 panels manually grouped into 709 clusters (Avg. 2.7 tests/cluster, SD=2.5). Algorithmic clustering decreased number to 358, while subsequent expert curation, taking into account intended use of the test, increased the number to 478, (Avg. 4.1 tests/cluster, SD=5.85). To assess clustering efficacy, we calculated the mean intra-cluster vertex connectivity, observing a 33% increase in relatedness within clusters following the two-step process (Manual: 1.45, WV: 1.45, DBSCAN: 2.1, DBSCAN+Manual: 3.0). Conclusion: This approach is effective in grouping panels similar clinically available panels. As the genetic testing market expands, completely manual curation of genetic testing products becomes increasingly impractical and error-prone. Semi-automated methods of organizing, sorting, and selecting the thousands of tests available in the wide range of available testing products, thus improving clinical workflows and, potentially, diagnostic outcomes.

2421S
Cost-effective NGS based BRCA-TP53 screening panel for hereditary breast cancer in India. M. Sen1,2, P. Agrawal3, N.S.N. Swetha1, V. Vittal3, V. Pathak1, G. Deshpande1, D. Vishwanath1, P. Ramamoorthy1, K. Subramanian1, V. Gupta1,2. 1) Strand Centers for Genomics and Personalized Medicine, Strand Life Sciences Pvt. Limited, Bangalore, India; 2) Strand at Mazumdar-Shaw Center for Translational Research, Bangalore, India.

Introduction: Hereditary BRCA1/2 mutations in breast cancer show high correlation with Triple Negative Breast Cancers (TNBC). In India, the incidence rate for TNBCs is approximately 25-30%, significantly higher than the rest of the world, creating a need for a pan-parallel, high throughput, cost-effective screening method for hereditary breast cancer patients. Cost analyses for current screening methods indicate the supplemental yield of panels was 7.4% for ovarian cancer (2/27), 11.7% for endometrial cancer (2/17) and 22.2% for gastric cancer (2/9). In the overall cohort of 317 patients, 7.6% (n=24) had a mutation identified on a panel that would not have been detected as a standalone test. In most cases, these mutations led to enhanced surveillance recommendations for the patient and family members. We conclude that multiplex panel testing increases the yield of mutation detection and adds to the capability of providing individualized cancer risk assessment.

2421M
Cost-effective NGS based BRCA-TP53 screening panel for hereditary breast cancer in India. M. Sen1,2, P. Agrawal3, N.S.N. Swetha1, V. Vittal3, V. Pathak1, G. Deshpande1, D. Vishwanath1, P. Ramamoorthy1, K. Subramanian1, V. Gupta1,2. 1) Strand Centers for Genomics and Personalized Medicine, Strand Life Sciences Pvt. Limited, Bangalore, India; 2) Strand at Mazumdar-Shaw Center for Translational Research, Bangalore, India. Materials and Methods: We designed a comprehensive BRCA-TP53 NGS panel covering all known cancer susceptibility genes. We performed an in-solution capture of genomic DNA libraries using Agilent SureSelect kits. We have used 15 known samples (Hapmap, cell lines and pre-validated clinical samples) and evaluated the performance characteristics of the panel. Results: Our enrichment based assay panel covers 100% of the targeted region. The overall N ONA hits were identified from 1.25 million reads. Exome sequencing identified de novo compound heterozygous mutations in exon11 and exon19 of POMT2 (NM_013382), i.e., c.1248C>G p.His416Gln and c.1912C>T p.Arg571Pro, respectively. The present observation indicates the extent of exome sequencing using preserved umbilicus in undiagnosed fetal structural abnormalities. The correct molecular diagnosis is ruled out to provide the parents with the recurrence risk of 25% together with the possibility of prenatal genetic diagnosis for a future pregnancy. In light of the above findings, we recommend that the use of umbilical cord should be considered in fetal cases with undiagnosed structural defects, particularly when a Mendelian disorder is highly suspected.

Diagnostic applications of next-generation sequencing have been impeded by complex experimental workflows associated with target DNA enrichment methods and errors and biases caused by PCR during sample preparation. Here we demonstrate single-molecule Oligonucleotide-Selective Sequencing (OS-Seq) for simplified and clinical-grade targeted sequencing. OS-Seq consolidates capture and sequencing of DNA in situ flow cell of a next-generation sequencing system. We developed a programmable and automated OS-Seq technique using the MiSeq sequencing system and implemented a PCR-free sample preparation method for single-molecule OS-sequencing. In addition, we introduced batch indexing, which utilizes several index sequences per sample and improves the uniformity of the sequencing yields between pooled samples. We developed two OS-Seq assays targeting coding exons, exon-intron boundaries and known pathogenic variants in introns of 136 and 49 genes (495,567 and 159,198 bases, respectively) implicated in cardiovascular diseases. The median sequencing depths were 420 and 739 and 99.35% and 99.8% of the bases on target regions were covered >15x. We validated SNPs, INDELs and CNVs using whole genome sequenced reference samples. High-quality genotypes for NA12878 were obtained from the Genome-in-a-bottle consortium and INDELs were confirmed with Sanger sequencing. Sensitivity to detect CNVs was >99% and 100%, respectively, and specificity was >99.9% for both assays. Sensitivity and specificity to detect short INDELs was 100% and sensitivity to detect CNVs of >1kb was >99%. We showed that single-molecule OS-Seq omits sequence duplicates, reduces sequencing errors by two-fold and improves uniformity of the coverage. Our results demonstrate the usability of OS-Seq for clinical sequencing applications.


Genetic disorders and birth defects account for a high percentage of the admissions in children’s hospitals. Congenital malformations and chromosomal abnormalities are the most common causes of infant mortality. However, highly heterogeneous causes of the disorders in the cases of children’s hospital provide low yield of molecular diagnostic procedures. Next generation sequencing (NGS) has a basically broad impact on many facets of biological and clinical research. To date the NGS is not routinely used in genetic diagnosis. Advent of NGS in the clinical laboratory setting requires the adoption of many processes and procedures. We developed diagnostic flows of molecular diagnosis using NGS to identify mutations in several malformation syndrome and related disorders in Kanagawa Children’s Medical Center. This panel approach was successfully established in the cases with unknown sequence alterations. Depending on the clinical diagnosis, we identified causative mutations in the patients. NGS-based panel analysis is reliable and cost-effective approaches in clinical setting of heterogeneous diseases. However, the panel analysis is so limited in the case of application for the extremely high heterogeneous disorders, such as autism and unknown multiple anomalies. We present current clinical utility of exon sequencing of Mendelian disorders (TruSight One) in the clinical context in our children’s hospital. These results indicated the clinical exome sequencing is useful for pediatric practice, and bioinformatic process will become essential for the clinical practice.

Active organic solvent-free paraffin removal is the key to efficient extraction of NGS-quality DNA from FFPE tissues. H. Khoja, E. Rudd, J. Han, A. Purdy, S. Kakumanu, A. Palmer, G. Durin, J. Laughman. Covaris Inc., Woburn MA.

Formalin Fixation and Paraffin Embedding (FFPE) of tissues, a mainstay of clinical histological analysis for the past century, is rapidly being adopted for targeted and whole genome sequencing. The extreme formaldehyde fixation and tissue dehydration of FFPE preserved tissue presents not only a technical challenge to reproducible DNA and RNA extraction and molecular analysis, but also a workflow challenge in a clinical setting. In this poster we present truXTRAC™, a novel and highly reproducible method of extraction and purification of DNA and RNA from FFPE tissue utilizing Covaris Adaptive Focused Acoustics™ (AFA). Our novel method utilizes highly controlled acoustic energy for effective removal of paraffin from FFPE cores, sections, and slides enabling efficient and rapid tissue rehydration, tissue digestion, crosslink reversal, and nucleic acid release. This critical step is carried out without the use of dangerous organic solvents or messy mineral oils. We will present quantitative fluorescent microscopy data illustrating the efficiency of active paraffin removal from tissues utilizing this novel approach. We will also illustrate the significant effect of efficient active paraffin removal on downstream applications and analyses of DNA and RNA from FFPE tissues. Our results, utilizing replicates of different tissue types indicate high dsDNA yields, and improved DNA and RNA quality as indicated by commercial qPCR based kits designed to assess DNA quality. To quantitate the quality of the extracted DNA, as a function of sequence coverage across the entire genome, we compared whole genome sequencing data from FFPE samples to sequencing data obtained from matched frozen tissue DNA. Our data clearly indicate significant improvement of sequence coverage and uniformity in FFPE samples processed using Covaris AFA. Our data also indicate that the high quality controllable DNA extraction allows for preparation of large insert (1-6kb) libraries which has previously been considered very difficult to achieve with FFPE stabilized tissues. The simplified single-tube method and rapid workflow allows for the parallel processing of 8 to 96 FFPE samples in a batch format easily adaptable to the throughput requirement of medium to high volume clinical labs. We present data showing high reproducible RNA and DNA extraction results regardless of the throughput of FFPE samples being processed.


Hereditary disorders of connective tissue are fairly common in the general population with Ehlers-Danlos syndrome (EDS), Marfan syndrome, Stickler syndrome, and other vasculature typically dominantly inherited and genetic testing is offered for several of the more common genes involved. We have developed and validated a next-generation (Ion Torrent) panel that analyzes 23 dominant genes that cause EDS (types I, II, IV), Marfan syndrome, Loeys-Dietz syndrome (types I-IV), Stickler syndrome (types I-III), and hereditary aneurysms (7 genes). The Ion AmpliSeq software was used to design a custom primer pool to amplify exonic and 25 bases of flanking intronic regions from 23 genes (1274 amplicons). Two Ion Torrent 318 chips were utilized for 15 blinded samples (4 with known mutations in different genes). The average depth across the 23 gene target regions (>40 reads) was 474. All positive samples were accurately identified. Reproducibility was assessed to be 99.2% for 257 SNP’s (2 duplicate samples on each chip). 117 amplicons were Sanger sequenced to improve overall coverage to greater than 99% of the targeted regions. Separate MLPA assays were performed for deletion/duplication analyses of those genes reported to have these pathogenic variants. CONNECT1 is the optimal cost-effective first line test for those with suspected dominant undiagnosed hereditary connective tissue disorders or those with a dominant family history of aneurysms.
2428T  
Xeroderma Pigmentosum (XP): Single nucleotide variants (SNV) that probably affect function in two Indian families identified using next generation sequencing (NGS) of 8 XP genes. K. Reddy1, M. Reddy2.  
1Institute of Bioinformatics, Whitefield, Bangalore 560066, India; 2Ophthalmology, St Johns Medical College, Koramangala, Bangalore 560034, India.  
Two Indian families with Xeroderma pigmentosum (XP) were tested using next generation sequencing (NGS) of 8 autosomal recessive XP genes: XPA, XPC, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5 and DDB2. Family 1: had a novel substitution variation in the ERCC5 gene, c.1A>G and a loss of the start site. Each parent had one copy of the gene with this variant and the two affected sons carried variants on both copies of ERCC5. The siblings presented with XP with no apparent neurological abnormality (IQ85 and IQ70 is probably due to impaired vision). In Nature there is also a read through fusion transcript BIVM-ERCC5 [Homo sapiens], an important paradigm of ERCC5. The expression of the conjoined genes may be another mechanism for gene regulation in eukaryotes and should be considered in correlating the phenotype with the genotype. Gene sequence and protein profile of ERCC5 and BIVM-ERCC5 may be necessary to assess the functional impact of SNVs in ERCC5 gene Family 2: had a c.1677C>G variation in XPC gene that has been reported in literature and creates a premature stop codon. One copy of the gene in the father, one copy in the mother and both copies in the proband carried this variant. These are the first Indian SNVs described in XP families and these are different from the frequent mutations from other parts of the world. Hence, variants that affect gene function in XP families from all parts of India will be assessed and haplotyped to identify a founder mutation. A focus on the mutations in areas of high consanguinity could potentially simplify the test. Also, understanding the distribution of SNVs in XP gene would aid in counseling families about consanguineous marriage and provide prenatal diagnosis option for carriers. Early eye sight corrections in patients would enhance performance in school. In the long run we hope to decrease the incidence of XP.

2429S  
Steroid resistant nephrotic syndrome (SRNS): NGS panel testing to direct therapy and intervention. L.J. Yarram-Smith1, A. Bierzyńska2, D. Smith3, M. Saleem1, M. Williams1.  
1Bristol Genetics Laboratory, North Bristol NHS Trust, Southmead Hospital, Bristol; 2Bristol Genetics Laboratory, Bristol, United Kingdom; 3Academic Renal Unit, Dorothy Hodgkin Building, Whiston Street, Bristol, BS1 3NY.  
Steroid Resistant Nephrotic syndrome (SRNS) is a disorder of the glomerular filtration barrier. It is characterised by massive proteinuria, hypoalbuminaemia and oedema, and is managed by non-specific heavy immunosuppression. An invasive biopsy is often required and the majority of patients progress to end stage renal failure. SRNS is associated with over thirty genes expressed at the glomerular filtration barrier. Rapid genetic diagnosis is important for therapy and intervention as genetic SRNS is non-responsive to immunosuppression, and has a lower rate of post transplant recurrence. A genetic result enables appropriate family counselling including choice of transplant donor and has the potential to avoid invasive renal biopsy. 

Bristol Genetics Laboratory has received >120 worldwide diagnostic referrals (80% paediatric and 20% adult) for NGS clinical panel testing of 37 SRNS in 4813 clinically relevant genes (broad, deep sequencing in the context of CONTRA). The panel was validated using 24 patients and 440 variants identified in the University of Bristol Academic Renal Unit RADAR study (Nimblgen/exome). Likely pathogenic variants have been identified in 30% of cases, most commonly occurring in NPHS1, NPHS2, WT1, COL4A3 and COL4A4 with a different mutation spectrum in adult and paediatric patients. A further 30% of cases have candidate variants. Variant stratification is challenging with a global service due to the presence of rare variants in genes expressed at the glomerular filtration barrier. Rapid genetic diagnosis is important for therapy and intervention as genetic SRNS is non-responsive to immunosuppression, and has a lower rate of post transplant recurrence. 

A genetic result enables appropriate family counselling including choice of transplant donor and has the potential to avoid invasive renal biopsy.

2430M  
1Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Hong Kong; 2Clinical Genetic Service, Department of Health, Hong Kong; 3PerkinElmer Health Sciences Inc., 940 Winter Street, Waltham MA 02451, USA.

Sizing of FMR1 trinucleotide repeats requires the use of capillary sequencer instruments, or by a labor Southern blot. To validate a simple, robust PCR assay for quantification of CGG repeats. We validated a new PCR-only method for quantifying Fragile X CGG repeat that utilizes a low-cost capillary electrophoresis instrument and the FragilEase™ reagent kit. Analytical performance was first demonstrated on 12 Coriell reference samples comprising normal through full mutations. Subsequently, a set of 112 archived clinical DNA samples, enriched for premutation and full mutations was analyzed. All samples were amplified successfully. Quantification of repeat numbers was interpreted by the use of standards with known repeats. The repeat numbers from the new assay were concordant with those obtained with the reference method. The intra-assay (CV < 2.5%) and inter-assay imprecision was within 1 CGG repeat. Our result demonstrated this new method is fast, robust that facilitates Fragile X testing in a clinical laboratory.

2431T  
1Medical Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom; 2Dept of Medical Genetics, University of Cambridge, Cambridge, UK; 3PHG Foundation, Cambridge, UK.

Clinical diagnostic next generation sequencing requires cost effective, efficient and robust methodology. Targeted gene panels are readily available, but significant resources are required to design, optimise and validate each panel, thus restricting the number of panels, number of diseases that can be tested for, and number of patients who can benefit from these tests in an individual laboratory. Whole exome sequencing (WES) overcomes this hurdle, but is considered impractical or too expensive in many health care settings. We have therefore investigated a single test which sequences 4813 clinically relevant genes and offers the best attributes of both panel and WES approaches. We first undertook a technical validation of the Illumina TruSight One sequencing panel by testing 195 samples with mutations in 74 different genes, previously identified by Sanger sequencing. All 188 single nucleotide variants, 20 deletions and 15 insertions were identified correctly. This was followed by sequencing all 4813 genes in 300 patients chronologically referred for diagnostic sequencing of either a large single gene, or a number of genes associated with a genetically heterogeneous condition. Sequencing results were filtered so the analysis focused only on genes determined by Clinical Geneticists to be relevant to the clinical presentation. Among the first 110 newly referred patients, 82 were analyzed for one of 36 different gene panels, ranging from 2-28 genes. 28 patients were analyzed for just single genes. Analysis of results from the remaining 290 samples is ongoing. To date, clear pathogenic mutations were detected in 57% (16/28) of patients analyzed for single genes, and in 28% (23/82) analyzed for gene panels. This allows confirmation of diagnoses for 35% of patients. Mutations were detected in 32 different genes; 8 patients (7.3%) had mutations in 7 genes for which genetic testing is currently not available in UK.

A single sequencing test covering a wider range of genes than currently available by panel or Sanger sequencing has been rapidly validated and introduced into the diagnostic service. Sequencing of this clinical exome offers versatility and simplicity, with comparable mutation detection rates to full exome sequencing but at significantly less cost. This test will enable more patients with a wide range of rare genetic diseases to receive rapid molecular confirmation of their diagnosis.
2432S

Clinical exome sequencing is increasingly used for solving diagnostic odyssey cases in children with suspected genetic syndromes and also cancer. The complex process of going from DNA sample to clinical report involves multiple, technologically, scientifically, and medically complex steps; despite early success, significant improvements can be made to increase the overall diagnostic yield of exome sequencing tests. We have developed an augmented exome sequencing workflow that boosts accuracy and coverage in over 7000 genes that are medically relevant to Mendelian disease, cancer, and pharmacogenomics as referenced by multiple databases and literature sources. The performance of this augmented exome was assessed using three methods: coverage over all exonic bases in our ACE exome assay compared to other standard exomes, comparison against the NIST standard genome, and examination of the structure and coverage of the 7000 medical genes in the latest reference assembly (GRCh38). Of the over 7000 targeted medical genes, our augmented clinical exome completely covers 50% more genes than standard exomes at comparable coverage. Further we demonstrate increased sensitivity and specificity as evaluated against the NIST standard.

In evaluating against the reference, over 650 of these medical genes are on ‘Fix’ patches released as part of incremental reference improvements by the NCBI, suggesting that the structure of these genes were not well described in GRCh37 and need to be addressed as part of clinical informatics pipeline. Over 150 of the patched genes are on existing panels, which may need to be re-assessed in light of new assembly information. An augmented exome sequencing and informatics approach enables clinical grade performance over a large percentage of medically important genes. We report on specific examples of pathogenic variants that would have been missed by standard exomes in clinical cases processed through our testing laboratory.

2433M
Contextualization and recommendation: How doctors and patients discuss whole-genome sequencing results. J.L. Vassy1,2, K. Davis2, J. Oliver Robinson3, J. Blumenthal-Barby4, K.D. Christensen3, R.C. Green1,2, A.L. McGuire5, P.A. Libet3 for the MedSeq Project. 1) Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 2) VA Boston Healthcare System, Boston, MA; 3) Fuqua School of Business and Sanford School of Public Policy, Duke University, Durham, NC; 4) Baylor College of Medicine, Houston, TX; 5) Partners HealthCare Center for Personalized Genetic Medicine, Boston, MA.

Background: There are concerns that non-geneticist physicians are not prepared for the role genomics will play in clinical medicine. It is unknown how they will talk to their patients about whole-genome sequencing (WGS) and its complex results. Methods: We have enrolled primary care physicians and cardiologists to participate with their patients in the MedSeq Project: a randomized trial of WGS in clinical care. Physicians undergo brief CME in genetics prior to patient enrollment. Each patient’s WGS data are interpreted and reported on a Genome Report, including results for monogenic disease variants, carrier status, complex polygenic traits, and pharmacogenomics. The Genome Report is delivered to the patient’s physician, and the two meet for a disclosure visit to discuss its findings. Discussions are recorded, transcribed, and coded with thematic content analysis to identify emergent themes. Results: 16 of 100 planned WGS disclosure visits have occurred to date: 11 in generally healthy primary care patients and 5 in cardiology patients. We identify 5 major themes. 1) Contextualization: Physicians use additional clinical information, such as age, family history, lack of symptoms, or other test results, to interpret the significance of WGS findings for each patient. 2) Prioritization: Among the many results listed on each patient’s Genome Report, physicians tend to emphasize explicitly the 1 or 2 most clinically relevant WGS findings for each patient, identifying “the single biggest thing to come out of this.” 3) Limitations of WGS: Some physicians discuss the technological limitations of sequencing and the unknown penetrance of certain variants in cautioning patients against over-interpreting certain WGS results. Some discuss their own inability to interpret certain WGS findings. 4) Misinformation: We have identified a few instances of inaccuracies in physicians’ interpretations of WGS results, such as underestimating a patient’s risk of having a child with cystic fibrosis. 5) Reconciliation: Physicians ultimately synthesize WGS results and all other clinical information to make recommendations for each patient, including genetic testing for family members, additional diagnostic testing, and referrals to subspecialists. Conclusions: Physicians and their patients discuss WGS similarly to other tests in clinical medicine. Concerns about their unpreparedness may be overstated, but efforts to reduce misinformation are warranted.

2434T
Towards highly sensitive diagnostic exome sequencing without the need for confirmations by Sanger sequencing. K.L.I. van Gassen, M.G. Efferink, P.H.A. van Zon, I.J. Nijman, B. van der Zwaag, J.K. Plooe van Amstel. Department of Medical Genetics, University Medical Centre Utrecht, Utrecht, Netherlands.

Current best practice genetic diagnostics by massive parallel sequencing includes confirmations of all reported findings by Sanger sequencing. Such confirmations eliminate false positive variant calls and exclude potential sample swaps, but often need the design of new PCR primers. Designing and ordering new PCR primers for confirmations create a laboratory bottleneck, adds substantial capital costs and leads to an increase in test turn-around-time. The proposed workflow utilizes the power of an independent biological replicate without the need for extra target enrichments or sequencing resources and is applicable to most enrichment techniques that support multiplexing. This replicate eliminates the need for Sanger confirmation of most variant calls. Additionally, this workflow should theoretically decrease false positive and false negative variant call rates. Here, we present the details of this workflow. We report on false positive and false negative variant call rates using this approach.

2435S
Biological assays to predict the functional impact of missense mutations: the case of the tumor suppressor gene BRCA1. G.A. Millot1,2, P. Thouvenot1, B. Ben Yamin1, L. Fournié1, C. Houdayer1,2, D. Stoppa-Lyonnet1, D.E. Goldberg3, A. Nicolas1. 1) UMR3244, institut Curie, Paris, Paris, France; 2) Sorbonne Universités, UPMC Univ Paris 06, 4, Paris, France; 3) Service de Génétique Oncopédiatrique & INSERM U830, Institut Curie, Paris, Paris, France; 4) Université Paris Descartes, Paris, France; 5) Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah.

Missense mutations generally come from single nucleotide polymorphisms (SNPs) present in the coding regions of genomes. These mutations trigger the weakest sequence modification in a protein, since a single amino acid is replaced by another one. The resulting effect of this amino acid substitution can range from “no effect” to “complete alteration” of protein function. For this reason, it is difficult to predict the functional consequences of missense mutations. Moreover, the ever-growing amount of exome data will dramatically increase the number of detected missense mutations, and likewise increase the challenge of interpreting their functional impacts. Our work is focused on the tumor suppressor gene BRCA1, whose germline inactivation leads to hereditary breast and ovarian cancers. To date, 668 different germline missense mutations have been identified in BRCA1. However, only a small fraction has been characterized as pathogenic (37) or neutral (103) using classical genetic methods. The 528 additional missense mutations remain unclassified, due to the lack of familial/populational information and the inability to predict their functional impacts. To address this issue, we designed and validated four experimental assays. Using 40 missense mutations, 25 previously classified as pathogenic and 15 previously classified as neutral, by genetic methods, we showed that these assays are efficient in predicting the pathogenicity of BRCA1 missense mutations. We also propose a method to improve the computation of sensitivity and specificity, which are two parameters critical for functional assays. Altogether, these results contribute to the high-throughput classification of missense mutations, a major challenge for the next decade.
Complete APTX deletion in a patient with ataxia with oculomotor apraxia type 1. R. van Minkelent, M. Guitart, C. Escotet, G. Yoon, P. Effros, G.M. Bolman, R. van der Helm, R. van de Graaf. A.M.W. van den Ouweland. 1) Department of Clinical Genetics, Erasmus Medical Centre, Rotterdam, Netherlands; 2) Genetic Laboratory, UDIAT-Centre Diagnostic, Neuropediatrics Unity, Corporació Sanitària Universitària Parc Taulí, Sabadell, Spain; 3) Division of Clinical and Molecular Genetics, Department of Pediatrics, The Hospital for Sick Children and University of Toronto, Canada.

Background Ataxia with oculomotor apraxia type 1 is an autosomal-recessive neurodegenerative disorder characterized by a childhood onset of slowly progressive cerebellar ataxia, followed by oculomotor apraxia and a severe primary motor peripheral axonal motor neuropathy. Ataxia with oculomotor apraxia type 1 is caused by homozygous or compound heterozygous mutations in APTX (chromosome 9p13.3). Case presentation: Our patient has a clinical presentation that is typical for ataxia with oculomotor apraxia type 1 with no particularly severe phenotype. Multiplex Ligation-dependent Probe Amplification analysis resulted in the identification of a homozygous deletion of all coding APTX exons (3 to 9). SNP array analysis using the Illumina Infinium CytoSNP-850K microarray indicated that the deletion was about 61kb. Based on the SNP array results, the breakpoints were found using direct sequence analysis: c.-5+1225_144991del67512, p.?.

Since parents were heterozygous for the deletion, homozygous complete APTX deletions have been described in literature for two other patients. We obtained a sample from one of these two patients and characterized the deletion (156kb) as c.-23379_+115566del155489, p.? The more severe phenotype reported for this patient is not observed in our patient. It remains unclear whether the larger size of the deletion (156kb vs 61kb) plays a role in the phenotype; no extra genes are deleted besides the non-coding APTX exons 1A and 2. Conclusion Here we described an ataxia with oculomotor apraxia type 1 patient who has a homozygous deletion of the complete coding region of APTX instead of homozygous or compound heterozygous for APTX mutations. We were unable to confirm a more severe phenotype for ataxia with oculomotor apraxia type 1 patients with a complete deletion of APTX, however, more research is needed to study the exact breakpoints/sizes of these kind of APTX deletions.
2440T


Non-synonymous missense changes that result in amino acid substitutions in the protein product represent the majority of variants of uncertain clinical significance (VUSs) identified by genetic testing. To analyze whether commonly used in silico tools can accurately characterize the possible disease association of missense mutations, we compared the accuracy of six commonly used algorithms (Align-GVGD, SIFT, PolyPhen-2, MAPP-MMR, SIFT, Grantham Analysis and Condel) using a dataset of 1,118 BRCA1, BRCA2, MLH1, and MSH2 variants previously classified as clinically deleterious or benign by our laboratory's variant classification program. For all algorithms, except Align-GVGD, the false-positive (FP) rate compared to the reference classification was substantially higher than the traditionally accepted threshold for clinical confidence, with a range from 30.6% - 58.5% for BRCA1, 27.1% - 40.1% for BRCA2, 17.9% - 67.9% for MLH1, and 17.1% - 56.1% for MSH2. Although the FP rates using Align-GVGD for all four genes were lower, including values of 2.2% for BRCA1 and 7.9% for BRCA2, the sample size was too small to provide robust analysis due to exclusion of 750 variants that were used to train the algorithm. The high FP rates for Condel, which classifies variants based on a weighted average of scores from five in silico tools, suggests that the use of multiple models is not significantly more accurate than any of the individual models in isolation. The results of this study suggest that none of the commonly used in silico tools achieve the use of a minimum threshold of specificity for the clinical use of predictive tools.

2441S

Detection of Pathogenic Mutations in Moderate Penetrance Breast Cancer Genes Significantly Increases the Number of Patients Identified as Candidates for Increased Screening. E. Rosenthal, H. McCoy, K. Moyes, B. Evans, R. Wenstrup. Myriad Genetic Laboratories, Salt Lake City, UT.

Hypothesis/Purpose: We sought to establish the clinical utility of multi-gene hereditary cancer panels for the identification of patients who may benefit from interventions to reduce cancer risk, focusing on 3 genes in which pathogenic mutations are estimated to carry a >20% lifetime risk for breast cancer: CHEK2, PALB2 and ATM. This level of risk meets professional society recommendations for initiating breast screening at younger ages and the use of MRI in addition to mammography. We determined the proportion of women identified as candidates for modified screening through genetic testing who would not have been identified with family history. Methods: We used the Claus tables to evaluate the reported family histories of women in whom pathogenic mutations were found in CHEK2, PALB2 and ATM through clinical testing with a 25-gene hereditary cancer panel. We determined the proportion of these women who would have been identified as having a >20% lifetime breast cancer risk based on family history. Results: Among 9201 patients tested between 09/04/2013 and 04/17/2014, 174 female patients were identified with a single mutation in either: CHEK2 (n=71), PALB2 (n=39) and ATM (n=64). Three patients carried mutations in 2 of the genes. Excluding 5 women who also carried a mutation in BRCA1 or BRCA2, only 15 (8.7%) of the remaining 172 women reached the >20% threshold for lifetime breast cancer risk using the Claus tables. By comparison, among this same group of 9201 tested individuals, 275 had a pathogenic mutation in BRCA1 or BRCA2. Therefore, inclusion of the 3 moderate penetrance genes improved the sensitivity of the testing, as defined by the identification of a genetic finding associated with an established breast cancer screening recommendation, by 62.5%. Conclusions: Clinical testing including the moderate penetrance genes CHEK2, PALB2 and ATM significantly increases the likelihood of identifying women who can benefit from modified medical management strategies that would not have been applied based on family history alone.

2442M


In March, 2013 The American College of Medical Genetics and Genomics (ACMG) issued recommendations to include reporting of incidental findings (IF) on all individuals who receive clinical whole exome sequencing (WES) or whole genome sequencing (WGS). These recommendations advised actively reviewing 56 genes primarily associated with an increased risk for cancer or cardiac disease for the presence of known pathogenic (KP) or expected pathogenic (EP) mutations, since early intervention is likely to reduce or prevent accrued morbidity or early mortality. The ACMG estimated that ~1% of patients undergoing WES will have reportable incidental findings. At GeneDx, IF analysis has been offered since July 1, 2013. Since then, 2148 unrelated probands underwent WES, 433 (20.2%) of whom opted out of receiving IF. The recommended 56 genes were assessed for KP and EP mutations present at ≤1% frequency in the 1000 Genomes database or our internal GeneDx annotation database, excluding previously identified polymorphisms and large copy number mutations. All variants were evaluated for evidence of pathogenicity in the literature (HGMD database), and consistency with the ACMG reportable mutation spectrum for the gene. Of the 975 completed tests, we have reported 86 IF in 81 (8.3%) cases. Of the 86 pathogenic variants, 61 (70.9%) were in cancer-related genes, 9 (10.5%) were in cancer-related genes, and 16 (18.6%) mutations were identified in genes associated with either hypercholesterolemia or malignant hyperthermia. IF were most commonly reported in MYBPC3 (19 mutations), SCN5A (10), KCNQ1 (8), and KCNH2 (7). No reportable variant was found in 34 (61%) of the 56 genes recommended by ACMG. Although previous pilot studies indicated that probands undergoing WES may have a high number of reportable ACMG IF (16-24%), these frequencies have not been replicated in our current study. Based on a much larger dataset, we have now found that the frequency of reportable ACMG IF is ~7-9% of cases undergoing WES, which is ~7-9x higher than the initial estimate provided by ACMG. This number is likely more accurate estimate of IF but may increase slightly in the future as more variants are evaluated for pathogenicity and the literature continues to expand.

2443T

Targeted gene panel sequencing using multi-parallel single-plex PCR amplification for the detection of somatic mutations. K. Yap, K. Kyotani, K. Tamura, M. Montoya, Y. Nakamura. Department of Medicine, Section of Hematology/Oncology. The University of Chicago, Chicago, IL, USA.

The use of whole exome sequencing has been instrumental in the discovery of novel variants and somatic mutations that may be involved in the pathogenesis of human disease. Generally, non-synonymous variants which give rise to protein coding changes or protein loss are more likely to be the culprit, indicating the positive utility of exome sequencing. However, the broad target regions of whole exome sequencing usually dictates the sacrifice of overall sequencing read coverage, reducing the ability to detect low frequency variants, which is especially relevant for somatic mutation detection in cancer. In this study, we selected 10 bladder cancer samples that were previously whole exome sequenced, for deeper sequencing using a targeted gene panel of 60 cancer genes (1070 amplicons) developed on the Wafergen Smartchip TE system. Using a multi parallel single plex PCR-based approach, the design avoids the undetectable fall-outs and primer inefficiencies that typically occur in a multiplex primer pool PCR reaction. The next generation of chip-based technology will allow for the development of multi-parallel PCR reactions that is possible in a plate-based design. We found that all the variants (33 somatic mutations) that were called on exome sequencing can be confirmed by this targeted gene panel. Furthermore, 10 somatic mutations were additionally detected by the targeted gene panel due to the significantly improved overall read coverage. Conversely, there were a number of false positive variants that arose in this targeted gene sequencing approach due to the PCR-based amplification steps. This issue can be overcome by conservative filtering criteria and disregarding variants that are not supported by sequencing reads with strong signal. Overall, this targeted gene panel sequencing will significantly reduce sequencing cost and increase coverage in target gene regions of interest. This will allow for the detection of low frequency somatic mutations in heterogeneous cancer samples, which is instrumental for clinical and research purposes. Additional bioinformatics processing will need to be carried out on data sets derived from targeted gene panels to account for the erroneous variants that may arise due to PCR amplifications.
Clinical and technical evaluation of NGS based testing for hereditary cancer syndromes.

**Introduction:** Next-generation sequencing (NGS) of gene panels and exomes is gaining clinical acceptance, although questions remain about the sensitivity, specificity and clinical implications of these tests. Expanding on our recently published work (Kurian et al, J Clin Oncol, 2014) we considered whether NGS can both replace and supplement traditional BRCA1/BRCA2 testing in patients indicated for hereditary breast/ovarian cancer testing.

**Methods:** Over 900 patients indicated for BRCA1/BRCA2 testing under clinical management guidelines were recruited and tested with a 220-gene NGS panel. In this analysis we focused on both sequence variants and copy-number variants (CNVs) detected by NGS in 29 moderate to high penetrance cancer risk genes. These patients also had traditional genetic testing results (e.g. from Sanger sequencing and QPCR) available for comparison.

**Results:** 261 alterations (196 pathogenic and 65 others) were reported in the traditional data, and all were detected by NGS when the corresponding test was ordered. In this set are 141 insertions, deletions, complex events and CNVs, with particularly technically challenging variants including indels up to 126bp and CNVs as small as one exon. Specificity was also high: all NGS variants for which we sought orthogonal confirmation (n>2000) were confirmed, including 42 pathogenic variants not previously reported. No single calling algorithm achieved this performance but rather a combination did (GATK, Freebayes, PolyMNP, CNV/tae and split-read detection). By testing for genes beyond BRCA1/2, diagnostic yield increased over 50%. As expected, clinical actionability of these findings varied, with various changes in care recommended when we were able to recontact patients. Communication of this information proved clinically feasible and was appreciated by the patients.

**Conclusions:** NGS can be a viable replacement for traditional genetic testing techniques for hereditary cancers and other syndromes, with the additional benefit of cost-effectively increasing diagnostic yield. Orthogonal confirmation of clinical NGS results remains a strongly recommended practice. Although the concordance of traditional and NGS data suggests that the cost-benefit of confirmation merits careful consideration over time.

**Note:** All of the variants in this study and their interpretations will be released to public databases by the time of the meeting.

---

Beyond BRCA1 and BRCA2: results from screening 94 genes in 200 patients with familial breast and ovarian cancer using panel sequencing and custom array-CGH.

**Background:** Breast and ovarian cancer (BC/OC) predisposition has been associated with a number of high- and low-penetration susceptibility genes. With the advent of NGS-based panel sequencing, testing all these genes became feasible in addition to the analysis of BRCA1, BRCA2, RAD51C and CHEK2. Here we report on the results of custom array-CGH for deletion/duplication analysis and panel-based screening of 94 genes that have been associated with hereditary cancer predisposition.

**Method** Selection criteria for the 200 patients included in this study were defined by the German Consortium for Breast and Ovarian Cancer. High-risk patients with previously excluded mutations in BRCA1 and BRCA2 were also included. Target enrichment was performed with the Illumina TruSight cancer panel, which includes 94 genes associated with both common (e.g., breast, colorectal) and rare cancers. All 94 genes were additionally analyzed with a customized, high resolution array CGH. **Results** In 28 % of the patients, BRCA1 and BRCA2 variations have been found. These were either clearly pathogenic protein truncating mutations or rare, unclassified missense variations with high probability of effect. In 30 % of the patients we found nonsense-mutations or unclassified missense variants in low penetrance susceptibility genes, especially NBN, CDH1, ATM and PALB2. TP53 mutations were revealed in 2 % of our cases. Interestingly, one of these TP53 mutations was found as a 10 % mosaic in blood cells from a patient with no familial history of cancer but late-onset of both BC and OC. The frequency of the TP53 mutation in these tumours is currently being investigated. In one patient with a familial history of both BC and colon cancer a pathogenic mutation in MLH1 could be identified, along with a frame-shift mutation in BRCA1. Complementary custom array-CGH in all 200 patients identified deletions in ATM, BRCA1, CHEK2 or RAD51C in 3 % of the cohort.**Despite comprehensive testing, 37 % of the patients did not reveal any convincing mutation, neither on nucleotide level nor on genomic level.**

**Conclusion** The extension of mutation screening beyond BRCA1 and BRCA2 reveals disease-causing mutations in high-penetrance genes, like TP53, as well as mutations in low-penetrance susceptibility genes, especially CDH1, ATM and NBN. However, the enormous number of unclassified sequence variants and the detection of mutations in “non-breast-cancer” genes pose a huge challenge for genetic counselling.
2446T


Whole Exome Sequencing (WES) has become an effective diagnostic tool for the identification of the molecular basis of human genetic disorders. WES allows the investigation of a patient’s clinical phenotype through the evaluation of specific variants found by the simultaneous analysis of approximately 20,000 genes in the human genome. The process begins with assessing the over 6000 genes already associated with human diseases. Resources including the Human Gene Mutation Database (HGMD), population frequency databases, gene-specific databases, and in silico pathogenicity prediction tools can be used to evaluate variants. New potential genetic etiologies not yet linked with human genetic diseases (candidate genes) may also be interrogated. The clinical utility and research potential of reporting out variants in candidate genes has yet to be established. During the initial 28 months of our WES program, approximately 7% (157/2242) of the reported variants have been in candidate genes. All reported candidate genes had animal models or pathway, expression, or functional studies tying the phenotype to the gene. Findings reported in candidate genes were retrospectively reviewed to determine whether any new information linking these genes to a genetic disorder became available after the initial report. These candidate genes were crosschecked against HGMD to assess how many of these genes had been newly added to the database since the report date. Of these, 3% (13/431 newly added genes) have since been added to HGMD as genetic causes of human disease. The associated disorders include neuropsychiatric, intellectual disability, kidney, cardiac, and metabolic disorders, among others. Autonomous disorders such as Noonan, Prader-Willi, and Joubert syndromes. This result suggests there is a value to reporting findings in candidate genes. Reevaluation of candidate genes on a periodic basis may be helpful in clarifying the cause of a phenotype with unknown genetic etiology. With this additional information, patient counseling should emphasize that genetic information is improving over time as a component of the benefits and limitations of WES. Ongoing review of already reported results is important to maximize the benefits of the existing technology and to ensure appropriate genomic diagnosis is provided.

2447S

Diagnostic sequencing in integrated clinical and research laboratory setting for 100 families at the Dorrance Center for Rare Childhood Disorders, S. M. Remmers, E. L. Shiend, J. J. Corneveaux, J. L. Schrauwen, J. J. Corneveaux, A. A. Kurdoglu, R. M. Wapner, M. de Both, R. Richholt, M. de Both, S. Rangasamy, M. J. Huettelman, V. Narayanan, D. W. Craig, 1) Neurogenomics Division, Translational Genomics Research Institute, Phoenix AZ, USA; 2) Dorrance Center for Rare Childhood Disorders, Translational Genomics Research Institute, Phoenix AZ, USA; 3) Molecular and Cellular Biology Interdisciplinary Graduate Program, School of Life Sciences, Arizona State University, Tempe AZ, USA; 4) Department of Medical Genetics, University of Antwerp, Antwerp, Belgium; 5) University of Arizona College of Medicine-Phoenix, Phoenix, AZ, USA.

Integration of a clinical diagnostic center and a genomic research lab provides a unique opportunity to improve the life of children with rare, undiagnosed genetic disorders. To date, we enrolled 225 families and sequenced the whole-exome, whole-genome of 150. Study design included sequencing whole-exome of family trios, constant re-analysis of genomic data with updated annotation and prioritization methods, application of mRNAsq from whole blood for challenging cases and for validation of candidate causal variants, and the establishment of patient specific fibroblast cell lines for functional studies. We report a concise summary of the first 100 families enrolled. Enrollment criteria included a previously undiagnosed condition, ambiguous genetic origin, and negative, or inconclusive genetic tests. Most patients exhibited some form of neurological phenotype and were characterized as one of following condition: Neurodegeneration, Epilepsy, Movement disorder, Mitochondrial disorder, Aicardi Syndrome, Aicardi Goutieres Syndrome, Mitochondrial Leukodystrophy, GI dysmotility, Hearing loss, and Neuromuscular dysfunction. We applied Illumina TruSeq Exome Enrichment capture method, HiSeq2000 platform with a 2x100bp paired-end sequencing set up as standard. In summary, we sequenced 60 family trios, 15 parent-proband duos, 9 singletons, and 16 families with multiple affected and/or siblings. In 16 families, whole genome sequence of the proband and/or the parents were also obtained. In 28 families Illumina TruSeq RNA sequencing was additional. We performed RNAseq and from 11 families, we obtained fibroblast tissue for cell culture. After filtration and prioritization, we categorized variants by their likelihood as: 1. Presumed Causal 2. Likely Causal/Major Contributor, 3. Candidate, 4. Unknown. In approximately 35% of cases a presumed causal or likely causal variant was found, and in about 9% of cases we identified a candidate variant. Post diagnosis, we helped facilitate enrollment of a number of diagnosed families in NIH funded clinical trials, or connected them with experts in the field, or provided alternative treatment options. We developed mouse models, and large-scale drug screens with previously approved FDA drug compounds for selected causal genes. Continuous outreach to the families and close integration of diagnostic and research approach was able to provide conclusion to the diagnostic odyssey for many families.

2448M


Vascular type of Ehlers-Danlos syndrome (vEDS), also known as EDS type IV (MIM#130050), is a life-threatening autosomal dominant inherited connective tissue disorder mainly caused by mutations in type III collagen, COL3A1 gene. The main types of mutations in vEDS are single amino acid substitutions for glycine in the GLY-X-Y repeat of the triple helical region or invariant splice sites. Recently, premature translation termination mutations are reported affecting nonsense-mediated mRNA decay of mutated allele. Here, we described a 26 y/o female with clinical features and family history of vEDS and characterized the functional consequences of an intronic COL3A1 deleted mutation. The coding regions and flanking intronic bases of the COL3A1 gene were screened by high resolution melting curve analysis method using genomic DNA. The patient is a heterozygote for a four bp deletion at 5’ splice site of intron 8 with in genomic DNA (c.636+5+4del GTAA/IV8) that do not affect the consensus GT or AG splice sites and changed from G to C transversion at Position +5 of the splice donor site in the intron. A minigene splicing assay demonstrated that this deletion is sufficient to cause a skipping of exon 8. G at Position +5 of the splice donor site is supportive of an importance for splicing of COL3A1. Intronic mutations may be responsible for vEDS in some families with otherwise negative mutation screening of COL3A1.
2449T
The Relationship between Blood Index and Thalassemia Disease. N. Ghazavi, M. Behnam, F. Hosseini, A. Hejazifar, A.R. Salehi Chaleshori, M. Salehi, N. Mansouri. Medical Genetic Center of Genome, Isfahan, Iran. Background and Aims: Thalassemia is a severe hemoglobin disease, recognizing by CBC and hemoglobin electrophoresis primarily. It is possible that CBC and hemoglobin electrophoresis were performed for the index case who recognized as αThalassemia but finally experimental studies clearly demonstrated that it was different type of thalassemia. In the present study, we report a case of β-thalassemia mutations which is called -101 C>T relative to the transcription start site of β-globin gene, and is a β+-thalassemia case, but it has been observed with different indices. Methods: The patient was admitted based on hematologic indices as α-Thalassemia, and diagnostic tests for α-Thalassemia, including GAP PCR, ARMS-PCR for α-Thalassemia were performed and all examinations were normal. For increased confidence, sequencing for patient’s β-globins’ gene was performed. We also observe more cases with the same index. Results: Our findings, present the incidence of mentioned mutation in β-globin gene in β+ patient, in contrast to the hematologic indices for β-Thalassemia heterozygote. Moreover, this case has no resemblance to other β+-Thalassemia cases and it can be mistaken as α-Thalassemia. (For example: The - 88 C>A mutation relative to the transcription start site, which several of our indices present beta-thalassemia). Conclusion: As mentioned above, our outcomes are extremely similar to the results of CBC and hemoglobin electrophoresis of α-Thalassemia, but the patient was β-Thalassemia. Such conditions as cited above are hazardous and end to affected neonates if the other partner is β-Thalassemia patient. In this condition, Sequencing of β-globins for one of the couples could be helpful.

2450S
Investigation of Genomic Deletions and Duplications by Custom MLPA in a Cohort of 338 Patients with Obesity, Developmental Delay, Hyperphagia, and Additional Features. C.S. D’Angelò1, C.V. Monica1, I.E.C. Claudia1, A.K. Chong2, R.B. Deborah3, R.B. Deborah3, M.P. Charles1, P.K. Celia3. 1) Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of Sao Paulo; 2) Department of Pediatrics, Children Institute, School of Medicine, University of Sao Paulo; 3) Department of Medical Genetics, School of Medicine, University of Sao Paulo. Obesity is highly heritable and a significant health problem for people with intellectual disability (ID). The recent implementation of whole-genome chromosomal microarray analysis (CMA) has resulted in the discovery of novel and rare CNVs implicated as risk factors for obesity, often also present with ID. However, the strong phenotypic overlap between syndromic forms of obesity poses challenges to accurate diagnosis, and many different individual cytogenetic and molecular approaches may be required. Multiplex dependent-probe amplification (MLPA) enables the simultaneous analysis of multiple targeted loci in a single test, and serves as an important investigation tool in situations where specific microdeletion and/or microduplication syndromes are suspected. Our aim was to design a synthetic probe set to screen for deletions and duplications at previously described loci associated with obesity in a cohort of 338 patients with syndrome of obesity of unknown etiology who have had a normal methylation test for PWS or in whom no 1p36 deletion could be found. Eighteen alterations were detected using the synthetic MLPA probe set. Ten of the detected alterations were delineated by chromosomal microarray, while the remaining alterations were fine-mapped by additional MLPA probes incorporated into commercial kits. We identified nine known CNVs representing clinically well defined microdeletion syndromes with obesity as a clinical feature: in 2q37 (4 cases), 9q34 (1 case) and 17p11.2 (4 cases). We also identified three deletions and a reciprocal duplication in the DiGeorge syndrome locus at 22q11.2. Additionally, we have identified two recurrent small deletions within the 22q11.2 ‘distal’ locus associated with a variable clinical phenotype and obesity in some individuals. Three individuals were found to have a recurrent CNV of one susceptibility loci at 1q21.1 ‘distal’, 16p11.2 ‘distal’, and 16p11.2 ‘proximal’. The overall detection rate with the synthetic MLPA probe set was about 5.3% (19 out of 338). Our experience leads us to suggest that MLPA could be considered as a first line diagnostic tool in the detection of the chromosomal microarrays for syndromic obesity, allowing for a number of microdeletion syndromes loci (e.g. 1p36, 2p25, 2q37, 6q16, 9q34, 11p14, 16p11.2, and 17p11.2) known to be clinically relevant for this patient population to be interrogated simultaneously. Financial Support: CAPES-FAPESP, CNPq.

2451M
Prenatal detection of chromosomal aberrations and its reflection at adult age. B.B. Ganguly1, S. Mandal1, N.N. Kadami2. 1) MGM Center for Genetic Research & Diagnosis, MGM Institute of Health Sciences, Mumbai, India; 2) Department of Pediatrics, MGM Institute of Health Sciences. Heritable chromosomal abnormalities have been reported in ~50% spontaneous abortion and ~5-10% of live pregnancies with prevalence of trisomy 21. Intervention of prenatal diagnosis is generally considered in suspected cases with high risk triple screen, significant family history and/or advanced maternal age. In our previous data, Down syndrome was recorded in ~4% cases, while balanced translocation and inversion were present in 5% among 140 pregnancies. The cases with balanced rearrangements might achieve a long life span since such alterations do not contribute to major phenotypic or clinical manifestation. Therefore, chromosome aberrations are expected in ~10% of children including 5% of adults; however, these aberrations can be prevented from transmission through generations with a straightforward approach of prenatal chromosomal analysis by conventional techniques. In the present report, we present data on chromosomal status in 269 fetal samples of live pregnancies. CVS or amniotic fluid was processed for FISH and/or conventional analysis of karyotypic status. Interface cells were considered for FISH whereas metaphases were processed for TG-banding and karyotyping. Mitotic metaphases were obtained from long-term culture of the specimen using standard technique. Ikaros and ISIS software (MetaSystems, Germany) was utilized for conventional karyotyping and FISH analysis. The cases were categorized into three age groups of 10 years interval and the abnormalities detected by FISH and G-banding was presented in different age groups. The incidence abnormality was maximum in the advanced age group as expected; however, the youngest age group showed higher number of structural aberrations over group II. Conventional G-banding result appeared with 71.5% structural aberrations, which was not detectable by FISH. The present report highlights the importance of karyotyping over interphase FISH. Also the study recommends the intervention of karyotyping at fetal age as well as before the termination of pregnancy and reduction of chromosome aberrations among adults which would further reduce the burden of birth defects and/or reproductive failure.

2452T
45,X(30%); 46,X,i(Xq)(60%) mosaicism. Case report. M. Pérez Sánchez1, 3, M. Aora Guijosa1, 3, M. Lopez Melchor1, 3, A.R. González Ramírez1, 2. 1) UGC Análisis Clínicos, Hospital Virgen de las Nieves, Granada, Spain; 2) Fibao. Hospital Clínico San Cecilio, Granada, Spain; 3) Instituto de Investigación Biosanitaria de Granada, Granada, Spain. The common structural rearrangement of the X is 46,X,i(Xq). An i(Xq) chromosome theoretically consists of wo copies of the entire long arm and no short arm. The gross banding patterns of most isocromosomes fits with this interpretation but in detail the situation is more complicated. Some isocromosomes have one centromere and are perfectly symmetrical, while others are asymmetrical but have two centromeres close to each other, and still others have a clearly asymmetrical central portion, raising the possibility of a more complex abnormality. On the other hand, Turner mosaicism are present in an important group of Turner syndrome with a wide clinical features ranging from full-blown Turner syndrome to normal phenotype. The karyotype is the election technique for these cases, but for a finest detection of the rearrangement point, the microarrays CGH will be of election. Here we present a case of a 19 years old woman that was referred for genetics studies. The clinical findings were amenorrhea with treatment response, normal secondary sexual characteristics, hormonal status at normal levels, the right ovary was not detected and the left one was normal when magnetic resonance studies were done. No appreciable mental retardation were detected. Chromosome culture and karyotyping were realized by standard thecnic and showed karyotype of 45,X (30%); 46,X,i(Xq) (60%) mosaicism. The rearrangement of the Xp region deleted and the Xq region duplicated, GH-array 180 K (PerkinElmer Platform) was performed, with a result of total loss of 55,17 Mb in the Xp22.33-p11.21 region (genomic coordinates: chrX:296520- 55468477), a partial duplication of 2.77 Mb in the Xp11.21-p11.1 (genomic coordinates: 55556815-55634786), and partial duplication of the entire Xp arm. The patient was concordant with CGH-array in the rearrangement of the karyotype results, but with the possibility of detect the exact rearrangement point. The presence of the Xp region near to the centromere can explain the mild Turner clinical features in this patient. As a conclusion, when a rearrangement is detected by karyotype, the CGH-Array can be indicated to detect the exact rearrangement point.
MALBAC WGA. By subsampling the generated sequences, we also were able to determine minimal coverage to achieve a 10 MB resolution. Results show a comparable performance between the PicoPLEX and the MALBAC WGA for their ability to produce optimal PGD by MPS results from a limited number of cells.

Whole genome amplification (WGA) is currently performed on biopsied blastocyst trophectoderm cells to subsequently perform aCGH to screen for large chromosomal aberrations in a Pre-implantation Genetic Diagnosis (PGD) setting. The different currently available WGA methods lead to amplification bias resulting in over- and under-represented regions in the genome. Current WGA methods, such as PicoPLEX (Rubicon Genomics) and subsequent aCGH analysis, make it possible to detect duplications and deletions at a 10 Mb resolution.

An alternative to aCGH is low coverage Massive Parallel Sequencing (MPS). Genomic regions for which a significantly higher or lower number of sequences are generated can be called as duplications and deletions. Unlike with aCGH, there is no reference DNA used within the MPS assay that can be used for a within-assay normalization of the representation bias. WGA representation bias might thus be more critical when using MPS.

Recently, a new Multiple Annealing and Looping Based Amplification Cycles (MALBAC) WGA method has been published, claiming unparalleled performance. This method has not yet been studied in a PGD setting. The goal of this study was to compare the well established PicoPLEX and MALBAC WGA for their ability to produce optimal PGD by MPS results from a limited number of cells.

Six repeats of samples consisting of 1 cell, 3 cells and 5 cells were collected from the Loucy lymphoblastoid cell line using micromanipulation techniques. Three repeats were amplified using PicoPLEX (Rubicon Genomics), and three using MALBAC. The amplification products were used to create Illumina sequencing libraries and were sequenced on an Illumina HiSeq1500. Results show a comparable performance between the PicoPLEX and the MALBAC WGA. By subsampling the generated sequences, we also were able to determine minimal coverage to achieve a 10 MB resolution.

With the emergence of next generation sequencing technology, the quantity of sequencing data being produced by genetics laboratories is exponentially increasing. In the last decade, not only has the volume of clinical genetic tests been on the rise, but the average number of genes covered by those tests has also increased dramatically. The expanding scope of genetic tests is leading to more data being generated and an increased rate of variants of unknown significance being returned to clinicians. Siloed data sets restrict our understanding of the spectrum of variation for a particular disease and result in higher rates of variants of uncertain significance. Data sharing is proving to be an essential element for overcoming these challenges of the genomic era. VariantWire, a consortium supported by Genelinisght, provides a data-sharing network that allows clinical genetic testing labs to share variant and gene interpretations, and the evidence behind those interpretations, in real time. Data sharing on VariantWire occurs in a secure environment and all case information in the network is de-identified. VariantWire is governed by the VariantWire Committee, which is comprised of one member from each participating entity (listed in author affiliations). The VariantWire committee guides the content of the data sharing policy and reviews all applications from laboratories seeking to join the network. As of May 21st 2014, VariantWire is sharing 23,318 variants across 336 genes. Of the shared variants 2,973 are classified as clinically significant (pathogenic or likely pathogenic) for over 100 diseases. Another 3,907 variants are classified as unknown significance and 16,438 were determined to be either likely benign or benign. Finally, 24 somatic variants in the EGFR gene are involved with drug response. Two hundred and seventeen variants have been identified in more than one lab, with 61 of those variants being classified as either pathogenic or uncertain significance. Of those 61 clinically significant variants seen more than once, only two variants had inconsistent classifications across labs. By enabling real-time data-sharing, labs using VariantWire are assured that they have the most up-to-date knowledge from other participating labs. Finally, the Genelinisght team aims in helping labs submit their variant data to ClinVar, a freely available database, ensuring the broadest sharing of data to further our clinical understanding of the human genome.
Developing Exclusion Datasets for Genome Filtering in the MedSeq Project. K. Machini 1,2, R. Shakkatyan 1, H. McLaughlin 1,4, O. Ceyhan Birsoy 2,1, D. Mettlerville 1, M. Lebo 2,1, R.C. Green 1,2,3, H.L. Rehm 1,2,5, 1) Partners Healthcare Personalized Medicine, Cambridge, MA; 2) Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 3) Division of Genetics, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA.

A major obstacle in the clinical application of whole genome sequencing (WGS) is the difficulty in systematically interpreting the enormous amount of data generated. This includes analyzing a genome for the return of incidental findings based upon the interpretation of previously published variants as well as the interpretation of novel loss of function variants in disease-associated genes. Within the framework of the MedSeq Project, 100 participants, half healthy and half with cardiomyopathy, are undergoing WGS. Sequencing is performed at Illumina’s CLIA-certified laboratory and alignment, variant calling, annotation and variant review are performed at the Partners’ Laboratory for Molecular Medicine. First, annotated variants with a minor allele frequency (MAF) <5% are filtered to identify those classified as DM or DM7 in the Human Gene Mutation Database (HGMD). A separate filter identifies nonsense, frameshift, and canonical splice-site (+/-1,2) variants with a MAF <1% from a list of disease-associated genes. Although this is a common strategy employed by laboratories performing incidental findings analysis, it is hindered by the plethora of variants with false claims of pathogenicity or in genes with insufficient evidence for disease causality. After full interpretation of the first 27 genomes, involving review of 1022 rare variants with claims of pathogenicity, only 74 variants (average of 2.7 per genome) reached the evidence level to be clinically reported. In the course of excluding the majority of variants, we have developed a series of filters. First, 19% of variants can now be excluded using a dataset of common variants encountered in >10% of individuals in our cohort. Second, 14% can be excluded using quality metrics derived from analyzing Sanger confirmation data to improve quality thresholds. Third, 27% of variants can now be excluded using a dataset of genes with claimed medical associations that have insufficient evidence. Finally, 15% of reported pathogenic variants can now be excluded because of evidence for a benign interpretation. This has allowed a dramatic reduction in genome review time which will continue to be reduced as quality thresholds improve.

Using an average coverage of 100x to 200x, it was possible to identify pathogenic variations in 62% of individuals with skeletal dysplasias and 65% of patients diagnosed with a form of neuromuscular disorder. Among the first 500 reported DES families, 168 (34%) have implemented this methodology through a customized panel of genes covering >20x and 69.63%; of the bases >10x. Despite a much less efficient coverage, a heterozygous de novo variant in the PRKAG2 gene, not reported in HGMD, was identified and confirmed by Sanger sequencing. The p.Phe293Val (c.877T>G) variant was found uniquely in the index case and was classified as ‘likely pathogenic’ in accordance with current recommendations. PRKAG2 pathogenic variants influence cardiac metabolism but pathological glycogen storage in heart due to mutated PRKAG2 may be prevention or significantly reversed pharmacologically. Conclusions: Despite being superior to WES in coverage efficacy and facility of result interpretation, particular NGS panels risk to be too narrow for providing with proper genetic service in the genome era, when effectiveness of testing influences not only genetic counselling, but also therapeutic options.
2459S
Prevalence of medically actionable findings: a summary of clinical Whole Exome Sequencing cases. Z. Niu1, J. Beuten1, M. Leduc1, W. He1, J. Zhang, P. Ward1, A. Braxton1, T. Vaughn1, D. Munzy2, F. Xia, R. Person1, S. Plon1, J. Lupski1, R. Gibbs1, A. Beaudet1, Y. Yang1, C. Eng2, 1) Molecular Human Genetics, Baylor College of Medicine, Houston, TX; 2) Texas Children's Hospital, Houston TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.

Clinical Whole Exome Sequencing test is a powerful and effective tool to identify the genetic etiology in patients with a variety of clinical presentations. This genome wide analysis has also drawn great attention to the returning of incidental findings. In this study, we prefer the term actionable finding, which may greatly impact the care of the patient (and potentially family members), if the findings can lead to prevention, early diagnosis, and medical interventions. The American College of Medical Genetics and Genomics (ACMG) issued guidelines for the return of incidental findings on a core group of 56 actionable genes from clinical whole exome or genome sequence to patients. From the 2000 sequential clinical Whole Exome Sequencing cases completed at the Medical Genetics Laboratory and Whole Genome Laboratory of Baylor College of Medicine, 95 medically actionable findings were reported for an overall rate of ~4.8%. Of these, 60 findings were reported in genes on the ACMG recommended list, and 35 findings were reported in genes which met the criteria for medically actionable by expert opinion from the clinical and diagnostic team. With the consent form was updated with an opt-out from actionable findings in non-ACMG recommended gene, 2 patients out of 190 opted out of additional reporting. Majority of families requested to receive all aspects of actionable findings. 24 findings were reported in hereditary cancer predisposition genes in patients without cancer history in the family. 2,699 pathogenic and 1,957 likely pathogenic variations were identified in the 95 patients with incidental findings. Followup testing of the parents for the medically actionable findings can be ordered after disclosure of the proband's result at no cost. 33 requests from 19 families of the 95 patients with incidental findings have been received to date. In multiple cases, parental testing was requested for one parent only, typically when the phenotype related to the actionable finding was noted by review of family history. Our early experience exemplified the extent of medically actionable findings uncovered from clinical Whole Exome Sequencing. Studies are needed to further evaluate the long term impact on patient care and at-risk family members when a medically actionable finding is reported.

2460M
Identification of Complex Rearrangements of the MECP2 Gene Requires a Combination of Molecular Diagnostic Techniques. S. Ordica, D. Lahey, O. Jaronina*, N. Carson*. Molecular Genetics, Regional Genetics Program, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada.

The methyl-Cpg-binding protein 2 gene (MECP2; MIM#300005) is a chromosome binding protein that acts as a transcriptional repressor, located on the X chromosome. Mutations in this gene can cause mental retardation, Rett Syndrome (RTT; MIM#312750) and have also been associated with Autism Spectrum Disorders. While mutations in this gene are the most common, complex rearrangements also occur. Here we describe three cases with complex rearrangements involving the MECP2 gene that were identified using MLPA analysis and required further molecular investigations. Case 1: MECP2 MLPA analysis in a female patient showed a reduced signal (~75% signal) for all MECP2 probes and a heterozygous deletion of CDKL5 (MIM 300203) and ARX (MIM 300382) that reside on the p-arm of the X chromosome and are implicated in Rett Syndrome and Mental Retardation. Consistent with these findings, microarray analysis showed a loss of a portion of the p-arm of the X, and a mosaic loss of the remaining X chromosome. Case 2: MECP2 MLPA analysis in a male patient showed a triplication of exons 1 and 2 and a duplication of exon 3 and 4 of the MECP2 gene. Consistent with these findings, microarray analysis showed a triplication and a duplication of the corresponding regions on Xp28. Since the patient’s phenotype was indicative of the presence of two copies of MECP2 (ie: MECP2 duplication syndrome, MIM 300260), a third partial copy is likely to be non-functional. Case 3: MECP2 MLPA analysis in a female patient showed a partial duplication of exon 4. Consistent with the MLPA findings, targeted sequencing of the MECP2 gene revealed a complex rearrangement involving exon 4 of the MECP2 gene: c.[523_1075delinsCG:1161_1188del], p.[Lys175_Ala358delinsArgfs; Pro388_Ser396del]. This mutation event is consistent with the presence of deletion hot spots in this gene and is predicted to be deleterious to MECP2 gene function. The three cases presented here illustrate that while MECP2 analysis is commonly conducted using Sanger sequencing and MLPA, a combination of molecular techniques is sometimes required to detect complex rearrangements involving the MECP2 gene and provide a clear diagnostic result.

2461T
Pathogenic mutations in genes responsible for Maple Syrup Urine Disease type 1A (BCKDHA), type 1B (BCKDHB), and type 3 (DLD) determined in a large pan-ethnic population of families referred to Genetic Counseling Clinic, University of Ottawa, Ottawa, Ontario, Canada.

Maple Syrup Urine Disease (MSUD) is an autosomal recessive metabolic system disorder that is due to dysfunction of the branched-chain alpha-keto acid dehydrogenase (BCKD) enzyme complex. This complex is composed of subunits encoded by 4 different genes. Founder mutations in 3 of these genes, BCKDHB and DLD in the Ashkenazi Jewish (AJ) and BCKDHA in the Morrocanites, lead to an increased prevalence of MSUD in these populations. We sequenced the coding region and intron-exon borders of these three genes in a pan-ethnic population of nearly 23,000 individuals referred to us for carrier screening using next generation DNA sequencing (NGS) to identify unique variants (number of different variants detected among the carriers), known variants (those previously cited in a publication or public database), and novel variants predicted to be pathogenic (nonsense mutations, conserved splice site mutations, or indels not divisible by 3). We found 16 carriers of 12 unique variants (7 known and 5 novel) in BCKDHA, 40 carriers of 15 unique variants (7 known and 8 novel) in BCKDHB and 22 carriers of 6 unique variants (2 known and 4 novel) in DLD.

Overall, 78 carriers of a pathogenic MSUD mutation were identified. The AJ founder mutations, p.Arg138Pro and p.Gly278Ser in BCKDHB and p.Gly229Cys in DLD, accounted for the majority of carriers in each subtype, even though only 6 carriers reported AJ descent. In contrast, only three carriers of the BCKDHA founder mutation, p.Tyr438Asn, were identified; most carriers for this MSUD subtype had other mutations. Interestingly, we identified about twice as many carriers of pathogenic BCKDVB variants compared to the other two MSUD subtypes. The non-BCKDHB/DLD variants were novel and we predict to be pathogenic on the basis of their truncating nature. In contrast, the groups of BCKDHA and BCKDHB unique variants consist of about half known and half novel truncating variants. In summary, we have determined the mutation prevalence and spectrum of MSUD types 1A, 1B and 3 in a large pan-ethnic population and show that a large number of novel pathogenic variants would have gone undetected using an alternative technology that is not as comprehensive as NGS.

2462S
Molecular Genetic Determination of Maple syrup urine disease cases from southwest Iran. A. Sedaghat1, G. Shariatii2,3, A. Saberi2,3, M. Hamidi4, M. Mohebbi5, H. Galehdari6.
1) Endocrinology Dept., Jundishapour University of Medical Science, Ahvaz, Iran; 2) Narges Genetic Diagnosis Lab, Ahvaz, Iran; 3) Genetic Dept., Jundishapur University of Medical Science, Ahvaz, Iran; 4) Pasteur Institute, Tehran, Iran; 5) Internal Medicine Dept., Jundishapur University of Medical Science, Ahvaz; Iran; 6) Genetic Dept., Shahid Charnam University, Ahvaz, Iran.

Maple syrup urine disease (MSUD) affects an estimated 1 in 185,000 infants worldwide, and is most frequent metabolic disorder in southwest Iran that ranges from very mild to severe form. The frequent occurrence of MSUD in this area is possibly due frequent consanguineous marriage. In last twelve month we analyzed 20 families with MSUD affected children. Most our cases have been preliminarily diagnosed by tandem mass as helpful tool for this propose. Mutations in four genes BCKDHA, BCKDHB, DTD, and DLD can cause MSUD. But we find recurrent changes in the two genes BCKDHA and BCKDHB responsible for outcome of the disease. In summary, we found in 8 cases mutations in the BCKDHA gene and in other 9 cases nucleotide changes in the BCKDHB gene. Three cases were negative, despite screening of all 4 responsible genes as have been mentioned. The negative results can be explained by two reasons: wrong diagnosis or existence of the 5th undiected gene in the pathogenesis of the MSUD. However, some detected mutations were novel in analyzed cases. In cases of novel changes in one gene, we additionally screened the other three genes to assert our results. Further, we used some predicting program such as “Predict SNP”, “I-Tasser”, and “Mutation Tester” to underline the pathogenicity of determined changes.
A homozygous frameshift insertion in the MRPS34 gene identified in a family with two affected boys Suffering from progressive retinal dystrophy, G. Shariati, A. Saberi, H. Galehdari, H. Hamidi, M. Sedaghat, M. Mohabbi, H. Galehdari. 1) Narges Medical Genetics Lab, Ahvaz, Iran, PhD, MD; 2) Genetic Dept., Jundishapur Medical science University, Ahvaz, Iran, PhD, MD; 3) Pasteur Institute, Tehran, Iran, PhD; 4) Endocrinology and Metabolic Medicine Dept, Jundishapur Medical science University, Ahvaz, Iran, MD; 5) Internal Medicine Dept, Jundishapur Medical science University, Ahvaz, Iran, MD; 6) Genetic Dept, Shahid Chamran University, Ahvaz, Iran, PhD.

Progressive retinal dystrophy covers a wide range of eye disorders including syndromic and non syndromic retinitis Pigmentosa. To date, more than 100 disease causing genes have been identified in different populations. Recently, a couple (first cousin) from southwest Iran referred us with two affected boys (8 & 6 years old) suffering from progressive retinal dystrophy as preliminary diagnosis. Genomic DNA of affected individuals was subjected for next generation sequencing (NGS) leading to the detection of numerous candidate genes including a homozygous insertion (c.320_321insAGGT) causing frameshift at codon 107 (p.Q107fs). Targeted sequencing of the MRPS34 gene exon 1 showed that both parents are heterozygous for this mutation. The MRPS34 gene codes for the mitochondrial ribosomal protein 34 that is highly expressed in eye. No clinical assertion has been found for the mentioned mutation, but we consider it as a strong candidate causative for progressive blindness in the family.


Ascertainment bias is a sampling bias in genetics, where a non-random sample is collected in such a way that some members of the intended population are less likely to be included than others. In genetics, the consequence is that the most severely affected individuals (e.g. those who present in the emergency rooms) are the first to be described in the literature for most genetic disorders. This bias impacts the development of diagnostic criteria and effectively removes patients with milder or atypical presentations from diagnosis and restricts the phenotypic spectrum of disease, literally marginalizing many individuals with the same disorder and consequently leaving them without a diagnosis. This approach was dictated in part by the cost associated with Sanger sequencing, where clinical diagnostic justification was required to order a gene test. Recently, massive parallel sequencing brought both a radical drop in cost-per-gene and the advent of large comprehensive gene panels, allowing individuals who do not classically fit diagnostic criteria to be evaluated for genes with or without high clinical utility. In patients without the classic clinical criteria, or those with a milder phenotype, genetic testing may provide the definitive diagnosis. Large gene panels also allow primary care physicians to order genetic tests on patients who only meet a subset of diagnostic criteria, moving genetics to the beginning of the testing process and effectively turning the traditional paradigm on its head. For the patient, this leads to a shortcut to what was otherwise a long maze of referrals and circumscribes the need for a clinical diagnosis prior to testing in many cases. For the geneticist, this provides much more information regarding the clinical spectrum of genetic disease. We report the results from over twenty cases of diagnoses confirmed by genetic panels (most de novo) identified by very large gene panels targeting patients with epilepsy, developmental delay, and mitochondrial disorders and the implications of identifying the genetic cause earlier than would have been achievable if clinical diagnostic criteria were used alone. We will also highlight how early identification has important clinical management implications for patient care.


Recent breakthroughs in next generation sequencing technologies have generated a major transformation in service delivery of genetic medicine. The ability to generate high quality sequence from thousands of genes in one experiment enables the development of comprehensive gene panel tests for genetically heterogeneous disorders, which are unattainable by Sanger sequencing. We have developed and clinically validated an in silico gene panel of 48 genes for hereditary spastic paraplegia (HSP) using exome sequencing. HSP represents a heterogeneous group of neurodegenerative disorders characterized by stiffness (spasticity) and progressive weakness (paraplegia) of the lower limbs, and may be associated with ataxia, intellectual disability, seizures, peripheral neuropathy, and visual defects. An accurate genetic diagnosis will reduce the number of patients who undergo serial investigations (such as multiple brain MRIs) or invasive investigations (such as muscle biopsy) in an attempt to establish a diagnosis. Confirmation of a genetic diagnosis is essential for comprehensive testing and genetic counseling for at-risk family members, and will make possible the availability of prenatal testing for at-risk couples. The previous strategy in the Molecular Diagnostic Lab has been to perform Sanger sequencing for nine genes with a diagnostic yield of ~ 44% (Yoon et al. 2013 Neurogenetics. 14:181-8). With implementation of exome sequencing we are able to widen our scope to all 48 genes known to be associated with HSP, and can readily expand the diagnostic panel as new HSP-related genes are discovered. This strategy also facilitates the development of additional disorder-specific gene panels with the same standardized workflow. We evaluated the performance characteristics of exome sequencing, including reproducibility and accuracy, with 100% detection of HSP genetic variants in our validation samples. We also explore the advantages and technical challenges of this strategy to further develop clinically relevant gene panels for genetically heterogeneous disorders.


Identification of the genetic basis for neurological disorders is made difficult by the large number of genes that can lead to such phenotypes. Further exacerbating the situation is the highly variable presentation different mutations in the same genes can show. These issues make it challenging to decide which gene(s) to analyze in detail for proper diagnosis. Sequencing individual candidate genes frequently results in a frustrating cycle of negative results that is both time consuming and expensive. Gene panels currently available typically look at dozens of genes at a time but these are difficult to maintain given the complex phenotypes observed in neurological disorders and rapid pace of new gene discovery. To circumvent these issues, Claritas developed an approach utilizing complementary sequencing technologies to analyze the exons and adjacent intronic sequence from an exome-based region of interest. This ROI includes hundreds of genes previously demonstrated as causal in a variety of neurological disorders. The phenotype-driven set of genes makes analysis faster and more economical than analysis of a whole exome and also minimizes the challenges of presenting incidental findings. By taking advantage of the power of combining the knowledge of clinical symptoms with the scale of massively parallel sequencing technology, patients are able to benefit with more directed and rapid results. DNA variants from the two platforms are compared to confirm novel pathogenic variants. In cases where coverage is missing on one of the platforms, possibly pathogenic variants are confirmed using Sanger sequencing. Confirmed variants are then analyzed by a team of medical directors and genetic counselors to determine clinical relevance with results reported back to physicians and patients.
2467T
RYR1 [Arg2454His] gene mutation identification in a family associated with malignant hyperthermia - an undergraduate research project. J. Li, D. Closemale. Biology, University of Wisconsin-Stevens Point, Stevens Point, WI.

Malignant Hyperthermia (MH) is a rare life-threatening dominant disorder that causes extreme fever, muscle rigidity, acidosis, and rapid heart rate, when exposed to general anesthesia. Researchers have previously identified four amino acid replacement mutations on the RYR1 gene in MH patients: [Arg44Cys], [Arg533Cys], [Val2117Leu] and [Arg2454His]. The RYR1 gene, located on chromosome 19, codes for membrane proteins on the surfaces of muscle cells that regulate calcium flow, which is essential for muscle contraction and relaxation. These mutations can cause symptoms of MH because they cause significant changes in the structure of the calcium channel proteins, rendering them nonfunctional. The purpose of my undergraduate research study was to screen for the aforementioned four point mutations of the RYR1 gene that may be the cause of MH in a German family that has an individual with this disorder and to identify family members that carry the mutation. With IRB approval, 15 participants filled out a survey describing whether or not they have experienced any signs of MH. DNA was extracted and purified from 10 hair follicles each. The four segments of the RYR1 gene were amplified in separate reactions by the polymerase chain reaction (PCR), cycle-sequenced using dye-labeled diodeoxynucleotides, and then separated by capillary electrophoresis. The DNA sequences of the family members, represented as electropherograms, were compared to a wild type DNA sequence provided by NCBI GenBank in order to detect the presence of any of the four point mutations mentioned above. Sequence comparisons revealed that four family members tested positive for the Arg2454His mutation. These findings correlated with all four experienced symptoms of MH. A recommendation was made to family members to be screened for the RYR1 [Arg2454His] mutation, to verify these results. Currently, there are only two genotyping companies in the U.S. that provide genetic testing of the RYR1 gene, which is very costly. However, since the specific mutation has been identified, family members would only need to have their DNA screened for this mutation, which would help lower costs considerably.

2468S
An exome sequencing strategy to diagnose lethal autosomal recessive disorders. S. Ellard1,2, E. Kivuva3, P.D. Tumpenny4, M. Parker5, A.-M. Bussell6, K. Stals6, R. Caswell7, H. Lango Allen7. 1) University of Exeter Medical School, Exeter, United Kingdom; 2) Royal Devon & Exeter NHS Foundation Trust, Exeter, United Kingdom; 3) Sheffield Children’s Hospital. Sheffield, United Kingdom.

Rare disorders resulting in prenatal or neonatal death are genetically heterogeneous. Some affected fetuses can be diagnosed by ultrasound scan but often not until mid-gestation and limited fetal DNA may be available. We used a strategy for diagnosing autosomal recessive lethal genetic disorders in non-consanguineous pedigrees with multiple affected fetuses. Exome sequencing identified genes with rare heterozygous non-synonymous coding or splicing variants in both parents and putative mutations were tested for co-segregation. In 8 couples of European ancestry we found on average only one gene where both parents were heterozygous for different rare potentially deleterious variants. A proof of principle study detected heterozygous DYNC2H1 mutations in a couple whose 5 fetuses had short rib polydactyly. Prospective analysis of three couples with multiple pregnancy terminations from 51 genes. Four diseases were determined to be responsive to diazoxide, a tier-2 approach that includes targeted next-generation sequencing and array-CGH deletion/duplication analysis of all ten known genes causing HI which can be performed. To date, we have identified the genetic etiology in 3/19 (16%) patients studied and likely causal mutations were identified in ABCB8, GLUD1 and GCK. Variants of uncertain clinical significance (VOUS) were also identified in 3 patients. Parental analysis will be performed to clarify mode of inheritance of the mutations and provide insights on the nature of the observed HI. Studies are ongoing to assess the impact of genetic testing results on patients’ medical management and clinical course.

2469M
Molecular diagnosis of congenital hyperinsulinism improves medical management and long-term outcome. G. Alkorta-Aranburu, A. Knight Johnson, L. Ma, V. Nelakuditi, S. Das, D. Del Gaudio. Human Genetics, University of Chicago, Chicago, IL., USA.

Congenital hyperinsulinism (CHI) is the most common cause of persistent hypoglycemia in children, which if unrecognized, may lead to severe brain damage. A prompt diagnosis and initiation of appropriate treatment is essential to prevent long-term neurological complications. Over the past 20 years there has been remarkable progress in understanding the molecular basis of hyperinsulinism (HI) and mutations in ten genes are currently known to cause HI in >90% of the cases suggesting that continuing research will identify new genes. Inactivating mutations of beta cell ATP sensitive K+ channel genes ABCB8 (SUR1) and KCNJ11 (Kir6.2) cause the most common and severe type of HI, which is unresponsive to diazoxide, a channel agonist. Focal HI, also diazoxide-unresponsive, is due to the combination of a paternally-inherited ABCB8 or KCNJ11 mutation and a paternal isodisomy of the 11p15 region, which is specific to the islets cells within the fetal lesion. It is important to differentiate these two types as surgery can cure focal but not diffuse hyperinsulinism, which frequently requires pancreatectomy. Molecular genetic testing and imaging with 18F-DOPA-PET/CT can help to diagnose diffuse or focal forms of HI. We developed a comprehensive 2-tiered testing algorithm for the molecular diagnosis of CHI to direct the clinicians along a process of exclusion, so that rarer forms of HI would only be considered once commoner and diazoxide-unresponsive forms are ruled out. Tier-1 includes mutation analysis of genes associated with diazoxide-unresponsive HI, ABCB8, KCNJ11 and GCK for which results are obtained within a week, allowing for rapid surgical intervention when needed. If a mutation is not identified, a tier-2 approach is required, if a mutation is known to be responsive to diazoxide, a tier-2 approach that includes targeted next-generation sequencing and array-CGH deletion/duplication analysis of all ten known genes causing HI can be performed. To date, we have identified the genetic etiology in 3/19 (16%) patients studied and likely causal mutations were identified in ABCB8, GLUD1 and GCK. Variants of uncertain clinical significance (VOUS) were also identified in 3 patients. Parental analysis will be performed to clarify mode of inheritance of the mutations and provide insights on the nature of the observed HI. Studies are ongoing to assess the impact of genetic testing results on patients’ medical management and clinical course.

2470T
Carrier Screening in the Sephardic/Mizrachi Jewish Population for Genetic Disorders with Known Founder Mutations. X. Cai, G. Akler, L. Shi, J. Zhang, G. Diaz, L. Edelmann, R. Kornreich. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

The Sephardic/Mizrachi Jewish (SMJ) population includes descendants of Jews from Spain, Southern Europe, North Africa and the Middle East. They are a heterogeneous group defined by a combination of geography, religion, and common linguistic and religious community. Similar to the Ashkenazi Jewish (AJ) population, recurrent founder mutations associated with autosomal recessive and X-linked genetic diseases have arisen in the SMJ population due to geographic and cultural isolation. Current genetic SMJ screening panels are limited mostly due to heterogeneity. The purpose of this study was to select diseases/mutations that may be appropriate for an SMJ carrier screening panel, and to establish accurate carrier frequencies among distinct SMJ subgroups in the New York metropolitan area. A comprehensive panel of mutations and recurrent founder mutations associated with autosomal recessive and X-linked genetic diseases were identified in the SMJ population. Saliva samples were collected with informed consent and screening was performed by Sequenom MassARRAY iPLEX platform on over 500 individuals including the Iraqi, Iranian, Syrian, Moroccan, Yemenite, Tunisian, Bukharian, and Egyptian communities. To explore the differences between AJ and SMJ populations, approximately 1500 AJ anonymized DNAs were also screened using the same panel, and the carrier frequencies were compared between these two groups. Forty nine autosomal recessive and two X-linked diseases were selected for screening, including 130 mutations from 51 genes. Four diseases were determined to be responsive to diazoxide, a tier-2 approach that includes targeted next-generation sequencing and array-CGH deletion/duplication analysis of all ten known genes causing HI which can be performed. To date, we have identified the genetic etiology in 3/19 (16%) patients studied and likely causal mutations were identified in ABCB8, GLUD1 and GCK. Variants of uncertain clinical significance (VOUS) were also identified in 3 patients. Parental analysis will be performed to clarify mode of inheritance of the mutations and provide insights on the nature of the observed HI. Studies are ongoing to assess the impact of genetic testing results on patients’ medical management and clinical course.
Clinical whole exome sequencing in a group of pediatric heterogeneous disorders: Yield and new gene discoveries. J. Gauthier1,2, J. Thil-fault1,2, J.F. Soucy1,2, F.H. Hamdan1, P. Campeau2,3, M.E. Samuels3, E. Lemire4, A.-M. Laberge5, C. Brunel-Guittotron6, S. Nizard7, G.A. Mitchell1,2,2,2, G.A. Rouleau1,2,1, JL. Michaud1,2,2,2, 1) Medical Biological Unit, Molecular Diagnostic Laboratory, Sainte-Justine University Hospital Center, Montreal, QC, Canada; 2) Division of Medical Genetics, Department of Pediatrics, Sainte-Justine University Hospital Center, Montreal, QC, Canada; 3) CHU Sainte-Justine Research Center, Montreal, Quebec, Canada; 4) Montreal Neurological Institute and Hospital, Montréal (Québec), Canada; 5) Department of Medicine, Université de Montréal, Montréal (Québec), Canada; 6) Montreal Diagnostic Laboratory, Sainte-Justine University Hospital Center, Montreal, Quebec, Canada.

It is now increasingly recognised that whole-exome sequencing (WES) is a useful diagnostic approach for the identification of disease mutations responsible for various genetic disorders but the yield of WES and the optimal methodology for its use in the clinic are not well defined. In the process of implementing WES in a clinical setting in a universal health care system, CHU Sainte-Justine, a Mother and Child University Hospital Center (Montreal, Quebec, Canada), recently launched a WES pilot project for individuals with rare conditions, unexplained despite extensive investigations. A total of 96 patients, mainly children, were recruited from genetics clinics for whom WES was performed and analysed. For patient recruitment, priority was given to clinical presentations known to have great genetic heterogeneity. We focused on rare variants known strongly predicted to alter gene function in genes previously associated with monogenic disorders. To date, the overall diagnostic yield in 70 patients with completed analysis is 36% (25/70). It is 37% (18/49) for patients with global developmental delay/intellectual disability, 33% for other patients representing a heterogeneous group of diseases (skeletal dysplasia, retinopathy, deafness, neurodegenerative diseases, etc.). We did not identify any incidental findings but we fortuitously discovered strong candidate genes for several conditions, including Megacystis-microcolon intestinal hypoperistalsis syndrome (MIM 2472M), Neuronal ceroid lipofuscinoses (MIM 2471S), amongst others.

For many children with rare, potentially genetic syndromic disease, a clear diagnosis remains elusive, and the diagnostic odyssey continues. The successful application of whole exome sequencing (WES/WGS) in causative gene discovery for rare Mendelian disease has the potential to vastly reduce or eliminate such situations. A very significant barrier to the routine clinical use of WES/WGS is the lack of standardized, validated and repeatable and scalable bioinformatics solutions that can facilitate the rapid annotation, filtering and classification of the large number of potentially biologically relevant variants detected by these technologies. We are currently evaluating QIAGEN’s new evidence-based clinical decision support software, Ingenuity Clinical, as a potential tool for clinical interpretation and reporting by analyzing a set of whole genome data generated on 30 family trios centered on probands, affected with severe congenital malformation(s) that are not consistent with a known clinical diagnosis. There is no clear diagnostic or disease increase, and standard genetic testing was negative. The Ingenuity Clinical software platform classifies variants using draft ACMG 2.0 assessment guidelines based on a pre-curated knowledge base of biomedical literature and clinical evidence that allows standardized and scaled annotation, filtering, classification and reporting of variants and clinically relevant variants. For cases in which a genetic cause for disease is suspected based on family history, but no causal variant has been previously identified, we use QIAGEN’s Ingenuity Variant Analysis to identify gene variants that may perturb biological pathways/networks upstream of known causal genes or overlaps cases with overlapping phenotypes. Initial analysis, biological interpretation and clinical assessment of variants through these Ingenuity platforms has unequivocally identified a mutation in one proband that is clearly consistent with the clinical presentation, and facilitates a diagnosis. For other cases, QIAGEN’s Ingenuity Clinical shows good candidates in all of them, but further assessment is required to determine their validity. The analyses on all 30 probands continue and findings will be presented.

Population diversity and the genetics of hypertrophic cardiomyopathy. A.K. Marra1,2, B.H. Funke1,2, H.L. Rehm2,3, M.S. Oleens1, B.A. Maron1, P. Szelovits1, D.M. Margulies1, J. Loscalzo1, I.S. Kohane2,8,1, 1) Health Sciences and Technology, Harvard-MIT, Cambridge, MA; 2) Center for Biomedical Informatics, Harvard Medical School, Boston, MA; 3) Laboratory for Molecular Medicine, Partners HealthCare Personalized Medicine, Cambridge, MA; 4) Department of Pathology, Harvard Medical School and Massachusetts General Hospital, Boston, MA; 5) Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 6) Laboratory of Molecular Cardiology, Department of Cardiology, the Heart Center, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; 7) Division of Cardiovascular Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; 8) Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

BACKGROUND—Risk stratification for hypertrophic cardiomyopathy (HCM) is an exemplar of the clinical gains attainable by targeted genetic testing. After decades of research, sequencing laboratories now identify a causal mutation in over one third of referred patients. Using these findings, clinicians routinely assess risk for the patient’s relatives and may even tailor therapy for rare patients. However, the clairvoyance offered by genetic testing comes with a cost—when variants are misclassified, there is potential for harm. METHODS—Using publicly accessible large-scale exome data, we identified variants previously considered causal of HCM that were overrepresented in the general population. We studied these variants in diverse populations using complementary sequencing data, and reevaluated their initial ascertainment in the medical literature. We reviewed patient records at a leading genetic testing laboratory for all occurrences of these variants during the near-decade-long history of the laboratory. RESULTS—Multiple patients, notably all of African or unspecified ancestry, received incorrect positive reports of variants classified as “pathogenic” or “presumed pathogenic” and later revised to “benign.” All identified variants were significantly more common in African Americans than European Americans (P < 0.001). We show that if moderately diverse control sequencing data had been available to test previous pathogenically assertions, these variants would likely have been classified sooner as benign, possibly avoiding multiple misclassifications in African-ancestry individuals and their families. We identify issues in the ascertainment process that may have led to these errors in the medical literature. CONCLUSIONS—Diverse population data can be used to improve the accuracy of variant interpretation. These findings expand upon current guidelines, which recommend using ethnically matched controls to interpret variants. As diverse sequencing data becomes more widely available, we expect variant reclassifications to increase, particularly for groups that have historically been less well studied.

Background: Hereditary nephrotic syndrome (HNS) is a genetically heterogeneous disorder representing a spectrum of inherited renal diseases that cause podocyte anomalies and resultant proteinuria. Mutations in the NPHS1, NPHS2, WT1, PLEC1, and LAMB2 genes cause the onset of HNS in neonates, infants, and children, with progression to end-stage renal disease in later life. The relative frequencies of causative gene variants in children with HNS are unknown. Methods: We analyzed the frequencies of pathogenic variants (PVs) and variants of unknown significance (VUS) in a targeted panel of five genes in 273 patients referred to a clinical laboratory for HNS testing. Mutation analysis was carried out by direct sequencing of the WT1, LAMB2, NPHS1, NPHS2, and PLEC1 genes in all cases. The study included 10 prenatal samples, samples from various childhood age groups (82 neonatal, 146 infant, 16 child), and 15 syndromic cases that were consecutively submitted. Results: Sequence variants (total 100 PV, 188 VUS) were detected in 159 patients. Variants were most frequent in the NPHS1 gene (50 PV, 82 VUS), followed by the PLEC1 gene (14 PV, 12 VUS). Variants were less frequent in the WT1 (15 PV, 14 VUS), NPHS2 (17 PV, 11 VUS), and LAMB2 (4 PV, 42 VUS) genes. VUS were found in multiple genes in 26 cases. Conclusions: An analysis of 273 samples submitted for HNS testing shows that over half have variants of interest. This study confirms the clinical value of genetic testing for HNS using this panel.


The prevalence of idiopathic nephrotic syndrome (NS) is approximately 16 cases per 100,000 individuals. Recent studies show that much of the observed disease incidence can be accounted for by mutations in just 29 genes, most of which are expressed in the kidney and which play a variety of biological roles. Until recently, children with NS, steroid-resistant NS (SRNS), proteinuria, and focal segmental glomerulosclerosis (FSGS) underwent sequential, gene-by-gene testing because there were no clinical tests. The observed disease incidence can be accounted for by mutations in just 29 genes in all cases. The study included 10 prenatal samples, samples from various childhood age groups (82 neonatal, 146 infant, 16 child), and 15 syndromic cases that were consecutively submitted. Results: Sequence variants (total 100 PV, 188 VUS) were detected in 159 patients. Variants were most frequent in the NPHS1 gene (50 PV, 82 VUS), followed by the PLEC1 gene (14 PV, 12 VUS). Variants were less frequent in the WT1 (15 PV, 14 VUS), NPHS2 (17 PV, 11 VUS), and LAMB2 (4 PV, 42 VUS) genes. VUS were found in multiple genes in 26 cases. Conclusions: An analysis of 273 samples submitted for HNS testing shows that over half have variants of interest. This study confirms the clinical value of genetic testing for HNS using this panel.

2477S A novel EDA splice site mutation cause hypohidrotic ectodermal dysplasia in a heterozygous female with severe phenotype. C. Weng1, T.Y Wei1, F. Jun2, P. Yu1, M. Qi1, M. Li1. 1) School of Medicine, Zhejiang university, Hangzhou, China; 2) James D Watson Institute of genome Sciences, Hangzhou, China; 3) Department of pathology and laboratory Medicine, University of Rochester, NY, USA; 4) Department of Reproductive Endocrinology, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China.

Hypohidrotic Ectodermal dysplasia (XHED[MIM 305100]) is a type of X-linked genodermatosis characterized by the abnormal development of sweat glands, teeth and hair. Our study aims to characterize the pathogenesis of a hypohidrotic female diagnosed with typical HED at molecular level. The proband underwent complete physical examination. Mutation screening was performed by Sanger sequencing on associated genes. In silico analysis was used to predict putative effect of the identified mutation, and in vitro exon trapping assay to analyze mutant transcript processing. PCR-based human androgen receptor gene assay (HUMARA) was applied to study the skewed X chromosome inactivation pattern. A novel mutation c.925-G>A-G at the end of exon 7 of EDA gene(MIM 300451) was identified. This mutation is expected to influence the function of that splicing acceptor. The exon trapping assay indicated the first 32 bp sequence of exon 8 had been skipped abnormally in the splicing process. The resulted premature stop codon truncated 71 amino acid including the tumor necrosis factors (TNF) domain at C-terminal-end. This may lead to the deformation of the inner anti-parallel β-strand, weakening the hydrophobic interfaces and losing a number of hydrogen bonds. X chromosome HUMARA assay for the affected proband and her mother of proband, who is also a mutation carrier but with normal phenotype, showed an obvious skewed X chromosome inactivation pattern.

Conclusion: A splice site mutation resulting in partial exon skipping was found in a typical hypohidrotic Ectodermal dysplasia patient. The mutation probably leads to a truncated protein with part of C-terminal skipping out, affecting the interaction properties of the EDA protein. The skewed X chromosome inactivation pattern is associated with the severity of the phenotype.

2478M Follow-up of the first 250 clinical WES cases: periodic re-analyses revealed additional molecular diagnoses. Y. Yang1, F. Xia1, R. Person1, N. Niu1, M. Leduc1, J. Beuten1, D.M. Muzny2, S.E. Pion2, J.R. Lupski3, A.L. Beaudet1, R.A. Gibbs2,3, C.M. Eng1. 1) Dept Molec & Human Gen, Baylor College of Medicine Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine Houston, TX.

We recently reported clinical whole exome sequencing (WES) results for the first 250 unrelated patients received between October 2011 and June 2012 (Yang et al NEJM 2013;369:1502). The patients mostly had neurologic manifestations suggesting potential genetic etiology but no definitive diagnosis could be rendered despite extensive workups. Molecular diagnoses by WES were made in 62/250 patients (25%), including 4 patients with 2 molecular diagnoses resulting in blended phenotypes. The initial 25% diagnostic rate was the highest rate observed among other commonly ordered diagnostic genetic tests including chromosomal microarray analysis and most targeted gene tests. Further, more diagnoses from the existing WES dataset are expected as new disease genes are published. We have been following up the 250 cases since the completion of the clinical report in 2012 performing re-analyses of the original WES variants every six months using a list of newly published disease genes. This effort, which was free of charge, resulted in additional molecular diagnoses in 4 AD disorders (genes: AHDC1, ASXL3, KCTN1, KMT2A), 2 AR disorders (genes: SERAC1, SFXN4) and 2 X-linked (XL) disorders (gene: WDR45, seen in 2 unrelated patients). In addition to findings in new disease genes, molecular diagnoses were also confirmed in 2 more cases by add-on parental studies and phenotype correlations. The follow-ups improved the WES diagnostic rate from 25% to 28% by adding 10 more diagnoses. Interestingly a case previously reported to carry a de novo truncating mutation in ANKRD11 was now found to carry an additional de novo truncating change in a new disease gene AHDC1, making this case the first time with two diagnoses out of the 250 samples. It should be noted that the current process, which focused on truncating mutations and only considered missense changes with compelling evidence, was not a full-scale re-analysis. Nevertheless, every new diagnosis required 2-3 hours on activities such as literature review, inheritance pattern and disease mechanism verification, phenotypic correlation, report update and communication with referrals. There is also significant time spent on evaluating and later ruling out other candidate cases. For a full-scale re-analysis, more automation is needed and the test may be offered for a fee. In our laboratory, WES patients without a clinical diagnosis were also offered the option of entering research studies for novel disease gene discoveries.
2479T
Identification of 290-bp deletion as a first report on the beta-globin gene in South of Iran. M. Hamidi1, GH. Shariati2, L. Dawoody Nejad1, D.L. Gahidian2, AH. Saberi3, B. Kaikhaei3, M. Mohammadi -Anea3. 1) Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran; 2) Narges Medical Genetics & PND Laboratory, No. 18, East Mihan Ave, Kianpars, Ahvaz, Iran; 3) Research Center of Thalassemia and Hemoglobinopathies, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.

In this study we describe the first report of 290-bp deletion (c.-176_92+25del) in four individuals from three unrelated families in Khuzestan province with Arabic ethnic background. They were referred for carrier detection as part of a national program for the prevention of thalassemia with anisocytosis, microcytosis and hypochromia. All the Individuals had elevated Hb A2 and normal Hb F levels. One of the individual, Offspring of K.B, was 5 years old girl inherited both defects from her parents. Physical examination of this patient showed paleness, slightly hepato- and splenomegaly and she was not transfusion dependent. The 290 bp deletion characterized by DNA sequencing and MLPA test. The deletion removes the promoter region, the entire exon 1 and 5 end of IVS-1. The MLPA results confirmed the deletion by probes ranging from prob 21(Promotor) to prob 1(HBB intron 1). Diagnose of this β-thalassemia deletion can be extremely useful for prenatal diagnosis in the couple especially in the consanguineous and ethnic marriage.

2480S

Human leukocyte antigen (HLA) is a group of genes that are extremely polymorphic among individuals and populations and has been associated with more than 100 different diseases and adverse drug effects. HLA typing is accordingly an important tool in clinical application, medical research, and population genetics. To date, several high-throughput HLA typing methods using next-generation sequencing (NGS) have been developed. We have previously developed a phase-defined HLA gene sequencing method using the MiSeq sequencer. Our sequencing protocol and pipeline provided essentially complete phase-defined HLA gene sequences without referring IMGT/HLA database; however, it required complicated and labor-intensive workflows especially in the library preparation step. Most importantly, the method was not well adapted for processing multiple samples. Here, we developed a Beaded Normalization for Uniform Sequencing (BeNUS) procedure using three steps of bead purification. BeNUS can easily and precisely normalize the molar concentrations of up to 96 samples. We applied long-range PCR to amplify HLA-B for 96 samples followed by transposase-based library construction and 300 bp paired-end sequence reads with the MiSeq. After sequencing, we observed low variation in read percentages (0.2% to 1.55%) among the 96 demultiplexed samples. On this basis, all the samples were amenable to haplotype phasing using our phase-defined sequencing method. Our HLA sequencing method optimized for 96 multiplexing samples is highly time and cost effective and is especially suitable for automated multi-sample library preparation and sequencing. In addition, we established a sequence capture method for the 3.8 Mb entire HLA region for sequencing and targeted-bisulfite sequencing to detect genetic (SNVs, small and large indels) and epigenetic (methylation variable positions) information together with the HLA haplotype map information. Only one capture step was enough to prepare the two different types of libraries, targeted sequencing and targeted-bisulfite sequencing, for 12 samples as a pooled sample. After one MiSeq run of 350 bp + 250 bp paired-end reads, more than 97% of the target region was covered by at least 20-fold in all samples, although the variation in read percentages among 12 samples was observed as 4.8% to 11.5%. Our future plan is to set these methodologies for practical application of medical research and clinical application.

2481M
Identifying disease causing variants ranging from SNVs, small InDels, single exons to whole gene deletions in the RB1 gene through a novel next-generation sequencing based approach. With the rapid development of test design and analysis tools a single NGS test can identify genetic abnormalities arising from single nucleotide variations (SNVs) and structural variations (SVs) including small and large indels and copy number variations (CNVs). At Strand Centers for Genomics and Personalized Medicine, we have run hundreds of diagnostic tests ranging from whole exome to disease specific panel based NGS tests. We illustrate the clinical utility of combining efficient panel based tests with the right analysis tools in providing quick and economic diagnostics in a representative set of retinoblastoma cases. Materials and Methods - The samples were tested using either Illumina’s TruSight Cancer panel (consisting 94 genes) or our custom designed Eye Disorders panel. In both panels, the RB1 gene is included. Paired end sequencing was done at an average coverage of >100X, the analyses were done using AVADIS® NGS software and StrandOmics™ was used for the interpretation and reporting of the variants. Results - We were able to provide a positive diagnosis in all retinoblastoma cases identifying the causative variants ranging from SNVs to small deletion to whole gene deletion. Performing CNV analysis on the NGS data normalized to the single exome data previously run on the panel, we are able to pick up statistically significant changes causing heterozygous deletion of the whole RB1 gene as well as the deletion of a single exon of the gene. This NGS based method was more sensitive than FISH (fluorescence in situ hybridization) in identifying deletions. In one case, previous analysis by FISH had identified a 13q14 deletion, however the NGS test revealed a heterozygous splice site variation as the causative variant. Conclusions - With the combination of panel based sequencing and the availability of the in silico tools to quickly identify disease causing variants which range from SNVs to whole gene deletions. These methods can be extended to other multi-gene diseases where sequential testing could cause significant delays in diagnosis.

2482T
Diagnoses by Clinical Exome Testing Suggest Wider than Expected Phenotypic Spectra on New Disease Genes: Implications for Choosing Testing Strategy and Interpretation of Results. F. Xia1, E.P. Simpson1, C.V. Knowlton1, M.F. Wang1, J. Zhang2, J. Zhang1, M.S. Leduc1, Z. Niu1, R. Person1, D.M. Muzny1, A.L. Beaudet1, R.A. Gibbs1, C.M. Eng1, Y. Yang1. 1) Baylor College of Medicine, Houston, TX; 2) Houston Methodist Hospital, Houston, TX.

Recently discovered disease genes account for a significant portion (~25-30%) of diagnoses made by clinical WES. However, some of the new genetic disorders do not have characteristic symptoms and the phenotype spectra are yet to be expanded with additional studies. The lack of knowledge and the complexity of disease phenotypes in disease genes that attribute to a more extended genetic diagnosis in disease genes. This highlights the importance of disease genes that attribute to a more extended genetic diagnosis in disease genes.

In this study we describe the first report of 290-bp deletion (c.-176_92+25del) in four individuals from three unrelated families in Khuzestan province with Arabic ethnic background. They were referred for carrier detection as part of a national program for the prevention of thalassemia with anisocytosis, microcytosis and hypochromia. All the Individuals had elevated Hb A2 and normal Hb F levels. One of the individual, Offspring of K.B, was 5 years old girl inherited both defects from her parents. Physical examination of this patient showed paleness, slightly hepato- and splenomegaly and she was not transfusion dependent. The 290 bp deletion characterized by DNA sequencing and MLPA test. The deletion removes the promoter region, the entire exon 1 and 5 end of IVS-1. The MLPA results confirmed the deletion by probes ranging from prob 21(Promotor) to prob 1(HBB intron 1). Diagnose of this β-thalassemia deletion can be extremely useful for prenatal diagnosis in the couple especially in the consanguineous and ethnic marriage.

In this study we describe the first report of 290-bp deletion (c.-176_92+25del) in four individuals from three unrelated families in Khuzestan province with Arabic ethnic background. They were referred for carrier detection as part of a national program for the prevention of thalassemia with anisocytosis, microcytosis and hypochromia. All the Individuals had elevated Hb A2 and normal Hb F levels. One of the individual, Offspring of K.B, was 5 years old girl inherited both defects from her parents. Physical examination of this patient showed paleness, slightly hepato- and splenomegaly and she was not transfusion dependent. The 290 bp deletion characterized by DNA sequencing and MLPA test. The deletion removes the promoter region, the entire exon 1 and 5 end of IVS-1. The MLPA results confirmed the deletion by probes ranging from prob 21(Promotor) to prob 1(HBB intron 1). Diagnose of this β-thalassemia deletion can be extremely useful for prenatal diagnosis in the couple especially in the consanguineous and ethnic marriage.

Recently discovered disease genes account for a significant portion (~25-30%) of diagnoses made by clinical WES. However, some of the new genetic disorders do not have characteristic symptoms and the phenotype spectra are yet to be expanded with additional studies. The lack of knowledge and the complexity of disease phenotypes in disease genes that attribute to a more extended genetic diagnosis in disease genes. This highlights the importance of disease genes that attribute to a more extended genetic diagnosis in disease genes.

In this study we describe the first report of 290-bp deletion (c.-176_92+25del) in four individuals from three unrelated families in Khuzestan province with Arabic ethnic background. They were referred for carrier detection as part of a national program for the prevention of thalassemia with anisocytosis, microcytosis and hypochromia. All the Individuals had elevated Hb A2 and normal Hb F levels. One of the individual, Offspring of K.B, was 5 years old girl inherited both defects from her parents. Physical examination of this patient showed paleness, slightly hepato- and splenomegaly and she was not transfusion dependent. The 290 bp deletion characterized by DNA sequencing and MLPA test. The deletion removes the promoter region, the entire exon 1 and 5 end of IVS-1. The MLPA results confirmed the deletion by probes ranging from prob 21(Promotor) to prob 1(HBB intron 1). Diagnose of this β-thalassemia deletion can be extremely useful for prenatal diagnosis in the couple especially in the consanguineous and ethnic marriage.
1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, United States; 2) Department of Medical Genetics, Oslo University Hospital, Norway; 3) Center for Human Immunobiology, Immunology Allergy Rheumatology, Texas Children's Hospital, Feigin Center, 1102 Bates, Suite 303, Houston, TX, United States.

Background: Congenital immunodeficiency is an inherited disorder of the immune system. Severe combined immunodeficiency (SCID), characterized by profound deficiencies of T cells and/or B cells at the time of birth, is the most severe form of congenital immunodeficiency. If not treated promptly, affected patients usually do not live beyond infancy due to severe, recurrent infections. Genetic heterogeneity of SCID frequently delays the diagnosis, which is crucial for life-saving treatment and optimal management. Methods: All targeted exons and at least 20 bp of flanking genomic sequences of 46 genes involved in SCID are enriched using SeqCap EZ solution-based capture and are subjected to Massively Parallel Sequencing (MPS) on Illumina HiSeq2000. Results: The target gene capture/MPS provides an average coverage of 1000X. Exons with insufficient coverage (<20X) or high sequence homology (pseudogenes) are complemented by PCR/Sanger sequencing using gene-specific primers to ensure the 100% coverage of all targeted regions. In a pilot study, deleterious mutations were detected in fourteen out of seventeen patients analyzed (82%). Identified mutations include compound frameshift mutations in CORO1A, a novel indel mutation in JAK3, a novel 70-bp duplication in a RAG gene, RMRP, novel missense mutations in TBX1, ZAP70, and X-linked FOXP3, as well as reported mutations in DCLRE1C and RAG1. Deep coverage of next generation sequencing revealed a c.361dupA frameshift mutation of IL7R at 82% of heterozygosity in the blood sample of a female patient. Further analysis of a buccal sample from the same patient after hematopoietic stem cell transplantation confirmed the presence of the wild-type allele at a low level (~2%), indicating possible somatic mosaicism for that patient (both in blood and in other tissue). Conclusion: High throughput deep sequencing analysis greatly increases the diagnostic yield of congenital immunodeficiency. Establishing a molecular diagnosis of immunodeficiency enables early immune reconstitution through proper treatment and guides a better management for improved long-term quality of life.


The advent of next generation sequencing (NGS) panels has become common place in clinical genetic testing strategies. Clinicians have the ability to order a single test that will interrogate tens to hundreds of genes in a time period, and at a cost, that was typical for single gene testing only a few years ago. Like all diagnostic tests, however, NGS panels should be evaluated for their clinical utility, a measurement that weighs all benefits and drawbacks of a particular test. Some general aspects of clinical utility include: analytic validity (how well the test can predict the presence or absence of a variant), clinical validity (how well the variant being analyzed is related to the presence of a specific disease), clinical utility (whether the test can provide information about diagnosis, treatment, management, or prevention of the disease), as well as the ethical, legal, and social implications (ELSI) of the test. Our laboratory has performed NGS testing for patients with autism, intellectual disability, and/or multiple congenital anomalies for several years. During this time, we have utilized two separate panels, a 49-gene panel, and then later an 86-gene panel. The larger panel increased our pathogenic variant rate by approximately 4-fold while only increasing our variants of uncertain significance (UCS) rate by approximately 1.5-fold. These results suggest that the clinical utility of the larger panel is increased over that of the smaller panel.
2486S

Proband whole-exome sequencing as a cost-effective diagnostic strategy for suspected Mendelian disorders. J. Thevenon1,2,3, J. St-Onge2,3,4, Y. Duffourd1,2,3, A. Massurel-Paulet1,2, S. El Chehadeh-Djebar1,2,3, C. Juge1,2, P. Callier2, M. Lefebvre1,2, M. Chouciane6, V. Darmency-Stamboul5, M. Mith2,7,8, F. Feillet2, C. Thauvin-Robinet1,2,3, L. Faivre1,2,3, J-B. Rivière1,2,3,4, 1) Centre de Génétique, CHU de Dijon, Dijon, France; 2) 1.Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (TRANSLAD), Centre Hospitalier Universitaire Dijon, F-21079 Dijon, France; 3) Équipe d’Accueil 4271, Génétique des Anomalies du Développement, Université de Bourgogne, F-21079 Dijon, France; 4) Laboratoire de Génétique Moléculaire, Plateau Technique de Biologie, Centre Hospitalier Universitaire Dijon, F-21079 Dijon, France; 5) Laboratoire de Chromatographie, Plateau Technique de Biologie, Centre Hospitalier Universitaire Dijon, F-21079 Dijon, France; 6) Laboratoire de Cytogénétique, Plateau Technique de Biologie, Centre Hospitalier Universitaire Dijon, F-21079 Dijon, France; 6) Service de Pédiatrie 1, Centre Hospitalier Universitaire Dijon, F-21079 Dijon, France; 7) Service de Neurologie Pédiatrique, Hôpital de la Timone, Assistance Publique des Hopitaux de Marseille, F-13005 Marseille, France; 8) Unité Mixte de Recherche 910, Institut National de la Santé et de la Recherche Médicale, Aix-Marseille Université, F-13005 Marseille, France; 9) Service de Médecine Infantile 1, Centre de Référence des Maladies Héréditaires du Métabolisme, Centre Hospitalier Universitaire Brabois-Enfants, F-54511 Vandoeuvre-lès-Nancy, France.

Background: Whole-exome sequencing (WES) is a powerful tool to identify the molecular basis of clinically and genetically heterogeneous suspects. Mendelian disorders. When trio-based WES strategy was preferred for diagnostic gene identification, we considered proband WES as a cost-effective diagnostic strategy. Methods: We present a sample of 41 subjects with epileptic encephalopathy and/or severe intellectual disability (18 females and 23 males), with no diagnosis after the realization of all the conventional investigations. Subjects and relatives were informed and consented for exome testing. WES was performed on index cases. We developed a standardized bioinformatic, interpretive, and validation pipelines for diagnostic WES. Variants were filtered when a gene referred in OMIM as related to the subject disorder. Prioritizations of the variants lead to control familial segregation of pathogenic or probably pathogenic variants, with a maximum of 5 per subject. If no plausible variant was evidenced and if an inherited intellectual disability was found from familial history, the whole data were considered for interpretation. Results: Overall, median coverage was of 86 fold and 93% of the RefSeq exons were covered at least by 10 reads. This strategy allowed the identification of 4 de novo variants (namely affecting SHANK3, ARID1B, DYRK1A, and TBPH), 2 X-linked maternally inherited variants (affecting CUL4B and SLC16A2), 2 autosomal recessive variants with consistent familial segregation (SCN10A and GFER). The diagnostic yield of this strategy was of 19% (8/41). Interpretation of the whole data in the remaining subjects led to the identification of variants recurrently affecting SLC13A5 gene in 2 multiplex families. Besides, candidate genes with consistent familial segregation were identified but not considered as diagnostic results. At last, anonymized pathogenic and candidate variants were uploaded on dedicated repositories (namely ClinVar, www.ncbi.nlm.nih.gov/clinvar/ and Phenome Central, https://phenomecentral.org/). Conclusion: In this sample, proband WES was performed as a last resort for 41 subjects. Overall diagnostic yield of this study was of 24% (10/41) and will increase with the scientific community using international repositories. Finally, a medico-economic study is being performed on this sample to evaluate this strategy’s cost-effectiveness.

2487M

Whole exome sequencing in patients with intellectual disability. M. Miña1,2, A. Alvarez-Mora3,4, L. Lijedahl2, O. Karlberg2, L. Rodriguez-Reyenga1,2, A. Muñoz4, J. Rossell3,2, E. Guillen2,3, M. Barrantes2, AC. Syvänen2, I. Madrigal1,2 1) Biochemistry and Molecular Genetics Department, Hospital Clínic and IDIBAPS, Barcelona, Spain; 2) for Biomedical Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain; 3) Department of Medical Sciences, Molecular Medicine, Uppsala University, Uppsala, Sweden; 4) Servicio de Pediatría, Hospital del Mar, Departamento de Pediatría y Obstetricia de la UAB Barcelona, Spain; 5) Universitario Son Espases, Palma de Mallorca, Spain; 6) Genetic Unit, Hospital Clínico Universitario Virgen de la Arrixaca, Murcia, Spain.

Intellectual disability is characterized by substantial limitations in intellectual, functioning and adaptive behavior, diagnosed before 18 years of age. It affects 1-3 per cent of the general population with 50-60 per cent of patients remaining undiagnosed due to the great complexity and the high heterogeneity of the genetic basis. The aims of this study were: to identify mutations in known intellectual disability genes, to identify new genes responsible for intellectual disability and to provide genetic counseling to families. We sequenced 32 individuals of self-reported European ancestry from eight families presenting with moderate to severe intellectual disability using the Illumina HiSeq 2000 Sequencing System. We found four families with novel indel mutations in genes associated with intellectual disability: a compound heterozygous mutation (c.5998delCT; c.10475delAA) in the VS313B gene; a deletion in the UBE3A gene (c.2013delG) resulting in an imprinting defect; an autosomal dominant mutation in the DYNC1H1 gene (c.4461dupC) and one splicing mutation affecting intron 11 of the IQSEC2 gene (ChrX:53267480_53267491delGT). We also identified two nonsense mutations in the X chromosome: a novel mutation in the SMC1A gene (c.1405C>T) and one previously described in the ORC1 gene (c.1567G>A). In the remaining 2 families, several candidate variants were identified and require additional studies. In conclusion, these results show the high heterogeneity and difficulties in intellectual disability and demonstrate the high performance, reliability and cost-effectiveness of this methodology (75% per cent diagnostic success) which is now currently applied in clinical diagnosis. Acknowledgements: FP7/2007-2013, grant agreement n°262055 (ESGI project) Premio discapacidad 2012 Fundació Agrupació Mutua.

2488T

Does increase in genomic microarray resolution result in increased diagnostic yield? S. Costa, A. Kropischi, C. Rosenberg. Genetics and Evolutionary Biology Dpt., University of Sao Paulo, Sao Paulo, Brazil.

We present data from 1,098 children with intellectual disability, obtained using different microarray platforms with increasing resolution (Table). Genomic imbalances were classified as causative (>1Mb or fulfilling criteria from Genetics in Medicine (2011) 13, 680-685) or Vous (Variants of Uncertain Clinical Significance).*

<table>
<thead>
<tr>
<th>Array resolution</th>
<th>Patients (n)</th>
<th>Frequency of normal array</th>
<th>Total frequency of alterations &gt;1Mb</th>
<th>&lt;1Mb</th>
<th>&lt;1Mb</th>
</tr>
</thead>
<tbody>
<tr>
<td>44K</td>
<td>57</td>
<td>52.6%</td>
<td>47.4%</td>
<td>35.1%</td>
<td>10.5%</td>
</tr>
<tr>
<td>60K</td>
<td>780</td>
<td>69.6%</td>
<td>30.3%</td>
<td>21.9%</td>
<td>5.1%</td>
</tr>
<tr>
<td>180K</td>
<td>191</td>
<td>66%</td>
<td>34%</td>
<td>18.8%</td>
<td>5.2%</td>
</tr>
<tr>
<td>850K</td>
<td>70</td>
<td>54.3%</td>
<td>45.7%</td>
<td>21.4%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Total</td>
<td>1,098</td>
<td>67.1%</td>
<td>32.9%</td>
<td>22%</td>
<td>5.4%</td>
</tr>
</tbody>
</table>

Except for 44K platform (more stringent patient selection) we detected higher frequency of alterations with increased resolution. The difference was due to higher detection of Vous, rather than of causative variants. These results question the benefits obtained by higher resolution arrays in the diagnosis of these patients. Grants: FAPESP/CPES.
In traditional cytogenetic methodologies, karyotyping is used to identify aneuploidy and mosaic trisomies associated with incomplete correction. However, the genotyping array also identifies the presence of maternal complete moles identifies individuals at risk of developing gestational trophoblastic neoplasias. The genotyping array completely moles and differentiation of digynic vs. diandric XXY karyotype complete moles and differentiation of digynic vs. diandric XXY karyotype. Accurate diagnosis of diandric partial moles and karyotyping also markedly improved the detection of chromosomal abnormalities (38% vs. 20% of all samples received). Another advantage over conventional cytogenetics is that it can detect sex chromosome abnormalities in nearly complete balanced gender ratios. Microarray analysis also artifactual normal female karyotypes (71.1% female vs. 28.9% male) result-de differentiation and display intellectual disability, suggesting genomic rearrangement. The purpose of this study is to establish the molecular basis of SRD that may be caused by genomic rearrangements using aCGH technology to complete the characterization of patients. Among our cohort of SRD Spanish families, 14 unrelated cases were selected for this study. All of them had undergone previous mutation screening for genes involved in SRD and displayed intellectual disability, suggesting genomic rearrangement. Agilent SurePrint G3 human CGH+SNP microarray 2x440k was performed for all patients, followed by a custom Agilent CGH+SNP 8x60k and other techniques to confirm the results positives. Our study allowed us to preliminary characterize 4 of 14 (29%) with a suspected duplication, mechanism previously unreported for SRD. Microduplications in 1q21.1, 15q13.3, 14q24 and 11q13.2 involving genes associated with intellectual disability, neuropsychiatric disorders and retinal dystrophy were detected. No genomic rearrangement was observed in 6 (43%) patients. Furthermore large loss of heterozygosity (LOH) regions were observed in 2 additional patients (14%) on different chromosomes and different sizes (one patient showed a LOH of 17Mb and the other patient showed 8 different LOH pointing to possible involvement of different genes. The discovery of a genetic basis for some of these conditions is allowing a change in the genetics diagnosis of various diseases. The purpose of this study is to establish the molecular basis of SRD that may be caused by genomic rearrangements using aCGH technology to complete the characterization of patients. Among our cohort of SRD Spanish families, 14 unrelated cases were selected for this study. All of them had undergone previous mutation screening for genes involved in SRD and displayed intellectual disability, suggesting genomic rearrangement. Agilent SurePrint G3 human CGH+SNP microarray 2x440k was performed for all patients, followed by a custom Agilent CGH+SNP 8x60k and other techniques to confirm the results positives. Our study allowed us to preliminary characterize 4 of 14 (29%) with a suspected duplication, mechanism previously unreported for SRD. Microduplications in 1q21.1, 15q13.3, 14q24 and 11q13.2 involving genes associated with intellectual disability, neuropsychiatric disorders and retinal dystrophy were detected. No genomic rearrangement was observed in 6 (43%) patients. Furthermore large loss of heterozygosity (LOH) regions were observed in 2 additional patients (14%) on different chromosomes and different sizes (one patient showed a LOH of 17Mb and the other patient showed 8 different LOH pointing to possible involvement of different genes. The discovery of a genetic basis for some of these conditions is allowing a change in the genetics diagnosis of various diseases.
An alternative method for the analysis of deletions/duplications with MLPA®using the QIAxcel® Advanced System. S.B. Fischer1, S. Herm2,4, M. Attenhoefer1, K. Heinimann1,2, I. Spier3, S. Aretz2, S. Cichon1,2, P. Hoffmann1,2. 1) Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland; 2) Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland; 3) Institute of Human Genetics, University of Bonn, Bonn, Germany.

Multiplex Ligation-Based Probe Amplification (MLPA) is the gold standard for the detection of exonic deletions/duplications in genetic diagnostic testing. MLPA is based on a multiplex PCR, but different from a classical multiplex PCR it is a two-step process involving a hybridisation/ligation reaction for generating the amplification template as the first step. Every exon of a specific target gene is covered by one or more probe-pairs which, when both probes bind to the exon, are ligated. In the second step, the ligated probe-pairs will then be amplified and subsequently analysed using a capillary electrophoresis system. The resulting electropherogram shows a specific peak pattern for each sample. The ratio of the peak height between target and reference probes or even reference samples is used for the detection of exonic deletions/duplications. The fragment analysis is usually performed using capillary (Sanger) sequencing, which is a costly and time consuming method. In this pilot-study we present a faster and more cost effective alternative approach for the MLPA analysis by using the QIAxcel® Advanced System from Qiagen®. The QIAxcel® Advanced System is an automated DNA and RNA analyser, replacing the traditional gel analysis of DNA and RNA. For our pilot study we used two different clinical samples: a) 12 FAP (familial adenomatous polyposis coli) patients with known mutation status as well as 5 healthy controls; b) 5 HNPPC (hereditary non-polyposis colorectal cancer, Lynch syndrome) patients with known mutation status as well as 3 healthy controls. For the MLPA analysis (hereditary non-polyposis colorectal cancer, Lynch syndrome) patients with known mutation status as well as 3 healthy controls. For the MLPA analysis (hereditary non-polyposis colorectal cancer, Lynch syndrome) patients with known mutation status as well as 3 healthy controls; b) 5 HNPCC clinical samples: a) 12 FAP (familial adenomatous polyposis coli) patients with known mutation status as well as 5 healthy controls; b) 5 HNPCC clinical samples; c) 12 FAP patients with unknown mutation status; d) 5 HNPCC patients with unknown mutation status. In this study we used the SALSA MLPA P043 APC ProbenMix® for the FAP samples and the SALSA MLPA P003 MLH1/MSH2 ProbenMix® for the HNPCC samples. The visualisation of each MLPA-generated amplification product was performed using a Sanger sequencer as well as the QIAxcel capillary electrophoresis system. Utilizing a newly developed protocol we were able to detect exonic deletion/delletions with the same resolution and accuracy as with Sanger sequencing. Based on our preliminary results using the QIAxcel Advanced system allows us to have reliable, fast and cost effective MLPA analysis.

Objectives Noninvasive prenatal testing (NIPT) of Mendelian disorders has been reported as case study this year. However, most previous publications focused on sex-linked disorders and autosomal dominant and recessive genetic diseases by assessing the fetal gender and determining paternal inherited mutation in the maternal plasma. Only few studies showed the feasibility to detect maternal inherited mutations and practical application for autosomal recessive monogenic diseases. Our study aims to report the efficiency of a robust method of NIPT to detect sex-linked, autosomal dominant and recessive Mendelian disorders. Methods In 2012-2013 we recruited eleven families who have a proband child and genetic counseling was provided to each couple. Informed written consent was obtained from each couple. Participants as follows: 1) two families are carriers of carrier loss mutation; 2) seven families have both with DMD; 3) one family has a 2-year-old girl with maple syrup urine disease (MSUD); 4) one family with congenital adrenal hyperplasia (CAH); 5) one family with ichthyosis. Blood samples were obtained at 10-16th gestational weeks. NIPT of Mendelian disorders was performed with a haplotype-assisted strategy previously described, to determine whether the fetus is affected with the same mutation as the proband child or not. Amniocentesis was performed to get the fetal tissue and Sanger sequencing was used to confirm the results. Results Totally in five families, the fetuses were identified to inherit the same alleles as the proband. In three families, the fetal escaped from inheriting any mutations from the parents. In the rest three families, the fetal was identified to inherit the same mutations either from the mother or the father. NIPT was finished in 10 days and the results were accordant with Sanger sequencing. Conclusions Noninvasive prenatal testing by maternal plasma DNA sequencing has great potential to supplement the current workflow for prenatal diagnosis of Mendelian disorders, not just for sex-linked, autosomal dominant, but also autosomal recessive. NIPT prior to invasive procedures in the second trimester will offer earlier and trustworthy information for the pregnant women and help to ease maternal anxiety at very early time.
2498S
Miami Otogenetic Program: Implementing genomic medicine in care of patients with impaired hearing. X.Z. Liu1,2, D. Yan1, D. Tekin1, M. Grant1, S. Blanton1, M. Tekin1. 1) Departments of Otolaryngology-Head and Neck Surgery, University of Miami Miller School of Medicine, Miami, Florida 33136, USA; 2) Dr. John T. Macdonald, Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL 33136, USA.

The Miami Otogenetic Program has provided a unique platform to carry out translational research on delivering genetic services to deafness patient care. Using target-enrichment/next generation sequencing (NGS), we are determining (1) the overall frequencies of different forms of genetic deafness, (2) identifying new genes for autosomal recessive non-syndromic hearing loss (ARNSHL) and autosomal dominant (NSHL), and (3) creating important Genomic Deafness Database (GDD) and Personalized Sequence Profile (PSP) for the clinical care of deaf patients where data is ranked based on its clinical validity and utility. We have collected a unique cohort of multiplex families derived from three unique sources from USA, China, and Turkey, suitable for determination of molecular epidemiology of hereditary deafness and for new gene identification using “target-enrichment” methods and NGS. Our interdisciplinary and collaborative team will conduct outcome evaluation of genetic service on deafness patient care in our diversity populations. We have established the Miami Otogenetic Program including the research and the clinical components. The infrastructure of our multidisciplinary otogenetics team has been presented along with our utilization of testing algorithms when evaluating patients with sensorineural hearing loss (SNHL). We have collected DNA samples from over 800 probands from multiplex families with no mutations in the common deafness genes from the three unique cohorts. A total of 60% of small multiplex families are identified to have mutations in the known deafness genes in a pilot study and the remaining 40% have mutations in other yet-unidentified deafness-causing genes. Creation of the Genomic Deafness Database (GDD) and Personalized Sequence Profile (PSP) is in process. Using these high-throughput technologies, we have identified several new genes for non-syndromic deafness. Hearing rehabilitation and counseling of patients with genetic causes of hearing loss are provided. The combined target-enrichment/NGS and whole exome sequencing (WES) is a powerful tool in the identification of new deafness genes in small multiplex families. The multidisciplinary otogenetic panel team approach is an effective way to bring the sequencing data to clinical practice for the clinical diagnosis and management of deaf and hard-of-hearing families.

2499M
Whole-Exome Sequencing to decipher the genetic heterogeneity of hearing loss in a Chinese family with deaf by deaf mating. J. Qing1,3, D. Yan1, Y. Zhou1, Q. Liu1, W. Wu1, Z. A. Xiao1, Y. Y. Liu1, J. Lu1, L. Lu1, D. Hu1, D. H. Xie1, X. Z. Liu1,2,3. 1) Departments of Otolaryngology-Head and Neck Surgery, University of Miami Miller School of Medicine, Miami, Florida 33136, USA; 2) Dr. John T. Macdonald, Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL 33136, USA; 3) Department of Otolaryngology-Head and Neck Surgery, Institute of Otology, the Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China.

Hereditary deafness is one of Mendelian diseases with a highly genetically heterogeneous trait. For many decades, linkage analysis combined with candidate genes approach have been the main tool to elucidate the genetics of hearing loss. However, this study design is costly, time-consuming and unsuitable for small families, mainly due to the inadequate numbers of available affected individuals, locus heterogeneity, as well as assortative mating. Whole exome sequencing (WES) has now become technically feasible and cost-effective method for detection of disease variants underlying Mendelian disorders due to the recent advances in next-generation sequencing (NGS) technologies. In the present study, we have combined both the Deafness Gene Mutation Detection Array and WES to identify deafness causative variants in a large Chinese composite family with deaf by deaf mating. The simultaneous screening of the 9 deafness common mutations (GJB2 c.35delG, c.176del16, c.225delG, c.299-300delAT; GJB3 c.538C>T (p.R180X); SLC26A4 c.IVS7-2A>G, c.2168A>G (p.H723R); miDNA 12S rRNA c.1555A>G, c.1494C>T) in Chinese populations, using the allele-specific PCR based universal array, resulted in the identification of the c.1555A>G in the mitochondrial DNA 12S rRNA in affected individuals in one branch of the family. We then subjected the mutation-negative cases to WES and identified novel causative variants in the MYH14 and WFS1 genes. This report confirms the effective use of a NGS technique to detect pathogenic mutations in affected individuals who were not candidates for classical genetic studies.

2500T

While next generation sequencing (NGS) is a valuable tool in cancer genomics, most cancer panels target variant hotspots in multiple cancer types. This often compromises the sensitivity and specificity required to accurately identify low-frequency mutations in specific cancers within their clonal contexts. To mitigate this limitation, we designed MyAML™: a comprehensive and specific NGS sequencing strategy used to identify mutations in, and stratify patients with, acute myeloid leukemia (AML). MyAML utilizes oligonucleotide baits that target coding exons of 193 genes and potential genomic breakpoint hotspots within 36 somatic gene fusion partners known or predicted to be involved in AML pathogenesis. We sequenced anonymized AML patient samples from Invivoscribe’s biobanks using the MyAML panel. Following DNA hybridization, target loci were sequenced on an Illumina MiSeq utilizing v3 chemistry with the 600-cycle kit to an average depth >50x. Using customized bioinformatics, we performed thorough mutation detection analyses to precisely identify and characterize single nucleotide variants (SNVs), indels, and structural variant breakpoints. We also calculated variant allele frequencies to investigate potential aneuploidy, loss of heterozygosity and clonality. Within the AML patients, we detected known pathogenic mutations, including known and novel missense SNVs, internal tandem duplications (ITD) in FLT3 over 300bp in size, and insertions in NPM1. Variants found in FLT3, NPM1, CEBPA, DNMT3A, IDH1, IDH2, and KIT were validated using either capillary electrophoresis or targeted amplification and sequencing. We also identified potentially pathogenic gene fusions in patient samples, which were confirmed by FISH or cytogenetic microarrays. In order to detect unique or novel mutations, diagnostic assays must be designed to maximize sensitivity and specificity such that they accurately identify all driver mutations in their clonal context. This is critical for predicting response and recurrence with various therapies. MyAML is the only comprehensive approach that characterizes AML by specifically targeted the genes and fusion genes known and predicted to be involved in AML pathogenesis. We demonstrate the utility of MyAML as a highly sensitive and accurate sequencing strategy for the comprehensive analysis of AML patients.
2501S  
Mutations in the PNPLA8 gene encoding the mitochondrial calcium-independent phospholipase A2 in a patient with lactic acidosis, spasticity, abnormal gait, dystonia and complex partial seizures. I. Thiffault1, C.J. Saunders1, F. S.H. Moon5, X. Liu1, K. Coffman3, J.B. LePichon1, E. Taboada6, L.D. Smith1, E.G. Farrow1, N. Miller1, S.F. Kingsmore1, R.W. Gross3. 1) Center for Pediatric Genomic Medicine, Children’s Mercy Hospitals, Kansas City, MO, USA; 2) Division of Bioorganic Chemistry and Molecular Pharmacology, Department of Medicine, Washington University School of Medicine, St-Louis, MO, USA; 3) Division of Neurology, Dept Pediatrics Children’s Mercy Hospitals, Kansas City, MO, USA; 4) Division of Pathology, Dept Pediatrics Children’s Mercy Hospitals, Kansas City, MO, USA.

Mitochondrial disorders are a group of clinically heterogeneous diseases, commonly defined by a lack of cellular energy due to OXPHOS defects. The majority of proteins involved in mitochondrial metabolism and maintenance are encoded by nuclear genes, with many yet to be associated with human disease. We performed whole exome trio sequencing on a young girl with mitochondrial myopathy manifested by progressive muscle weakness, hypotonia, spasticity, seizures, poor weight gain, and lactic acidosis. She was found to be compound heterozygous for two frameshift mutations, p.Asn112HistX29 and p.Leu639AlafsX4 in the PNPLA8 gene, which encodes mitochondrial calcium independent phospholipase A2γ (iPLA2γ; MIM#612123). iPLA2s functions in a variety of pathways, including cellular growth, lipid homeostasis, and second messenger generation, exerting their function through the catalysis of the cleavage of acyl groups on glycerophospholipids. The clinical presentation of our patient is reminiscent of the phenotype reported in the mouse null for the Pnpla8 gene. The Pnpla8-/- mouse has cognitive dysfunction, decreased exercise endurance, mitochondrial dysfunction, enhanced exercise endurance, and cardiac hypertrophy. As observed in the Pnpla8-/- mouse model, there are unique morphologic changes observed in mitochondria of this patient’s muscle biopsy, which showed an abnormal concentric disarray of internal cristae and globular dense osmiophilic inclusions. In addition, there was a significant increase in secondary lysosomes, many of which contained degenerating mitochondria. Variable fiber size was found with isolated small atrophic fibers of all subtypes; no ragged red fibers were identified. Electron transport chain function was not affected. Oxidative enzyme reactions (SDH and COX) showed no alterations in activity. These findings are consistent with a mitochondrial myopathy. Western blots showed dramatic decreases in multiple bands known to be absent in the Pnpla8-/- mouse in muscle biopsy tissue from this patient. This could be expected from the small decrease in protein translation appositely in frameshifts. Although other iPLA2-related diseases have been identified, such as infantile neuroaxonal dystrophy (MIM#256600) and neutral lipid storage disease with myopathy (MIM#610717), this is the first report of PNPLA8-related disease in a human. We suggest PNPLA8 join the increasing list of genes involved in lipid metabolism associated with neuromuscular diseases.

2502M  

Next generation sequencing (NGS) allows investigation of multiple genes in a single test at a reduced cost compared to traditional sequencing methodologies. We design a targeted OncoHeme NGS Panel of 37 genes known to have mutations in AML, MDS, MPN, CLL, and DLBCL diseases. The genes were selected based on current literature and potential for clinically actionable findings. Here we evaluate the analytical components of the test life cycle process for developing this panel. Using a cross-platform sequencing strategy we sequenced 31 samples on Illumina MiSeq and Ion Torrent PGM instruments. To assess the effect of targeted enrichment, 20 and 11 samples were sequenced using AmpGiseq PCR product library with TruSeq Nano adapters and Agilent Sure Select capture methods respectively. Ampiseq PCR products with Ion Torrent adapters were also sequenced on the PGM for all 31 samples. CLCbio Server software was used to analyze all the sequencing data. PGM data was also analyzed using TorrentSuite software to compare performance with CLCbio. We obtained about 5 million sequenced MiSeq reads per sample (average coverage ~4000x) yielding ~55 variants including Single Nucleotide Variants (SNV) and Insertions/Deletions (INDEL). The variants were detected at a minimum coverage of 100x and a minimum variant coverage of 20x. An average of ~1.7 million sequenced reads per sample from Ion Torrent (average coverage ~2000x) yielded ~38 variants using TorrentServer and ~300 variants using CLCbio. In-frame Internal Tandem Duplication (ITD) events ranging from 3 bp to more than 400 bp have been reported in the literature in 10-20% of AML patients, most ITDs are no more than 100bp. However, medium sized INDEL (20-70bp) are challenging to detect in NGS data. CLCbio identified these medium sized INDEL (up-to 42bp) for Ion Torrent sequenced samples but not for the MiSeq sequenced samples. We therefore also analyzed MiSeq data using PINDEL tool that successfully reported these medium sized INDEL. We also evaluated the effect of PCR duplication bias on variant calling using the functionality of ‘mark duplicates’ and ‘remove duplicates’ in Samtools, Picard and CLCbio plugin. To allow for testing multiple mutations in a single test, we developed the CLCbio Server filter module which will allow this assay to determine tumor heterogeneity at the single-cell level. We designed a well-defined approach for analysis and interpretation of an NGS clinical panel.

2503T  

Evaluation of copy number variation (CNV) in tumor tissues typically involves SNP microarray or aCGH. These methods have high whole-genome resolution, but require large amounts of input material, have high fixed costs, and have sub-optimal performance on formalin-fixed paraffin-embedded (FFPE) samples. Here, we employ a 250,000-plex SNP-targeted PCR methodology using next-generation sequencing (NGS), targeting chromosomes 1, 2, 13, 18, 21 and X, and regions 4p16, 5p15, 7q11, 15q, 17p, 22q11, and 22q13, to detect CNVs from fresh and FFPE samples. The assay was validated using genomic DNA from 96 human samples with known karyotype, including 71 with deletion or duplication of a region, using a minimum of 3 samples per region; 100% accuracy for copy number was observed. Single-molecule sensitivity was established by analyzing single cells. We determined the utility of this method to detect CNVs in breast cancer using both breast cancer cell lines and fresh frozen (FF) or FFPE breast tumor samples. Analysis of 5 breast tumor cell lines and matched non-tumor cell lines revealed multiple amplifications and deletions (median: 13) in all 5 tumor samples, whereas CNVs were absent in the matched non-tumor control samples. Of 31 FF and 2 FFPE breast tumor samples, 90.3% (28/31) of FF and 100% (2/2) of FFPE samples showed full or partial CNVs in at least 1 and up to 15 regions; of the 30 samples with detected CNVs, 93.3% had a CNV of either 1q or 17p. As two most common breast cancer CNVs reported on this panel. Significantly, CNVs were detected in a large proportion of samples across all evaluated tumor stages: 100% (7/7) of Stage I, 77.8% (7/9) of Stage IIa, 90.0% (9/10) of Stage IIb, and 100% (7/7) of Stage III tumors. Evidence of tumor heterogeneity was also observed. Although not investigated here, the very low DNA sample input requirements will allow this assay to determine tumor heterogeneity at the single-cell level. An assay, using our methodology, analyzing all chromosomes arms would require a sequencing cost of approximately $5 per sample. Thus, this method offers a powerful, efficient, and scalable approach for investigating large numbers of FF or FFPE samples.
2504S
Automated miRNA expression profiling in FFPE tissue using nucleae protection coupled with next generation sequencing. D. Thompson, I. Botros, M. Rounseville, H. Harrison, P. Roche. HTG Molecular Diagnostics, Tucson, AZ.

Diagnostic pathology primarily uses formalin-fixed paraffin embedded (FFPE) tissues. Large FFPE tissue collections - both current and archival - are available, often with links to patient outcomes. A limitation to the utilization of this resource for molecular pathology is the difficulty of nucleic acid extraction from FFPE samples. Crosslinking-induced fragmentation of nucleic acids in FFPE tissue makes extraction and downstream analysis time-consuming and difficult, especially when scaling up to routine screening.

We have developed EdgeSeq, a coupling of our RNA extraction-free nucleae protection assay (qNPA) with next generation sequencing (NGS) - mediated quantification. Library preparation occurs in two simple steps: automated nucleae protection, performed on the HTG Edge Processor, followed by limited PCR cycles to prepare libraries for NGS. No RNA extraction or enzymatic processing of the sample is necessary. Sequencing of the resulting libraries allows counting for quantification of RNA in the sample. The advantages of NGS for detection include a large dynamic range with high sensitivity. We initially used the EdgeSeq technique to develop a miRNA whole-transcriptome assay, based on Version 20 of miRBase. Application of the assay shows highly reproducible miRNA profiles are obtained from a variety of sample types including plasma, FFPE, frozen tissues, and cell lysates, with CVs of less than 10%.

EdgeSeq shows excellent correlation (R=0.97) when measuring matched frozen and FFPE cancer tissues. Importantly, poorly-expressed miRNAs are not “lost” when moving into FFPE. The high correlation demonstrates that FFPE tissue can be used in the EdgeSeq assay with full confidence that the results mirror expression in the original (generally unavailable) tissue sample. Of note, 20% of what we identify as FFPE samples demonstrate that tissue-specific miRNA markers are appropriately detected. The EdgeSeq miRNA assay is an excellent tool for NGS-based profiling of miRNA expression, especially for high-throughput screening of FFPE tissues.

2505M
Non-invasive cell-free tumor DNA-based detection of breast cancer-related copy number variations. B.G. Zimmermann1, E. Kirkizl1er1, M. Hill1, T. Constantin1, S. Sigurjonsdottir2, B. Hoang1, N. Chopra1, M. Rabinowitz1, 1) Research and Development, Natera, Inc., San Carlos, CA, USA; 2) Statistics, Natera, Inc., San Carlos, CA, USA.

Breast cancer screening involves mammography, which has high false positive rates and misses some cancers. Analysis of tumor-derived circulating cell-free DNA (ctDNA) for cancer-associated copy number variations (CNVs) may allow for earlier, safer, and more accurate screening. Here, we employed a single-nucleotide polymorphism (SNP)-based massively multiplex PCR (mmPCR) approach to screen for CNVs in ctDNA isolated from the plasma of breast cancer patients. The mmPCR assay targeted 3,168 SNPs on chromosomes 1, 2, and 22q, which often have CNVs in tumors. We have developed a computational tool, called EdgeSeq, that allows the analysis of high-throughput sequencing data to detect CNVs in the ctDNA. The tool identifies CNVs with high accuracy and low false positive rates. We have identified over 200 CNVs that are specific to breast cancer and are not present in normal tissues. This approach has the potential to revolutionize breast cancer screening and diagnosis.

2506T

High sensitivity combined with very low false positive variant calls are critical requirements in an NGS cancer panel. For the Oncomine® Cancer Research Panel targeting about 3800 known cancer specific variants covering 143 genes and using the Ion PGM™ Sequencer we develop an error model that achieves an average of 0.1 false positives per sample rate in the hotspots and a total of about 5 false positive variants per sample in 250 KB of non-hotspot positions when assessed with 1000 Genome samples. A DNA sample mixture was made by combining cell lines and synthetic linear DNA containing known cancer variants so that 334 hotspots were present in a single sample at 10%; allele frequency. This sample was used to create an Ion AmpliSeq™ library and was sequenced on a Ion S18™ Chip using the Ion PGM™ Sequencer. Using our new error model, a sensitivity of >97%; for SNPs and 75% for indels was observed.

2507S
Detection, estimation and correction of technical effects in copy number estimation using NGS in a targeted cancer panel. Y. Zhan1, B. Johnson1, Z. Zhang1, C. Van Loy1, P. Williams1, P. Wyngaard2, A. Carr1, D. Brinza1, A. McDaniels3, C. Liu4, S. Tomlin5, S. Sadis1, M. Andersen1, J. Veitch1, 1) ThermoFisher Scientific, South San Francisco, CA; 2) University of Michigan, Molecular and Cellular Pathology.

We demonstrate a metric “Median Absolute Pairwise Difference” applied to assessment of the per-sample reliability of copy number estimation using the Oncomine® Cancer Research Panel that includes amplicons targeting 49 genes implicated in cancer copy number gains and 26 tumor suppressor genes with cancer related copy number losses using the Ion PGM™ Sequencer. We demonstrate the need for correction of per-sample technical effects including those based on per-amplicon GC content and length. Using mixtures of ATCC cell lines with known copy number changes we observe that detection of genes with copy number >=8 and tumor cellularity >=30% can be identified with a sensitivity greater than 98% when the correction methods are applied.

2508M

As large-scale gene panels have emerged in the clinical genetic diagnostics space, it has become essential to simultaneously query the mutation status of large numbers of genes at a cost that is both affordable and low cost. Multiplex ligation-dependent probe amplification (MLPA) is the gold standard technique for detecting genomic copy number variation (CNV) in the clinic, but traditional MLPA is limited to detection of 40-50 targets per reaction because individual target peaks must be resolvable from one another on a capillary electropherogram. Here we report a novel extension of the original MLPA technology that enables detection of CNV events using high-throughput sequencing instrumentation, which we term MLPAseq. We demonstrate the simultaneous interrogation of ~1000 targets across 48 samples in a single Illumina Miseq run, and detect duplication and deletion events found in common reference samples and in clinical samples from our diagnostic laboratory. MLPAseq displays high sensitivity and specificity for single- and multi-exon events across >40 clinically relevant genes, including ATM, BRC1, BRC2, CFTR, DMD, FANCA, MEN1, MLH1, MSH2, NHP1, PKHD1, PLP1, PMP22 and SMN1. Specifically, we have used MLPAseq to identify CNVs in BRC1 and BRC2 with >99% sensitivity and >99.5% specificity, as verified by array comparative genomic hybridization (aCGH) and hybridization capture methods. MLPAseq is low cost, and both laboratory and data analysis workflows are easily automated, making it ideal for primary CNV detection or secondary CNV call confirmation for exome or Mendeliiome-sized gene panels in a clinical genetic diagnostics environment.
2509T
Identification of five G6PD common deficiency variants using a novel SNaPshot method in patients of the province of Chiriqui, Panama. O. I. Balsita1, 2, R. W. Allen2, 1 Ctr Geneticstest, David, Chiriqui, Panama; 2 CEGEN, Universidad Autónoma de Chiriqui, Panama; 3 Department of Forensic Sciences, Oklahoma State University, USA.

The clinical, biochemical and genetic consequences of the mutations of the glucose-6-phosphate dehydrogenase (G6PD) gene can be important physiologically. G6PD deficiency is an enzymopathy affecting about 400 million people worldwide and biochemical testing of neonates has shown that G6PD deficiency is the metabolic disorder with the highest prevalence in the world. G6PD deficiency is a clinically heterogeneous disorder that may be asymptomatic, and go through life without being aware of their deficiency. They are, however, at risk of having acute hemolytic crises in response to infection, eating fava beans, and to drugs having a high oxidation potential. A SNP typing strategy was designed to simultaneously distinguish six different mutations that defined five G6PD deficiency variants. The variant with its respective mutations are the following: 1) G6PD A (+) (c.376 A>G, p.Asp126Asn); 2) G6PD A (-)(c.376 A>G/c.202 G>A, p.Asp126Asn/Val68Met and c.968T>C, p.Tyr229His); 3) Mediterranean (c.563 C>T, p.Ser180Phe); 4) Canton (c.1376G>T, p.Arg459Leu) and the Canton A variants; 5) Kaiping (c.1388G>A, p.Arg463His). This approach is based on the single-base extension of an unlabelled minisequencing primer that anneals one base upstream of the relevant SNP. The presence of DNA polymers, a fluorescently labeled dideoxynucleotide (ddNTP) gets incorporated into the SNP site. A total of 24 adult individual of combination G6PD enzyme were screened for the six mutations mentioned before. Genomic DNA was extracted from buccal swabs of the deficient subjects using the NucleoSpin®Tissue Kit (Macherey-Nagel). PCR amplification reactions were performed using primers containing the SNPs from each DNA sample. A six-plex of minisequencing reactions with the SNaPshot multiplex kit (Applied Biosystems) was carried out and analysis of the extension products was performed using capillary electrophoresis. The results were confirmed by restriction fragment length polymorphism (RFLP) or sequencing. The results shown that 84% of the screened patients presented the selected mutations indicating they are highly informative. The G6PD A-variant (72%) is the most frequent, followed by the Mediterranean (16%) and the Canton (13%) variants. 16% of individuals did not present any of the studied mutations and further sequence is recommended for them. The approach used has a higher multiplexing capacity, robustness, and extreme sensitivity.

2510S

Over 70 liver transplants are performed in the US each year with similar numbers performed worldwide. While advances in immunosuppression therapy (IST) have occurred over the last two decades, rates of acute cellular rejection (ACR) still remain significant, and co-morbidities from ISTs such as nephrotoxicity and diabetes remain as major clinical issues. The ability to subclinically identify liver graft recipients who are on ACR trajectories, and those that can successfully withdraw from IST using minimally invasive, reproducible biomarkers with high specificity and sensitivity would represent a major advancement in personalized patient care. We performed miRNA profiling of 752 transcripts on 318 serum samples from 90 liver recipients transplanted from the NIH (ITN)-030 and CTOT03 studies. 48 recipients were randomized to supervised minimization and IST withdrawal, prior to a clinically indicated biopsy event, and were analyzed for their rejection and for identification of those who may tolerate low, or no, IST. Serum miRNA profiles at time of biopsy from 104 samples with and without biopsy proven ACR were compared in a two-stage study. 15 miRNAs were observed to be significantly associated with ACR diagnosis after multiple testing corrections (FDR-adjusted p-value < 0.05). A Logistic regression model consisting of a 3 miRNA panel to differentiate ACR from non-ACR with an AUC of 0.90 (95%CI: 0.84-0.95), 92.6% sensitivity & 84.2% specificity [p=0.0001]. This 3 miRNA ACR signature was tested in an independent validation cohort (sera samples from 19 ACR and 16 non-ACR patients), confirming the performance of the model to differentiate ACR from non-ACR (AUC of 0.89 [95% CI: 0.83 - 0.94], 84% sensitivity and 75% specificity, p = 0.01). Statistically significant alterations in this three miRNA ACR panel preceded rejection events by up to 40 days. The enhanced accuracy and expansive targeted gene sequencing tests, leaving an open question as to what are the best practices for mutation screening and clinical reporting of sequence data. Here, we present data on the development and validation of a low cost, high throughput test using the Fluidigm Access Array™ and Illumina HiSeq™ 2500 Rapid Run technology to create a multi-gene sequencing based carrier screen/ pharmacogenetic laboratory developed test for cystic fibrosis, spinal muscular atrophy, 18 common Ashkenazi Jewish Disorders, four cytochrome p450 genes important in common drug metabolism pathways and Factor II and Factor V Leiden. Deep multiplexing (up to 768 samples per run) and use of the rapid run technology has made high throughput sequencing on the Illumina HiSeq2500™ cost effective and fast. By combining multiple tests and providing results of unceded tests in the bioinformatics pipeline, we can achieve additional operational flexibility and efficiencies of scale in a production genomic setting. Further, we present a comparative analysis of three different reporting algorithms for cystic fibrosis carrier screening. We provide data on the use of a limited, phenotypically validated set of 154 variants in a production setting and validated concordant results between the two platforms, and a general variant calling protocol. The data from the first 10,000 de-identified samples is evaluated to give guidance as to the benefits and drawbacks of the three different reporting paradigms.
2513S
Next-Generation Sequencing of the "Clinome" in a hypertensive newborn identifies novel mutations in ACE of the renin-angiotensinogen system. H. Davoud1, C. Beaurel1, O. Dar1, N. Carson1, B. Potter2, C. For R. Consortium1,4, J. Majewski4, D. Bulman1, P. Chakraborty1, T. Lacazze6, K.M. Boycott1,2, D.A. Dyment1,2, J. Richer1, 1) Department of Genetics, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada; 2) Children’s Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada; 3) Department of Epidemiology, University of Ottawa, Ottawa, Ontario, Canada; 4) McGill University and Genome Quebec Innovation Centre, Montréal, Québec, Canada; 5) Department of Pediatrics, University of Ottawa, Ottawa, Canada.

Next Generation Sequencing (NGS) has significantly improved our ability to make molecular diagnoses for rare disease and there is significant opportunity for its application in the neonatal intensive care setting (NICU). Here we present the case of a premature newborn with anaemia and refractory hypotension. After no response to catecholamines, the neonate was able to maintain blood pressure and kidney function on vasoopressin. Next generation sequencing of a comprehensive sequencing panel, targeting >4800 clinically-relevant disease genes (termed the "clinome"), identified compound heterozygous pathogenic mutations in the Angiotensin I Converting Enzyme gene (ACE, NM_000789.3; c.820_821delAG [p.Arg274Glyfs*17] and c.3521delG [p.Gly1174Alafs*12]). Mutations in this gene are known to cause Renal Tubular Dysgenesis (RTD; OMIM 267490), that results in failure of lung development and fetal, or neonatal, demise. The few long-term survivors with RTD have experienced chronic or end-stage renal disease. Our patient’s prematurity (27 weeks) and favorable response to vasoopressin potentially has spared the child significant renal pathology. At 7 months, chromosomal age, the patient’s renal ultrasound was normal. One month later, his potassium was normal on the kayexalate previously prescribed and he appeared to have nicely recovered from his neonatal acute kidney disease secondary to hypovolemia. The clinome approach provided an answer in 5 days and its focus on disease-related genes permitted pathogenicity-status of the variants to be readily evaluated, without further research study, which has obvious benefits for patients in the NICU setting.

2514M

Multiplicom’s MASTR (Multiplex Amplification of Specific Targets for Resequencing) assays enable multiplex PCR amplification of all required gene regions in a limited number of PCR reactions, resulting in the generation of highly efficient, low cost assays to establish a wide range of clinical and diagnostic applications. MASTR assays are based on a simple two-step protocol enabling specific amplification of the regions of interest followed by the incorporation of molecular barcodes in each amplified product to unambiguously link each read to the sample it originated from. This simple and straightforward two-step protocol enables direct compatibility with all commercially available, bench top massively parallel sequencing (MPS) platforms for cost-effective sequencing. Multiplicom currently offers sixteen targeted gene panels for diagnostic predisposition testing of cancer (e.g. breast and colon cancer) and inherited diseases (CFTR, DMD, Marfan...). Also, MASTR based cancer panels for detection of somatic mutations in tumor tissue that target important cancer therapy related genes linked to many cancers, including melanoma, lung, colorectal, gastrointestinal, prostate, breast, brain and pancreatic cancer. MASTR MOPS protocol enables high multiplexability of the complete coding region of cancer genes: the TP53 gene panel; panel containing BRAF, NRAS, KRAS and a panel containing EGFR, ERBB2, PTEN, PIK3CA and PIK3R1. The ability of these assays to use minimal amounts of precious FFPE material allows unlocking a wealth of genetic information for many tumor types using a reproducible and cost effective method. Furthermore, we developed and employed large targeted gene panels, comprising more than 500 exons, for mutation analyses in conditions such as sudden death. Pre-MPS mutation analysis of these conditions was virtually impossible due to large workload, cost and turn around time. Our data show that these large panels allow any diagnostic lab to perform these tests. Also, we showed that a substantial number of patients contain more than one disease causing mutations, not discovered before as a result of a cascade of testing enabling better genetic counseling. For several of these MASTR panels large validation studies have been performed or are ongoing. Results from these studies will be presented and show that multiplex PCR based gene panels enable comprehensive testing of diagnostic relevant genes which can be readily applied to large cohorts of patients.

2515T
Next-generation sequencing as a genetic diagnostic tool to improve the detection of tuberous sclerosis complex (TSC) causative variants. H. Lui1, T. Wu2, P. Chan1,3,4, 1) Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan; 2) Institute of Molecular Medicine, National Taiwan University College of Medicine, Taipei, Taiwan; 3) Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan; 4) Graduate Institute of Medical Genetics and Proteomics, National Taiwan University College of Medicine, Taipei, Taiwan.

Tuberous sclerosis complex (TSC) is an autosomal dominant disorder characterized with benign tumor growth in multiple organs including skin, brain, heart, lung, kidney, and other organs causing clinical symptoms such as skin abnormalities, developmental and behavioral problems, and lung and kidney diseases. Identification of pathogenic mutation in either TSC1 gene at chromosome 9q34 or TSC2 gene at chromosome 16p13 has been determined to cause TSC and is sufficient to make a definitive diagnosis of TSC. The molecular genetic testing of TSC1 and TSC2 has been incorporated into the diagnostic criteria of TSC. However, the fact that TSC1 gene consists of 23 exons and TSC2 gene consists of 41 exons spanning in 53 and 45 kb of genomic DNA regions leads the genetic testing of TSC by traditional Sanger sequencing a time and labor-consuming work. In this study, we tested the use of next generation sequencing as the TSC genetic testing tool by taking the advantage of its high throughput to raise the sensitivity of mosaicism. To enhance the detection of TSC causative variants including large deletion, insertion and inversion, we performed next generation sequencing with target enrichment strategy. The whole genomic regions containing all the exons, introns and 10 kb of both 5’ untranslated region (UTR) and 3’UTR of the TSC1 and TSC2 were captured by costumer designed Roche NimbleGen SeqCap EZ Choice Library. The enriched libraries were paired-end sequenced (2’300cycles) by Illumina Miseq system. Bioinformatics tools including BWA, SAMtools, Picard, GATK and ANNOVAR were applied for variants analysis. SIFT and PolyPhen2 were used to predict the biological significance of the genetic variants. Sequence data were further analyzed by Pindel and Breakdancer software to reveal the precise chromosomal breakpoints of the structural variations such as large insertion/deletion (indel) and inversion. All results were visualized by integra- tive genomics viewer (IGV). Other than 13 exons TSC1 and 14 exons TSC2 causative variants in TSC1 or TSC2 genes including 51 single nucleotide variants (SNVs), 20 small indels and 8 large indels were identified in total 79 families. Moreover, 2 lineal relatives from 2 TSC families turned out to be TSC1 and TSC2 mutations with different combinations, respectively. As examined, causative variants in TSC1 or TSC2 genes including 51 single nucleotide variants (SNVs), 20 small indels and 8 large indels were identified in total 79 families. Moreover, 2 lineal relatives from 2 TSC families turned out to be TSC1 and TSC2 mutations with different combinations, respectively. These data suggest that next generation sequencing serving as a genetic diagnostic tool is economical and sufficient to detect the tuberous sclerosis complex causative variants.

2516S
Comparative study for the evaluation of a new technology for cystic fibrosis screening. M. Majolini1, M. Rongioletti1,2, P. Papa1, C. Vaccarella1, A. Luciano1, C. Centrone3, B. Minuti3, V. Mazzucchi1, M. Belli1, F. Torricelli1, G. Liambruno1, 1) Clinical Pathology Department, San Giovanni Calibita Hospital, Rome, Italy; 2) SOD Diagnostica Genetica, AOU Careggi, Florence, Italy.

Introduction Screening for Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene mutations, is strongly recommended in infertile couples planning a pregnancy by assisted reproductive technology (ART). This study evaluated the performance of the new Nanochip CF70 kit (Savoy Diagnostic, Israel), a microarray assay, and compared it with the Innolipa kit (Innogenetics, Belgium) Methods We analyzed 382 blood samples with Innolipa and Nanochip technologies that identify respectively 70 and 56 CFTR mutations. Discordant results were analyzed with the Dewyser CFTR Core Kit (Dewyser, AB, Sweden) based on PCR allele specific technology, and Sequenom’s MassArray system (Diatechpharmagenetics, Italy) Results Innolipa and NanoChip were concordant for 373/ 392 samples, 21/392 (0.5%) discordant results were tested with the aforementioned technologies: DS confirmed Innolipa results in 18/21 samples and 1/2 in NanoChip results, while Dewyser and Sequenom did not recognize identification of two different homozygous deletions; although they were not present in Innolipa panels, in 2/21 samples Innolipa indicated a mutation with the warning no interpretation possible Conclusions In this study the Innolipa assay confirmed its reliability and Nanochip showed that it could become competitive with slight changes to the software.
DNA methylation is associated with gene regulation as a component of epigenetic mechanism. Abnormal methylation status alters gene expression, which results in a subset of congenital disorders including Prader-Willi syndrome (PWS) and Angelman syndrome (AS). Methylation-specific PCR (MSP) assay has currently been used for early diagnosis of PWS, which allows early start of growth hormone therapy to prevent obesity and diabetes mellitus. An HPLC column has been used to differentiate amino-acid sequences, and a new anion exchange HPLC column has recently been developed to differentiate DNA sequences, which is used for assessment of SNPs. Using the new column, we tried to differentiate DNA methylation status in PWS and AS patients using their sodium bisulfite-converted DNA samples which were PCR amplified with a primer set of SNRPN promoter region. As a result, control normal individuals showed a bimodal peak pattern (an early peak for the methylated allele and a late peak for the unmethylated allele), whereas PWS patients showed a single peak pattern (an early peak for the methylated allele only) and an AS patient showed a distinct single peak pattern (a late peak for the unmethylated allele only). These results were consistent with the results of MSP assay and bisulfite sequencing in the SNRPN region. Taken together, the new assay can clearly distinguish between 0, 50, and 100% methylation status. We are now investigating to define its sensitivity for methylation. Since the anion exchange HPLC column-based assay is a rapid and inexpensive method that can be automated on a standard type of HPLC devise, the assay will be a best high-throughput screening method of DNA methylation for a large number of patients with epigenetic-associated common diseases such as diabetes mellitus, and thus it will contribute to preemptive medicine of various diseases.

**2518T**

**Development and verification of a Noonan genes Ion AmpliSeq™ panel.**

M. Nelen1, H. Yntema1, F. Pantalone1, J.C. Machado2, C. Rossi3, M. Tartaglia1, J.L. Costa4, K. Nveiling1, 1) Dept. of Human Genetics, Radboud UMC, Nijmegen, The Netherlands; 2) Dept. of Hematology, Oncology and Molec. Medicine, Istituto Superiore di Sanità, Rome, Italy; 3) Ipatimp, Porto, Portugal; 4) Unità Operativa Genetica Medica, Policlinico S. Orsola-Malpighi, Bologna, Italy.

Noonan syndrome is a relatively common autosomal dominant congenital disorder with a high phenotypic variability. It is a clinically and genetically heterogeneous disorder that belongs to the group of Rasopathy diseases, caused by mutations in genes dysregulating the RAS/MAPK pathway. Currently, mutations in 14 genes have been described. Standard analysis using Sanger sequencing is expensive and time consuming. Prenatal analysis of all the genes is often not possible due to limited amount of material the small amount of available fetal DNA often allows the analysis of a maximum of five genes. Here we describe the development of a Noonan genes Ion AmpliSeq™ panel for prenatal diagnosis. It includes the coding regions of 14 known Noonan genes, plus 5 basepairs to cover exon-intron boundaries. In total, two pools including 269 amplicons were designed, allowing amplification of all the target regions with only 20 ng DNA. The amplicon sizes range between 125 and 275bps, covering 100% of the desired regions of interest. In a first pilot experiment, pre-and postnatal DNAs carrying known mutations have been sequenced. The average chip loading was 81%, and the average read length was 184bp. Coverage analysis revealed that 266 of the 269 amplicons were covered completely, whereas 3 amplicons show no coverage. The failing amplicons concern (parts of) exon 1 of the genes BRAF, CBL and MAP2K2. No pathogenic mutations have been described for these regions. Coverage depth for the other regions was good with 98% of all amplicons being covered above 100X. Mutation analysis was performed for all samples, and the described mutations (25 different mutations in 12 different genes) could be identified in all of them. The sequencing data of the prenatal samples is very much comparable to the postnatal samples. Further sequencing is ongoing, to reach a total cohort of 100 analyzed samples carrying known mutations, including both prenatals and postnatal DNAs. These preliminary results demonstrate that the Noonan Gene AmpliSeq™ panel may be used in the clinical research setting for prenatal as well as postnatal samples. The simultaneous analysis of all genes is not only attractive in case of limited DNA, but also offers a faster analysis at reduced costs.

Recent advances in molecular genetics have allowed the determination of the genetic cause of non-syndromic hearing loss, but a large percentage of patients still remain with unidentified cause. It points an immedial need for new methodological strategies for the detection of an increased number of mutations in multiple genes. In this work we developed a panel of 98 mutations, previously identified, in 20 different genes for screening of hearing loss, using mass spectrometry system MassArray, Sequenom® and also evaluated the contribution of selected mutations in the etiology of deafness in Brazilian individuals. To minimize erroneous results, mainly caused by the interaction of primers in multiplex reactions, we tested and standardized the assays using 21 control samples with different mutations in GJB2, SLCO2A4 and MT-RNR1 genes. For every mutation that failed, primers were redesigned and the assays were grouped into different wells. After standardization, the values of sensitivity and specificity of the technique proved to be above 84%. A total 150 unrelated individuals affected with non-syndromic hearing loss were then screened using the developed muta-

tion panel. The possibility of environmental factors causing HL was ruled out by historic information and medical examinations. Mutations were identif-

ied in 43 patients (28.7%) and the etiology could be concluded in 16% of the cases. All positive results were validated by other techniques. The c.356delG mutation in the GJB2 gene was the most prevalent, identified in a total of 24 individuals in at least one allele. Other mutations in the GJB2 were identified in 17 subjects, and the p.M34T mutation was the second most prevalent detected for cases. Mutations in SLCO2A4 gene were found in five individuals, and p.V609G mutation was the most common of this gene, detected in three individuals. Mutations in MT-RNR1, MYO15A, OTOF and CDH23 genes were also identified in one patient each. The genotyping of mutations in mass spectrometry system has been shown to be faster and cheaper than Sanger sequencing and allowed to analyze a larger num-

ber of genes than those currently assessed to diagnose hearing loss. Thus, the panel developed presented promise for unraveling the etiology of hearing loss and for genetic counseling.

2522S Classification of Incidental Finding Variants in 6503 Participant’s Exomes. L. Amendola1, P. Robertson2, D. Nickerson2, M. Dorschner2, J. Salama1, EVS 6500 annotation consortium1, G. Janyk1, 1 Division of Medical Genetics, University of Washington, Seattle, WA; 2 Genome Sciences, University of Washington, Seattle, WA.

Two challenges in genomic medicine implementation are the classification of variant pathogenicity and identified variants (IFs). The ACMG has proposed discovery and return of IFs for 56 genes from genomic tests. In order to explore expert crowd sourced variant classifications and estimate the rates of IFs, we classified all putatively pathogenic variants in 6503 participants in the ESP (exome sequencing project). The 6503 participants in the ESP included African (n=2,203) and European (n= 4,300, including 187 Ashkenazi) ancestry. We considered variants in 117 genes associated with medically actionable conditions which may be undiag-
nosed in adults. Briefly, the pathogenicity criteria included the requirement for published studies with multiple individuals and/or segregation or de novo mutation data or novel truncations, where that mechanism is known to cause disease. Using consistent criteria, 48 experts reviewed 628 variants that were classified as ‘disease causing mutations’ in the Human Genetic Muta-
tion Database (HGMD) and had minor allele frequency < 0.005 for autosomal dominant and x-linked (612) and <0.1 for autosomal recessive (16) disorders. 67 stop and splice variants found in ESP but not in HGMD were also considered. Initially, a random 156 (25%) of the HGMD variants were blindly double-reviewed; of these 81 (52%) did not match and 53 mismatches involved pathogenic or likely pathogenic classification. We then blindly re-
classified all missense variants classified as pathogenic or likely pathogenic: 56% changed classes (44/79) with 42 of 44 reclassifications moving from pathogenic to likely pathogenic or variant of uncertain significance. The fre-

quency of pathogenic, likely pathogenic, and expected pathogenic novel disruptive mutations were 0.6, 1.0, and 0.1% in those of European ancestry and 0.2, 0.5, and 0.4% in those of African ancestry. Relationships of classifi-
cation to allele frequency and in silico scores were also explored. The defi-

cient of findings in subjects of African ancestry may reflect the lack of relevant literature. 87% of HGMD ‘disease-causing’ mutations did not meet our criteria for pathogenicity. Even with expert reviewers and clear criteria, there was substantial inconsistency in the classification of variants that was marked by over interpretation of variant pathogenicity. This suggests that crowd sourcing of variant annotation has limitations.


Introduction: The importance of understanding gene fusions in cancer biology is increasingly important, and may have profound impacts on next-
generation drug discovery efforts. Historical testing methods such as fluores-
cence in-situ hybridization (FISH) are limited by scalability and subjective interpretation. Growing numbers of fusion genes have been identified, and other oncogenes in hematological malignancies have further complicated matters. Targeted sequencing technologies, such as Anchored Multiplex PCR (AMP), have the ability to simultaneously interrogate hundreds of tar-
genome regions to identify novel and known gene fusions from small amounts of input material in a single reaction. In this study, we will demonstrate the detection, breakpoint characterization, and concordance between current technologies and the targeted Archer(TM) FusionPlex(TM) Heme next-generation sequencing panel. Materials and Methods: Heme fusion samples were sourced from CAP-accredited tissue repository (Asterand) and through collaboration with Massachusetts General Hospital. Samples with known/unknown Hematological malignancies, as determined by existing technology, were included for the study. The Archer FusionPlex Assay was performed according to Enzymatics package insert, and sequenced on the Illumina MiSeq platform. Sequencing analysis was performed via the Archer Analysis Pipeline, and reported fusion results were compared to current methods. Results: Concord-
cance between current methods and the FusionPlex NGS results is high, with all positive samples also having a positive NGS result. In contrast to current methods, the Archer FusionPlex assay was able to characterize fusion partners previously unidentified via RT-PCR or other methods with-

out prior knowledge of fusion partner or the laboratory and high cost of RNA-

seq screening. Conclusions: The Archer(TM) FusionPlex(TM) Heme panel is a powerful new tool for identifying disease relevant gene fusion events from small amounts of input material in Hematological cancers. Our initial results support that this technology is a robust and higher throughput tool for the detection of “gold standard” of FISH testing, while overcoming the obvious limitations of reflex testing. Furthermore the novel AMP-based sequencing method allows for cost-effective high-throughput identification of fusion breakpoints and fusion partners to the relevant oncogenes without prior knowledge of the partners.
2525T
Sensitive and Comprehensive Method to Detect Mutations in RB1 Gene Improves Care for Retinoblastoma Patients and Their Families. W. Li, J. Buckley, P. Sanchez, D. Maglinte, L. Vidueysky, T. Triche. Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, CA. Retinoblastoma is childhood eye cancer caused by inactivating mutations in both alleles of the tumor suppressor gene, RB1. RB1 is the only gene in which mutations are known to cause heritable predisposition to retinoblastoma. Retinoblastoma can lead to the loss of vision and, sometimes even life. Molecular identification of a germline RB1 mutation in a timely manner is very important for the effective care and management of retinoblastoma patients and their families. However, current methods to detect disease-causal mutations in RB1 are lengthy, usually taking weeks or months. In this study, we report a next-generation sequencing (NGS) based method that is capable of capturing pathogenic mutations in the entire 200 kb RB1 gene, both exons and introns. The process of DNA extraction, sequencing data generation, and completion of data analysis requires only 3 days. In addition, the associated bioinformatic pipeline can take NGS data and accurately measure the copy number of the RB1 gene and MYCN gene, thus enabling the NGS method to simultaneously detect point mutations, small InDels and large deletion/duplications in the RB1 gene and MYCN gene amplification of the MYCN gene on a single test platform, which was impossible previously. Because of the deep sequencing coverage of the entire RB1 gene, this method can detect even a low percentage of mosaic mutations directly from blood samples, which is essential for genetic counseling of retinoblastoma patients and their families. To assess the feasibility and impact of mutation and other RB anomalies like promoter hypermethylation, we are developing a novel targeted NGS assay to simultaneously assay both RNA and DNA, in order to measure the transcript level of expression and DNA methylated status of the RB1 gene. The assay will be developed by development of a targeted methylation assay to complete DNA, RNA, and epigenetic profiling of the RB1 gene. Taken together, we believe that these combined NGS based methods can lead to rapid, inexpensive, comprehensive, and accurate detection of RB1 mutations and their expression. This in turn will enable detection of germ line carriers who are at lifetime increased risk of a second malignancy like osteosarcoma, a decidedly more aggressive and more often fatal disease compared to retinoblastoma.

2526M
A multi-platform amplicon sequencing method for fast and reliable variant detection. M. Toloue, L. Matzat. NGS, Bio Scientific Corp, Austin, TX. The application of NGS to clinical work and patient diagnoses has been limited by several factors including cost, time, and limitation of samples. Rapid benchtop sequencing as well as multiplexing reduce both cost and time limitations of sequencing; however, sample preparation for these platforms remains a rate limiting step. Whereas target capture methods require constraint of a DNA-seq library and a target capture process, amplicon libraries can be constructed by simple PCR steps that sequentially build platform-specific and barcode sequences onto target sequences. Here we present an amplicon sequencing library construction method that produces libraries in highly multiplexed reactions and is compatible with multiple platforms. Our library preparation allows accurate targeting and even coverage across amplicons. Data will be presented demonstrating amplicon coverage and SNP detection as well as the flexibility for customized amplicon panel design.

2527T
A Comprehensive Profile of Hereditary Myopathies by Next Generation Sequencing on 43 Early-Onset Patients and Subsequent Development of a Biomarker Assay by Liquid Chromatography-SRM-Mass Spectrometry. S. Hahn1,2,3, V. Vasta4, S. Jung4, Q. Zhang4, S. Eun5, A. Cho6, B. Lim7, J. Chae7. 1) Department of Pediatrics, University of Washington School of Medicine, Seattle, WA; 2) Seattle Children's Hospital, Seattle, WA; 3) Center for Developmental Therapeutics, Seattle Children's Research Institute, Seattle, WA; 4) Fred Hutchinson Cancer Research Center, Seattle, WA; 5) Department of Pediatrics, Korea University Research Center, Seoul, Korea; 6) Department of Pediatrics, Ewha Woman's University School of Medicine, Seoul, Korea; 7) Department of Pediatrics, Seoul National University School of Medicine, Seoul, Korea. Hereditary myopathies encompass a significant proportion of patients with chronic muscle disease. The diagnostic approach for hereditary myopathy is often difficult due to its heterogeneous genetic background and similar clinical presentations. Most patients require a muscle biopsy for the diagnosis, which is particularly challenging in infants or newborns. Unfortunately, many patients remain undiagnosed due to the lack of specific markers for many myopathies or due to ambiguous pathological results. Furthermore, Sanger sequencing of individual genes is challenging because one gene can cause a wide variety of clinical and/or pathological features, while similar clinical features can be caused by mutations in different genes. Here, we explored Next Generation Sequencing panel (~600 genes) on 43 patients with early-onset myopathy, targeting known pathogenic genes for hereditary myopathies or secondary muscle weakness. Fourteen novel/rare variants, fifteen known pathogenic variants and five VUS in 17 genes were identified in thirty two patients. Three had myasthenic myopathy (DOK7, AGRN, and GFTP1); two had peripheral neuropathy (GARS, DYN1L1). Incomplete penetrance was strongly suspected in four cases with variants in COL1A1, FYR1, CCDC78 and MYBPC3 genes, indicating that careful interpretation is required in the context of clinical, laboratory and pathological findings when the rare variant is inherited from asymptomatic parents. As COL6A4 genes were most commonly affected in our cohort while no specific biomarkers are currently available, we explored a Liquid Chromatography Selected Reaction Monitoring Mass Spectrometry (LC-SRM-MS) to identify and quantify the proteotypic peptides of Collagen VI in skin fibroblasts. We identified several candidate signature peptides for Collagen VI and are currently evaluating the method to see if the abundance in patient’s fibroblasts is reduced or absent. This assay can open a valid experimental model to study the candidate drugs that are becoming available. Although our study indicates that simultaneous sequencing of multiple genes for various types of hereditary myopathies is clinically relevant, a substantial number of patients still remained without molecular diagnosis in our cohort implying that new causative genes for hereditary myopathy are to be identified. Further development of biomarker assay for hereditary myopathies is necessary and could be feasible with LC-SRM-MS in selected conditions.
Genetic Diagnosis of Duchenne Muscular Dystrophy by Clinical Exome, Whole Genome and Transcriptome Sequencing, H. T. Wang1,2, H. Lee1, A. Eskin1, V. Arbedo1, K. Squire1, J.L. Deignan2, N. Khanlou2, P.B. Shieh2, S.F. Nelson1,2, 3, 1) Human Genetics, University of California Los Angeles, Los Angeles, CA; 2) Center for Duchenne Muscular Dystrophy, University of California Los Angeles; 3) Department of Pathology and Laboratory Medicine, Los Angeles, CA; 4) Department of Neurology, University of California Los Angeles.

Duchenne muscular dystrophy (DMD; OMIM #310200) is the most common form of muscular dystrophy characterized by a progressive loss of skeletal muscle function. Roughly 1 in 5000 male births are affected worldwide. The X-linked DMD gene is one of the largest genes in the human genome and mutations within the 2.2 Mb gene are the unequivocal cause of DMD. The mutational spectrum of DMD is heterogeneous with approximately 60-65% of affected individuals with large exonic deletions, 5-10% with large duplications and 25% with single base, small indels, or splice site mutations typically resulting in absent dystrophin expression. Although the gene responsible for DMD has been known for over 25 years, as many as 30% of affected boys have not received a molecular diagnosis in the United States, demonstrating some barriers to clinical diagnosis. Multiple ligation dependent probe amplification (MLPA) or array CGH, reliably identify exonic deletions and duplications through probe hybridization targeted to specific exons. When MLPA fails, Sanger sequencing of exons and splice sites is undertaken. The large size of the DMD gene, however, makes comprehensive assessment of all exonic mutations challenging and expensive, and a small portion of histologically diagnosed patients do not have a detected mutation. Here we report the use of whole exome sequencing as a clinical test to detect point mutations, small insertion and deletions, and single or multiple exonic deletions and duplications in the DMD gene simultaneously, with high sensitivity and specificity. By sequencing to high depth of coverage, we reliably identified the disease causing point mutations, exonic deletions and duplications for 27 of 30 samples. We utilized paired-end whole genome sequencing to identify duplicated breakpoints for two samples and further employed RNA-seq on biopsied muscle tissue coupled with whole genome sequencing in two refractory cases to uncover a novel genomic rearrangement and an unusual change in splicing with base-pair resolution.

Hi-Plex: a flexible, streamlined and cost-effective approach to targeted massively parallel sequencing, T. Nguyen-Dumont1, B.J. Pope1, P. Hammel2, M. Mahmoodi3, H. Tsimiklis4, M.C. Southey5, D.J. Park1. 1) Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Parkville, VIC, Australia; 2) Victorian Life Sciences Computation Initiative, Carlton, VIC, Australia; 3) Department of Computing and Information Systems, The University of Melbourne, Parkville, VIC, Australia. Purpose: Massively parallel sequencing (MPS) has revolutionised biomedical research and offers enormous capacity for clinical application. We previously reported Hi-Plex, a streamlined highly-multiplexed PCR-based target-enrichment system for MPS. Here, we present its optimization to 1,000 amplicon-order multiplexing and gene panel screening of thousands of specimens.

Methodology: Hi-Plex library-building consists of a single-step, highly-multiplexed PCR, followed by size selection, using relatively inexpensive, readily available reagents in a simple half-day protocol. A given library can be sequenced with both the TruSeq and Ion Torrent chemistries. Hi-Plex does not require normalisation of barcoded libraries before pooling and sequencing. The Hi-Plex system relies on the development of a novel design software that allows the uniform definition of library size, so that subsequent paired-end sequencing can achieve complete overlap of read pairs. Variant calling from Hi-Plex-derived datasets relies on the identification of variants appearing in both reads of a read pair, permitting stringent filtering of sequencing chemistry-induced errors using our ROVER software. Comparable sequencing efficiency was achieved using material derived from lymphoblastoid cell lines and formalin-fixed paraffin-embedded tumours.

Results: Initially reported in a 60 amplicon multiplex (60-plex) setting, Hi-Plex’s chemistry was expanded to perform in a 1,000-plex level. Using equal amounts of all primers, we observed that 95% amplicons were represented within 20-fold of the median coverage. Using the improved Hi-Plex chemistry and expanding on a published demonstration of Hi-Plex capabilities to perform high-throughput population sequencing, we developed a next generation sequencing assay to detect point mutations, small insertions and deletions, and single or multiple exonic deletions and duplications in the DMD gene simultaneously, with high sensitivity and specificity. By sequencing to high depth of coverage, we reliably identified the disease causing point mutations, exonic deletions and duplications for 27 of 30 samples. We utilized paired-end whole genome sequencing to identify duplicated breakpoints for two samples and further employed RNA-seq on biopsied muscle tissue coupled with whole genome sequencing in two refractory cases to uncover a novel genomic rearrangement and an unusual change in splicing with base-pair resolution.

Conclusions: Hi-Plex is simple, accurate, low cost, modular and flexible in terms of target region design and sequencing platform. These features render the approach highly attractive for an extensive range of clinical and research applications.
Nijmegen breakage syndrome detected by newborn screening for T cell receptor excision circles (TRECs). A.N. Adhikari1, J. Patel2, R.A. Gatti3, C. Brown4, U. Sunderam5, K. Kundu6, R. Srivasan4, S.E. Brenner4, J.M. Puck7, J.A. Church2. 1) Department of Pediatrics, Children’s Hospital of Los Angeles, Los Angeles, CA; 2) Departments of Human Genetics and Pathology & Laboratory Medicine, David Geffen University of California Los Angeles School of Medicine, Los Angeles, CA; 3) Innovations Labs, Tata Consulting Services, Hyderabad, AP, India; 4) University of California, Berkeley, CA; 5) Department of Pediatrics and Institute for Human Genetics, University of California, San Francisco, CA; 6) Department of Pediatrics, Keck School of Medicine, University of Southern California and Children’s Hospital, Los Angeles, CA.

Severe combined immunodeficiency (SCID) is a group of disorders characterized by reduced or absent T-cell number or function. Newborn screening for SCID utilizes quantification of T cell receptor excision circles (TREC). These circular DNA byproducts of thymic T cell receptor rearrangements correlate with numbers of circulating T cells. Although TREC screening was developed to identify patients with SCID, it has identified T lymphopenia in newborns with disorders that may not otherwise present until later in life. We present a case of an infant patient with Nijmegen breakage syndrome, an autosomal recessive disorder characterized by microcephaly, a distinct facial appearance, short stature, immunodeficiency, radiation sensitivity, and a strong predisposition to lymphoid malignancy. Low newborn TREC led to confirmatory testing that demonstrated persistently low T cells from birth. Whole exome sequencing (WES) was performed on the infant and his healthy parents to look for causative genetic variant(s) potentially associated with the observed phenotypes in the infant.

The sequencing data were processed through our analysis pipeline. The first several steps employed standard tools for mapping and variant calling, but integrated two different carefully-tuned callers to yield high quality sets of variants. Comprehensive variant annotation was performed to flag the compound heterozygous, de novo, uniploidal disomy, and X-linked recessive variants with potential clinical importance. After filtering the variants and prioritizing genes based on an immunological disorder associated gene list, we found two previously unreported compound heterozygous nonsense mutations in the gene for nibrin (NBN) (c.1146C>T;p.Gln344*, c.952T>G;p.Leu281*), predicted to cause nonsense-mediated mRNA decay of transcripts from both alleles in the infant patient. The identified NBN variants were confirmed experimentally using Sanger sequencing in a CLIA-approved laboratory. Immunoblotting showed absence of nibrin protein and a colony survival assay demonstrated radiosensitivity comparable to patients with ataxia telangiectasia. Thus, WES after abnormal SCID newborn screening made the early diagnosis of Nijmegen breakage syndrome.
2534S  
Detection of Low Level Mixed Chimerism Using High Throughput SNP Genotyping. A.A. Nakorchevsky1, E. Flores1, X. LF, T. Hong2, A.O.H. Nygren1. 1) Agena Bioscience, San Diego, CA, USA; 2) Hackensack University Medical Center, Hackensack, NJ, USA.

Patients diagnosed with blood malignancies often receive allogeneic bone marrow transplant or stem cell transplants following the regimen of chemotherapy. Status of the post-transplant patients has to be monitored carefully to allow for early diagnosis of such post-transplant adverse effects as transplant rejection, graft vs. host disease or hematologic relapse. Triaging of the transplant recipients in clinical and research settings is achieved by monitoring the Minimal Residual Disease (MRD) or measuring the amount of mixed chimerism in peripheral blood lymphocytes (PBL) of the patients. Mixed chimerism is a phenomenon where in addition to the donor cells the recipient white blood cells are detected post allogeneic transplant indicating that the possibly malignant recipient cells have evaded cytoreductive treatment and are undergoing hematopoiesis. While MRD molecular techniques target disease-specific markers such as mutations and translocations, mixed chimerism is detected via PCR-based typing techniques such as Short Tandem Repeat (STR) and Variable Number Tandem Repeat (VNTR) analysis. We created a single nucleotide polymorphism genotyping method to detect mixed chimerism in PBL and circulating cell free DNA. The panel targets 92 independent SNPs with a minor allele frequency 0.45-0.55 via multiplex PCR followed by single base extension into the SNP site and detection with MALDI-TOF mass spectrometry. Identification of low percentage mixed chimerism is achieved by analyzing the cumulative skew in genotyping data across a cohort of 92 markers. The feasibility of the panel and algorithm was verified using artificially created mixed chimeric samples from mice. Using this model system we achieved a 1% limit of detection with an analytical sensitivity and specificity of 0.95 and 0.9 respectively. Finally we evaluated a set of pre-validated clinical chimeric samples and obtained 100% concordance with an orthogonoal algorithm. The advantages of this SNP-based method compared to already established methodologies is that it has higher sensitivity than STR-based methods and it does not require disease specific markers or prior knowledge of either donor or recipient genotypes. Results are obtained within 8 hours with no library preparation and the experimental workflow can be used with other genotyping or sequencing technologies.

2535M  
Silver-Russell syndrome and segmental UPD(7q) detected by array CGH. P. Tavares1, A. Vaglio2, R. Lemos3, C. Ventura4, A. Pereira1, A. Sousa1, J. Sai5, J. Pinto Basto1, R. Quadrelli1, F. Rendeiro1. 1) CGC Genetics, Porto, Portugal (www.cgcgenetics.com); 2) Instituto de Genética Médica, Hospital Italiano, Montevideo, Uruguay.

Introduction: 20 to 60% of cases of Silver-Russell syndrome (SRS) are caused by the epigenetic changes of DNA hypomethylation at the telomeric imprinting control region (ICR1) on chromosome 11p15, involving the H19 and IGf2 genes. About 10% of cases are due to maternal uniparental disomy (UPD) of chromosome 7 [Penaherrera et al 2010]. A few cases are due to segmental UPD7 [Hannula et al 2001 and Eggermann 2008]. Objective: We report a SRS clinical case due to a maternal segmental UPD7 and will compare the clinical and molecular data with the published segmental UPD7 SRS clinical cases. Results: The clinical diagnosis was confirmed by the identification of a run of homozygosity at 7q11.22 to 7q31 with 38.7 Mbp (Cytoscan 750k, Affymetrix) with abnormal methylation pattern on genes GRB10 and MEST (MS-MLPA ME032-A1, MRC Holland), confirming a maternal UPD7. Conclusions: This clinical case supports the association between segmental UPD7 and SRS. Despite the fact that segmental mUPD7 is restricted to 7q arm, imprinting genes on 7p arm are also involved, enlightening the complexity of this disease. This clinical case can contribute for a better understanding of the molecular mechanisms of SRS.

2536T  

Leber Congenital Amaurosis (LCA) is the most common cause of incurable childhood blindness. Non syndromic LCA is characterized by a clinical, genetic and physiopathological heterogeneity. Hitherto 15 genes are identified which mutations cause a stationary and dramatically severe cone-rod dystrophy with very poor visual performances or a progressive, yet severe, rod-cone dystrophy with low but measurable visual acuity in the first two decades of life. LCA genes are involved in variable retinal functions; cilia dysfunctions are the leading cause of the disease (7 cilia genes; ca. 30% of the cases). Patients with LCA occasionally develop renal, neurologic and or skeletal symptoms defining a series of syndromes which hallmark is cilia dysfunction. So far, 19 cilia genes are identified which cause syndromic type 1 and type 2 LCA. With the ever expanding number of mutations in the known LCA genes not all patients with cilia dysfunction can be assigned to one of these syndromes. We performed whole genome sequencing on LCA patients with mutations in a custom designed panel that captures 18,179 known and 1,666 known genes selected from cilia databases. Biallelic disease-causing mutations in known genes were identified in 17/610 patients: CEP290 (n = 3), CRB1 (n = 4), RPE65R1 (n = 2), ICG2 (n = 1), IFIT140 (n = 2), AHI1 (n = 1), ALMS1 (n = 1). Single or biallelic disease-causing mutations were detected in 9/610 individuals but Sanger sequencing failed to detect a mutation on the second allele. Finally, 36/610 patients harbored biallelic mutations in three novel genes. In summary, ciliome resequencing allowed identifying the disease-causing gene in 33 % of the patients addressed for LCA making targeted sequencing an interesting alternative to expensive whole exome sequencing. The identification of mutations in genes responsible for Mainzer-Saldino syndrome (IFIT140), Joubert syndrome (AH1) and Alström syndrome (ALMS1) before extraocular expression become obvious demonstrated the importance of NGS-based molecular diagnosis to set-up a rational and efficient follow-up of patients.
expanding disease phenotype. Further demonstrates the utility of WES in confirming clinical diagnosis and had megalencephaly, abnormal brain imaging, and global developmental mutations in the components in the PIK3-pathway. Although all three patients megalencephaly. In summary, we report three additional individuals with PIK3CA de novo 3 toe syndactyly, and recent accelerated weight gain. Brain MRI revealed plasia, and umbilical hemangioma. The second proband is a 2.6 year old with congenital megalencephaly, somatic asymmetry, connective tissue dys- we identified a AKT3 and PI3KCA since birth, stiffness of extremities, megalencephaly, communicating hydrocephalus-polymicrogyria-polydactyly-hydrocephalus (MPPH - MIM 1255X/750X for phase I II validation, respectively. Phase I samples have an average of 97 variants in the comprehensive panel, which are 100% concordant with amplicon/Sanger sequencing results. In phase II samples, all known point/indel mutations in BRCA1, MEN1, MUTYH, VHL, APC and PTEN were successfully detected. Furthermore, pathogenic exonic CNVs in LGMD and HPCD, and APC intronic CNVs in APC were identified by Sanger confirmation/complementation allows accurate identification of a point mutation (c.65C>A) and an exon 14 deletion of PMS2 gene in one case. This case and the other 43 cases were confirmed. In the next step, a number of single gene diseases, including neuromuscular, cardiovascular, developmental, and metabolic diseases. The sequencing data were produced using MiSeq, HiSeq2500 and MiSeq were compared to identify mutation that were found at high depth of HiSeq2500 data but not present in MiSeq Data. We evaluated the accuracy of each panels to validate NGS technologies for clinical diagnosis. We show that inherited disease panels are able to detect rare mutation in the disease samples, and our data provides good foundation for further clinical studies.

2539T Expanding the phenotype of genes encoding the components of the PI3K pathway, M.A. Walkiewicz1, A. Dang Do2, N.L. de Macena Sobreira1, J. Beuten1, J. Zhang1, Z. Niu1, F. Xia1, J.N. Bodurtha2, B.H. Graham1,4, F. Scaglia1,4, D.M. Muzny1, A.L. Beaudet1, R.A. Gibbs1,2, Y. Yang1, C.M.英格1), Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX. 2) Department of Pediatrics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD. 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX. 4) Texas Children’s Hospital, Houston, TX. (100% coverage. Mutations/unclassified variants detected by NGS were further confirmed by amplicon(s), are complemented by amplicon/Sanger sequencing to reach 100% coverage. Mutations/unclassified variants detected by NGS were further confirmed by Sanger sequencing. Potential copy number variation (CNV) suggested by our NGS assay are confirmed by either array CGH or MLPA methods. 10 phase I samples (commercially available plus internal control gDNA from different types of tissues) and 13 phase II samples (with known mutations) were used for validation. Results: The mean coverage is 1255X/750X for phase I II validation, respectively. Phase I samples have an average of 97 variants in the comprehensive panel, which are 100% concordant with amplicon/Sanger sequencing results. In phase II samples, all known point/indel mutations in BRCA1, MEN1, MUTYH, VHL, APC and PTEN were successfully detected. Furthermore, pathogenic exonic CNVs in LGMD and HPCD, and APC intronic CNVs in APC were identified by Sanger confirmation/complementation allows accurate identification of a wide spectrum of mutation types, including point mutations, small indels, mosaicism. In addition, CNVs at the exon level can be analyzed simultaneously. Our data underscores the important clinical utility of NGS-based analy- sis in the molecular diagnosis of hereditary cancer.

2541M Mutations in STK11 identified exclusively in individuals with clinical histories suggestive of Peutz-Jeghers syndrome. J.S. Dolinsky1, S.B. Keeler1, C. Chang1, M. Umali1, K. McGillinick1, S. Li1, E. Chao1,2, 1) Ambry Genetics, Aliso Viejo, CA, USA; 2) University of Califor-nia, Irvine, School of Medicine, Irvine, CA. Pathogenic mutations in STK11 are detected in an estimated 80% of individuals clinically diagnosed with Peutz-Jeghers syndrome (PJS). STK11 is characterized by gastrointestinal PJS-type hamartomatous polyposis, mucocutaneous pigmentation, and cancer predisposition including breast cancer, which can occur at early ages in PJS with up to a 57% breast cancer risk by age 70. Due to the risk of early onset breast cancer, it has been proposed that family members could develop breast cancer prior to polyposis symptomatology. Thus, STK11 has been included on a number of commercially available and academic research hereditary breast cancer (HBC) multigene panels. In this study, we sought to identify individuals with a STK11 mutation detected on HBC multigene panels offered by Ambry Genetics from March 2012 through May 2014. We hypothesized that some individuals with HBC in the absence of other PJS features will harbor a STK11 patho-genesis mutation. 12,928 individuals underwent panel testing with BRCAplus (BRCA1, BRCA2, CDH1, PTEN, STK11, TP53) or BreastNext (BRCAplus genes and ATM, BARD1, BRI1, CHEK2, MRE11A, MUTYH, NBN, NF1, PALB2, RAD50, RAD51C, RAD51D). Clinical data submitted via test requisition form and subsequently verified by the clinician was assessed for individu- als with a STK11 mutation. Clinical data from an additional 73 individuals with mutations in STK11 identified through single gene testing or another multigene panel was also assessed. Several probands had histories consistent-ent with HBC, however those probands met clinical criteria for PJS. Remaining probands had exclusions for proper assessment of the clinical data. There were 12 probands (9 males, 3 females) with a STK11 mutation who were segregated by a family history of PJS or features consistent with PJS, including one individual identified through a HBC panel. The remaining 12,927 individuals referred for HBC multigene panel testing did not harbor a STK11 pathogenic mutation. While STK11 is included on many high weight 12,927 had a STK11 mutation identified in approximately 0.44% of the HBC multigene cohort, there are significantly more VUS than mutations identified in STK11 on HBC panels (p=0.001). This analysis does not support our hypothesis, but rather suggests that individuals not be screened for STK11 on HBC panels, introducing more uncertain results without improving diagnostic yield. STK11 testing should be reserved for broad spectrum hereditary cancer panels, polyposis panels or single gene analysis in individuals with a history suggestive of PJS.
2542T
Risks associated with utilizing molecular data to guide tumor surveillance in BWS, C. Shuman1,2,3, S. Choufani4, L. Steele4,5, N. Parkinson2,4, J. Lauzon6, J. Tapoe1, P. Ray5,6, P. Weksberg1,2,3,7,8, 1) Clinical & Metabolic Genetics, Hosp for Sick Children, Toronto, ON, Canada; 2) Genetics and Genome Biology Program, Research Institute, The Hospital for Sick Children; 3) Department of Molecular Genetics, University of Toronto; 4) Paediatric Lab Medicine, Molecular Genetics, The Hospital for Sick Children; 5) Department of Medical Genetics, Alberta Childrens Hospital; 6) Genetic Services, Alberta Health Services; 7) Institute of Medical Science, University of Toronto; 8) Department of Paediatrics, University of Toronto.

In general, it is well accepted that tumor risk assessment can be significantly improved by incorporating epigenomic/genomic data. In this regard, recent publications (Brioude et al. 2013; Scott et al. 2006) recommend tailoring tumor surveillance for children with Beckwith-Wiedemann syndrome (BWS) based on the specific molecular etiology, when identified. However, technical limitations of the test in determining somatic epigenetic mosaicism can lead to negative health outcomes. BWS is an etiologically heterogeneous, pediatric overgrowth syndrome associated with an increased risk (~4 - 21%) for embryonal tumor development. This risk led to tumor surveillance recommendations including 3 monthly abdominal ultrasounds, (Epi)genetic alterations involving the telomeric imprinting centre, IC1 (i.e. gain of methylation [GoM]at IC1 or paternal uniparental disomy [UPD] of 11p15.5) have the lowest risks for tumor development. Those with alterations involving the centromeric imprinting centre, IC2 (i.e. loss of methylation [LoM] at IC2 or mutations in CDKN1C) have significantly lower risks and no cases of Wilms tumor (WT) have been reported. These findings led to new molecular group specific recommendations to discontinue ultrasound screening for WT in children with demonstrated IC2 LoM. We therefore present 3 cases involving BWS and IC2 alterations on initial MS-MLPA testing who subsequently were found to have renal tumors/nephrogenic rests detected by abdominal ultrasound. Pt 1: Nephrogenic rests (stable) were detected in a 16 month old female with IC2 LoM in DNA from saliva and blood, on repeat testing. Pt 2: WT was found in a 15 month old female with LoM in DNA from blood. Retesting on a new blood sample revealed borderline findings of IC2 LoM and IC1 GoM consistent with UPD. Pt 3: Unilateral WT was detected in a 9 month old female with IC2 LoM in DNA from saliva and blood, on repeat testing. Pt 2 and Pt 3 showed borderline IC2 LoM. No evidence of UPD was found. Pyrosequencing these cases in the research lab demonstrated methylation indices most consistent with low level mosaicism for UPD. Therefore, we would advocate caution in revising medical management guidelines especially when the etiology involves mosaicism and/or the testing technology generates data that are not definitive. Technological advances may resolve some of these issues to support stratification of tumor risk assessment and surveillance by molecular test results.

2543S
Universal Lynch syndrome screening in an integrated health care system: Assessment of patient perspectives on benefits and barriers. J.E. Huntford1, F. Fergusson1, A. Davis1, E. Barber1, H.K. Reiss1, S.K. Peterson2, S. Syngal1, L. Acheson1, G. Wiens1, K.A.B. Goddard1, J. Reiss1, 1) Center for Health Research, Kaiser Permanente Northwest, Portland, OR; 2) Department of Behavioral Science, University of Texas MD Anderson Cancer Center, Houston, TX; 3) Dana-Farber Cancer Institute, Boston, MA; 4) Departments of Family Medicine and Community Health, Reproductive Biology, and Oncology, Case Western Reserve University, Cleveland, OH; 5) Vanderbilt Hereditary Cancer Program, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN.

Lynch syndrome (LS) is the most common form of hereditary colorectal cancer (CRC) and accounts for 2-3% of all CRC cases. A diagnosis of LS has significant impact on clinical management of patients and their at-risk family members. Universal screening for LS using either microsatellite instability (MSI) or immunohistochemistry (IHC) testing among CRC patients is increasingly implemented in clinical practice. The aim of this study was to assess knowledge of and attitudes toward MSI screening among newly diagnosed CRC patients. Participants were recruited through Kaiser Permanente Northwest and were administered a phone survey regarding MSI screening. The study population consisted of 74 men and 50 women age 39 to 87 years. Only 3% of participants had any prior knowledge of MSI screening, and 8% anticipated having an abnormal screening result. However, most participants expressed interest in MSI screening: 92% wanted to know their risk of hereditary CRC and 93% agreed that MSI screening should be available to anyone. Overall, participants endorsed the benefits of MSI screening for themselves and their families: 91% would be glad that their doctor ordered this test, 94% would like information that could help their family, and 99% responded they would share their screening results with their family. Older participants were less likely to be worried about additional costs associated with an abnormal screening result. Another potential barrier to screening was a concern about privacy and confidentiality of the test result, endorsed by 72% of participants. In general, younger and lower income participants were more likely to be worried about additional costs associated with an abnormal screening result. Although CRC patients in our study had little a priori knowledge about MSI screening, they responded positively to undergoing screening and recognize the potential benefits to themselves and their families. However, concern regarding additional out-of-pocket healthcare costs may be a barrier to MSI screening, particularly among younger and lower income patients.

2544M

Diagnostic exome sequencing is currently the most comprehensive clinical genetic testing option available to patients with neurological disorders. Historically, individuals presenting with features suggestive of an underlying genetic syndrome had limited options in the diagnostic process, especially in regards to their subsequent management. With the recent advancement of Next Generation sequencing techniques, and the rapid growth in clinical utility of diagnostic exome sequencing (DES) in the clinical neurology setting, the diagnostic and treatment options available for epilepsy patients is progressively changing. Here we present our institution’s DES data to show the effectiveness of DES in diagnosing ion channel related epilepsy, which therefore lays the foundation for potential pharmacogenomic treatment protocols in the near future. In the first 496 cases sent to our institution for DES, 131 (27.93%) patients were diagnosed with epilepsy as a major clinical feature upon referral. Approximately 43 (32.8%) of these patients were found to have positive, pathogenic alterations, and an additional 11 (8.4%) patients were found to have “likely positive” alterations in a variety of genes. In both of these cohorts combined (n=54), 1 patient had an alteration in a calcium-activated chloride channel gene (ANO3[610110]), 3 patients had alterations in sodium channel genes (SCN3A[MIM 182391], SCN1A[MIM 182389]), and 3 additional patients had alterations in potassium channel genes (KCNG2 [602235], KCNC3 [176284]). As the current literature suggests, there are rapidly increasing therapeutic possibilities for epilepsy, especially for disorders involving major ion channels such as SCN1A (MIM 182389). We anticipate the continued and increased use of DES in identification of pathogenic alterations for patients with ion channel-related epilepsy, and recommend the integration of DES technology into research and clinical settings for the long-term development of targeted therapies.
Next-generation sequencing diagnostics for inherited arrhythmogenic cardiac disorders. C. Marschall, I. Vogi, HG. Klein. Center for Human Genetics, 82152 Martinsried, Bavaria, Germany.

Inherited arrhythmogenic cardiac disorders are composed of primary arrhythmia syndromes such as long QT syndrome (LQTS), catecholaminergic polymorphic ventricular tachycardia (CPVT) and Brugada syndrome (BrS) and the cardiomyopathies with structural changes leading to an arrhythmia risk such as hypertrophic cardiomyopathy (HCM), dilatative cardiomyopathy (DCM) and arrhythmogenic right ventricular cardiomyopathy (ARVC). All these diseases are associated with significant morbidity and mortality and are a known risk factor for sudden cardiac death. The presymptomatic diagnosis is essential as for most of the cases a therapy is available. Mutation detection in the index cases followed by cascade screening of the family members could be life saving. As some cardiomyopathies may present with only little structural symptoms, the differential diagnosis of these arrhythmogenic disorders may be difficult. Unfortunately some arrhythmogenic disorders are extremely heterogeneous. A Nextera sequence capture in solution target enrichment assay was designed for the capture of coding regions including splice sites of 61 genes known to be associated with arrhythmogenic disorders. Target enrichment followed by re-sequencing on the Illumina MiSeq was used for mutation detection. More than 40 clinical samples with a suspected inherited arrhythmogenic disorder were analyzed with a minimum coverage of 30-fold. Data analysis was performed with the CLC Genomic Workbench software. Potentially disease causing variants and regions with an insufficient coverage were re-analyzed with Sanger sequencing. Disease causing mutations were detected in over 50% of the cases and less than 20% showed ambiguous variants.

Genetic testing leads clinical care in neonatal diabetes: a new paradigm. E. De Franco1,5, S.E. Planagan1, J.A.L. Houghton1, H. Lango Allen2, D.J.O. Mackay15, I.K. Temple5, S. Ellard5, A.T. Hattersley1. 1) University of Exeter Medical School, Exeter, United Kingdom; 2) Wessex Regional Genetics Laboratory, Salisbury Health Care Trust, Salisbury, UK; 3) Human Genetics and Genomic Medicine, Faculty of Medicine, University of Southampton, Southampton, UK.

Traditional genetic testing focusses on analysis of one or a few genes according to clinical features; this testing paradigm will now change as next-generation sequencing allows simultaneous analysis of multiple genes. Neonatal diabetes is the presenting feature of many discrete clinical phenotypes defined by different genetic aetiologies. The genetic subtypes differ in treatment requirements and associated clinical features. We investigated the impact of early, comprehensive testing of all known genetic causes of neonatal diabetes. We studied 1020 patients with neonatal diabetes diagnosed before 6 months of age and referred from 79 countries between 2000 and 2013. Mutations were identified by comprehensive genetic testing including targeted next-generation sequencing of 21 neonatal diabetes genes. A genetic diagnosis was obtained in 840 patients (82%). The most common cause were potassium channel mutations (n=390), which were less common in consanguineous families (46% vs 12%, p<0.0001). Transfer from insulin to sulfonylurea therapy with improved glycaemic control was observed in most of these patients. Median time from diabetes diagnosis to referral for genetic testing decreased from over 4 years before 2005 to less than 3 months (2011-2013). Early referral altered the clinical phenotype at time of genetic testing. In patients with a genetic diagnosis of transient neonatal diabetes remittance was seen only in 10% with early referral (less than 3 months from diagnosis) compared to 100% with late referral (more than 4 years from diagnosis, p<0.001). Similarly in patients with genetically diagnosed Wolcott-Rallison syndrome 11% with early referral had syndromic diabetes compared to 82% (p<0.001) with late referral when skeletal/liver involvement was common. Patients are now referred for genetic testing close to presentation with neonatal diabetes. Comprehensive testing of all aetiologies identifies causal mutations in over 80% of cases. The genetic result predicts optimal diabetes treatment and development of related features. This represents a new paradigm for clinical care with genetic diagnosis preceding development of clinical features and guiding clinical management.

Identification and characterization of CFTR deletion and duplication mutations among nonwhite patients with cystic fibrosis. I. Schrøver2, L. Pique1, S. Graham2, M. Khairoun2. 1) Department of Pathology, Stanford Medical Center, CA; 2) Department of Pediatrics, Stanford Medical Center, CA; 3) Sequoia Foundation, La Jolla, CA; 4) California Department of Public Health, Richmond, CA.

Even with the implementation of newborn screening programs across the United States, there remains an inequitable identification of cystic fibrosis (CF [OMIM #219700]) in nonwhite groups. Diagnosis can be improved by determining the type and estimated frequency of individual mutations in nonwhite populations. We therefore investigated the spectrum of deletion and duplication mutations, which are not detectable by Sanger sequencing, in CF patients in the U.S. with African, Native American, Asian, East Indian or Middle Eastern backgrounds. Of the approximately 500 eligible nonwhite CF patients in the Cystic Fibrosis Foundation (CFF) Patient Registry data-base with zero or one CFTR (OMIM *602421) mutation(s) instead of the expected two, 142 probands were enrolled into our study. Of the enrolled probands, 89 were fully genotyped by direct sequencing of CFTR. However, 26 had only a single mutation and 27 had no mutations detected. These patients were subsequently analyzed by Multiplex Ligation-dependent Probe Amplification (MLPA) and novel deletion/duplication breakpoints were further characterized by a walking PCR technique. MLPA testing of the 52/53 probands for whom DNA remained available identified 15 del/dup mutations in 13 probands: two were identified in two probands each, one of whom already had a pathogenic mutation identified by sequencing (c.1521_1523delCTT, p.Hes508del). Thus, only 14/15 of these mutations accounted for unidentified alleles (14/78 = 18%). Six of the rearrangements were novel and included one single bp deletion, three complex deletions, three complex deletions containing 2-12 bp insertions, one duplication and a multigene deletion of exons 19-24 of CFTR that extends into the neighboring CTTNB2 gene on chromosome 7. Two previously reported rearrangements - a deletion of exons 2-3 and a complex deletion of exons 17a-17b that includes a 62 bp insert - were detected in four probands each (~5% of unidentified alleles). Deletions and duplications in the CFTR gene appear to be relatively common in nonwhite CF patients and account for ~18% of unidentified alleles after Sanger sequencing. MLPA testing is currently routinely applied in diagnostic CFTR testing, and is an especially valuable part of the testing algorithm in nonwhite individuals. Inclusion of such mutations facilitates early diagnosis and can help improve the quality of life of nonwhite and mixed ethnicity individuals affected with this condition.
2549S Comprehensive evaluation of the FBN1, LTBP2 and ADAMTS4 genes in 667 patients with ectopia lentis. B. Callewaert1, L. De Potter1, S. Van Nuffel1, E. Debaes2, J. De Baeker2, O. Vanacker3, H. Van Esch4, M. Bien- vle1, J. van den Ende1, J. Desis2, A. De Paep1, P. Coucke1. 1) Center for Medical Genetics, Ghent University Hospital, Ghent, Ghent, Belgium; 2) Center for Medical Genetics, University Hospital Leuven; 3) Center for Medical Genetics, University Hospital of Brussels; 4) Center for Medical Genetics, University Hospital of Antwerp; 5) Department of Medical Genetics, Hôpital Erasme, Université Libre de Bruxelles.

Introduction: Ectopia lentis (EL), a displacement of the lens of the eye, has an estimated prevalence of 6.4 cases per 100,000 individuals. EL can present as an isolated phenomenon or within a syndromal constellation with multisystemic involvement. The most prevalent congenital cause of EL is Marfan Syndrome (MFS), which requires a lifelong cardiovascular follow-up. Because of the high prevalence of MFS, EL should be evaluated at young age and FBN1 testing is often requested. When no detectable FBN1 mutation is found, the absence of a definite diagnosis may cause psychological distress and sometimes unnecessary examinations. Recently, mutations in LTBP2 have been implicated in congenital glaucoma and microphthalmos with EL. ADAMTS4 mutations have been implicated in EL and EL et pupillae. This study evaluates the contribution of mutations in FBN1, LTBP2 and ADAMTS4 mutations in isolated EL in a large study group. Methodology: Over a 22 month time period a total of 667 probands with EL were referred for FBN1 analysis. We evaluated the FBN1, ADAMTS4 and LTBP2 genes in these patients using a step-wise next-generation approach. Results: 616 out of 667 probands harbored an FBN1 mutation (of which 44 did not fulfill the MFS revised revised revised criteria). 51 did not (of which 19 patients did fulfill clinical criteria for MFS) resulting in a detection rate of 92.4% (616/667). In the remaining 51 patients, homozygous or compound heterozygous mutations in ADAMTS4 were found in four patients. (c.767_786del; c.2021_2022delCT, c.2977C>T, c.963dup). In one proband showed a homozygous mutation in LTBP2 (c.4964A>G). In another proband only one mutation (c.3850C>T, p.Arg1284Cys) was found. Conclusions: FBN1 is by far the primary gene to screen upon the diagnosis of EL. However, our results show that the clinical spectrum resulting from FBN1, LTBP2 and ADAMTS4 mutations includes isolated EL. Therefore, if no FBN1 mutation is found, screening of ADAMTS4 and LTBP2 is advised as this may reveal a final diagnosis for 12% of the remaining patients and even up to 20% if the FMS criteria are not met. In addition, our data indicate that LTBP2 mutations are no-likely cause of FBN1 negative MFS with EL.


Microarray-based comparative genomic hybridization (aCGH) is a powerful technique used in cytogenetic and cancer research for genome-wide detection of copy number variations (CNVs). As a consequence of recent improvements in technology heterogeneity, there has been an increasing focus in single cell analysis in tumors and micrometastases characterization, and in pre-implantation and noninvasive prenatal (fetal cells circulating in maternal blood) investigation. Traditional FISH and PCR based techniques, and more recently BAC arrays, have been used to provide insights into a single cell’s genome, with limited resolution. Here we describe a same day, cost-effective, analysis workflow that combines whole genome amplification (WGA), with CN profiling using high-resolution oligo CHG microarrays for single cell research. To assess the accuracy of aberration detection, a single cell model system was built from a set of normal and aberrant cell lines with varying sizes of known CN changes. As references, lymphocytes isolated from normal male and female individuals were used. Due to the minute amounts of genetic material contained within each cell, all samples and references were subjected to a multiple displacement amplification-based WGA to increase the amount of DNA while maintaining its genomic representation. Amplified samples were then differentially labeled and combined in 7 pairs of test vs. test samples and 1 pair of reference vs. reference per 8-pack microarray slide. This eliminates the use of a reference for every array and allows the processing of 14 samples on one slide. Combined samples were hybridized for 2-6 hours to an 8x60K CHG microarray containing probes optimized for single nucleotide resolution, microarray slides were washed and scanned. The data were extracted and analyzed for CN alterations using algorithms and a single cell specific analysis method implemented in Agilent CytoGenomics 2.9. Using an optimized workflow for sample preparation, hybridization, microarray analysis and CN data processing, we present a workflow that allows the analysis of 14 single cells per day. We are currently assessing the accuracy of this workflow to be implemented in the single cell model system. Aberrations affecting whole and portions of chromosomes, as small as 6 Mb, were accurately identified and confirmed by qDNA CN profiling. This short workflow enables researchers to obtain reliable results at a high resolution while remaining cost-effective.

2552S Performance comparison between low coverage semiconductor sequencing and array comparative genomic hybridization to analyze copy number variation. B. Min, M. Seo, J. Kim. Seoul National University, Seoul, South Korea.

Copy number variations are alterations of the DNA and have an important role in the pathogenesis of human disease including cancer and genetic disorders. Comparative genomic hybridization is the most widely used analytical method for genetic identification of copy number. Next generation sequencing technologies provide an opportunity to detect copy number by comparing the number of sequence reads between patients and control samples. To evaluate the use of semiconductor sequencing for copy number variation diagnosis, Detection of copy number using the Agilent 4 X 180k aCGH platform was compared with low coverage semiconductor sequencing on the patients with genetic disorders. Detected changes in copy number using low covarate sequencing were validated by using aCGH results data and Sanger sequencing. Most copy number variants detected among patients using the aCGH method were successfully detected by low coverage sequencing. In the diagnostic setting, a balance of low coverage sequencing cost and diagnostic sensitivity and specificity should be considered. Turnaround time of experiment also should be considered to challenge array comparative genomic hybridization method. semiconductor sequencing method is alternative sequencing platform with a reduced turnaround time and cost of sequencing. We propose that low coverage semiconductor sequencing can be used as a diagnostic procedure for the patients with CNVs.
2553M
Effects of clinician guided genomic risk assessments: HelloGene. H. Jin1, T. Kim1, K. Ahn2, J. Bhak3, H. Joo1, S. Dong2, 1) TheragenEtex, Suwon, Gyeonggi, South Korea; 2) Personal Genomics Institute, Genome Research Foundation, Suwon, South of Korea; 3) Kyung Hee University School of Medicine, Seoul, South of Korea.

Genome wide association studies (GWASs) on common diseases have been reported increasingly since 2005. These results were applied for the personalized genomic risk assessments tests by several corporations, such as 23andMe, Navigenics, DeCODEEme and so on. These companies offer direct-to-consumer (DTC) tests without the obligatory involvement of the health care provider. Since DTC is illegal in Korea, the test results of personalized genomic risk assessments tests, HelloGene by Therabio Inc., should be delivered and explained by clinician. HelloGene service was launched in January, 2013 in Korea. However, genomic risk assessments test were not popular in the healthcare facilities, and social awareness about risk assessment test was limited among clinicians. To overcome these circumstances, Therabio Inc. has initiated the clinical trial with 18 university hospitals via academic society, the Korean Society of Health Screening and Promotion, in January, 2014. The trial consists of two survey questionnaires which are given before and after the HelloGene service. The first questionnaire consists of the ten questions regarding basic understanding and personal impression about the test such as previous experience, knowledge and so on. After the test, the subjects were surveyed with the second questionnaire, which have 20 questions mainly about the satisfaction and change in perspective about the genomic risk assessments test. Up to now, 55 subjects were completed both questionnaires. Interestingly, 54 out of 55 subjects were responded that it was helpful to have genetic counseling by clinician. In addition, 54 out of 55 subjects showed willingness for changing their lifestyle according to the results to prevent diseases. Even though the trial is on-going, it is already producing important information regarding general populations’ impression and perspective towards genomic risk assessment test, and it is predicted to showing positive aspects for genomic risk assessment tests in Korea.

2555S
Screening Results From 79424 Patients Tested by CFnx, a 147 Mutation Cystic Fibrosis Screening Assay Built on the Illumina BeadXpress Platform. C. Holland, R. Tinawii-Aljundij, M. Weindel, J. Steecker. Progenity, Ann Arbor, MI.

Introduction: In March 2011, the American College of Obstetricians and Gynecologists (ACOG) updated its screening recommendations for cystic fibrosis: include offering CF screening to all women of reproductive age. The current ACOG mutation panel offers a comparatively limited risk-reduction for racial minority patients. We developed a CF screening assay that interrogates 147 mutations and 4 variants and provides greater risk reduction across all ethnicities. Methods: The assay consists of 12 multiplex PCR reactions, treatment of the product with a SAP/Exo I mixture, 12 multiplex allele-specific primer extension (ASPE) reactions with biotinylated CTPs, consolidation of ASPE reactions, hybridization onto the VeraCode® Bead set (Illumina), binding of the streptavidin-Aluox Fluor® (Invitrogen) conjugate, and scanning of the beads by the BeadXpress®. 96 samples and controls are run per batch in 384-well plates. Sample transfer is assisted by automated liquid handling. All putative positive samples are repeated for confirmation. The run-time for CFnx is approximately 12 hours. Results: Through 2 June 2014, Progenity has completed testing on 79,424 patients. A total of 2484 heterozygotes and 17 compound heterozygotes were confirmed, resulting in an overall confirmed positive rate of 3.15%. Of the heterozygotes detected, 1944 were part of the ACOG panel (77.7%), and 557 were non-ACOG panel mutations (22.3%). A total of 22 out of 23 ACOG mutations were identified, along with 67 non-ACOG mutations. This includes 557 patients that would not have been otherwise identified. The most common mutations observed were deltaF508 (1368), R117H (257), W128X (62), G551D (60), G551V (36) and 3849+10kbC>T (27). In addition, out of the top 20 mutations identified, 5 were non-ACOG mutations. A total of 251 samples resulted in failed analysis (0.32%). Conclusions: The CFnx assay provides robust screening across a broad range of CF mutations, provides significantly increased carrier detection for all ethnicities, compared to the ACOG panel, and an improvement in identification over other expanded panels.

2554T
Rare finding of non-transient congenital extra and intra-medullary acute myeloid leukemia with Beckwith-Wiedemann syndrome. G.A. Jervis, H. Monforte, R. Ortega, L. Harris, M.J. Sutcliffe. Department of Pathology and Laboratory Medicine, All Children’s Hospital, John Hopkins Medicine, St. Petersburg, FL.

A male neonate presented with macrosomia, hepatosplenomegaly, right ventricular hypertrophy, macroglossia, neutropenia, and “blueberry muffin rash”. Extradrenal myelomonocytic leukemia cutis was diagnosed by pathology from skin biopsy with a concomitant involvement of acute myeloid leukemia (AML), FAB subtype M5 in the bone marrow. Constitutional SNP HD array revealed a terminal 4.8 Mb gain of 11p15.5, a terminal 3.7 Mb deletion of 8p23.3, and a mosaic 10.5 Mb deletion between 17q21.31-q22. Cyto genetics and FISH confirmed an unbalanced der(8)(8;11)(p23.3;p15.4) with duplication of the Beckwith Wiedemann syndrome (BWS) locus, consistent with the diagnosis of BWS, KCNQ1OT1 imprinting abnormality with hypermethylation of the locus and normal H19 overgrowth and tumor predisposition syndrome. MS-MLPA detected an allelic imbalance with duplication of the BWS region; an incidental finding with no clinical relevance. The wide phenotypical variability is associated with the type of beta-globin mutation, the co-inheritance of alpha-thalassemia and the ability for persistent production of fetal hemoglobin (HbF) in adult life. For the latter, three different quantitative trait loci, accounting for 20-50% of HbF variation, have been identified by now. Single nucleotide polymorphisms (SNPs) in the gamma-globin gene promoter (HBG2), in the BCL11A gene and the HBS1L-MYB intergenic region lead to increased residual HbF levels in adults. Method: A teststrip-based reverse-hybridisation assay was developed for the simultaneous detection of SNPs in the HBG2 (g.-158 C>T), BCL11A (rs1447407, rs10189857), HBS1L-MYB (rs28384513, rs9399137) genes. Results: The new StripAssay enables the concomitant identification of genetic variants known to influence beta-thalassemia disease severity. Based on the presence of positively modifying alleles, and combined with alpha- and beta-globin genotyping, it allows the prediction of patients likely to display less severe phenotypes. Favourable properties, such as the rapid DNA extraction protocol, ready-to-use reagents and test strips, as well as the potential for automation of the hybridisation/detection and interpretation steps, make the StripAssay convenient and easy to perform within less than six hours. Conclusions: Testing for genetic modifiers influencing disease severity will lead to more specific and effective treatment, and support clinical decisions regarding the beginning of transfusion therapy in beta-thalassemia patients. Furthermore, the knowledge about prognostic markers has implications for genetic counselling and prenatal diagnosis.
A multidisciplinary approach to the evaluation and care of patients with inherited lung disease. B. Raby1,2, D. Toledo2, L.M. Yonker2, E. Henske1, T.B. Knane4. 1) Division of Pulmonary and Critical Care Medicine, Brigham & Women's Hospital, Boston, MA; 2) Channing Div. Network Medicine, Brigham & Women's Hospital, Boston, MA; 3) Partners Center for Personalized Medicine, Boston, MA; 4) Massachusetts General Hospital, Boston MA.

Like other medical disciplines, clinical pulmonary medicine is entering the post-genome era, where increased accessibility to genetic testing, converging with greater public awareness of the role of genetics in health, poses new challenges to pulmonologists. Moreover, most clinical training programs provide little exposure to monogenic conditions other than alpha-1-antitrypsin deficiency and cystic fibrosis. To address this, we established the Brigham and Women’s Hospital Pulmonary Genetics Center, a multidisciplinary program that centralizes the diagnosing, managing, and counseling of patients and at-risk family members with genetic lung diseases. All patients are evaluated by a pulmonary geneticist in conjunction with a certified genetic counselor. Patients are triaged for appropriate genetic testing, and for subsequent evaluations by specific subspecialists, including pediatric pulmonologists, adult medical geneticists, a diverse team of pulmonary subspecialists with expertise in rare lung diseases, and the lung transplant team. Clinical fellows from both the pulmonary and medical genetics fellowship programs formally attend the clinic as part of their training. Partnering with the Laboratory for Molecular Medicine (LMM) of the Partner’s Center for Personalized Medicine, we developed a CLIA-certified next-generation sequencing panel capable of evaluating more than 60 genes related to diverse inherited lung diseases, including familial forms of cystic disease, pulmonary fibrosis, bronchiectasis, and pulmonary hypertension. All patients are counseled both prior to and following testing. Here we describe the first 75 referrals to the center, outlining the spectrum of patients evaluated, their experiences undergoing genetic testing, and clinical outcomes resulting from this testing.

The most common reasons for referral were for evaluation of patients with fibrosis or cystic lung disease with a family history of similar, and for young adults presenting with unusual constellations of clinical findings suggesting rare genetic disease. Pathogenic variants have been identified for the majority of patients tested. The most common diagnosis, made in 10% of referred patients, has been fibrosis caused by short telomere syndrome. The initial case series also includes families with Brit-Hogg-Dube Syndrome, Blau Syndrome and Primary Ciliary Dyskinesia. The implications of confirmed diagnoses are described, highlighting the utility of this approach.

Identification of truncating mutations in the CHD8 gene in patients with autism spectrum disorders by clinical whole exome sequencing (WES). J. Zhang1, M.S. Leduc1, Z. Niu1, R.E. Person1, W. Alcaraz1, E.E. Roeder3, T. Moss4, J.M. Nguyen5, M.R. Wagle6, P.A. Ward1, A.A. Braxton1, T.R. Vaughn7, D.M. Muzny2, A.L. Beaudet1, R.A. Gibbs1,2, C.M. Eng1, F. Xia1, Y. Yang1. 1) Dept Molec & Human Gen, Baylor Coll Med, Houston, TX; 2) The Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 3) Department of CHSA Pedi-Genetics, Texas Children’s Hospital, TX; 4) Genomic Medicine Institute, Cleveland Clinic, OH; 5) Department of Pediatrics, University of Texas at Houston, TX; 6) Department of Pediatrics-Neurology, Texas Children’s Hospital, TX.

Autism spectrum disorder (ASD) is a complex neurodevelopmental disability characterized by impaired social interactions, communication difficulties and restricted and repetitive stereotyped behaviors, with an overall prevalence of 62 cases per 10,000 people. Genome-wide studies such as chromosomal microarray and whole exome sequencing (WES) have identified de novo changes in numerous genes in patients with ASD. While most of those genes have yet to be validated as candidate autism genes, a few such as CHD8 and ADNP have been better defined. CHD8 encodes chromodomain helicase DNA binding protein 8 and was previously linked to autism in nine patients in one report in 2012 (PMID 23160955). Here we report four additional cases referred for clinical WES who were found to harbor novel truncating mutations in CHD8. Patient 1 was a 5.5 year old female with clinical phenotypes including expressive language delay, dyspraxia, motor delay, hypotonia, dysmorphic features, tall habitus, macrocephaly, obesity, strabismus, and scoliosis. She was found to carry a heterozygous de novo c.4611dupA (p.V1538fs) mutation in the CHD8 gene. Patient 2 was a 14 year old male presented with autism, intellectual delay with regression, attention deficit disorder, hypotonia, ataxia, tremors, tics, dysmorphic features, nausea, constipation, seizures, retinal disease and scoliosis. This patient had a heterozygous de novo c.2565del (p.N855fs) mutation in the CHD8 gene. Patient 3 was a 4.5 year old female with autism, global developmental delay, speech delay, hypotonia, dysmorphic features and macrocephaly and a heterozygous c.4468dupT (p.C1490fs) mutation in the CHD8 gene. The parental samples were not available for study. Patient 4 was a 5 year old male with autism, intellectual disability, speech delay, hypotonia, dysmorphic features, macrocephaly, and hyperextensibility. WES revealed a heterozygous c.6518C>A (p.S2173X) in the CHD8 gene, Sanger sequencing showed that the mother is negative for this mutation; the father was not available for study. Overall, the four cases with truncating CHD8 mutations identified by WES further expand the phenotypic spectrum of CHD8-related disorders. The occurrence of CHD8 mutations in our clinical WES samples is four out of a total of 1800 unrelated patients referred for autism spectrum disorder and intellectual disability with or without other organ system involvement.
2559M
Use of a Patient-Centered Conceptual Framework to inform the development of the ClinGen Resource. M.S. Williams¹, S. Aronson², H.L. Rehm³, S. Goehringer¹, A. Milosavljevic¹, E.M. Ramos³, Clinical Genome Resource Consortium. ¹Genomic Med Inst, Geisinger Health System, Danville, PA; ²Executive Director of Information Technology, Partners HealthCare Personalized Medicine Boston, MA; ³Dept. of Pathology, Brigham & Women’s Hospital and Harvard Medical School Director, Laboratory for Molecular Medicine, Partners Personalized Medicine Boston, MA; ⁴Department of Molecular and Human Genetics Baylor College of Medicine Houston, TX; ⁵Division of Genomic Medicine, National Human Genome Research Institute, NIH Bethesda, MD.

Background Healthcare systems represent a complex ‘system of systems,’ resulting in challenges in coordination, access, communication and integration that need to be addressed in order to improve the quality and decrease the cost of care. One such system is the Clinical Genome Resource, a product of the NIH-funded Clinical Genome Resource consortium (ClinGen). The ClinGen resource will provide information about actionable genomic information to clinicians at the point of care. As the integration of the resource into the healthcare system was being contemplated, the consortium agreed to explore a patient-centered conceptual framework. The key insight in this framework is that the only constant in the care delivery system is the patient and the patient’s caregivers. Following this framework, we reimagined the ClinGen resource information delivery process from the patient/caregiver perspective—so-called patient-centered care. Methods The patient-centered framework created by the UK National Health System Heart Improvement Programme was used. The interaction of the patient/caregivers with their providers within this framework is termed the clinical microsystem. A workgroup of ClinGen was convened to use this framework to conceptualize the organization and relationship of the elements that would interact with the ClinGen resource either directly or indirectly. Results The workgroup used an iterative approach to develop a ‘system map’. This was presented to the ClinGen Consortium steering committee for comment and discussion which were incorporated into a final draft. This draft defined the critical applications in the ecosystem such as Electronic Health Records and Laboratory Information Systems, dbGaP, ClinVar, and other external data sources and identified key interactions and data flows that needed to occur to achieve optimal utility. The final design was approved by the steering committee. Conclusion The approved system map will be used to construct and implement the ClinGen resource. This means that the project is committed to efficient clinical microsystem applications and data flows that will help the goal of improving health through the use of evidenced-based genomic information can be realized. Currently the various ClinGen workgroups are taking ownership of the different interactions and data flows to determine the specifications and standards that will be needed to optimize each interaction and system dependency.

2560T
Mutation analysis of CYP21A2 and correlation between genotype - phenotype in 163 Vietnamese patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. V. Duong¹, K. Tran², P. Le², K. Nguyen¹, H. Le¹, M. Fukami³, V. Ta⁴. ¹Medical Genetics and Metabolism, National Hospital of Ped, Hanoi, Viet Nam; ²Hanoi Medical University, Research Center of Gene and Protein, Hanoi, Vietnam; ³National Research Institute for Child Health and Development, Department of Molecular Endocrinology, Tokyo, Japan.

Background: Congenital adrenal hyperplasia (CAH) is one of the most common inherited metabolic disorders. The most common form of CAH (classical) is caused by mutations in CYP21A2, the gene encoding the adrenal steroid 21-hydroxylase enzyme (21-OHD; OMIM 2019110). Objective: To identify the mutations in the CYP21A2 gene in Vietnamese patients with CAH and attempt a genotype-phenotype correlation. Methods: Molecular analysis was performed using PCR, multiplex ligation dependent probe amplification and direct sequencing of PCR products of the CYP21A2 gene in 163 CAH patients. Correlation between phenotype and genotype was evaluated based on identified mutations and clinical manifestations. Results: Mutations were identified in 321 alleles. Twelve different causative mutations with 29 different genotype were identified in CYP21A2 including two novel mutations. The most frequent genetic defect was c.293-13A>C (115 alleles; 35.8% mutation, followed by Large deletion (88 alleles; 27.4%); c.1066C>T (p.R356W) (49 alleles; 15.3%); c.515T>A (p.L172N) (33 alleles; 10.3%). The rarer mutations were c.952C>T (p.Q318X) (9 alleles; 2.8%); c.1276C>T (p.R426C) (7 alleles; 2.2%); c.920_921insT (L307fs) (7 alleles; 2.2%); c.1447_1448insG (p.R446fsX4) (6 alleles; 1.9%); c.372delA (E246fs) (1 allele; 0.3%); c.438G>A (p.M147) (1 allele; 0.3%); and novel mutations c.160T> G (1 allele; 0.3%). The majority of patients (127 cases; 77.9%) were homozygous. Sixteen cases were compound heterozygous. Genotype accurately predicted phenotype in 97.3 and 95.3% of patients with salt-wasting (SW) and simple virilizing (SV), respectively. 44 cases with homozygous mutation c.293-13A>C-G are associated with SV phenotype and 3 patients present with SV form. Conclusions: The spectrum of mutations of the CYP21A2 gene in Vietnamese patients is comparable to the reported in other populations. Large deletion accounts for nearly one-third of the genetic defects. Therefore, laboratory should include methods for detecting point mutations as well as large deletions. Genotype-Phenotype correlation was high in the studied patients.

2561S
Preventive effects of α-tocopherol on telomere shortening in human buccal cells. S. Yabuta, Y. Shidoji, M. Masaki. Graduate School of Human Health Science, University of Nagasaki, Nagasaki, Nagayo, Japan.

[Background and Objectives] Telomere shortening is considered as a solid index for senescence of human cells. However, telomere length may be affected by both physical aging and oxidative stress. In as much as telomere shortening is caused by oxidative stress, it should be preventable with meal intervention (Blackburn, E et al., JAMA 2010;303:250-257). According to the previous results, oxidative stress may cause genotoxic damages as well as suppression of telomerase activity, which is also active in certain somatic cells. In this context, habitual intake of antioxidant nutrients such as β-carotene and vitamin E may be important to prevent telomere shortening. The present study addresses interactions between the relative telomere length (RTL) in buccal cells and anti-oxidant vitamin-related genes including BCMO1, SR-B1, and ISX genes. These three gene products are involved in intestinal uptake and metabolism of fat-soluble vitamins. Finally, we analyzed a relationship between buccal RTL and daily intake of antioxidant nutrients. [Methods] The subjects consist of Japanese men and women (n=70, aged 20-59y) voluntarily recruited with a written informed consent. Telomere length was measured by monochrome multiplex quantitative PCR method with buccal cells. The telomere length was expressed as a relative telomere length (RTL) comparing with a reference sample, BCMO1 (rs6564851), SR-B1 (rs2278986), and ISX (rs362090) were typed by real-time PCR method with TaqMan probes. The data of nutrient intake were collected by food frequency questionnaire (FFQ). [Results and Conclusions] Allele frequencies of G and T in rs6564851 were 0.836 and 0.164, respectively and these values fit into those of Japanese population uploaded in SNP database. In association of RTL with the gene polymorphisms, there was a positive and significant correlation between buccal RTL and daily intake of α-tocopherol in men of BCMO1 carriers, 1.2% (p<0.05). When BCMO1 T-carriers were further classified by ISX AA type and ISXG-carrier type, only the ISX AA type showed a positive correlation between buccal RTL and α-tocopherol intake (p<0.05). Further in vitro experiment is under way to validate a protective effect of α-tocopherol on the telomere shortening, where the role of hydrogen peroxide as oxidative stress may be modified by combined genotype of BCMO1 T-carrier and ISX AA homozygote in a human cell line.

Introduction: A number of founder and recurrent mutations are prevalent in the Ashkenazi Jewish population. Screening for these mutations has reduced the incidence of many autosomal recessive diseases. Yet, only 22 conditions are available on current Ashkenazi Jewish screening panels. The purpose of this study is to expand the knowledge base of pathogenic mutations in this population and develop an expanded screening panel.

Methods: Whole-genome sequencing was performed by Complete Genomics on 128 disease-free Ashkenazi Jews, who identified 13,768,157 variants. These variants were passed through a pipeline which scored the Online Mendelian Inheritance in Man (OMIM) and ClinVar databases. The pipeline yielded 201 variants that were called pathogenic in these databases. These variants were manually curated through literature review and scored for pathogenicity. The literature was also reviewed for known Ashkenazi Jewish mutations. The results were combined and weighted by clinical utility and allele frequency.

Results: In addition to the 22 conditions that are currently screened for, an additional 69 conditions that have pathogenic alleles were identified. In total, 168 significant pathogenic mutations were established for 52 autosomal recessive, 34 autosomal dominant and 5 X-linked disorders. The diseases include those that inform risk to offspring, which are generally autosomal recessive, and range from very severe, such as glycine encephalopathy and Leigh syndrome, to milder, such as hyperoxaluria and ichthyosis. Some of these disorders are informative for adult-onset conditions that may convey risk of future cancer (BRCA1/2 and mismatch repair testing), obesity, or macular degeneration or may have atypical presentations. Some mutations convey different phenotypes in the monoallelic and biallelic states; thus, testing would offer information pertaining to both personal risk as well as risk to offspring. Conclusions: As the number of identified conditions increases, for a larger number of mutations becomes possible, we need to reframe our goals of population screening and genetic counseling. An expanded panel can further reduce the incidence of autosomal recessive conditions, reduce morbidity and mortality by offering earlier diagnosis and support need for information and enable planning for future life events. The number of mutations identified could be feasibly screened for on a population-wide basis and are individually selectable following genetic counseling.


Technological advancements in next-generation sequencing have allowed whole exome sequencing to become feasible as a clinical test in cases of suspected genetic disease when initial tests have failed to yield a molecular diagnosis. Although many single case reports exist, there is limited data regarding molecular diagnostic success rates of exome sequencing in a medical genetics clinic setting. Initial estimates have suggested a diagnostic yield of 15-30% in select cases. Recent publications relevant to incidental findings (pathogenic variants in genes unrelated to the patient’s phenotype) have also shown that this test has a broad and complex potential impact. We quantified the diagnostic yield and incidental findings in exome sequencing cases in the University of Michigan Pediatric and Biochemical Genetics Clinics, from October 2012 to the present. Cases were identified in which clinical exome sequencing tests were sent by UM Pediatric Genetics faculty. Clinical notes were reviewed for phenotype information, previously-ordered studies, and family history. Research-based testing and next-generation sequencing-based focused gene panel testing were excluded. To date, results have returned for 39 tests. In 12 cases (31%), pathogenic variants in disease-causing genes related to the patients’ phenotypes were identified. New diagnoses included Bainbridge-Ropers syndrome (MIM 615485), Kabuki syndrome (MIM 147920), Myle syndrome (MIM 139210), Opitz-Kaveggia syndrome (MIM 305450), triosephosphate isomerase deficiency (MIM 615512), and Van Maldergem syndrome (MIM 615546). 23 of the 27 remaining cases had variants of uncertain significance (VUSs) reported. No patients or families opted out of receiving a report of incidental findings. 5 cases (13%) had incidental findings in one of the 56 genes recommended to be reported by the ACMG, with pathogenic variants identified in BRCA1, BRCA2, MYBPC3, and MYH7. Early results suggest that clinical exome sequencing is a promising new technology that can lead to a molecular diagnosis in a subset of cases, when applied properly, such testing may lead to a diagnosis more quickly and likely at less cost in both monetary and psychological terms. An understanding of the limitations of this testing and the possibility of incidental findings is necessary for its proper application.
Implications of Massively Parallel Sequencing in screening for Autosomal Recessive conditions: the risk of being a "genetic wallflower". L. Burnett, J.D. Chesher, L. Nguyen, P.M. Lew, A. Proos, L.C. Ding, L. Koe1. 1) NSW Health Pathology North, Sydney, NSW, Australia; 2) Sydney Medical School, University of Sydney, NSW, Australia; 3) School of Information Technologies, University of Sydney, NSW, Australia; 4) Department of Obstetrics and Gynaecology, QEI Institute Research Institute for Mothers and Infants, University of Sydney, NSW, Australia.

Introduction Genetic screening for autosomal recessive (AR) genetic carriers is available in many communities “at risk” due to high prevalence of pathogenic variants. Examples include Tay-Sachs disease, Althasmearia, and cystic fibrosis. The range of tests included in screening programs is broadening, and will accelerate with the introduction of Massively Parallel Sequencing. As more genes are included in screening programs, a risk emerges that a wallflower will be rejected by every analysis. We recently modelled the frequency of autosomal dominant (AD) genetic conditions arising as Incidental Findings (IF) in Whole Genome Sequencing. We found that the proportion of tested individuals with significant IFs plateaus to a limit even as the number of genes tested increases beyond those in the ACMG IF Recommendations.

Aim To model the rate of increase in AR genetic carriers with increasing numbers of AR genes in test menus. We also explore the likelihood of genetic wallflowers emerging.

Methods A mathematical model based on binomial distribution was used to predict the number of AR carriers based on prevalence of pathogenic variants and number of tested conditions. Monte Carlo simulation was used to calculate the probability of any two individuals in this screened population being carriers of exactly the same AR genetic conditions.

Results The model was validated using known rates of AR genetic carriers found in Australian Ashkenazi Jewish community screening programs. We have found that the number of AR genetic carriers will rise as the number of tested conditions is increased. We also report the likelihood of genetic wallflowers emerging in AR genetic carrier screening programs.

Conclusions As the number of AR conditions included in testing panels is increased, the number of AR genetic carriers identified in a population will increase. This behaviour is in marked contrast to that for AD conditions, where the proportion of a population with a reportable IF will plateau. These findings have significant implications for health economic evaluation and planning. Finally, we report on the ethical and social implications of genetic wallflowers emerging.

Analyses of TSC genes among Brazilian tuberous sclerosis complex (TSC) patients. L.G. Duher-Almeida1, J.P.G. Almeida2, L. Masulik1, M. Richardson1, M. Miranda2, S.A. Antoniou1, S. Rosenberg1, L.A. Haddad1. 1) Department of Genetics and Evolutionary Biology, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil; 2) Department of Pediatrics, Division of Child Neurology, Faculdade de Medicina da Santa Casa de São Paulo, São Paulo, Brazil; 3) Department of Pediatrics, Universidade Federal do Paraná, Curitiba, Brazil.

Tuberous sclerosis complex (TSC) is a multisystem disorder, with variable expression and autosomal dominant inheritance. Clinically it is due to hamartia and brain tumour development in different tissues, notably in the brain, kidneys, heart, skin and lungs, causing organ dysfunction. Mutations in either tumor suppressor genes TSC1 or TSC2 are responsible for TSC. TSC1/TSC2 protein heterodimer inhibits the mammalian target of rapamycin (mTOR) complex 1, controlling cell growth and proliferation. Although TSC diagnosis is basically clinical, since the 2012 specialist panel review, the finding of a TSC1 or TSC2 pathogenic mutation has been considered sufficient for the definite diagnosis of the disease. In addition, mTOR inhibitors have been clinically approved to treat under specific guidelines three hamartomas that afflict TSC patients. To set up the basis for molecular diagnosis of TSC and functional analysis of TSC1 and TSC2 mutations in São Paulo, Brazil, an ongoing project aims to establish and compare at long term Sanger sequencing and massively parallel technologies for TSC1 and TSC2 DNA sequencing. As a preliminary step, we conducted Sanger sequencing of the TSC1 gene encompassing its full coding sequence, an average of 132 bp of intronic segments next to exon boundaries, in addition to the gene core promoter of 28 Brazilian patients with definite TSC diagnosis. Seven patients (25%) displayed TSC1 nonsense/frameshift mutations. Among 31 other DNA variants identified, 27 were DNA polymorphisms. Two and one additional DNA point variants from the same patient flanked a putative, specific transcription factor binding site, 5' to the TSC1 core promoter. In addition, a novel DNA variant resulting in the TSC1 noncoding exon 2 was predicted to change the sequence potential to behave as a splicing enhancer. In summary, we describe 25% of TSC patients with pathogenic mutations in the TSC1 coding sequence. Moreover, our data disclose four novel DNA variants in TSC1 potentially regulatory regions that are likely to unravel novel pathogenic mutations, and thus need to be experimentally tested.

Conclusions The Implementing Genomics in Practice (IGNITE) Network: A Coordinated Effort to Study and Improve Implementation of Genomics in Clinical Practice. S.E. Kimmel1, G.S. Ginsburg2, E.P. Bottinger2, J.C. Denny3, D.A. Floaehart5, T.I. Pollin1, E.B. Madden2, K. Weitzel3, M. Alexander4, C. Horowitz2, L.A. Orlando2, N.S. Calman3, P.R. Dexter2, M. Levy4, A.R. Shuldiner6, H.A. Jenkins4, J.A. Johnson8, IGNITE Network. 1) Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA.; 2) United States; 2) Duke University Medical Center, Duke University, Durham, NC; 3) Icahn School of Medicine at Mount Sinai, Mount Sinai School of Medicine, New York, NY; 4) Vanderbilt University School of Medicine, Vanderbilt University, Nashville, TN; 5) Indiana University School of Medicine, Indiana University, Indianapolis, IN; 6) University of Maryland School of Medicine, University of Maryland, Baltimore, MD; 7) College of Pharmacy, University of Florida, Gainesville, FL; 8) National Human Genome Research Institute, National Institutes of Health, Bethesda, MD.

Purpose: Despite great progress in understanding the impact of genomics on patients’ health and outcomes from therapy, implementation of genomics in practice remains challenging. There is an urgent need to better understand the methods of incorporating patients’ genomic information into clinical care, evaluating implementation efforts, and measuring outcomes of genomics implementation. Methods: The aims of the Implementing Genomics in Practice (IGNITE) Network, funded by the National Human Genome Research Institute and the National Cancer Institute of the NIH, are to assess methods of implementation of genomic information in practice within specific projects, rigorously evaluate these methods across projects in order to create generalizable knowledge, and to provide proofs of concept of the value of genomics in medical care. Special interest has been placed on partnering with communities and laboratories to both influence and understand the practice of genomics in medical care. Results: IGNITE is a network of 48 projects and a Coordinating Center that have been funded to incorporate genomics into clinical practice (genetics, pharmacogenetics, and family history) into the EMR and provide CDS for implementation of appropriate interventions or clinical advice across multiple and diverse sites. IGNITE includes incorporation of structured family history tools, genotyping protocols for more accurate diagnosis and custom treatment and prophylaxis, standardized guidelines and somatic pharmacogenomics. With the assistance of a Coordinating Center, the projects are using established frameworks for implementation research to define barriers to implementation, develop and disseminate solutions to these barriers, and assess outcomes, especially those most important to ensuring rapid and sustained adoption, such as patient and provider satisfaction, clinical outcomes, and cost effectiveness. Summary: The IGNITE Network’s goal is to determine what works and what does not in genomic medicine implementation in practice and to provide generalizable knowledge that can be used to foster similar implementation in other settings, for other diseases, and for the myriad of potentially useful genetic-based testing and histories. Details of the IGNITE projects and the challenges and lessons learned from implementation can benefit the scientific and clinical community.
2567S
Clinical Whole Exome Sequence analysis of an 8-year-old Caucasian female affected with intellectual disability (ID), failure to thrive, myopia, microcephaly, seizures, dysmorphic features, and language regression with aphiogenic identified two compound heterozygous alterations in LINS, the human homolog of the Drosophila segmentation lis protein, which is involved in the Wnt signaling pathway. The paternally-inherited c.2020dupA frameshift alteration and maternally-inherited c.1394+1G>T splice site alteration are both expected to result in a deleterious effect on the LINS gene. The patient evaded diagnosis through clinical evaluation and extensive genetic testing over many years including negative chromosome analysis.

2568M
Genetics of beta thalassemia in Iran: Still requires consideration? N. Mahdieh1, B. Rabbani2, R. Shiri Heris3. 1) Deputy of Research and Technology, Ministry of Health and Medical Education, Tehran, Iran; 2) Medical Genetic Group, Qazvin University of Medical Sciences, Qazvin, Iran; 3) Faculty of Paramedicine, Ilam University of Medical Sciences, Ilam, Iran.
Beta thalassemia is observed in many countries. Prevalence of thalassemia-causing mutations is different among various parts of the world and even within different ethnic groups living in Iran. In addition to clinical, biochemical and hematologic diagnostic methods, molecular diagnostics is used for prenatal diagnosis of thalassemia. We analyzed all published data about the frequency of beta globin mutations among Iranian subjects. Mutation frequency was calculated in more than 5000 at risk couples studied in Iran. Ten following mutations were common: IVS I-5 (G>C), IVS II-1 (G>A), FSC36/37 (-T), IVS I-1 (-G), CD30 insertions were common: IVS I-5 (G>C), IVS II-1 (G>A), FSC36/37 (-T), IVS I-1 (-G).

2569T
Diagnostic Yield of Genetic Testing in the Children's Hospital of Colorado Autism Genetics Specialty Clinic. N.J. Meeks1,2, K. Brown1,2, B. Miller1,2. 1) Children's Hospital Colorado, Aurora, CO; 2) University of Colorado, Anschutz Medical Campus, Aurora, CO.
Background: Autism spectrum disorders (ASDs) are a group of neurodevelopmental disorders characterized by deficits in social communication, restricted interests, repetitive patterns of behavior and normal language development onset early in development. Several well-defined genetic disorders have been identified in patients with ASDs; however, clinical genetic testing is diagnostic in only a minority of cases. The purpose of this study was to determine the diagnostic value of genetic testing in the population of children with ASDs seen at the Autism Specialty Genetics Clinic, part of the Autism Treatment Network, at the Children's Hospital Colorado. Methods: A retrospective chart review was performed of patients seen in the Autism Genetics Clinic. Patients who met inclusion criteria had the following: an initial or follow-up genetics evaluation within the past six years; a diagnosis of an ASD through formal evaluation; and at least one molecular genetic test completed and resulted. Results: Of 130 charts reviewed, 107 patients met study criteria. The average number of molecular tests ordered per person was 3.02 (5% CI: 2.74-3.32). Eighteen patients (16.2%) had molecular genetic testing that was diagnostic. Copy number changes found on chromosomal microarray (CMA) accounted for 67% of all pathogenic findings with a yield of 13.9%. Other pathogenic findings included mutations in the PTEN, MECP2 and PTPN11 genes, positive FISH testing for a familial mutation, and two patients with abnormalities detected on karyotype. Biochemical screening for metabolic conditions, FMR1 testing, and multi-gene intellectual disability panels evaluating for more than one genetic condition did not yield any pathogenic findings. Of 73 published studies, almost every patient (95%) had a chromosomal microarray as part of their evaluation. These results continue to support the use of CMA as a first line evaluation in children with ASDs. Also, in 1 molecular or cytogenetic tests resulted in a variant of uncertain significance. Incorporating this information in clinic visits will improve pre-test genetic counseling in this population.

2570S
Whole exome sequencing (WES) provides a straightforward detection of rare variation including de novo mutations in parent-child trios, but a systematic atic interpretation of the diagnostic yield and an assessment of the costs associated with the testing is currently lacking. We implement this project to determine the increased diagnostic yield and provide a thorough cost analysis of the clinical care with and without the use of WES. This study will serve as a guide for centers intending to implement WES and provide a clear benefit versus the associated costs of obtaining this increased diagnostic yield. Twenty children of healthy, unrelated parents were randomly selected from the 2011 patient population at the Sylvia Toth Center (STC) in Utrecht, the Netherlands. The STC is a specialized center for children with complex neurological and other chronic disorders. DNA libraries were prepared using Kapa Biosystems reagents, enriched using Agilent SureSelect All exon V5 with a custom pooling protocol, and sequenced on the Illumina HiSeq 2500. A list of exonic variants was obtained by filtering against public and in-house database according to the expected de novo inheritance model as well as for recessive and compound heterozygote variants. In parallel, the clinical records of patients were obtained and a cost summary of medical treatments, hospital visits, care, and all other resource use was calculated per patient. This cost was then compared to the cost of care using WES, assessed retrospectively on each patient. Comparing the diagnosis and costs with and without the use of WES gives a clear picture of the economic feasibility of putting WES into standard diagnostic practice at the STC and similar centers. The diagnostic yield for the 13 patients sequenced thus far is 23%, confirming past studies diagnostic yields on intellectual disability cohorts. The three variants found are in genes recently associated with intellectual disability (ANKRD11, C7NBB1, ANPD), and the variants are all frame-shift deletions resulting in protein truncation. On average, patients were seen 13 times to the hospital, overnight stays and different diagnostics to unravel the genetic cause of their neurological disorder. It is therefore deemed very plausible that the total cost of the current diagnostic pathway is many times higher compared to WES, indicating the economic feasibility of implementing WES early in the diagnostic pathway.
2571M
Genetic investigation of cystic fibrosis transmembrane regulator mutations in a cohort of consecutive patients candidate for assisted reproductive techniques. F. Papa1, M. Rongioletti1, MB. Majolini1, C. Vaccarella1, I. Simonelli2, V. Mazzucc1, A. Luciano1, P. Pasqualetti2, G. Liumbruno1. 1) Clinical Pathology Department, San Giovanni Calibita Hospital, Rome, Italy; 2) Medical Statistics & Information Technology Fatebenefratelli Association for the Research, Rome, Italy.

Introduction. The present study, investigated the frequency of mutations in the CFTR gene, in a group of consecutive patients candidate for assisted reproductive techniques with the aim of identify subjects carriers of the most severe ones. Methods. 22,416 alleles were screened for 56 CFTR gene mutations utilizing the CFTR INNO-LiPA Results. CFTR mutations were detected in 6.2% of the screened alleles. In the large group of alleles analyzed 93.4% were wt, 4.4% were characterized by mild mutations, and 1.7% by severe or severe/mild mutations. Indeed, the most common severe mutation was JF508N observed in 192/22,416 (0.86%) of all alleles analyzed, followed by the N1303K mutation with the frequency of 36/22,416 (0.16%). Whereas regarding mild mutations, the most frequent was the 5T polymorphism present in 916/22,416 (4.1%). Conclusions. Our results analyzed, followed by the N1303K mutation with the frequency of 36/22,416 (0.16%).

2572T
Next-generation sequencing for the diagnosis of autism spectrum disorders using genomic capture targeting multiple candidate genes. L. Rodriguez-Revega1, 2, 3, M. Alvarez-Mora1, 2, I. Madrigal1, 2, 3, R. Calvo4, O. Puig5, M. M11a, 2, 7, 1) Biochemistry and Molecular Genetics Department, Hospital Clinic, Barcelona, Spain; 2) CIBER de Enfermedades Raras (CIBERER), Barcelona, Spain; 3) IDIBAPS (Institut d’Investigacions Biomèdiques August Pi i Sunyer), Barcelona, Spain; 4) Psychiatry and Child and Adolescent Psychology Service, Hospital Clinic, Barcelona, Spain.

Autistic spectrum disorders (ASD) is one of the most common neurodevelopmental disabilities, with an average estimated global prevalence of 62 cases per 10,000 children and an approximate 4:1 male to female ratio. Over the past few years, the genetic basis of ASDs has been pursued aggressively using all kind of high-throughput genomic analysis technologies: SNP-array, CGH-array, next generation sequencing based techniques and genome-wide association studies. Despite the progress in the identification of several candidate genes and causative genomic copy number variations (CNVs), the vast majority of ASD cases still remain unexplained. An obstacle for molecular diagnosis of ASD patients has been clinical and genetic heterogeneity of patient cohorts in combination with a recently pointed out multihit model of the disease. This model is based on the observation that most of the abnormalities identified have been associated with highly variable phenotypes and seems insufficient to cause ASDs on their own, supporting the hypothesis that CNVs contribute to ASDs in association with other CNVs or point variants located elsewhere in the genome. In this study, we developed a next-generation sequencing-based screening based on the capture of a panel of genes involved, or suspected to be involved in ASD, on pooling of indexed DNA and on paired-end sequencing in an Illumina MiSeq platform, followed by confirmation by Sanger sequencing. A cohort of 44 ASD patients with negative result for arrayCGH was screened to evaluate this strategy in terms of sensitivity, specificity, practicability and cost. In silico analysis was performed using DNAexes® software. Sequencing data provided, on average, 99.75% coverage of the 44 genes selected at more than 100-fold mean depth of coverage. Disease causing mutations were identified in 10 patients resulting in a molecular diagnosis rate of 22.7%. These results demonstrate the efficiency of NGS in performing molecular diagnosis of ASD. The failure of finding recurrent mutations highlights the genetic heterogeneity of ASD. Acknowledgments: This study was supported by Fundación Alicia Koplowitz (AKOPOWITZ11_006). The CIBER de Enfermedades Raras is an initiative of the ISCIII.

2574M
Fabry disease diagnosed through family screening. G. Sarca1, C. Dragonini2, E. Severini1. 1) Professor C.Angelescu Hospital, Bucharest, Romania; 2) Genetic Lab, Bucharest, Romania; 3) Carol Davila University of Medicine and Pharmacy, Bucharest, Romania.

Background. Early recognition of Fabry disease (FD) is difficult due to the clinical variability, the non-specific signs and symptoms and the rarity of the disease. In the same time, early recognition involves early initiation of enzyme replacement therapy (ERT) known to delay the time and prevent the disease progression. Our study describes the screening for FD using genetic testing after a diagnosis of Fabry disease in a male family member.

Objectives. To detect the disease causing mutation and to clarify the genetic status of asymptomatic family members. Subjects and methods. All 7 family members (both males and females) were genetically tested. Genetic testing included isolated DNA from blood samples and sequence analysis of all coding exons and all intron-exon boundaries of the GLA gene. Enzyme activity levels were evaluated too. Results. The disease causing mutation in the family was the pathogenic GLA mutation (c.485G>A) present in exon 3. Two males were found hemizygotes having one copy of GLA gene mutation and no enzyme alpha GAL activity. All five females were found carrier of the same mutation and had alpha-Gal A activity in the normal range. Younger male received early ERT and expressed a mild disease phenotype (no cardiac and renal involvement) comparing with his older male relative. Conclusion. Measurement of alpha-Gal A enzyme activity is unreliable for carrier detection. Early genetic testing should be considered in younger persons with a family history of Fabry’s disease.
2575T Translating allelic heterogeneity of GJB2 gene to clinical practice in Romanian population with congenital isolated hearing-loss. E. Savereid1, C. Dragomir2, A. Stanciu3, D.T. Stefanescu2, L. Savur1. 1) Carol Davila University of Medicine and Pharmacy, Bucharest, Romania; 2) Genetic Lab, Bucharest, Romania.

Background. Different alleles within the same gene can cause a similar variant phenotype. Previously published studies showed the allelic heterogeneity of GJB2 gene as main genetic cause of isolated congenital hearing-loss phenotype. The proportional distribution of the different mutations within GJB2 gene varies in different ethnic groups. Objectives. The aim of the present study was to provide a complete and updated spectrum of mutations in GJB2 gene and to identify the most prevalent mutations in Romanian population. Subjects and Methods. To overcome our aims, we used consecutive sample selection in order to obtain a better representation of entire Romanian population. All available persons with congenital hearing-loss were included based on the inclusion criteria (clinical records of congenital sensorineural hearing loss, non-syndromic hearing loss, no related findings, patients with affected siblings but unaffected parents). Testing protocols included ARMS-PCR and DNA sequencing techniques for detection of known mutations or identification of mutations within two genes associated with hearing loss, GJB2 and GJB6. Results. Most prevalent mutation was c.35delG (40.0%) in both homozygotic and heterozygotic forms. The second known mutations or identification of mutations within two genes associated with isolated congenital hearing-loss phenotype. The proportional distribution of the different mutations within GJB2 gene and to identify the most prevalent mutations in Romanian population. Current ties, University of Washington, Seattle, WA; 3) Department of Pediatrics, Herman and Walter Samuelson Children's Health Enterprises, Children's National Medical Center, Washington, DC; 4) Genetic Engineering and Biotechnology Institute, TUBITAK Marmara Research Center, Kocaeli, Turkey; 5) Department of Genetics, Afyon Kocatepe University, Afyon, Turkey; 6) Information Technologies Institute, TUBITAK Marmara Research Center, Kocaeli, Turkey; 7) Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX, USA; 8) Genetics and Biotechnology Institute, TUBITAK Marmara Research Center, Kocaeli, Turkey.

Methods. We conducted semi-structured interviews with 13 parents of children diagnosed with ASD, and 15 providers from multiple specialties involved in the care of patients with ASD to identify how parents and providers define the benefits of and barriers to genetic testing for ASD. Results: Although parents were open to testing, it was often a low priority compared to other aspects of their child's care. Some parents expressed confusion about the test and reasons for conducting it, but they were willing to have their child tested, based primarily on the provider's recommendation. The majority of parents, “all they could” for their affected child. Wanting a causal explanation for ASD or information for reproductive decision making were motivators for some parents, but for the majority the motivator was “crossing it off the list” of things tried. While some providers were supportive of testing, many providers were ambivalent. Providers often made the decision to offer testing on a case by case basis. Barriers to testing mentioned by both providers and families included difficulty in conducting the blood draw, concern that information might create guilt or blame, and the cost of the test, especially when insurance would not cover it. Additional concerns from providers were stigma, variants of uncertain clinical significance, incidental findings, the added burden to the family with possible waste of time and energy, and the fact that the testing is too complex for many families to understand. The biggest barrier for families was not being aware of it. Discussion: Currently, there is variability in provider attitude and practices toward genetic testing for ASD, as well as parental uncertainty about the value of testing. When providers recommend testing, parents generally follow their advice but they do so to ensure that they are providing all recommended care. Further discussion among stakeholders is needed to clarify the purpose of genetic testing and its appropriate use in ASD care.

2576S Where’s the Benefit? Views on Genetic Testing for ASD. K.B. Shutske1, L.K. Willig1, L.D. Smith2, J.B. LePichon3, N.A. Miller1, D.L. Dinwiddie4, G. Twist1, A. Noell1, B.A. Hesse1, L. Zeilmer1, A.M. Atherton1, A.T. Abdelmoty1, J.E. Petrikin1, N. Safina1, S.S. Nyp1, B. Zuccarel1i1, I.A. Larson1, A. Modrich1, S. Herli1, M. Creed1, Y. Zhaohui1, Y. Zuan3, R.A. Brodsky1. 1) Center for Pediatric Genomic Medicine, Department of Pediatrics, and Department of Pathology, Children's Mercy Hospital, Kansas City, KS 64108; 2) University of New Mexico Health Science Center, Albuquerque, NM 87131; 3) Department of Medicine, Johns Hopkins University, Baltimore, MD 21205.

Pediatric neurodevelopmental disorders refractory to traditional diagnosis: Diagnostic rate, cost and change-in-care of whole genome versus exome sequencing. S.F. Kingsmore1, S.E. Sokol2, J. Saunders3, E.G. Farrow1, L.K. Willig1, L.D. Smith2, J.B. LePichon3, N.A. Miller1, D.L. Dinwiddie4, G. Twist1, A. Noell1, B.A. Hesse1, L. Zeilmer1, A.M. Atherton1, A.T. Abdelmoty1, J.E. Petrikin1, N. Safina1, S.S. Nyp1, B. Zuccarel1i1, I.A. Larson1, A. Modrich1, S. Herli1, M. Creed1, Y. Zhaohui1, Y. Zuan3, R.A. Brodsky1. 1,2,3,4, H.K. Tabor1,2,3,4, W. Burke1,2, 1) Institute for Public Health Genetics, University of Washington, Seattle, WA; 2) Bioethics and Humanities, University of Washington, Seattle, WA; 3) Department of Pediatrics, University of Washington, Seattle, WA; 4) Treuman Katz Center for Pediatric Bioethics, Seattle Children's Research Institute, Seattle, WA. 2577M Pediatric neurodevelopmental disorders refractory to traditional diagnosis: Diagnostic rate, cost and change-in-care of whole genome versus exome sequencing. S.F. Kingsmore1, S.E. Sokol2, J. Saunders3, E.G. Farrow1, L.K. Willig1, L.D. Smith2, J.B. LePichon3, N.A. Miller1, D.L. Dinwiddie4, G. Twist1, A. Noell1, B.A. Hesse1, L. Zeilmer1, A.M. Atherton1, A.T. Abdelmoty1, J.E. Petrikin1, N. Safina1, S.S. Nyp1, B. Zuccarel1i1, I.A. Larson1, A. Modrich1, S. Herli1, M. Creed1, Y. Zhaohui1, Y. Zuan3, R.A. Brodsky1.
Whole exome sequencing (WES) is expected to revolutionize the practice of clinical genetics with a paradigm shift in how clinicians reach a diagnosis in patients with an unknown etiology. It holds tremendous potential thereby raising expectations of physicians and families to provide an etiologic diagno-

We report on our experience in a small subset of patients offered WES to identify an etiology. WES was preauthorized and offered to fifteen patients in the past 18 months since its clinical availability. All families were eager for testing and some initiated conversation about WES. A rigorous consent process followed regarding pros and cons of WES and the potential for incidental findings. All the patients had negative single gene molecular testing and SNP microarray. Results from twelve patients are presented (three pending). A pathogenic variant was identified in five patients (41% yield). Two have changes in an established disease gene: ATRX, associated with alpha thalassemia X-linked MR syndrome and ARID1A, associated with Coffin-Siris syndrome, although the phenotype was not evident clinically. In an adult with ataxia, a variant in the KCND3 gene confirmed a diagnosis of SCA19. One patient had a change in a newly described gene DYN1C1H1, allowing a better genotype-phenotype correlation. Although the initial report in another patient suggested no clear pathogenic changes, discussion between the clinician and laboratory personnel resulted in attributing clinical significance to a variant of unknown significance (VOUS), with diagnosis of a rare type of congenital disorder of glycosylation. In two patients a de novo change in novel genes was identified but lack of supporting data or functional studies prevents attribution of clinical significance. In the remaining six patients four have a negative result and two have a VOUS. This report demonstrates some of the challenges and benefits of WES in a clinical setting and offers a learning opportunity to enable wider utility. The yield for an etiology in our sample is high (41%), although far from 100% as expected by families. We have not had an incidental finding reported to date, but the families are open to receiving this information. Despite the paradigm shift brought about by WES with test results guiding the phenotypic delineation, we highlight the importance of a thorough clinical evaluation which guides interpretation of sequence variants of unknown significance.

Diagnostic exome sequencing establishes molecular diagnoses among patients with gastrointestinal disease. L. Shahrizaidi1, K.F. Gonzalez2, D. El-Khechen1, Z. Powis1, C. Alamillo1, C. Mroske1, D. Shinde1, K. Radtke1, R.M. Baxter1, B. Tippin-Davis1, S.K. Gandomi1, E.C. Chao1,2, S. Tang1. 1) Amyb Genetics, Aliso Viejo, CA, 92656; 2) Department of Pediatrics, University of California, Irvine, Irvine, CA 92697.

Diagnostic exome sequencing (DES) has proven to be an effective and efficient diagnostic tool for patients with undiagnosed genetic diseases. This test often indicated when a patient has a suspected genetic diagnosis and where prior genetic testing has been uninformative, when there are multiple genes/diseases in the differential, or when the phenotype is not consistent with a known genetic disorder. DES has been instrumental in providing a diagnosis for patients among a broad range of indications. The five most common indications for testing in our cohort have included patients with neurological, musculoskeletal, craniofacial, ophthalmologic, and gastrointestinal phenotypes. In a retrospective analysis of the first 500 reported patients who underwent DES at one laboratory, we identified 125 patients who presented with syndromic gastrointestinal (GI) disease. Among these patients, 8 presented with GI symptoms as the main feature, while GI issues were among other indications for testing in the remaining patients. Some of the most common GI symptoms included irritable bowel syndrome (IBS), gastroesophageal reflux disease, constipation, diarrhea, intestinal hypo-
Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. T. Žemotjev1,2, S. Köhler1, L. Mackenroth1, M. Jäger1, J. Hecht1, P. Krawitz1, L. Graul-Neumann1, S. Doelken1, N. Emke1, M. Spielberg1, N.C. Oeien1,2, M.R. Schweiger1, U. Krüger1, G. Frommer2, B. Fischer1, U. Kornak1, A. Ardeshradavani1, Y. Moreau1, S. E. Lewis2, M. Haendel1, D. Smedley1,2, D. Horn1, S. Mundlos1,2, P. N. Robinson1,2, 1) Institute for Medical Genetics and Human Genetics, University Clinics Charité, Berlin, Berlin, Germany; 2) Institute of Biorganic Chemistry, Polish Academy of Sciences, Poznan, Poland; 3) Max Planck Institute for Molecular Genetics, Innsbr. 63-73, 14195 Berlin, Germany; 4) Berlin Brandenburg Center for Regenerative Therapies (BCRT), Charité-Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany; 5) Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany; 6) Agilent Technologies, Hewlett-Packard-Strasse 8, 76337 Waldbronn, Germany; 7) Department of Electrical Engineering, STADISU Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium; 8) Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA; 9) University Library and Department of Medical Informatics and Epidemiology, Oregon Health & Sciences University, Portland, OR, USA; 10) Mouse Informatics group, Wellcome Trust Sanger Institute, Hinxton, UK.

Less than half of patients presenting with suspected genetic disease receive an etiological diagnosis. Though next-generation sequencing (NGS), and in particular whole exome sequencing (WES), has enabled an unprecedented acceleration in the pace of Mendelian disease gene discovery, extensive challenges remain to integrate NGS, bioinformatic, and clinical data into effective diagnostic workflows. Because only variants in the ~2800 Mendelian disease genes can be interpreted in a clinical context, we have established an approach that targets variants in these causative genes, which we will refer to as the disease-associated genome (DAG), developing a targeted enrichment DAG panel (7.1 Mb), for which we have identified potentially pathogenic variation in 15 of the 58 exome-sequenced probands. While cytogenetic and microarray-based platforms have been used extensively for DD/ID diagnostics, causal genetic variation remains undetermined in a large proportion of DD/ID-affected individuals. Therefore, as part of the Clinical Sequencing Exploratory Research Consortium, we are conducting whole exome sequencing in ~450 parent-offspring trios over a period of 4 years to establish a genetic diagnosis for unexplained cases of DD/ID. Primary genetic results (DD/ID causative) as well as secondary findings (pathogenic variation detected in parents) are being returned to probands and their parents in a clinic setting, and follow-up studies are underway to determine how genomic test results affect the well-being of participants. To date, we have enrolled and consented 99 parent-offspring trios (some with DD/ID-affected siblings), and have completed exome sequencing and analysis for 29 of these trios. Among children with no diagnosis after standard clinical testing (e.g., clinical microarrays), we have successfully identified DD/ID pathogenic variants in 6 probands (21% of our total) including mutations in TCF4, SCN2A, MECP2, ARX and GRIN2B. We have also found variants that, while not definitively causal, are plausible candidates; for example, compound heterozygosity for impactfull variants in PCNXL2 shared by affected siblings, and de novo variants in EBF3, a conserved transcription factor known to be important in brain development. Further, we have identified potentially pathogenic variation in 15 of the 58 exome-sequenced parents (26%). Secondary findings include variation in AYR1 (malignant hyperthermia susceptibility), RYR1 (arhymogenic right ventricular cardiomyopathy), as well as carrier status for HEXA (Tay-Sachs disease) and CFTR (cystic fibrosis). The results of our clinical sequencing project demonstrate, thus far, that large-scale DNA sequencing improves the rate of definitive diagnoses in DD/ID-affected children and thereby meets a valuable clinical need.
2584T
Autosomal Dominant Hypertrophic Cardiomyopathy (HCM) is an important modifier of the cardiomyopathy of Fabry Disease (FD): Implications for α-Galactosidase A replacement therapy. D. Doheny1, R. Desnick1, D. Macaya2, C. Antolik2, H. Rehm3, A. Alfare3, T. Caliss4, J. DaRe5, S. Garmani1, 1) Dept Gen & Genomic Sci, Icahn School of Medicine at Mount Sinai, New York, NY; 2) GeneDx, 207 Perry Pkwy, Gaithersburg, MD; 3) Harvard Medical School, Laboratory for Molecular Medicine, 65 Lansdowne Street, Cambridge, MA; 4) Transgenomic, 5 Science Park, New Haven, CT; 5) Department of Biochemistry and Molecular Biology, University of Massachusetts, 710 North Pleasant street, Amhurst, MA.
Hypertrophic Cardiomyopathy (HCM) is a common autosomal dominant disorder of the myocardium that affects ~1 in 500 individuals world-wide and is the most common cause of sudden cardiac death in individuals under 35 years. To date, variants in >20 genes have been identified that cause HCM, most encoding sarcomere proteins. The cardiac features of HCM are indistinguishable from the cardiomyopathy of FD due to α-galactosidase gene (GLA) mutations. Screening panels that sequence the most common HCM-causing genes also include GLA, detecting pathogenic (P) and likely pathogenic (LP) variants. We hypothesized that previously unrecognized FD patients, particularly those with the “later-onset” phenotype, would be detected by these panels. In 6486 HCM patients referred for panel sequencing, GLA variants were identified and stratified by sex and phenotype, including “classic” or “later-onset” phenotypes, benign variants (BV), and variants of unknown significance (VUS). The presence of HCM variants was also recorded. Of the 6486 HCM patients tested, 96 (1.5%) were identified with a GLA variant. Of these 96, 27% also had a P or LP HCM variant, most being in the MYBPC3 genes. When stratified by Fabry phenotype and sex: 1) 13 patients had the “classic” phenotype, but only one, a heterozygote, had a P/LP HCM variant, indicating that the cardiomyopathy in most was due to FD; 2) among the other Fabry phenotypes, a higher percentage of patients had a P/LP HCM variant suggesting that the cardiomyopathy was due primarily to HCM variants: among 32 patients with the “later-onset” phenotype, 25% had a HCM variant; among 43 with GLA BVs, 35% had a HCM variant and among 8 with GLA VUSs, 25% (both females) had a HCM variant. Of those with GLA BVs, 79% had the pseudodeficient D313Y lesion of which 38% had a HCM variant. These findings are the first to indicate that variants in other genes may modify or underlie the pathogenesis of the major phenotype. Three patients were only affected left ventricular hypertrophy or HCM should be tested for HCM-causing genes, and those with HCM found to have GLA mutations: 1) should be evaluated for α-Galactosidase A replacement therapy. D. Doheny1, R. Desnick1, D. Macaya2, C. Antolik2, H. Rehm3, A. Alfare3, T. Caliss4, J. DaRe5, S. Garmani1

2585S
Characterization of Malaysian children with Beckwith-Wiedemann syndrome and Silver-Russell syndrome using methylation specific - multiplex ligation-dependent probe amplification. M. Thong1, F. Thuraga2, R. Y.Y. Poh1, I. Tauful1, H.B. Chew2, G.S. Ch'ng1, W.T. Keng1, 1) Department of Paediatrics, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia; 2) Department of Biomedical Science, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia.
Introduction: Beckwith-Wiedemann syndrome (BWS [MIM 130650]) and Silver-Russell syndrome (SRS [MIM 180860]) cause gigantism and growth retardation, respectively. Importantly, overexpression of IGF2 and H19 whereas IC2 controls the expression of CDKN1C and KCNQ1OT1. Methodology: Blood samples were collected from 13 BWS and 18 SRS patients based on clinical features. Genomic DNA was extracted. The DNA of patients and controls were subjected to multiplex ligation-dependent probe amplification (MLPA) to detect the copy number and methylation status in imprinted genes of chromosome 11 for BWS and SRS. The resulting PCR products were sent for fragment analyses. Results and discussion: Out of 13 BWS patients, three patients had IC2 hypomethylation with normal IC1 methylation. The IC2 hypomethylation may be implicated in KCNQ1OT1 activation and CDKN1C inhibition, leading to the gigantism. Six patients had IC2 hypomethylation and IC1 hypermethylation simultaneously which indicated normal UPD at chromosome 11p15.5. One patient had both IC1 and IC2 hypomethylation and two patients had IC1 hypomethylation with normal IC2. One mutation was identified in CDKN1C at position 78 in one BWS patient. One had normal methylation at both IC's. Among the 18 SRS patients, five showed normal methylation at both IC's. One patient had IC1 hypermethylation and IC2 hypomethylation, one had IC1 hypo-methylation with IC2 hypermethylation. Hypomethylation of IC1 causes the biallelic expression of H19 and biallelic silencing of IGF2, resulting in growth restriction. However, one had normal IC1 and IC2 hypermethylation and one had normal IC1 with IC2 hypomethylation. Three had hypermethylation at both IC1 and IC2 and five had IC1 hypermethylation with normal IC2. None of our SRS samples showed UPD in chromosome 7. Conclusion: The patients in this study presented with gain and loss of methylation, duplication and deletion in both BWS and SRS, in different imprinted genes, emphasizing the heterogeneity of these two growth disorders syndromes.

2586M
An age-based categorical framework to guide informed decision-making about next generation sequencing results in newborn screening. L.V. Milko, J. D'Aniello, K. Foreman, C. Turcott, J. Booker, L. Boshe, M. Gucsvas-Calikoglu, A. Aylsworth, J. Muenzer, D. Frazier, B. Powell, M. Roche, N. Strange, N. Vora, C. Powell, J.S. Berg, University of North Carolina Chapel Hill, School of Medicine, Chapel Hill, NC.
A tenet of newborn screening (NBS) is the potential for early detection and prevention of diseases prior to clinical manifestation of symptoms. Although next generation sequencing (NGS) offers great promise in increasing the number of conditions that can be presymptomatically screened, it also presents significant ethical complexity for conditions where efficacious treatment may not be available, or where avoiding the “diagnostic odyssey” might be the only benefit. The NC NEXUS (North Carolina Newborn Exome Sequencing for Universal Screening) project is developing an “binning” classification framework to characterize different categories of information that parents might be interested in learning, and to study the psychosocial impacts of their choices. The framework is divided into four quadrants that balance a semi-quantitative scale of “clinical actionability” (a measure of the likelihood and severity of disease outcomes, and the efficacy and potential harms of interventions), against the age of onset of disease or age of implementation of preventative intervention. The core panel (“NGS-NBS”) is composed of clinically actionable genes that are implicated in childhood-onset disorders, including conditions on standard newborn screens. All parents of screened newborns will receive results of the core NGS-NBS panel. Given the lack of empirical data on how to present other types of genetic information, parents will be randomized into either a control arm receiving only results of the NGS-NBS panel, or an experimental arm tasked with deciding whether to learn additional information from three categories: 1. clinically actionable conditions with typical disease onset or interventions beginning in adulthood (eg. Lynch syndrome); 2. childhood-onset, non-medicinally actionable disorders with no known effective interventions (eg. Rett syndrome); and 3. carrier status for recessive disorders. Information about adult-onset disorders with no known interventions (eg. Rett syndrome); and 3. carrier status for recessive disorders. Information about adult-onset disorders with no known interventions (eg. Alzheimer disease) will not be returned. We hypothesize that classifying the types of results that parents may opt to learn into clear categories will assist physicians and parents in engaging in informed decision-making in the complex setting of genome-scale sequencing of healthy newborns. It will also be useful in understanding how and why parents choose whether to participate in sequencing of their child's DNA, what results they elect to receive, and the impacts of their decisions.
2587T

Epilepsy is a common neurological disorder with a lifetime incidence rate of 3%. Intellectual deficiency affects 1-3% of children. Microarray CGH detects a chromosomal abnormality in 10-15% of individuals with intellectual disability and epilepsy. In 8% of patients with epilepsy, genetic testing has been uninformative. Extensive metabolic study and brain imaging techniques lead to a diagnostic in respectively 5% and 0.8% of cases. There are many conditions that associates epilepsy and mental retardation, caused by mutations in more than 100 genes, making genetic diagnosis expensive and challenging in the absence of other congenital anomalies. With the recent advances in genetics, it is becoming easier to analyse panel of genes responsible for a condition. In order to simplify the genetic diagnosis in epilepsy and intellectual deficiency, we have gathered in one analysis 150 genes. The coding regions are amplified by customised Ampliseq protocol and we sequence the fragments on an Ion Proton®. Mutation calling is done with Life Technologies variant caller and Cartagenia. Here, we present the preliminary results. To date, we have tested more than 120 patients. The inclusion criteria to this study are intellectual deficiency or developmental delay and epilepsy. We have confirmed mutations in few genes: STXBP1, FOXP1, KCNQ2, SCN2A, RFT1, ATRX, GLDC, NRXN1, NLGN3, SYNGAP1, MAGI2, SLC9A6, KDM5C. Overall, we are expecting a causal mutation rate of 15-25%. We will present the clinical and metabolic data of patients with causal mutations. Moreover, other substitutions detected were already described in the literature in patients with specific epilepsy with or without intellectual disability and were sometimes inherited from a healthy parent (SCN1A, KCNQ3, SCN9A, SCN1B). They could act as modifier genes inducing epilepsy in patients affected by a second mutation responsible for the intellectual phenotype. We need more data to confirm this hypothesis. In conclusion, gene panel analysis could be a first line diagnostic test in patients with intellectual deficiency and epilepsy in view of low cost and high mutation pick up rate.

2588S
Should the ACMG expand the required reportable disorders or findings on Exome Sequencing? Reporting a recent experience, R.M. Zambrano, Y. Lacassie. Department of Pediatrics, LSHUHC, New Orleans, LA. and Children’s Hospital of New Orleans.

Exome and Genome sequencing are becoming the standard in the Genetic clinics for the evaluation of patients with rare disorders. These technologies are allowing the identification of the molecular bases of many syndromes and also uncover incidental or secondary findings that could be of medical value for the patient and their family. The ACMG has published a statement listing the recommendations for reporting incidental findings in the exome and genome sequencing. Based in a recent experience in the use of incidental clinical Exome data supported the clinical diagnosis of severe primary hyperparathyroidism in a neonate, we propose to include a list of disorders that can cause severe neonatal diseases and recommend that the laboratories performing these tests do report all the recessive genes in which both parents are heterozygotes. We report a 5-day-old Caucasian female born to a 29-year-old father and a 32-year-old consanguineous G5P2 mother. Their last pregnancy was detected early. Only complications included polyhydramnios and premature labor requiring bed rest for 8 weeks. The baby was born NSVD at 3747 WGA. The newborn was discharged home at DOL 2 but had to be re-admitted after 2 days due to hypocalcemia and poor feeding, weight loss, hypotonia and lethargy. Laboratory work-up showed severe hypercalcemia with a serum calcium level of 33.7 mg/dL. Repeatedly and DOL 2 and PTH was elevated at 867 pg/mL (NI 15-65 pg/mL). We suspected neonatal severe hyperparathyroidism (NSHPT) caused by a mutation in the CASR gene. As the stress of delivery had been relieved, we contacted the laboratory which quickly confirmed that both parents were heterozygous for a mutation in the CASR gene: c.206G>A (p.R96H). This allowed us to further support our clinical diagnosis of NSHPT. Sanger sequencing in the parents confirmed the previous result. The parents were counselled on the genetic heterozygosity for genes causing severe neonatal disease or disorders with available treatment is recommended. This would allow genetic counselling, prenatal testing, and the prompt therapy after birth or later.

2589M
Exploration of the benefit of risk-stratified colorectal cancer screening based on common genetic variants - current status and future potential. C. Neuburger, K. Chumka, K.M. Kuntz, Z.K. Stadler, R.W. Burt, M. Wiells, N. Calonge, D.T. Zallen, T.G. Ganiats, E.P. Whitlock, E. Webber, K.A.B. Goddard, N.B. Henrikson, M. van Ballegooijen, A.C.J.W. Janssens, A.G. Zauber, I. Lansdorp-Vogelaar. 1) Department of Public Health Sciences, Mayo Clinic, Rochester, MN, United States; 2) Department of Epidemiology, Emory University, Atlanta, GA, United States; 3) Department of Health Policy & Management, University of Minnesota, Minneapolis, MN, United States; 4) Memorial Sloan-Kettering Cancer Center, New York, NY, United States; 5) Department of Medicine, University of Utah, Salt Lake City, UT, United States; 6) Intermountain Healthcare, Salt Lake City, UT, United States; 7) Genomic Medicine Institute, Geisinger Health System, Danville, PA, United States; 8) The Colorado Trust, Denver, CO, United States; 9) Department of Science and Technology, Society in Science VirginiaTech, Blacksburg, VA, United States; 10) Department of Family and Preventive Medicine, University of California, San Diego, CA, United States; 11) Center for Health Research, Kaiser Permanente Northwest, Portland, OR, United States; 12) Group Health Research Institute, Seattle, WA, United States.

Background. Common genetic variants contribute to colorectal cancer (CRC) risk and can be used to stratify the population into CRC risk categories. However, the discriminatory performance of such risk-stratification algorithms is currently limited. In this study, we investigated the current and potential future benefits of using risk-stratified colonoscopy screening, based on common genetic variants, versus uniform colonoscopy screening at ages 50, 60 and 70. Methods. We used the MISCAN-Colon simulation model to determine cost-effective colonoscopy screening strategies for people with a relative risk (RR) for CRC of 0.1, 0.2, ..., 9.8, 9.9 and 10. The costs and effects of risk-stratified screening in the population were determined based on the current discriminatory performance of common genetic variants (area under the ROC curve (AUC) of approximately 0.6) compared to uniform screening at ages 50, 60 and 70 for all. Because it is expected that the discriminatory performance of risk-stratification based on common genetic variants will increase in the future, we also estimated costs and effects of risk-stratified screening based on hypothetical common genetic variants with higher levels of discriminatory performance (AUC of 0.65, 0.70, ..., 0.90). Results. With current discriminatory performance, the optimal colonoscopy screening strategy ranged from no screening for people with a RR of 0.9 or less to a CRC screening every 3 years for people with a RR of 5.4 - 10. Screening at ages 50, 60 and 70 was optimal for people with a RR between 0.9 and 1.3. This stratification resulted in 1% more life years gained than uniform screening (less than 1 life year per 1,000 individuals) for the same overall costs. With increased discriminatory performance, the gain in life years increased from almost 4% for an AUC level of 0.65 to more than 18% for an AUC level 0.90. Conclusions. Given the very modest discriminatory performance of common genetic variants in risk-stratification for CRC, the current benefits of risk-stratified CRC screening based on these variants are limited. New variant discoveries are needed to yield a substantial improvement in discriminatory performance, and are necessary for risk-stratified screening to become clinically significant.

2590T
HCV infection and interferon-based treatment induce p53 gene transcription in chronic hepatitis C patients. J. Nowak1, B. Siwiak-Koscielna2, E. Kaluzna1, J. Rembowski1, I. Mozer-Lisewska2, I. Berezynska1, J. Wysocka-Leszczyńska, A. Kowała-Plaskowska1, J. Wysoki1, D. Januszewska-Lewandowska1,2,3, K.A.B. Goddard4, S. Kundu5, T.G. Ganiats1,2, R.W. Burt1, M. Wiells1, N. Calonge1, D.T. Zallen1, T.G. Ganiats5, E.P. Whitlock1, E. Webber1, K.A.B. Goddard1, N.B. Henrikson12, M. van Ballegooijen1, A.C.J.W. Janssens1, A.G. Zauber1, I. Lansdorp-Vogelaar1. 1) Inst Human Gen, Polish Academy Sci, Poznan, Poland; 2) University of Medical Sciences, Poznan, Poland; 3) Department of Medical Diagnostics, Poznan, Poland; 4) Department of Science and Technology in Society VirginiaTech, Blacksburg, VA, United States; 5) Group Health Research Institute, Seattle, WA, United States.

Common genetic variants contribute to colorectal cancer (CRC) risk and can be used to stratify the population into CRC risk categories. However, the discriminatory performance of such risk-stratification algorithms is currently limited. In this study, we investigated the current and potential future benefits of using risk-stratified colonoscopy screening, based on common genetic variants, versus uniform colonoscopy screening at ages 50, 60 and 70. Methods. We used the MISCAN-Colon simulation model to determine cost-effective colonoscopy screening strategies for people with a relative risk (RR) for CRC of 0.1, 0.2, ..., 9.8, 9.9 and 10. The costs and effects of risk-stratified screening in the population were determined based on the current discriminatory performance of common genetic variants (area under the ROC curve (AUC) of approximately 0.6) compared to uniform screening at ages 50, 60 and 70 for all. Because it is expected that the discriminatory performance of risk-stratification based on common genetic variants will increase in the future, we also estimated costs and effects of risk-stratified screening based on hypothetical common genetic variants with higher levels of discriminatory performance (AUC of 0.65, 0.70, ..., 0.90). Results. With current discriminatory performance, the optimal colonoscopy screening strategy ranged from no screening for people with a RR of 0.9 or less to a CRC screening every 3 years for people with a RR of 5.4 - 10. Screening at ages 50, 60 and 70 was optimal for people with a RR between 0.9 and 1.3. This stratification resulted in 1% more life years gained than uniform screening (less than 1 life year per 1,000 individuals) for the same overall costs. With increased discriminatory performance, the gain in life years increased from almost 4% for an AUC level of 0.65 to more than 18% for an AUC level 0.90. Conclusions. Given the very modest discriminatory performance of common genetic variants in risk-stratification for CRC, the current benefits of risk-stratified CRC screening based on these variants are limited. New variant discoveries are needed to yield a substantial improvement in discriminatory performance, and are necessary for risk-stratified screening to become clinically significant.
2591S
Assisted reproductive treatment is not a risk factor for chromosomal abnormalities in spontaneous abortion. s. Shim, Dept of Ob and Gyn, CHA Gangnam Medical Center, seoul, South Korea, MD.

Background: The causes of miscarriage include fetal and maternal factors. More than half first-trimester pregnancy losses are attributed to fetal chromosomal abnormalities, in which autosomal trisomy is known to be the most frequent cause (about 60%). Maternal risk factors for first-trimester pregnancy losses are maternal obesity, old age, and prior history of pregnancy loss. Recently, ART is more frequently utilized with an increase in infertility rate. However, there is a paucity of data about the frequency of embryonic chromosomal abnormalities in first-trimester miscarriage after assisted reproductive treatment (ART). We examined whether or not embryonic chromosome aberration in first-trimester pregnancy losses is more frequent in pregnant women with ART than in those with natural pregnancy. Methods: We conducted a retrospective study in CHA Medical Center. Study population consisted of 511 pregnant women (natural pregnancy group 3131% 160/511) vs. ART pregnancy group 68.49% <350/511) who underwent either spontaneous abortion or surgical dilatation and curettage between 2011 and 2012 with the diagnosis of missed abortion during first trimester (gestational age <14wks). We compared the frequency of embryonic chromosome aberration according to natural pregnancy group or ART pregnancy group, BMI >25 or BMI <25, the presence or absence of prior history of pregnancy loss, and maternal age >35 or <35. Results: The overall rate of aneuploidy was 51.6%, and mean maternal age 35.0 years. There was no significant difference in abnormal karyotypes between ART pregnant group and natural pregnancy group (P=0.154). However, abnormal karyotypes were significantly more frequent in patients with maternal age >35 years than those with maternal age <35 years, in patients with BMI >25 than in those with BMI <25, and in pregnant women with prior history of pregnancy loss than in those without prior history of pregnancy loss (each for P<0.05). Conclusions: ART treatment is not a risk factor for chromosomal abnormalities in first trimester miscarriage. However, maternal age more than 35 years, maternal obesity (BMI >25) and prior history of pregnancy loss are associated with abnormal karyotyping in first trimester miscarriage.

2592M
A workflow based information system infrastructure to support translational science: The NIH Undiagnosed Diseases Program experience. A.E. Links1, D. Draper1, E. Lee1, V. Lebedev2, M. Didenko3, D. Adams1, M. Brudno1, S. Dumitru1, M. Gridea1, W.P. Bone2, B. Coessens3, S. Verhoeven1, C.F. Boekeloo1, W.A. Gahl1, M. Sincan4. 1) Undiagnosed Diseases Program, National Human Genome Research Institute, Bethesda, MD; 2) RURO Inc., Frederick, MD; 3) Department of Computer Science, University of Toronto, Ontario, Canada; 4) Cartagenia NV, Leuven, Belgium.

The Undiagnosed Diseases Program (UDP) is a translational medicine initiative focused on bridging the gap between the patient bedside and the research bench. The application of translational medicine presents two major challenges: “translational research,” or bench-to-bedside integration, and conduction of complex and multidisciplinary research through collaborative research networks. Moreover, the UDP deals with many independent cases conduction of complex and multidisciplinary research through collaborative research networks. In an effort to overcome these two challenges and manage the complexities of the UDP, clinical and research processes were mapped and used for development of a Translational Research Information System (TRIS) known as the Undiagnosed Diseases Program Integrated Collaboration Server (UDPICS). The UDPICS ecosystem includes a workflow based translational research information management server, an ontology based phenotyping application, a biosample management system, an animal tracking system, a cloud-based genomic analysis platform, and an electronic laboratory notebook. It contains a complicated data structure with over 80 active workflows and over 2000 discrete data elements for efficient tracking of research projects and clinical services. In addition, the UDP has designed a collaborator instance of UDPICS to allow external collaborators to work with non-personally identifiable information (non-PII) specific to the individuals involved in the collaboration, upload reports for easier feedback to the clinical staff, and engage in recorded discussions with members of the UDP team. In conclusion, the UDP has developed a system that is capable of supporting both translational research and clinical services while providing a collaboration platform to allow the research consortia to tackle the complex and diverse scientific problems posed by the many different undiagnosed diseases.

2593T
Assessment of the variant annotation interpretive gap among major variant databases. M. Lee, B. Vecchio-Pagan, G.R. Cutting. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

The interpretive gap is a growing divide between variants which have been identified and those which have been annotated. This gap is a significant obstacle when seeking to assign disease liability to variants detected by exome sequencing and when interpreting sequencing results in diagnostic laboratories. The breadth of the interpretive gap may be assessed by evaluating the consistency and depth of data across traditional disease-focused as well as “normal” variation databases. We surveyed five databases: OMIM (curates disease phenotypes), HGMD (curates disease variants from the literature), ClinVar (inventories variants from user submissions), LOVD (inventories variants from user submissions), and The 1000 Genomes Project (inventories variants detected in “normal” individuals). There were 2780 genes curated in OMIM, 4733 in HGMD, 18645 in ClinVar, 20172 in LOVD, and 17406 in 1000 Genomes. Across the inventory databases (ClinVar, LOVD, and 1000 Genomes), 67.7% of genes were present in all three; however, only 10.4% of genes were found in all five. To assess variant depth, 2179 genes annotated in all five databases were extracted. Variant counts reported for each gene differed by as much as an order of magnitude per gene. To determine consensus and quality of variant annotation, we selected the 30 disease-implicated genes with the highest number of variants in OMIM and extracted variant data for each gene from OMIM, HGMD, ClinVar, and 1000 Genomes. MAF data from dbSNP and 1000 Genomes revealed that 60.0±34.2% (range: 17.1-99.1%) of variants reported in the 30 genes had a dbSNP rsID, and 26.8±45.9% (range: 0.114-95.0%) had a HGMD MAF, indicating that high quality feature data is missing for a substantial fraction of variants in genes with strong disease associations. Remarkably, only 41 (0.43%) of 9383 unique variants are annotated by all four databases. Furthermore, variant type (e.g. SVN, deletion, insertion, indel) distributions of the 30 genes in HGMD, ClinVar, and 1000 Genomes were strikingly different among the curated variant types. A suspected over-representation of SVNs was seen, and represented 72.6% of variants in HGMD, 82.6% in ClinVar, and 98.5% of variants in 1000 Genomes. These findings highlight that the interpretive gap is multidimensional, as there is an absence of consistency regarding the number, composition, and associated feature data of variants annotated in widely used variant databases.
Clinical whole exome sequencing (WES) production update at Baylor Whole Genome Laboratory (WGL): Improved procedures for faster TAT and better disease gene coverage. Y. Ding1, S. Matakas1, C.J. Buhay1, M. Wang1, N. Veeraraghavan1, T. Chiang1, A.C. Hawes1, W. Liu1, N. Saada1, J. Ma1, L.K. Dolores-Freiberg2, J. Chandaranana3, C.J. Qiu2, R. Najjar4, M.N. Bainbridge1, Y. Han1, H. Dinh1, J.V. Korchina1, Q. Wang1, E. Boerwinkle5, J.R. Lupski5, C.E. P Ion6, A.L. Beaudet5, C.M. Eng7, R.A. Gibbs1, D.M. Muzny2, Y. Yang7. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 3) Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX. Baylor WGL has been providing clinical WES services since October 2011. To date, approximately 3800 WES cases including 3650 germline WES, 150 cancer WES as well as ~3100 mitochondrial genomes have been sequenced on the Illumina HiSeq2500 or HiSeqX1000. The current sample volume is approximately 200 samples per month. NGS sequencing of mitochondrial genomes has been included in the clinical WES test starting with samples received on 10/15/2012. The mitochondrial genome was amplified by a single long-range PCR before subjecting to paired-end library constructions. Mitochondrial and WES libraries from the same sample were barcoded and sequenced in the same lane (1:30) on Illumina’s HiSeq platform. This process eliminates virtually all sequencing cost for the mitochondrial genome. Paired-End library constructions for WES and mitochondrial genome were automated using Beckman Biomek NXP/Span-8 robots, and show high throughput (up to 400 samples per month), data quality and reproducibility. The enrichment for exome sequences utilized a solution based whole exome capture probe set, VCRome 2.1, which covers approximately 200,000 coding exons (35 Mb). An average of >10 Gb data were generated for each WES and 95% of the targeted exome regions were sequenced at a depth of ≥20X (average coverage >140X). The mitochondrial genome was sequenced at >20,000X to facilitate the detection of large deletions. Further, we have validated and implemented additional procedures initially developed at the Baylor Human Genome Sequencing Center (HGSC) in order to achieve better sequencing coverage and shorter turnaround time (TAT) (Muzny et al. 2014 ASHG abstract). First, by applying the “spike-in” method (see text below); second, the enrichment procedure reduces the probe capture time from 72 to 24 hours. Further validation of additional quick WES (QWES) procedures developed at HGSC will enable the completion of the wet lab exome sequencing process within 5 working days at WGL. Improved clinical WES protocols and automation have enabled Baylor WGL to produce higher quality WES data with better disease gene coverage and shorter TAT and prepared the lab for more challenging services such as WES for prenatal or neonatal ICU samples.
The rapid pace at which new gene-disease associations are reported for monogenic disorders poses a tremendous challenge to the clinical interpretation of genome-scale sequencing data in patients and research participants. The NIH-funded ClinGen consortium is working with the NCBI to generate a publicly available resource for delineation of clinically relevant genes and variants. As part of this effort, a workgroup was assembled to develop standardized procedures for curating genes and their relationship to monogenic disorders, in order to provide well curated, up-to-date information for use in research and clinical analyses. The workgroup first addressed the problem of categorizing the clinical validity of gene-disease associations according to strength of evidence. The workgroup established seven categorical designations for strength of evidence: Definitive, Strong, Moderate, Limited, No Reported Evidence, Disputed, and Evidence Against. These categories are defined according to the type and strength of evidence provided in the published literature or other public sources, the presence or absence of contradictory evidence, and whether the initial report of disease-gene causality has been replicated.

This framework is initially being applied to genes relevant to three clinical domains: hereditary cancer, cardiology, and inborn errors of metabolism.
Public perceptions of disease actionability and severity and their potential utility for making decisions about Genomic Testing Results.  

P.S. Sinicrope¹, K.D. Graves², Y. Zhou², J.B. McCormick³, S.T. Vadaparampilli⁴, N.M. Lindor¹. ¹) Health Sciences Research, Mayo Clinic, Scottsdale, AZ; ²) Lombardi Comprehensive Cancer Center, Georgetown University, Washington D.C.; ³) Divisions of General Internal Medicine and Health Care Policy & Research, Biomedical Ethics Program, Rochester MN; ⁴) Moffitt Cancer Center, Tampa, FL.

Background: While much research highlights disease actionability and severity as important factors informing the return of genomic results, few studies actually address public perceptions of these concepts. We examined perceptions of disease actionability and severity and their relationship to clinical, attitudinal, and sociodemographic characteristics. We also explored individuals’ perceptions of the utility of disease actionability and severity in making hypothetical decisions about general genomic testing results.

Methods: We conducted a cross-sectional online survey with a representative sample of US adults using the YouGov opt-in panel. Our survey was informed by focus groups, the multi-disciplinary investigative team, and cognitive interviews. The survey contained 136 items and 3 videos covering the basics of genetics and disease actionability and severity. Sections included: 1) Overview and Motivations for Genetic Testing; 2) Perceptions of Severity; 3) Perceptions of Actionability; 4) Perceptions of Severity/Actionability combined; 5) You do the Binning; and 6) Assessing the Binning Process/Demographics.

Results: Respondents included 900 adults (51% Female, 65% White, 12% Black, and 15% Hispanic) with an average age of 45 (SD=17). About 64% reported knowing "something, but not very much" about genetic testing and 60% preferred to make their own decisions about genomic results they would want to learn about vs. letting medical experts decide. The majority (>85%) found the concepts of actionability and severity to be useful individually and combined; 46.6% indicated a preference for using actionability over severity. Over half (53.8%) reported being very/extremely confident in their ability to score for actionability and severity. Participants’ scoring of actual medical scenarios varied for both actionability and severity (p<0.0001). Conclusions: We found that actionability and severity are perceived quite variably, raising the question of how meaningful current binning efforts are for patients. Despite the challenges of scoring disorders, respondents indicated confidence in their ability to score, to make decisions for themselves versus expert decisions, and to receive results of all types, with only a minority want to be more selective. These results may inform development of patient-centered approaches to decision making in genomic testing.
HUMAN COMMUNICATION DISORDERS IN PATIENTS WITH DOWN SYNDROME.- A. Musso1, A. Unnigar1, K. Foley1, D. Mercer1, M. Marbie2, F. Tsien3; 1) Genetics, LSUHSC New Orleans, New Orleans, LA, USA; 2) Genetics, Children's Hospital of New Orleans, New Orleans, LA, USA.

According to the Centers for Disease Control and Prevention, an estimated 75% of patients with Down syndrome suffer from some form of hearing loss. The hearing loss can be conductive hearing loss (CHL), sensorineural hearing loss (SNHL), or mixed hearing loss, which is a combination of CHL and SNHL. CHL in Down syndrome patients is usually due to otitis media with effusion which is managed medically with antibiotics and pressure equalization (PE) tubes. Causes of SNHL in the Down syndrome population are congenital inner ear abnormalities including internal auditory canal hypoplasia, enlarged vestibular aqueducts, cochlear nerve hypoplasia, and others. Due to the irreversible nature of SNHL, hearing aids and speech therapy are the optimal intervention for patients with Down syndrome. The incidence of SNHL in Down syndrome children ranges from 4% to 55%. We are conducting a retrospective study of at least 200 patients evaluated in the Down syndrome clinic at Children's Hospital of New Orleans and their satellite clinics throughout the state of Louisiana. Based on audiograms conducted a retrospective study of at least 200 patients evaluated in the Down syndrome clinic at Children's Hospital of New Orleans and their satellite clinics throughout the state of Louisiana. Based on audiograms of the hearing loss, we will discern whether the hearing loss is conductive, sensorineural, or mixed. We will evaluate the following standard audiological procedures: otoscopy, tympanometry, otocoustic emissions, and pure tone audiometry. Hearing loss type and severity will be determined by comparing air conduction and bone conduction pure tone thresholds. This study will provide further data on the incidence of hearing loss in children with Down syndrome. Geneticists can play an important role in enhancing the communication outcomes of Down syndrome patients through referral for early hearing and speech intervention.

HUMAN COMMUNICATION DISORDERS IN PATIENTS WITH DOWN SYNDROME.- A. Romero-Diaz1, J.M. Aparicio-Rodriguez1; 2) Phoniatrics, Otoloneurology; 2) Genetics, Hospital para el Nino Poblan, Puebla, Mexico.

Down syndrome is one of the most frequent chromosomal abnormalities live births. Most patients with a trisomy of a portion or all of the chromosome. Its prevalence is approximately 1/800 live births. The human communication disorders are common in patients with Down syndrome. Morphological abnormalities of the head and neck frequently in patients with Down syndrome as midface hypoplasia, characterized by malformation of the eustachian tube, short palate, macroGLOSSIA, and narrowing of the oropharynx and nasopharynx, joined as systemic factors associated with muscular hypoplasia, results in a high incidence of recurrent otitis media and sleep apnea syndrome in these patients. HEARING.- The prevalence of hearing loss in children with Down syndrome is high, especially the transmission or conductive hearing loss since between 50-70% of cases have chronic serous otitis media between 3 and 5 years; the neuro-sensory deafness represents only 4%. LANGUAGE DEVELOPMENT.- Language development is usually delayed. This delay correlates with cognitive ability; Most oral communication, which may be more or less significant defects language is obtained; and in some cases can not develop oral language and require augmentative and / or alternative communications. CHANGES IN THE VOICE.- Typically have a high palatal vault, large tongue, generalized hypotonia and maxillary hypoplasia. No skeletal abnormalities affecting craniofacial features. Although his brain is structurally normal can be observed brachycephaly and a flattened shape in the center of the face due to a malformation of the frontal bone. DOWN SYNDROME AND STUTTERING.- Have certain degrees of learning difficulties that can affect the ability to understand and produce speech and language. One of the biggest problems for some people with Down syndrome is the unintelligibility of their speech toward others. Unintelligibility and disfluency often go hand in hand. MEMRY.- Children with Down syndrome usually show memory skills short than visual-temporal verbal memory. This means that they will learn more easily if the information presented to them visually if they are presented verbally. ATTENTION.- In Down syndrome there are alterations in the brain mechanisms involved when changing focus of attention. So often have difficulty sustaining attention for extended periods of time and ease of distraction from diverse and novel stimuli.
2604M
AMMECR1 gene disruption and expression impairment in a balanced X-autosome translocation patient. M.M. Oliveira, R.S. Guilherme, V.A. Melo, N. Kosyakova, T. Leitão, G. Carvalheiro, M.I. Melaragno. 1) Departamento de Morfologia e Genética, Universidade Federal de São Paulo, São Paulo, Brazil; 2) Instituto de Human Genetics, Jena University Hospital, Friedrich Schiller University, Jena, Germany.

Detailed molecular characterization of a woman with balanced X-autosome translocations has proven to be valuable in searching candidate genes for diseases. Patients with balanced chromosomal rearrangements involving an X-chromosome generally present a skewed X inactivation, with the normal X-chromosome being preferentially inactivated. Thus, a gene disruption in X-chromosome frequently results in absence of a functional copy of the affected gene. Here we report a seven year old female patient who presented with karyotype 46,X,t(X;9)(q23,q12)dn, disproportionate short stature, microsomia, minor facial dysmorphisms, septal atrial defect with thin aneurysmal septal tissue, scoliosis, diffuse bone demineralization with peripheral osteopenia, recurrent sinusitis, and normal cognition. Replication banding showed a preferential inactivation of the normal X-chromosome in lymphocytes, and genomic array indicated for a balanced translocation. The autosomal breakpoint affects a heterochromatic region. X-chromosome breakpoint mapping was performed by array painting technique using a glassneelld/ microdissection derived DNA probe, exclusively of the derivative chromosomes. Array painting determined the disruption of AMMECR1 gene at intron 1, within the genomic coordinates chrX:109,509,444-109,552,503 (GRCh37/hg19). The breakpoint localization was validated by FISH using the BAC probe CTD-3066N24. AMMECR1 mRNA whole blood expression was assessed by qRT-PCR in the patient and eight female controls. Two TaqMan assays were used and both showed absence of AMMECR1 expression in the patient and high gene expression levels in all controls. Therefore, AMMECR1 gene function impairment may be related to her clinical abnormalities. Although AMMECR1 gene function is unknown, its expression is ubiquitously in human cells and tissues, with exception of the nervous system. AMMECR1 gene expression occurs primarily in organs and cell types that were affected during the patient's development. AMMECR1 protein is highly conserved throughout evolution, presents nuclear localization, several phosphorylation sites and similar conformation as proteins that interact with nucleic acid, suggesting a regulatory role. This patient may represent a unique model for the evaluation of the phenotypic consequences of AMMECR1 impairment in humans and the claim of this gene as a candidate gene. Some variants of unknown clinical significance (VOUS) were also described by FISH. Genomic arrays identified 13 potentially pathogenic copy number variations (CNV): 7 deletions (4p16.3p15.33, 4q13.3q21.1, 8q13.3, 10q26.2q26.3, 16p13.3, 22q11.21, Xp22.33) and 6 duplications (2q32.1, 4p16.1.1, 16p13.11, 17q12.1, Xp22.33). The 4p16.1 duplication (~950 kb) was observed in two non-related patients. Two patients showed overlapping CNVs in Xp22.33 (~3.2 Mb deletion and ~9 Mb duplication). None of these CNVs were described in the Database of Genomic Variants (DGV), except the duplication 4p16.1 that was possibly relevant because involves HMX1 gene. Some CNVs were associated with OAVS phenotype, such as BAPX1 in 4p15.33, responsible for hemifacial microsomia; HMX1 in 4p16.1, involved in the oculoauricular syndrome; EYA1 in 8q13.3, important for the craniofacial development; YPEL1 and ERK1 in 22q11.2, involved in branchial arch development. Besides, larger deletions, duplications and VOUS found in our patients may represent new candidate regions for the phenotype. These candidate genes and/or regions will deserve further investigation. The identification of the OAVS genes and genomic position are important to understand better the molecular pathways and clinical heterogeneity of the syndrome. Financial support: FAPESP, Brazil.

2605T
GENETIC BASIS OF AGING, TELOMERES AND TELOMERASE, FOUR PEDIATRIC PATIENTS WITH PREMATURE AGING OR COCKAYNE'S SYNDROME. M. Barrientos-Perez, J.M. Aparicio-Rodriguez, 1) Endocrinology; 2) Genetics, Hosp para el Niño Poblano, Puebla; 3) Estomatología Benemérita Universidad Autónoma de Puebla, México.

There has been a direct relation between Telomeres, telomerase and aging due to multiple investigations that are carried out on them, so it has been associated with human youth and the cure of cancer. There is a direct relationship in the alteration of the Telomeres and Telomerase with stress and inbreeding. With respect to consanguinity there is a syndrome of premature aging called Cockayne syndrome. Four cases of cockayne syndrome (CS) from two families with first degree consanguinity living in a small town where radioactive uranium (UR) mines were found. The clinical manifestation of CS were analyzed in the Hospital Para el Niño Poblano (HNP) by a multidisciplinary study. The main clinical manifestations were dermatological alterations as photosensitivity to sunlight and predisposition to skin cancer “xerodermia pigmentosum” (XP); endocrinological alterations as dwarfism and senile appearance; ophthalmological findings from cataracts to pigmentary retinal degeneration; neurological alterations as mental retardation and senorial hearing loss; and upper limbs, lower limbs and vertebral column degeneration was also found. Four cases with Cockayne Syndrome were diagnosed in two Mexican families with consanguinity from an area with radioactive contaminants, showing that consanguinity is the etiological factor of this syndrome. Keywords. Consanguinity, autosomal recessive inheritance, Cockayne’s syndrome, Telomeres, telomerase, uranium, xerodermia Pigmentoisus.
**2607M**

A rare combination: mosaic Turner syndrome by isochromosome Xq with 17p13.3p13.2 microduplication and congenital cataract with autosomal dominant inheritance and incomplete penetrance in the same individual. J.A. Rojas Martínez, J.C. Acosta Guio, J.A. Rojas Martínez, J.C. Acosta Guio, 1) Instituto de Genética Humana, Pontificia Universidad Javeriana, Bogotá, Colombia; 2) Instituto de Ortopedia Infantil Roosevelt, Bogotá, Cundinamarca, Colombia; 3) Instituto de Investigación en Nutrición, Genética y Metabolismo, Universidad El Bosque, Bogotá, Cundinamarca, Colombia.

The combination of Turner syndrome with congenital cataract with autosomal dominant inheritance and incomplete penetrance has not been reported, since even this form of cataract in isolated presentation is exceptionally rare. There are also no reports presentations with rearrangements on chromosome 17. Next, we present the case of a twenty month old girl born in Bogotá (Colombia), product of first pregnancy maternal, non-consanguineous parents. She was diagnosed a month living with bilateral congenital cataract and also has a history of neurodevelopmental delay and background in the family of mother, grandmother and maternal uncle with bilateral congenital cataract. Physical examination found multiple minor facial anomalies, and other abnormalities. Neurological examination with generalized hypotonia with truncal predominance, and scarce and unclear language. Suspecting chromosomal abnormalities, diagnostic study begins with high resolution B-bridge karyotype reporting: mos 46,X,iso[Xq]82/45,X[18], with the presence of a mosaic given by major cell line with a isochromosome of the long arm of the X chromosome and another cell line with complete monosomy of the X chromosome. Comparative Genomic Hybridization (CGH) was performed, confirming the previous cytogenetic analysis. Additionally, a 274 Kb duplication was identified in 17p13.3p13.2 (with breakpoints in 3561131-3835801) that compromises 5 genes. Congenital cataract with clear autosomal dominant inheritance present in this patient exhibits incomplete penetrance, as the mother of the patient must have the mutated allele, because she transmits to her progeny, but is not expressed in her phenotype. Furthermore 8 genes were identified in the region corresponding to the duplicate 17p13.2p13.3 segment in this patient. One of these genes, GSG2 encodes a protein called haspina whose overexpression has been shown to be associated with genomic instability and aneuploidy, so the authors suggest that when absent can result in lesions of the excretory system and thyroid dysplasia.

**2608T**

Polycystic kidney disease and multiple malformations in a patient with tetrasomy 2q13q21.1. L. Dupuis, L. Dupuis, C. Roadhouse, R. Badilla-Forrás, D.J. Stavropoulos, R. Mendoza-Londono, 1) Clinical Genetics, Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; 2) Department of Paediatric Laboratory Medicine, The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada.

Autosomal tetrasomies are rare events which typically cause a severe clinical phenotype. Well characterized autosomal tetrasomies include cat-eye syndrome (tetrasomy 22), Pallister Killian (tetrasomy 12p) and tetrasomies of 9p and 18q. We describe a case of chromosome 2q13-q21.1 triplication, resulting in tetrasomy for this segment. This is the first report of a liveborn with a triplication of chromosome 2 of this magnitude (19.876 Mb).

The patient presented at day two of life with dysmorphic features, bilateral hip dislocation, overlapping toes and bilateral pulmonary stenosis. Abdominal ultrasound revealed dilated pyelus uteris, large polycystic ovaries and dysplastic cystic kidneys that led to end stage renal disease and kidney transplant. At her most recent evaluation at the age of four years, medical issues included global delay, lung hypoplasia, bilateral hip dislocation, G-tube dependence, strabismus and mild sensorineural hearing loss. MRI of the brain revealed appropriate myelination and prominent lateral and third ventricles but no evidence of hydrocephalus. Microarray analysis (44,000 oligo) showed that the area of tetrasomy involved 115 RefSeq genes and 13 OMIM Morbid genes. A thorough review of the literature reveals two previous cases of tetrasomy of chromosome 2q that partially overlapped the area involved in our patient. Mercer et al (2009) reported an adult female with a 7.28Mb triplication of 2q12.3-q13 with infertility and had left polycystic kidney, patent ductus arteriosus, left streak ovary, bicornuate uterus and deafness. Wang et al (2013) reported a study of tetrasomy 2q12-q21 with brain malformations, multicystic kidneys, absence of the right thumb and cleft lip and palate. The area of overlap between these three cases comprises a 3Mb region on 2q13 (chr2:111,158,401-114,416,131 hg18) that harbors four OMIM genes. These include MERTK a tyrosine kinase that when absent can result in retinitis pigmentosa, IL1B and IL1RN which participate in the immune system, and the transcription factor PAX8 that when mutated can result in lesions of the excretory system and thyroid dysplasia.

The case presented narrows the critical region for renal cyst and genital malformations and may be of great use in the diagnosis, treatment and counseling of 48, XXXX patients and families.
2610M


Introduction: Sex chromosome aneuploidies are the most common chromosomal abnormalities with an incidence of 1 in 400 births. Tetrasomy and pentasomy X, XXY, XYY, and XXXY, are a few examples. 48,XXYY is a condition with an incidence of 1:18,000-1:100,000 male births. 48,XXYY syndrome is the most common of these three occurring in 1:18,000-1:40,000 male births and is related to a wide spectrum of physical findings, congenital malformations, cognitive impairment, and behavioral problems. Case Presentation: This is a 3-year-old boy male patient who was born by cesarean section from non-consanguineous parents after a first full term uneventful gestation. Maternal and paternal ages were 29 y/o. No significant family history. He started to have myoclonic, atebile seizures at 4 months old. EEG showed generalized discharges. He was started on anticonvulsants with complete remission of seizures. At 2 and a half y/o aggressive behavior and motor stereotypes were noticed. Clinical examination revealed global developmental delay with playgaghypathy, hyperactivity, severe truncal hypotonia with head lag presentation. MRI spectroscopy revealed symmetric signal abnormality in thalamus, basal ganglia and brainstem. Urine organic acids were normal; pyruvate, plasma aminoacids and Acylcarnitine profile were normal. Mutations or deletions were not detected in mitochondrial DNA screen or in FOXX1/MEEC2 sequence analysis. Karyotype showed 48,XXYY chromosome abnormality. Discussion: This is a male patient with 48,XXYY syndrome who has physical and neurodevelopmental features consistent with it. Common findings include cardiac defects, dental problems, cinctacty and hypogonadism. The first two were evident in our patient as well as hypotonia, which occurs in 75% of patients. Congenital malformations like cleft palate, heart and kidney defects were ruled out in this patient. About 10-15% of individuals have seizure disorders, like this child. Elevated rates of autism (28-45%) and mood disorders (46.8%) and attention deficit and hyperactivity disorder (70%) have been seen. Therefore this could explain patients’ aggressive behavior and repetitive conducts. Conclusions: 48,XXYY syndrome has a wide spectrum of phenotypic presentation. It is important to identify the variability in physical features and cognitive functions to establish an early diagnosis, initiate management and to offer appropriate genetic counseling and therapies for rehabilitation.

2611T

Clinical features of Sp13 duplication syndrome. T. Kuchikata, S. Itoh, H. Yoshishahi. Division of Medical Genetics, Tokyo Metropolitan Child Medical Center, Tokyo, Japan.

Sp13 duplication syndrome (OMIM #613174) is newly recognized chromosomal microduplication syndrome characterized by developmental delay and intellectual disability, facial dysmorphism. NIPBL in the duplicated region is known as a part of causative gene. Loss-of-function mutations in NIPBL cause cohesin abnormality, which leads to chromosome segregation abnormality during the mitotic and meiotic cell cycles. A part of cohesinopathy causes Cornelia de Lange syndrome (CdLS). Here, we describe a Japanese case with Sp13 duplication syndrome. Subject: 4y-male. He was delivered at term with a complicated pregnancy due to maternal preeclampsia, for a birth weight of 2.1kg. His unrelated parents and siblings are healthy. He had a delay in psychomotor development with age. At the age of 3 months old, brain MRI revealed there was no definitive abnormality and myelination was not retarded. At the age of 2 year old, the weight was 11.2kg (25th percentile) and the height was 83.3cm (25th percentile). He had mild bilateral auditory disorder, strabismus, right cryptochidism, left accessory ear and dysmorphic features which are prominent forehead, downslating palpebral fissure, hypoplastic, broad nose bridge, and long slender fingers. The gestalt was suggestive of Kabuki syndrome. KMT2D and KDM6A mutations were not detected. Radiological survey for skeletal dysplasia revealed no definitive abnormality. A peripheral karyotype and their chromosomal microarray analyses were normal.

2612S

Klinefelter Syndrome (48, XXXY) in a Patient With Mild Mental Retardation and Psychotic Personality Traits. H. Pachajona1,2, M.F. Hernandez1. 1) Centro de Investigaciones en Anomalías Congénitas y Enfermedades Rasas (CIACER), Universidad Icesi, Cali, Colombia; 2) Fundación Clínica Valle del Lili, Cali, Colombia.

Introduction: Klinefelter syndrome is the most common aneuploidy in males with a prevalence of 1-0.2% in the general population, which rises up to 3% in males with fertility issues, although only 35% of cases are diagnosed. The affected males tend to be tall, have narrow shoulders, wide hips, sparse body hair, gynecomastia and small testis; they present androgen deficiency and azoospermia. Besides the previous physical characteristics, there have been reports on the expression of behavioral and cognitive traits that tend to be very variable, possibly according to the type of aneuploidy, with an established association between the number of extra X chromosomes and cognitive deficit, and a not so clear association of different chromosomal variants of Klinefelter and psychotic behavior, with some authors proposing the origin of the extra chromosome as a determinant of different behavioral traits. Case report: 13 year old male with karyotype 48,XXXXY, who besides presenting the classical physical features, is under psychiatric treatment because of presenting trouble at home and at school for being aggressive and impulsive. A review on the literature is made. Conclusion: When approaching a patient with Klinefelter syndrome is necessary to acknowledge the importance of an optimal triad of hormonal therapy and a multidisciplinary approach including endocrinology, pediatrics, genetics, neuro-psychology and psychiatry when needed.

2613M

Maternal Uniparental Dismy Prader-Willi Syndrome in an XXY boy. P. Phowthongkum1, J. Berkowitz2, A. Schneider3. 1) Internal Medicine, Einstein Medical Center Philadelphia, Philadelphia, PA; 2) Genetics Division, Einstein Medical Center Philadelphia, Philadelphia, PA.

Prader-Willi Syndrome (PWS) is a neurodevelopmental disorder characterized by neonatal hypotonia, failure to thrive, intellectual disability, hyperphagia, short stature, and characteristic behaviors. PWS is associated with chromosome 15q11-13 deletions, maternal UPD or an imprinting abnormality. 47,XY is a sex chromosome disorder which is usually subtle. We report long term follow up of a ten year old boy with PWS and 47,XY. He was born at term by emergency C-section for fetal distress to a 27-year old G1P0 mother and 30-year old father. At delivery he was hypotonic with poor gag reflex. Examination revealed cryptorchidism and subsequent feeding problems for a birth weight of 1.9kg. His unrelated parents and siblings are healthy. He had a delay in psychomotor development with age. At the age of 3 months old, brain MRI revealed there was no definitive abnormality and myelination was not retarded. At the age of 2 year old, the weight was 11.2kg (25th percentile) and the height was 83.3cm (25th percentile). He had mild bilateral auditory disorder, strabismus, right cryptochidism, left accessory ear and dysmorphic features which are prominent forehead, downslating palpebral fissure, hypoplastic, broad nose bridge, and long slender fingers. The gestalt was suggestive of Kabuki syndrome. KMT2D and KDM6A mutations were not detected. Radiological survey for skeletal dysplasia revealed no definitive abnormality. A peripheral karyotype and their chromosomal microarray analyses were normal.

[Copyright © 2014 The American Society of Human Genetics. All rights reserved.]
2614T
The Extended Phenotypic Spectrum of 7p14.3-15.3 deletion syndrome. M. Michelsson-Kernan1,2, D. Lev1,2, C. Vinkler1,2, L. Blumkin1,2. 1) Inst Med Genetics, Wolfson Medical Ctr, Holon, Israel; 2) Metabolic Neurogenetic clinic, Wolfson Medical Ctr, Holon, Israel; 3) Pediatric Neurology Unit, Wolfson Medical Ctr,Holon, Israel.

Structural aberration of chromosomes are associated with various syndromes. Microdeletion of the short arm of chromosome 7p is rare. Phenotypic presentation is variable and include craniofacial malformations, hypoplastic ears, heart defects, hypotonia, short fingers with tapering phalanges and severe developmental delay. Most of the cases are de novo deletions. Few cases are due to unbalanced rearrangements. We present a twenty months-old girl with severe developmental delay, dysmorphic features, hypoplastic ear crusies, tapering fingers, hypotonia, severe ataxia and congenital alacrima. Karyotype in leucocytes was normal, 46 XX. Affymetrix Cytogenetics Whole-Genome HD Array chip revealed a 10375.473 KBP loss in chromosome 7, between p14.3 and p15.3. This region harbors 96 known genes, including HOXA gene cluster and several cytogenetic regions associated with genetic disorders. Parental studies are normal. Although most of phenotypic features were previously described in patients with 7p microdeletion, our patient presents with distinct features including severe congenital alacrima and severe ataxia. These features have not been described previously.

2615S
Pigmentary mosaicism type Ito in a balanced X-autosome translocation with no gene disruption at the breakpoint. M. Melaragno1, R.S. Guilherme1, C.E. Steiner1, G.M. Carvalheiro1, N. Kosyakova1, T. Liehr2, M.M. Oliveira1. 1) Disciplina de Genetica, Universidade Federal de Sao Paulo, Sao Paulo, Brazil; 2) Department of Medical Genetics, Universidade Estadual de Campinas, Sao Paulo, Brazil; 3) Institute of Human Genetics, Jena University Hospital, Friedrich Schiller University, Jena, Germany.

Pigmentary mosaicism type Ito (PMI) is a heterogeneous symptom complex characterized by hypopigmented whorls and streaks forming Blaschko’s lines, with neurological deficits, epilepsy and/or asymmetric abnormalities in other organs. Although its etiology is not fully understood, PMI is typically associated to apparently balanced X-autosome translocations involving Xp11, especially the band Xp11.2. Several hypotheses have been proposed to explain this association, such as disruption of X-linked genes, effect position, mosaic functional Xp disomy and functional autosomal monosity due to a random X-inactivation pattern. We report the first case of PMI and balanced X-autosome translocation involving Xp11 that was submitted to a high-resolution breakpoint definition. An 18-year-old girl presented with karyotype 46,X,t(X;21)(p11.22;p11.1), Blaschko linear hypopigmentation in the trunk, neurodevelopmental delay, and electroencephalographic abnormalities. The normal X-chromosome was preferentially inactivated in lymphocytes, genomic array indicated for a balanced translocation, and the autosomal breakpoint affects a heterochromatic region. X-chromosome breakpoint mapping was performed by array painting technique using a glassneedle microdissection derived DNA probe, exclusively of the derivative chromosomes. Array painting determined the Xp11.22 breakpoint as intergenic, within the genomic coordinates chrX:51,760,718-51,776,770 (GRCh37/hg19). The breakpoint localization was validated by FISH using the BAC probe RP11-1129M7. There are no breakpoint neighboring genes that could be a candidate for pigmentary dysplasia. Although we observed a skewed X inactivation pattern in lymphocytes, the hypopigmented skin may present a random X-inactivation pattern causing a functional Xp disomy. Therefore, the X-inactivation pattern study only in lymphocytes is not sufficient to assess the pathogenic mechanism leading to PMI in patients with rearrangements involving X-chromosome. Most descriptions of balanced X-autosome translocation associated with PMI date back to the 80’s and 90’s, when the rearrangement investigation was limited to low resolution techniques and interpretation of a rearrangement was dependent on classic cytogenetic methods based upon a broad chromosome region. The report of a patient with PMI and balanced X-autosome translocation involving Xp11.22 with an intergenic breakpoint reduces the strength of the hypothesis of an X-linked loci disruption. Financial support: FAPESP.
2618S
Idic(15) syndrome: clinical studies of 32 new individuals. A. Battaglia, T. Filipi. Dev Neurosci., Stefania Maria nst/Univ Pisa, Pisa, Italy.
Chromosome region 15q11q13, known for its instability, is highly susceptible to clinically relevant genomic rearrangements, such as supernumerary marker chromosomes formed by the inverted duplication of proximal chromosome 15. Inv dup(15) results in tetrasomy 15p and partial tetrasomy 15q. The large ones, containing the Prader-Willi/Angelman syndrome critical region (PWS/ASCR), are responsible for the inv dup(15) or idic(15) syndrome. Diagnosis is achieved by standard cytogenetics and FISH analysis, using probes both from proximal chromosome 15 and from the PWS/ASCR.
Microsatellite analysis on parental DNA or methylation analysis on the probe DNA, are also needed in order to detect the parent-of-origin of the idic(15) chromosome. Array CGH has been shown to provide a powerful approach to detect the duplication and its extent. Differential diagnosis concerns the possible occurrence of double supernumerary isodicentric chromosomes derived from 15, resulting in partial hexasomy of the maternally inherited PWS/ASCR. Large idic(15) are nearly always sporadic. Antenatal diagnosis is possible. Here, we report on 32 new idic(15) patients; all sporadic, 1/32 had a double idic(15). Phenotypic features were quite variable, and 80% presented with a distinct "autistic-like" profile, providing a behavioral signature for ASD arising from the proximal 15q. Epilepsy occurred in 90%, and in only 20% could be controlled by AEDs. The interictal EEG showed slow/sharp waves, and/or bifaxic spikes-polyspikes, spikes-wave complexes, and an excess of fast activity over both fronto-temporal areas. Intellectual disability was severe-profound in 85%, with expressive language limited to dyslalic sounds/single words. Structural CNS defects were seen in 30%. Early central hypotonia was present in all. Follow-up ranged from 2 to 24 years.

2619M
Large distal duplication of chromosome 22q. D. Ortiz, L. Karger, M. Babcock, L. Edelmann, A. Babu, L. Mehta. Dept. of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY.
About 25 cases of live-born trisomy 22 have been reported in the literature with characteristic dysmorphic features and major organ anomalies typically resulting in demise soon after birth (Heinrich et al, 2012). Other microduplications of chromosome 22 include Emanuel syndrome [der(22)(11;22)], cat-eye syndrome (tri- or tetrasomy 22q11.2) and 22q11.2 microduplication syndrome. We describe a live-born female infant with a large 27.3 Mb duplication of chromosome 22q11.23q13.33 resulting in trisomy 22q11.23-pter. During the pregnancy there was persistant asymmetric IUGR with severe microcephaly (<3rd percentile) and oligohydramnios. Fetal echocardiogram was normal. Amniocentesis was offered and declined. The patient was born at 36.3 weeks gestation and had dysmorphic features, poor tone and poor respiratory effort. Additional features included a secundal atrial septal defect, bilateral hip dislocation, and tracheomalacia. No renal or brain anomalies were found. Karyotype revealed a derivative chromosome 22q11.23qter. During the pregnancy there was persistent asymmetry of normalized measurements supports the future use of ratios of distances extracted from images. Next, facial images of 16 persons (4 UPD, 12 deletion) from the same cohort were evaluated with automatic image analysis tools for UPD/deletion differentiation. This type of analysis provided good accuracy of differentiation between UPD and deletion individuals. Since UPD is the most prevalent etiology and has a subtly different phenotype, there is a risk of under diagnosis. In a future study we plan to test this hypothesis using blinded dysmorphologist assessment and automatic facial analysis tools. We would assess whether this risk could be reduced by training automatic systems to identify the individual etiologies.

2621S
Wiring the Brain in Down Syndrome: Unique Identical Twins Discordant for DS. J.R. Korenberg, T. O’Donnell, J.O. Edgin, C. Vacher, J. Anderson, G. Gerig. 1) Center for Integrated Neuroscience and Human Behavior, Department of Pediatrics, Brain Institute, University of Utah, Salt Lake City, UT; 2) School of Computing, University of Utah, Salt Lake City, UT; 3) Department of Psychiatry, University of Arizona, Tucson, AZ; 4) Department of Neuroradiology, University of Utah, Salt Lake City, UT.
Genetic disorders as Down syndrome (DS) provide a solution to the challenge of human neuroscience, to develop a unified view of the brain across structural, functional, and temporal levels integrated with properties such as cognition. However, subtle brain and cognitive phenotypes are obscured by genome-wide variation. We report a multimodal analysis including genetics, cognition, structural and functional MRI of the only living pair of identical male twins, age 22, discordant for DS. These resulted from an extra chromosome 21 in an otherwise identical genome, with the remainder of variance contributed by environmental and stochastic effects. We established MZ identity by Illumina genomic arrays, and established typical clinical and facial features. 1/2 of DS in one, >96% trisomy 21 in cord blood, blood and fibroblasts, and in the other a normalized phenotype, >92% normal karyotype in these tissues. The twins completed the Arizona Cognitive Test Battery for DS, an assessment including neuropsychological measures of hippocampal and prefrontal function. The 8% mosaic twin had a full-scale IQ of 123, typical performance on a frontal measure (CANTAB IDE), but impaired spatial working and associative memory (CANTAB Spatial Span and Paired Associates Learning (PAL)), markers of hippocampal dysfunction. Early language delay was noted; at baseline testing he presented with a 24 point split in verbal and performance IQ. The twins showed scores typical for the DS range (full scale IQ=42), with prominent impairments in (PAL). The twins were also scanned on high resolution MRI/DTI and compared with 22 typical DS and 17 controls. Overall, normalized volumes for the 96% mosaic twin in the right hippocampus to the 96% mosaic twin in the left hippocampus were similar to controls. An exception was DS-like volume in the 8% mosaic twin of the right fusiform gyrus, involved in visual memory and language. Integration of MRI with cognition in full-normal, revealed correlation hippocampal volumes with IQ results. Only the 96% mosaic twin demonstrated increased internetwork synchrony across networks seen in typical DS during functional MRI. The twins are a graphic vision of how a simply trisomy for 21 disturbs brain wiring and the results implicate the hippocampal and fusiform gyrus formation in the neural systems responsible for visual-spatial (PAL) and linguistic deficits of DS.
2622M
Congenital Myasthenia Syndrome: Uniparental disomy of chromosome 2 and homozygous mutation of GFPT1. S. Rangasamy1,2, R. Richholt1,2, K.M. Rainier1,2, S. Sniader1,2, J.J. Corneveaux1,2, J. Schuh5,2, R.B. Krat1,2, A. Kourdoglu1,2, M. De Both1,2, S. Szelinger2,1, B.E. Hijm1,2, S. Swaminathan1,2, M. Huentelman1,2, D. Craig1,2, V. Narayanan1,2.

1) Neurogenomics Division, Translational Genomics Research Institute, Phoenix AZ, USA; 2) Dorrance Center for Rare Childhood Disorders, Translational Genomics Research Institute, Phoenix AZ, USA; 3) Department of Medical Genetics, University of Antwerp, Antwerp, Belgium; 4) University of Arizona College of Medicine, Tucson, AZ; 5) Molecular and Cellular Biology Interdisciplinary Graduate Program, School of Life Sciences, Arizona State University, Tempe AZ, USA.

Congenital myasthenia syndrome (CMS) includes a group of genetic disorders involving genes encoding components of the neuromuscular junction. We describe an unusual case of a child with CMS who was found to have long stretches of homozygosity on chromosome 2. The female child was born full term but required emergency C-section, followed by prolonged intubation. She was later diagnosed with hypotonia, weakness with reduced spontaneous movement, and was areflexic. Genetic testing for spinal muscular atrophy, Prader-Willi syndrome, and congenital myotonic dystrophy were negative. Muscle biopsy study at 1 month of age showed normal fiber type size and grouping, but there was severe staining of the T-tubule system suggestive of a mitochondrial disease with reduced cytochrome C oxidase levels. EMG showed an electrodierecmenal response and she had a partial response to treatment with pyostimine. Gene testing for SURF1, SCO1, SCO2, COX10 and COX11 were negative and the chromosomal microarray test showed normal gene dosage. Chromosomal microarray identified a long (57.6 Mb and 90 Mb) contiguous region of homozygosity on chromosome 2: arr 2p21q11.2 (45,013,724 - 102,641,201) hmz; 2q22.1q36.3 (140,091,291 - 229,908,245) hmz. The Congenital Myasthenia Group at the Mayo Clinic confirmed a CMS diagnosis. At the time, CMS gene panel sequencing was negative. Family trio based whole-exome sequencing (WES) was performed at the Dorrance Center for Rare Childhood Disorders as a part of research study. Exome data identified a homozygous, frameshift, c.865-866insG insertion on chr2, at exon 9 of GFPT1 gene, a low coverage region by WGS. Insertion was confirmed by Sanger sequencing which demonstrated that father was homozygous reference and mother was heterozygous. Based on our findings, we propose that this child had two identical (maternal) copies of large segments of chromosome 2, implying maternal uniparental disomy, harboring the GFPT1 insertion that causes autosomal recessive congenital myasthenia. Previously reported patients with recessive GFPT1 mutation showed reduced GFPT1 levels and decreased protein glycosylation. Glycosylation at neuromuscular junctions is important for key intracellular signaling protein function and compromised glycosylation in GFPT1 mutation leads to altered neuromuscular phenotype. The precise pathogenic mechanism for our observed GFPT1 insertion will be determined with future studies.

2623T
CRANEOFACILAE AND FEATURES ASPECTS ORAL OF CHILDREN WITH THE POBLANO CHILD HOSPITAL GOLDENHAR SYNDROME IN THE PERIOD 2013-2015. D. Reyes Ramirez1,2, R. Aparicio1,2, G. Muñoz Quintana1,3, G. Mendoza1,4, J. Fleischer1,2, A. Zepeda1,2, N. Ramírez1,2, H. Pesqueira1,2, P. Ochoa1,2, P. Soto1,2, P. Quezada1,2, J. Gómez1,2, R. Hinojosa1,2, P. López1,2.

1) Pediatric Estomatology, Benemerita Universidad de Puebla, Puebla, Puebla; 2) Genetics, Hospital para el niño Poblano, Mexico.

The facial skull malformations and chromosomal aberrations are considered alterations in the phenotypic and secondary structural error of the birth. All these changes must be observed during childhood and confirmed by chromosomal karyotype or deficiency in DNA repair. Pediatric patients were evaluated and showed different congenital and chromosomal alterations. These changes at the level of its multiple system and organic structure, chromosomal analysis were analyzed and classified as neuromuscular and organ type of syndrome respectively. Another group of genetic disorders known as mutations and they are inherited in different generations. A wide range of pediatric patients with congenital and genetic diseases by alterations mainly associated with craniofacial development during pregnancy, prenatals, or at birth are described in this study to analyze characteristics clinical, medical, or surgical procedures and medical evolution according to the syndrome of malformation in study. The wide variety of Craniofacial anomalies often makes them unclassifiable. Goldenhar, also known as the first Syndrome and is second son of Arch or oculo-auriculo-vertebral spectrum, is a complex of Craniofacial and vertebral anomalies. These findings can be found in solitary or usually associated with microtia, mandibular hypoplasia, or congenital vertebral malformations. The incidence is limited and varies between 1 in 45,000 to 1 in 100,000 inhabitants. Currently it should be considered as a bilateral malformation, which would differentiate it from the Morsosima Hemifacial. Know the relationship of oral diseases in children with Goldenhar Syndrome, as well as craniofacial features will give us a broad overview in order to improve the orthodontic care of this patient. The earlier and excellent medical and surgical treatment is important to improve their life quality.

2624S

1) Pathology, Montefiore medical center, Bronx, NY; 2) Genetics, Children Hospital at Montefiore, Bronx, NY.

A 284kb microduplication at 7q21.11 involving SEMA3A possibly linked to agenesis of corpus callosum and visual impairment in a patient with spastic quadriparietal cerebral palsy. D. Ma1, N. Singh-Bhatia2, P. Levy3, R. Marion2, R. Naeem1.

1) Neurogenomics Division, Translational Genomics Research Institute, Phoenix AZ, USA; 2) Dorrance Center for Rare Childhood Disorders, Translational Genomics Research Institute, Phoenix AZ, USA; 3) Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO.

Background: Contactin 4, a cell adhesion molecule and a member of the immunoglobulin super family, is expressed primarily in the nervous system. Deletions and duplications encompassing the CNTN4 gene have been reported in children with autistic spectrum disorder and developmental delay. Intriguingly, all previously documented cases were inherited form unaffected fathers raising questions regarding reduced penetrance and pathogenicity of these rearrangements. Methods: Clinical and molecular cytogenetic characterization studies were performed on a 2-year-old boy who presented with severe speech apraxia, mild gross motor delay and absence of autistic features. The findings in the proband are compared with previously reported cases. Results: The proband is non-verbal, but displays age appropriate receptive speech skills, cooperative play, joint attention, and eye contact. Additional findings include intrauterine growth retardation, premature birth, a large secundum atrial septal defect, and mild facial dysmorphism. Chromosomal microarray analysis revealed a 128kb deletion within the CNTN4 gene corresponding to the coordinates chr3: 2,878,363-3,068,381 (hg19). The deletion takes out exons 7-12 and introduces a frameshift mutation at amino acid 5154 resulting in a termination codon 22 amino acids downstream the deletion site. Two splice variants (4 and 5) are not affected by this deletion. Fluorescence in-situ hybridization studies revealed a paternal origin of the deletion. Conclusion: Although the CNTN4 intragenic deletion in our proband is predicted to result in loss of CNTN4 function, the child did not manifest any of the clinical features associated with other known phenotype. Our data provide additional support for the significance of CNTN4 gene in language development and emphasize the phenotypic heterogeneity. We discuss potential factors contributing to reduced penetrance and attempt to find possible phenotype genotype correlation.

2625M
A 284kb microduplication at 7q21.11 involving SEMA3A possibly linked to agenesis of corpus callosum and visual impairment in a patient with spastic quadriparietal cerebral palsy. D. Ma1, N. Singh-Bhatia2, P. Levy3, R. Marion2, R. Naeem1.

1) Pathology, Montefiore medical center, Bronx, NY; 2) Genetics, Children Hospital at Montefiore, Bronx, NY.

A 284kb heterozygous duplication was identified by array comparative genome hybridization (aCGH) on chromosome 7q21.11 in a patient with spastic quadriparietal cerebral palsy, cerebral dysgenesis including agenesis of corpus callosum, visual impairment and severe global developmental delay. This duplicated region, involving the regulatory elements of SEMA3A along its exonic 1 and the transcriptional start point, has not been documented in any publicly available case or control databases. Within the family, another two brothers and one sister also presented with cerebral palsy and global developmental delay, suggesting a candidate gene in a possibly autosomal dominant inheritance mode with variable expressivity and penetrance. An in depth examination of the SNP data revealed five loss of heterozygosity (LOH) regions on 2q32.2-q24.1, 3q25.2-q26.31, 6q14.2-q14.3, 6q22.1-q22.31 and 8q11.21-q11.23 with various lengths from 3.1Mb to 16.8Mb. A search for potential disease-causing genes in a recessive inheritance mode did not reveal obvious candidates that could explain the clinical features. A literature review of SEMA3A indicated a high expression of the gene in all adult and fetal brain regions and heart as well as fetal skeletal muscle. Gene expression has been associated with the growth of the human corpus callosum and the effect on retinal pigment epithelial cell activity, which is quite well matched with the multiple congenital anomalies in this patient. In addition, a recent case report has linked the deletion of the similar region with novel type of syndromic short stature. Though the gene dosage of SEMA3A appears to be critical for the normal development of brain, the functional impact of the duplication on its regulation of gene expression remains to be elucidated.
2626T

Williams syndrome (WS) is caused by a microdeletion on chromosome 7q11.23 encompassing the elastin gene. WS is characterized by distinct facies, congenital cardiovascular malformations, connective tissue abnormalities and intellectual disability. Craniosynostosis has seldom been described as a complication of WS. We identified seven patients with craniosynostosis. They showed abnormal cranial morphology, prominent microcephaly or headache. Four had sagittal craniosynostosis (dolichocephaly), one had metopic craniosynostosis (trigonocephaly), and two had both. Two of them underwent cranioplasty. Patients with WS often show hyperactivity together with developmental delay. If craniosynostosis remains undiagnosed, disabilities may be worse. Increased intracranial pressure may modify behavioral abnormalities. Therefore, proper diagnosis and neurosurgical intervention are necessary. If patients with WS show prominent microcephaly or cranial deformation including dolichocephaly or trigonocephaly, neuroradiological evaluation including 3D-CT of the cranium is necessary. We suggest that craniosynostosis is an important complication of WS. Craniosynostosis should be considered in every patient with WS and cranial deformation.

2626M
Subtelomeric investigation in individuals with microform of HPE. L. Ribeiro-Bicudo1,2, V. do O2, B. Gamba2, A. Richieri-Costa2. 1) Genetics, Goleas Federal University, Saopto, Brazil; 2) Hospital for Rehabilitation of Craniofani Anomalies, Bauru, Sao Paulo, Brazil; 3) Genetics, Sao Paulo State University, Botucatu, Sao Paulo, Brazil.

Holoprosencephaly (HPE) is a malformation sequence where the cerebral hemispheres fail to separate into distinct left and right halves. Three levels of increasing severity are described in HPE: lobar, semi-lobar and alobar. Another milder subtype of HPE called the middle interhemispheric variant (MIHF) or syntelencephaly, has been recognized. A new phenotype was first described in a series of Brazilian patients, the so-called holoprosencephaly-like (HPE-L) phenotype. The etiologies and the variable clinical spectra of HPE and HPE-L, also called HPE minor form or microform, seem to be related and extremely heterogeneous. Data about frequencies of classic HPE are well known. It is considered one of the most common human malformations, with an estimated prevalence of at least 1/16,000 live births and in 1/250 conceptuses, and it is usually classified according to type of CNS involvement or facial anomalies observed. On the other hand, similar statistics concerning HPE-L and/or microforms are unknown, but certainly it represents a challenge in several areas, such as clinical findings, differential diagnosis, prognosis, and genetic counseling. HPE-L can be seen as a condition important to the understanding of many of the mechanisms involved in normal and abnormal craniofacial development. Attention should be given to newborn babies with isolated large cranialvault where the definitive diagnosis only will be possible with the phenotypic evolution. It is estimated that the cause of HPE is due to cytogenetic anomalies in 30-50% of individuals, to their environmental causes and unknown genetic factors, estimated in 15%, and unknown in 30-40%. In the present work it was based on the observation of previous data showing subtelomeric aberrations in HPE individuals. We selected individuals presenting the microform of HPE with normal MRI and no developmental delay. These individuals had facial characteristics such as hypothelorism, single central incisor, flat nose, and cleft lip/palate. We performed MLPA screening based on the P036 kit on 64 DNA samples. No subtelomeric aberrations was found in the samples analyzed. Even considering that the sample is small, it seems that the subtelomeric aberration is not a common event in the microform of HPE.

2627S
Neurodevelopmental and neurobehavioral characteristics in males and females with CDKL5 duplications. P. Szafarski1, S. Golla1, W. Jin1, P. Fang1, P. Hixson1, R. Matalon2, D. Kinney2, H-G. Bock2, W. Craigen1, P. Magoulas1, J.L. Smith1, W. Bi1, A. Patel1, S.W. Cheung1, C. Bacino1, P. Stankiewicz1. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Departments of Pediatrics and Neurology, University of Texas Southwestern Medical Center, Dallas, TX; 3) Division of General Academic Pediatrics, Department of Pediatrics, The University of Texas Medical Branch at Galveston, Galveston, TX; 4) Memorial Children’s Hospital, Galveston, TX; 5) Department of Pediatrics, University of Mississippi Medical Center, Jackson, MS.

Point mutations and genomic deletions of the CDKL5 (STK9) gene on Xp22 have been reported in patients with severe neurodevelopmental abnormalities, including Rett-like disorders characterized by early-onset seizures, infantile spasms, severe intellectual disability with absent speech, and microcephaly. The phenotypic resemblance to Rett syndrome has been proposed to result from similar functions or interactions in the neuronal molecular pathways between CDKL5 and MECP2. In contrast to MECP2, which is present only in females with CDKL5 duplications.

2627T
Neurodevelopmental and neurobehavioral characteristics in males and females with CDKL5 duplications. P. Szafarski1, S. Golla1, W. Jin1, P. Fang1, P. Hixson1, R. Matalon2, D. Kinney2, H-G. Bock2, W. Craigen1, P. Magoulas1, J.L. Smith1, W. Bi1, A. Patel1, S.W. Cheung1, C. Bacino1, P. Stankiewicz1. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Departments of Pediatrics and Neurology, University of Texas Southwestern Medical Center, Dallas, TX; 3) Division of General Academic Pediatrics, Department of Pediatrics, The University of Texas Medical Branch at Galveston, Galveston, TX; 4) Memorial Children’s Hospital, Galveston, TX; 5) Department of Pediatrics, University of Mississippi Medical Center, Jackson, MS.

Point mutations and genomic deletions of the CDKL5 (STK9) gene on Xp22 have been reported in patients with severe neurodevelopmental abnormalities, including Rett-like disorders characterized by early-onset seizures, infantile spasms, severe intellectual disability with absent speech, and microcephaly. The phenotypic resemblance to Rett syndrome has been proposed to result from similar functions or interactions in the neuronal molecular pathways between CDKL5 and MECP2. In contrast to MECP2, the clinical consequences of increased dosage of CDKL5 are poorly understood. To date, only larger-sized (8-21 Mb) duplications harboring CDKL5 have been described. We report six females and four males from seven unrelated families with CDKL5 duplications ranging in size between 540 kb and 935 kb. Three families of different ethnicities had identical 665,915 bp duplications ranging in size between 540 kb and 935 kb.

2629T
Discontinuous microdeletion at 1p13.3 involving NBPF4 but not ALX3 in a patient with severe frontonasal dysplasia. A. Ali1, R. Pateli1, S. Jaiswal1, S. Singh1, R. Raman1,2,3,4. 1) Centre for Genetic Disorders, Banaras Hindu University, Varanasi, Uttar Pradesh, India; 2) GS Memorial Plastic Surgery Hospital Hospital; Varanasi, India; 3) Department of Zoology, Banaras Hindu University, Varanasi, Uttar Pradesh, India.

Frontonasal dysplasia (FND) arises due to improper embryonic development of head and face. It is defined by at least two of the following features: hypertelorism, broad nose, clefts of nose, absent nasal tip, cleft lip and/or palate, small anterior nasal aperture, and bicornut, prominent forehead, and hairline. Three types of frontonasal dysplasia are reported: type-1 and type-2 which are correlated with mutations in ALX3 (1p13.3) and ALX4 (11p11.2) genes respectively, whereas type-3 is correlated with mutations in ALX1 gene (12p13.1). Here, we report a case of FND with some overlapping features of all the three types of FND characterized by hypertelorism, absent cupid bow, cleft of bilateral nostrils, lacking midline bony nasal septum, deviation of nasal septum to right side, frontal bossing with depression in left side, large anophthalmia, and eye microptalmia, left orbit without lens, bilateral epicanthal folds, overjet teeth, flexion deformity of little finger of left hand and unilateral cryptorchidism (left side). Cytogenetic microarray analysis using Affymetrix CytoScan® HD array was performed for the patient, both the parents, and 7 unaffected controls. A deletion of 125 kb region at locus 1p13.3 was detected in the patient only and none of the parents and unaffected controls. The deleted region contains two genes SLC25A24 and NBPF4. SLC25A24 encodes a carrier protein that transports ATP-Mg which exchanges it for phosphate. NBPF4 is a neuroblastoma breakpoint family gene known to be associated with a number of developmental and neurogenetic diseases such as microcephaly, macrocephaly, autism, schizophrenia, mental retardation, congenital heart disease, neuroblastoma, and congenital kidney and urinary tract anomalies. ALX3, a known candidate, is located in a deleted region but is in normal copies. This study suggests, deletion of NBPF4 as the cause of frontonasaldysplasia in the patient due to the following two reasons. One, NBPF4 is known to cause variable phenotypes including cranial abnormalities, deletion of the 5’ half of the gene is sufficient to cause neuroblastoma, and since NBPF4 is also known to cause neuroblastoma, we hypothesize that ALX3 probably regulates NBPF4 expression and therefore ALX3 mutation cause FND through downregulation of NBPF4. Therefore haploinsufficiency of NBPF4 causes FND. Investigation is underway to further confirm the role of NBPF4 in FND.
Growth Standards for Children and Adolescents with Smith-Magenis Syndrome. L.R. Fleming, F. Duncan, W. Introne, J. McGready, K. Schulze, A. Alady, J. Hoover-Fong, A.C. M. Smith. 1) Medical Genetics Branch, NHGRI/NIH, Bethesda, MD; 2) Office of the Clinical Director, NHGRI/NIH, Bethesda, MD; 3) Dept of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 4) Center for Human Nongenomic Medicine, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; 5) Greenberg Center for Skeletal Dysplasia, Johns Hopkins University, Baltimore, MD; 6) McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD.

Background

Smith-Magenis syndrome (SMS) is a complex genetic disorder, with an incidence of 1/15,000, characterized by distinctive craniofacial features, cognitive impairment, sleep disturbances, and neurobehavioral abnormalities. Poor linear growth and weight gain are common in infancy and childhood, and nutritional support is required. The majority of individuals have growth retardation (SGA). SMS-specific growth curves have not been available previously to allow clinicians to assess growth. Methods: Anthropometric data (height, weight and head circumference) were collected by clinician measurement, review of the medical record and parent submission for 143 individuals (82 Females, 61 Males) with SMS in Europe, the USA and Australia. Of these, 138 (98.5%) had deletions and 5 (3.5%) had point mutations in RA1. Penalized cubic smoothing splines were used to estimate SMS curves for height and weight for periods of 0-36 months and 2-12 years. These curves were compared to WHO (0-36 month) and CDC (2-18 year) norms. Results: At birth, the majority of term infants with SMS are within the clinically normal range for weight and length. Infants with SMS exhibit a decline in height velocity, with the median height tracking < 5th centile of average stature by 3 years of age. Females with SMS have a lower height. SMS females have delayed pubertal development. The median BMI of SMS females is less than females of age the same age. The median BMI of SMS males is higher than males. Relative short stature persists through early adolescence, with median height for individuals with SMS through age 12 near the 25th centile for sex specific CDC age norms. By 6 months of age there is substantial attenuation of growth velocity in SMS males. Median weight velocity over the first year of life is seen in SMS females. Weight curves for 2-12 years of age show accelerated weight gain beginning in late childhood (~7 years). In comparison to CDC norms, the median weight in SMS males and females falls above the median percentage for weight than height. Further research to better understand the natural history of SMS and the genotype-phenotype relationships in these patients.

One cannot presume siblings with similar clinical findings result from the same underlying genetic cause!—familial genomic instability, leaky proof reading mechanism, or true diagnosis? A. Tsai, K. Kovak, K. Keller, J. Kushner, C. Rogers, S. Moore. Molecular and Medical Genetics, OHSU, Portland, OR.

When siblings present with similar clinical findings, we usually presume they result from the same etiology. It is not uncommon for a clinical geneticist to order all testing on one child in the family, and subsequently offer siblings specific testing for the revealed genetic disorder. For a cytogenetic anomaly identified by microarray, a reference lab often suggests a FISH study for parents or at-risk individuals in the family rather than a full microarray. We present three families in which multiple children within the family are affected with similar clinical features. Family 1 is fraternal twins. They were born SGA with similar dysmorphic features. A 180k oligo array revealed a copy gain of 1.94 Mb: 5q35.2q35.3 (175,491,045-177,467,408)x3 and similar features. The male twin’s array showed a single copy gain of 433.7 kb (maximum size 458 kb) in 5q35.2q35.3 (176,700,522-177,136,262)x3 that resides within/overlaps with a portion of the sister’s, is predicted to disrupt several exons of the NSD1 gene and present with Sotos syndrome. Family 2 is fraternal twins, both with autism and small optic nerves. Microarray in one twin revealed a distal deletion 22q11.21 (19,408,946-19,790,658)x1 and the brother showed a 16p13.11 duplication (15126709-16292235)x3, 1.16-2.13 Mb. Family 3 is a maternal half sibling with dysmorphic features and intellectual disability. The brother has a 969.6 kb duplication at 1p13.3, and the sister has a 598 kb deletion at 16p11.2 consistent with the 16p deletion syndrome and the 969.6 kb duplication. Siblings only share 25-50% of genes that are subject to recombination, individual epigenetic exposures and stochastic events. On cannot presume siblings with similar clinical findings result from the same underlying genetic cause. Family 1 shows co-existence of 2 different CNVs from different parental origins. Families 2 and 3 reveal different CNVs and, if only FISH were offered, the anomalies in siblings of both families could have been missed. In familial cases, we recommend that all siblings undergo comprehensive testing so that familial events can be explored. Further investigation is required if the phenotype cannot be explained by known microarray anomalies. Appropriate genetic causes should be considered individually, especially in half siblings.

An interstitial microduplication in 17q11.2 encompassing the NF1 gene, in a girl presenting with a Prader Willi like syndrome: expanding the NF1 microduplication. C. Vinkler1, T. Lerman-Sagie2, D. Levi1. 1) Inst Med Gen, Wolfson Med Ctr, Holon, Israel; 2) Metabolic Neurogenetics Service, Wolfson Medical Center, Holon Israel; 3) Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.

Intersitial microduplication of chromosome 17q11.2 is rare condition with less than 40 reported cases. The first microduplication was described in 2004 in a patient with type 3 with 1Mb duplication. The clinical phenotype includes developmental delay and type 3 with 1Mb duplication. The clinical phenotype includes developmental delay, facial dysmorphism and sometimes seizures. Eating disorder with severe obesity has not been described in this group of patients. We present a 14y old girl who was followed in our clinic for several years. We have had dysmorphic features, obesity, hypotonia, global developmental delay and atioseptal heart defect. Prader-Willi syndrome was ruled out by molecular tests. Family history is unremarkable. At 14y of age she is still overweight (due to eating disorder) intellectually disabled and has verbal communication. Her dysmorphic features include round face, short and narrow forehead, small orbital fissures, broad and sparse eyebrows, short and small nose, widely spaced inverted nipples, camptodactyly and clinodactyly of the 5th finger bilaterally. Brain MRI is unremarkable except for a mild enlargement of the left atrium. CMA test showed a de novo 8,746Mb duplication. The chromosome 17q11.2 encompassing the NF1 gene along with other 14 OMIM morbid genes. None of the genes has been associated with eating disorders. Obesity has been mentioned in two other cases of 17q11.2 duplication. We propose that CMA test should be considered in cases of eating disorders and sometimes in the case where body mass index is elevated and no specific etiology is found. The common clinical area of eating disorder and perhaps help in understanding this abnormality.
2634M
Bilateral absence of the ulna in 4q terminal deletion syndrome. M. Marble1, T. Meaux2. 1) Dept Ped/Gen/LSUHSC, Children’s Hosp New Orleans, New Orleans, LA; 2) LSUHSC.

We report a patient born with bilateral absence of the ulna who was found to have a terminal deletion involving the long arm of chromosome 4. The deletion was identified by chromosome microarray analysis and determined to be 21.98 Mb in size and to involve 4q32.3-4qter. The molecular coordinates encompassed by the deletion were 4:169,154,141-191,133,858. The patient was born at 38 weeks gestation by spontaneous vaginal delivery. The pregnancy was complicated by intrauterine growth retardation, oligohydramnios, and maternal hepatitis C. Birth weight was 3000 grams. Other clinical features included atrial septal defect, patent ductus arteriosus, micrognathia, cleft palate, inner epicanthal folds and long philtrum. At 2 years of age, weight was 10.6 kg (10-25th percentile) and height was 80 cm (5th percentile). She had delay in language skills with an approximately 15 word vocabulary and no sentences. She could sit on her own and could pull up on furniture but was not yet able to walk. The atrial septal defect and patent ductus arteriosus closed spontaneously. Radiographs showed a short humerus and absent ulna bilaterally, but apparent absence of the mid phalanx. Past reports have suggested an association between chromosome 4q deletions and variable ray defects but most of these reports used traditional cytogenetic methods. We found a few cases in the literature with absent ulna associated with 4q terminal deletions but very few for which comparative genomic hybridization was documented. We compared the deletion in our patient to previous reports in the literature. The findings in our patient, along with other reports, suggest that the terminal 4q chromosomal segment contains a region important for ulnar ray development.

2635T
Congenital heart disease and Sturge-Weber syndrome in a young female with 22q11.2 triplication. L. Moita-Vieira1,2, L.M. Pires3, S. Vaz4, P. Macci5, I.M. Careira6, R. Pires7,8, 1) Molecular Genetics and Pathology Unit, Hospital of Divino Espirito Santo of Ponta Delgada, Azores Islands, Portugal; 2) Centre for Biodiversity, Functional and Integrative Genomics (BioFIG), Faculty of Sciences, University of Lisbon, Portugal; 3) Instituto Gulbenkian de Ciência, Oeiras, Portugal; 4) Laboratory of Cytogenetics and Genomics, Faculty of Medicine, University of Coimbra, Portugal; 5) Department of Pediatrics, Hospital of Divino Espirito Santo of Ponta Delgada, Azores Islands, Portugal.

In the scope of the population-based epidemiological study on congenital heart disease (CHD) performed in the Sao Miguel Island (Azores archipelago, Portugal), we are investigating the genetic basis of these complex diseases which are a major cause of serious morbidity and mortality in childhood. Since 22q11.2 deletions and duplications are frequently implicated in CHD, we carried out a systematic analysis of this genomic region in CHD patients from the Azorean Islands. The analysis consisted of clinical evaluation and searching for copy number variations (CNVs) by Multiplex Ligation-dependent Probe Amplification (MLPA), according to MRC-Holland protocol. The MLPA results detected in a female patient higher doses than 22q11.2 duplication, suggesting the presence of a double dosism (2:2, i.e., two chromosomes with 22q11.2 duplication) or a tetrasomy (1:2, one normal chromosome and a 22q11.2 triplication in the other). These two hypotheses were investigated by FISH analysis which confirmed a tetrasomy. In order to ascertain if the 22q11.2 triplication was inherited or acquired de novo. The results showed that her non-affected father presented 22q11.2 duplication (1:2). Therefore, the triplication here described could be explained by an unequal crossing-over of 22q11.2 region that triggered the expansion of the duplication to a tandem triplication during spermatogenesis, in a non-allelic homologous recombination mechanism mediated by the presence of low-copy repeats. To the best of our knowledge, the present study reports the second documented case of 22q11.2 triplication, being the first case reported by Yobb and collaborators (2005). Our patient displays the dysmorphic facial features, cognitive deficit and heart defect (restrictive interventricular communication and membranous sub-aortic stenosis) usually seen in patients with 22q11.2 deletion and microduplication syndromes, as well as phenotypic characteristics of Sturge-Weber syndrome (in particular the presence of nevus flammeus — port-wine stain —, behavioral disorders, and visual fields defects). Taking this into consideration, we propose that 22q11.2 triplication is a variation of 22q11.2 microduplication syndrome, with aggravated phenotype due to the major dosage of implicated genes.

2636S
Familial 17q12 duplication presented as SGA/IUGR and microcephaly during pregnancy: A counseling dilemma. A. Singer1, I. Maya2, C. Viniker3. 1) Dept Clinical Gen Unit, Barzilai Med Ctr, Tel Aviv, Israel; 2) Institute of Medical Genetics, “Rabin” Petahtikva, Israel; 3) Institute of Medical Genetics, “Wolfson” Holon, Israel.

Recurrent genomic rearrangements of chromosome region 17q12, ranging from 0.3 to 2.1 Mb, have been described to be associated with different clinical phenotypes. Patients carrying a 17q12 duplication present with intellectual disability, developmental delay of various degree, epilepsy, schizophrenia, autism, brain abnormalities, esophageal atresia and various renal and urinary tract abnormalities. In the normal population, duplication of 17q12 is quiet rare (<0.02%) and it penetrance estimate is around 21%. Hence, prenatal counseling poses a dilemma. We present a baby boy with prenatal and postnatal microcephaly and IUGR/SGA. He was born to a healthy nonconsanguineous couple. Late amniocentesis (32 week) because of symmetric SGA/IUGR, microcephaly and echocardiographic intracardiac focus found 1.4 MB 17q12 duplication. Parents’ CMA showed that it was inherited from his father. During counseling the father was found to have normal HCs but he has dyslexia and some speech disturbances. Counseling in this case posed a dilemma since 17q12 duplication is known to have a variable phenotype and incomplete penetrance. The parents decided to continue with the pregnancy. The boy was born at term. Examination was normal except for microcephaly (HC< -2SD). Prenatal microcephaly has not been previously described in 17q12 duplication syndrome. The clinical significance of this prenatal finding adds to the counseling dilemma.

2637M
Atypical 22q11.2 deletion at the distal end of the common 3Mb deletion. N. Bhattacharya1,2, T. Goldwasser2, R. Reingold2, S. Klugman2, P. Levy3, B. Morrow4, 1) Children’s Hospital at Montefiore, Bronx, NY, USA; 2) Montefiore Medical Center, Bronx, NY, USA; 3) Albert Einstein College of Medicine, Bronx, NY, USA.

Background: 22q11.2 deletion syndrome is a contiguous gene deletion characterized by congenital heart disease, palatal defects, immunodeficiency, hypocalcemia, renal anomalies, developmental delay and facial dysmorphism. Up to 64% of individuals with 22q11.2 deletion syndrome exhibit psychiatric disorders, including bipolar disorder and schizophrenia, with wide phenotypic variability. Deletions in the 22q11.2 region usually include a common 3Mb deletion (~90%) or a proximally nested 1.5Mb deletion (~7%). Other atypical deletions include distally nested deletions and deletions distinct from the 3Mb region. The 3Mb and 1.5Mb deletions both include the TBX1 gene, associated with the congenital heart disease seen in this syndrome.

Case: A 20-week primigravid was referred to our prenatal genetics department for counseling after this atypical deletion was noted to have a Dandy-Walker variant on 16 week anatomy scan. The proband was the product of a non-consanguineous union between a 23 year old mother and 37 year old father. The mother has a history of bipolar disorder, ADD, OCD, and developmental delay, diagnosed in adolescence. Aminioceintesis was performed and arrayCGH revealed a 0.369Mb maternally-inherited deletion in chromosome 22, specifically, arr[hg19] 22q11.21 (21081260-21449911) x 1 mat. At birth, arrayCGH confirmed the 0.369Mb deletion (~7%). Other atypical deletions include a common 3Mb deletion (~90%) or a proximally nested 1.5Mb deletion (~7%). Other atypical deletions include distally nested deletions and deletions distinct from the 3Mb region. The 3Mb and 1.5Mb deletions both include the TUBB1 gene, associated with the congenital heart disease seen in this syndrome.

Dysmorphisms. Further study will aid in an improved genotype-phenotype correlation and enhance our understanding of distal deletions of the 22q11.2 region.

Discussion: This case presents an atypical deletion of the 22q11.2 region, with notable phenotypic features of 22q11.2 deletion syndrome - namely, psychiatric issues as well as velocparyngeal insufficiency and craniofacial dysmorphisms. Further study will aid in an improved genotype-phenotype correlation and enhance our understanding of distal deletions of the 22q11.2 region.
2638T

A novel microdeletion affecting SNRPN but preserving distal gene expression leads to Prader-Willi Syndrome. T. Diallo1, R. Begay1, D. Stavov1, S.L. Graw1, T. Boyle2, M.R.G. Taylor3, P.R. Baker II3. 1) Adult Medical Genetics Clinic, University of Colorado Denver - Anschutz Medical Campus, Aurora, CO., United States; 2) Departments of Medicine and Pathology, University of Colorado Denver - Anschutz Medical Campus, Aurora, CO., United States; 3) Clinical Genetics and Metabolism, The Children’s Hospital, Denver, CO., United States.

Background: Prader-Willi Syndrome (PWS) is an imprinting disorder characterized by initial neonatal hypotonia, feeding problems, and failure-to-thrive, followed by hyperphagia, obesity, and developmental delay. The most common causes of PWS are 5-7 Mb class I or II deletions of paternally expressed genes at the 15q11.2 locus. Atypical microdeletions have implicated various genes as critical for the development of pathology and continue to shed light on our incomplete understanding of PWS. Methods and Results: A 23 year-old woman with typical features of Prader-Willi Syndrome was evaluated clinically. As a neonate she was severely hypotonic with early feeding problems and failure-to-thrive. By 8 years of age her parents noted lack of satiety, obesity, and food hoarding. She exhibited delayed puberty, mild developmental delay, unique behavioral characteristics, and her IQ was considered to be low-normal. Her overall PWS clinical score was 13 (based on Holm et al. 2002 criteria). Methylation analysis, classic PWS FISH, and clinical SNRPN expression studies were normal, but oligoarray studies revealed a microdeletion of up to 132 kb. SNP homozygosity analysis narrowed the deletion to approximately 70 kb between intron 2 and intron 4 of SNRPN, and to not extend into the SNURF transcript. RT-PCR analysis was compatible with loss of SNRPN and retention of SNURF transcripts. Importantly, the patient’s deletion did not include the SNORD116 cluster or the imprinting center, with preserved expression of transcripts distal to SNRPN. Conclusion: Rare cases of PWS with attenuated phenotypes have been associated with microdeletions in snoRNA clusters within the 15q11-13 locus. In these cases methylation testing and SNRPN expression testing were still diagnostic. Our proband, with a striking, albeit attenuated, clinical picture represents the first patient with a PWS phenotype and a demonstrable deletion involving the SNORD116 cluster or the imprinting center, with preserved expression of transcripts distal to SNRPN. A novel microdeletion affecting SNRPN but preserving distal gene expression leads to Prader-Willi Syndrome.

2639S

Overgrowth in association with 3q25 microdeletion. K. Enomoto1,2,3, N. Kurosawa1, R. Satomi1, Y. Sugawara1, A. Watanabe2, S. Watanabe2, K. Kurosawa2. 1) Department of Pediatrics, JA Toride Medical Center, Toride, Ibaraki, Japan; 2) Department of Pediatrics, Tsuchiura Kyodo General Hospital, Tsuchiura, Ibaraki, Japan; 3) Department of Pediatrics and Developmental Biology, Tokyo Medical and Dental University Graduate School, Tokyo, Japan; 4) Division of Medical Genetics, Kanagawa Children’s Medical Center, Yokohama, Japan.

Prevalence of congenital disorders associated with overgrowth is not so much compared to what with growth failure. Generally speaking, the rare and specific physical features could sometimes reduce the difficulty with diagnosis in some unknown cases. However, the diagnosis in association with overgrowth is often difficult because of few or other distinctive symptoms or nothing but nonspecific or confusing dysmorphic features. It’s not easy to detect disease-causing alterations of chromosome with traditional investigations such as karyotype, though some characteristic syndromes can be analyzed by using technic of molecular biology or genetics, e.g. fragile X syndrome, Beckwith-Wiedemann syndrome (BWS), and Simpson-Golabi-Behmel syndrome. We have limited cases showing overgrowth previously reported associated with chromosome anomalies, e.g. duplication in 4p, deletion in 5q35 (Sotos syndrome), duplication in 11p15 (BWS), duplication in 12p, 22q13 deletion syndrome, and Pallister-Killian syndrome (mosaic 12p tetrasomy). We herein present a new case of 3q25 microdeletion extending to 700kb involving SHOX2 and some other genes identified by microarray CGH examination: arr 3q25.32(157,649,680-1,323,314)x1 (hg19). The patient who is 5-year-old boy at present was born at 38 weeks gestation. His birth weight, length, and head circumference were 3,028g (+2.89 SD), 53.0cm (+2.66SD), and 36.0cm (+2.25SD), respectively. He was transferred to our hospital in early newborn owing to his systemic angioma including facial invasion without hemihypertrophy or limb asymmetry. He had not been complicated by any other problems in newborn to early childhood. He exhibits obvious overgrowth especially in his postnatal weight and head circumference (macrocephaly), mild developmental delay, and facial dysmorphism. Facies is characterized by frontal bossing, downsloping palpebral fissures, bulbous nose, and mildly dysplastic external ears. Angioma is now getting remarkable regression, though it has the dysmorphism. The smallest deletion range in the previously reported cases. As our observations delineate the phenotypic spectrum associated with a clearly defined deletion of chromosome 3q25, we suggest that the haploinsufficiency of the genes deleted region contributes to the above mentioned some characteristic features, and this part of chromosome can be the almost critical lesion of 3q25 microdeletion syndrome.

2640M

Diamond-Blackfan anemia and intellectual disability: a new contiguous gene syndrome at 15q25.2. M. Gorce1, A. Guichet1, A. Donzeau2, A. Ziegler1, M. Barth1, I. Pellier1, D. Bonneau1, E. Colin1. 1) Department of Biochemistry and Genetics, CHU Angers, ANGERS, France; 2) Department of Pediatrics, CHU Angers, Angers, France.

15q25.2 microdeletion is an emergent CNV locus for intellectual disability, dysmorphic features and congenital anomalies. Two distinct microdeletions have been described at this locus: 1) a distal deletion (11 cases) responsible for neurodevelopmental and neuropsychiatric disorders and 2) a proximal deletion (6 cases) which is a susceptibility locus for cognitive deficit, dysphragmatic hernia and possibly Diamond-Blackfan anemia (DBA). This proximal deletion is said to predispose to DBA because it contains the gene RPS17 encoding for a ribosomal protein, responsible for 2% of DBA. Until now, however, DBA has been diagnosed with certainty in only one case of 15q25 proximal deletion. The additional case reported here had a history of intrauterine growth retardation. At age 18 months, he had a moderate developmental delay, dysmorphic features and musculoskeletal anomalies. He had a normochromic macrocytic anagenerative anemia with elevated erythrocyte adenosine deaminase activity and elevated HbF (3,2%) highly suggestive of DBA. A 15q25.2 microdeletion of 2.2Mb including RPS17 was identified using SNP array and the deletion of RPS17 was confirmed by FISH using a specific probe. The patient’s deletions were present in father’s and was impossible to test in mother. To date, only few mutations in RPS17 have been reported in patients with DBA. Anemia was mentioned in 4 cases among the 8 previous reported cases of 15q25.2 proximal deletion but the definite diagnosis of DBA was made in only 1 case. The present report confirms that patients with 15q25.2 deletion involving RSP17 are at risk for DBA and possibly DBA-associated malignancies.
Expanding the Phenotypic Profile of Kleefstra Syndrome: A Female with Near-Normal Intelligence and Developmental Dyspraxia. D. Sisón1, D. Gibril2, A. L. Gropman1, C. Sprouse2, T. Sadeghin1, C. Samango-Sprouse1,2,3,4,1. The Focus Foundation, Davidsonville, MD; 2) Children’s National Medical Center, Washington, DC; 3) George Washington University School of Medicine, Washington, DC; 4) Neurodevelopmental Diagnostic Center for Children, Washington, DC.

Kleefstra syndrome is a rare neurogenetic disorder most commonly caused by deletion in the 9q34.3 chromosomal region and is associated with intellectual disabilities, severe speech delay and behavioral problems. To our knowledge, this is the first patient with a 9q34.3 deletion that is not intellectual disabled but has a developmental disorder with overlapping features expanding the phenotypic profile and prognosis for patients with Kleefstra syndrome. The patient is a 6-year-old female whose neurodevelopmental and cognitive profile was assessed using the Beery-Buktenica Visual Motor Integration tests (Beery-VMI), the Expressive and Receptive One Word Picture Vocabulary Tests-Revised (EOWPV-R) and Preschool Language Scale -4th Ed. (PLS-4), the Wechsler Intelligence scales (WPPSI-III) and the Social Responsiveness Scale (SRS-2). Verbal IQ was 81 and Global IQ was 94 on WPPSI-III.Brief subsets indicated normal intelligence and receptive and expressive vocabulary (standard score = 82, EOWPV-R) Auditory Comprehension and Expressive Communication were within normal limits for age but in the low-average range. SRS-2 indicated delays in social cognition (T-score = 89) social communication (T-score = 81) and autistic features (T-score = 90) but the patient did not meet the criteria for Autism Spectrum Disorder. TBR1 and SLC4A10, critical for early cortical development and regulation of the intracellular pH of neurons respectively, may have an influence on pathogenesis of psychomotor development. Motor planning deficits affecting her speech, oculomotor and graphomotor abilities for which the patient received early intervention services. These services may have contributed to the patient’s above average speech and intellectual capabilities than previously reported cases with a 9q34.3 deletion, expanding the phenotypic profile associated with Kleefstra syndrome. Thus, dyspraxia should be considered in the diagnosis and prognosis of patients with a 9q34 deletion as this case-report suggests that targeted intervention therapies may significantly improve patients’ cognitive outcome.

Craniofacial Dysmorphism And Mild Intellectual Disabilities In A Child With A Paternally Inherited 14q32.1 Deletion. Y. Wang, J. Martinez Dept Pathology, Univ South Alabama, Mobile, AL. 600 ClinC Dr. Mobile, AL 36688.

Deletions in chromosome 14q32 have been reported in association with intellectual disabilities and congenital anomalies. This is an imprinted genomic region with paternally and maternally expressed genes. We present a child with dysmorphic features and mild intellectual disabilities associated with a paternally inherited 14q32.1 chromosomal deletion. This is a 9 yr old white male seen in the Genetics clinic with history of developmental delay and unusual features. The child was born at term following a 20 yr old mother from a pregnancy complicated by oligohydramnios. Birth weight was 8 pound 6 ounces. Birth length was 20 and ¾ inches. He had no neonatal problems but he was later noted to have global developmental delay, feeding and sensory processing difficulties and speech abnormalities. He was attending school at the 3rd grade level within an I.E.P and he is having learning difficulties. He was also evaluated for constipation and GI studies revealed a “small colon”. A rectal biopsy was negative for Hirschsprung's disease. The patient received early intervention services. These services may have contributed to the patient’s average speech and intellectual capabilities than previously reported cases with a 14q32.1 deletion, expanding the phenotypic profile associated with Kleefstra syndrome.

In ten individuals, proximal regions PRODH to DGCR10 were found to be non-deleted by SNP array. These findings were not identified by MLPA, as these regions are not covered by MLPA array. These findings were not identified by MLPA as it primarily covers the low copy repeats that flank the deleted region. We analyzed DNA samples from 218 individuals with Affymetrix v6.0 SNP array and with MLPA P250 DiGeorge probemix. There was concordance between the two assays and MLPA results. We found that 201 patients (92.2%) had a 3Mb deletion, 5 (2.2%) had a 2Mb deletion and 12 (5.5%) had a 1.5Mb deletion.

Three common deletion sizes have been reported in the majority of patients with chromosome 22q11 microdeletion syndrome (22q11DS): 3Mb in more than 90% of patients, 2Mb in a small proportion (5%). The frequency of CHD was significantly higher in the 3Mb deletion in comparison with the 1.5 Mb deletion and no CHD was observed in the 2Mb deletion. The frequency of CHD was significantly higher in the 3Mb deletion in comparison with the 1.5 Mb deletion and no CHD was observed in the 2Mb deletion. However, the frequency of CHD increased significantly higher in the 3Mb deletion compared with the 1.5 Mb deletion, showing no correlation with the presence of CHD.

The frequency of CHD was significantly higher in the 3 Mb deletion in comparison with the 1.5 Mb deletion (58% vs 25%; p = 0.005; OR 4.38, 95% CI 1.05-20.787). The frequency of CHD was not observed for the 1.5 Mb deletion. We also analyzed other clinical data (Kurahashi et al 2007; Michaelovsky et al 2012). This resulted in a significant correlation of PRODH to DGCR10 with the presence of CHD. The frequency of CHD was significantly higher in the 3 Mb deletion compared with the 1.5 Mb deletion (p = 0.005; OR 4.0, 95% CI 1.46-31.30). This large cohort of Chilean patients confirms that the large 3Mb deletion is the most common one in this syndrome, and that there is an apparent role of PRODH and proline metabolism in the psychiatric phenotype. Funded by Fondecyt-Chile grant #1130392.
264SS
Unknown CNVs found in 52 Bulgarian patients with intellectual disability and congenital malformations. S.P. Hadijdekov1, D. Avdjieva-Tzavela1,2, R. Rukova1, P. Staneva1, D. Nosheva1, R. Ticheva1, D. Toncheva1,2. 1) Department of Medical Genetics, Medical Faculty, Medical University- Sofia, Bulgaria; 2) State University Pediatrics Hospital “Prof. Ivan Mitev”, Medical Faculty, Medical University- Sofia, Bulgaria.

The evergreen resolution of the DNA microarrays and the optimization of the technique allow detection of larger aberrations and a variety of small copy number variations (CNVs) whose clinical significance is unknown in some cases. Methods: We studied 52 patients with congenital anomalies and intellectual disability. Whole-genome oligo-array CGH was performed using the BlueGene CytoChip oligo 2×105K microarray, v.1.1. Results: We found a total of 247 CNVs, of which 15 pathogenic (7 deletions, 8 duplications), 124 benign (62 deletions, 62 duplications) and 108 with unknown clinical significance (68 deletions, 40 duplications). The unknown CNVs were detected among 34 patients with the following distribution in size : 52.7% less than 100kb; 37% from 100 Kb to 500kb; 9.2% from 500kb to 1 Mb; 0.9% larger than 1Mb. We applied a specific algorithm in the interpretation of the unknown variants. Conclusion: The rates of unknown CNVs in our study were notably high: This indicates that some of these variations may be probably benign for the Bulgarian population and cannot be found in the studied other populations. It is speculated that the Bulgarians are characterized by high genetic heterogeneity. This demonstrates the obvious need for large population studies and mapping of variations of unknown significance in the Bulgarian population for enabling the precise interpretation of unknown CNVs in the clinical praxis. The protocol and informed consent documents were reviewed and approved by the local ethics committee and prepared according the Declaration of Helsinki and local country laws. Acknowledgements: Grant 02/76-21.12.2009, National Science Fund, Bulgaria.

2646M
A rare case of speech sound disorder with a heterozygous BCL11A deletion. A. Huang1, B. Peter1, Center for Mendelian Genomics1, W. Raskind2,3,5, 1) Speech & Hearing Science, Arizona State University, Tempe, AZ; 2) Genome Sciences, University of Washington, Seattle, WA; 3) Psychology & Behavioral Sciences, University of Washington, Seattle, WA; 4) Psychology, University of Washington, Seattle, WA; 5) Medicine, University of Washington, Seattle, WA.

Speech sound disorder (SSD) interferes with a child’s ability to acquire clearly intelligible speech. One proposed subtype is childhood apraxia of speech (CAS), thought to result from faulty programming of movement sequences for speech. CAS is often comorbid with reading disorder (RD), a disorder of written language. We recently described a child with CAS with a de novo heterozygous microdeletion of BCL11A in 100 kb of surrounding noncoding sequences. Speech and gross motor development showed evidence of poor motor programming and low muscle tone. In eight of ten previously reported cases with microdeletions in 2p15-p16.1, multiple genes including BCL11A have been identified; phenotypes were severe. BCL11A regulates fetal hemoglobin and is located in an RD candidate region; its role in speech, reading/spelling, and motor development is not yet well understood. A previous study reported 16 CNVs in 12 children with CAS but causality could not be established (Laffin et al., 2012); none of the reported CNVs involved BCL11A. We investigated the role of CNVs of BCL11A and other regions, including candidate regions for CAS and RD, in 9 individuals with CAS and 3 individuals with RD. One control individual had long-term RD and a small BCL11A microdeletion that was not related to CAS or RD. Also included were three unaffected individuals, biologically related to two CAS participants, and two families with familial CAS, with 9 and 4 affected individuals, respectively. DNA was extracted from buccal coat. Genotyping was performed with 984 k SNP CGHScan and with cnHap and PennCNV algorithms. No participant had CNVs involving BCL11A. Overlaps with two CNVs in the Laffin et al. (2012) study were found to be common CNVs. Five participants (1 RD, 1 TT, 1 CAS; 2 controls) carried four copies of an identical region of the FOXP2 gene. Two participants (1 RD, 1 CAS) had a CNV in an identical region of CNTNAP2. There was no evidence that these CNVs segregated, however, and they are unlikely to be pathogenic. Two or more participants with CAS and RD shared 64 other CNVs. We conclude that BCL11A deletions are uncommon in CAS. In future studies, we will evaluate the coding and regulatory roles of this gene in RD and CAS and follow up on CNVs from the present study. - Genotyping and CNV analyses were provided by the University of Washington Center for Mendelian Genomics (1U54HG006493 to Drs. D. Nickerson, J. Shendure and M. Bamshad).

2647T
Congenital asplenia in a patient with chromosome 1p36 deletion. L. Pisani1, F. Zachariah2, A. Gomez3, J. Reiner1, N. Cohen1, L. Mehta1. 1) Dept. of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, NY; 2) Dept. of Pediatrics, Columbia University College of Physicians and Surgeons, NY.

Chromosome 1p36 deletion syndrome is a well-described microdeletion, with an incidence of about 1/5000 newborns. Common manifestations are typical facial features, intellectual disability, hypotonia, seizures, structural brain abnormalities, congenital heart defects, ocular anomalies and hypothroidism. Immunodeficiency has not been previously associated with this syndrome. We report a 2-year-old girl who presented with a complex partial seizure in the setting of an acute febrile illness. She had a history of global developmental delay and was noted to have frontal bossing, open anterior fontanel, low set ears, flat nasal bridge, short palpebral fissures and overlapping 2nd-3rd toes bilaterally. Blood and CSF cultures at admission were positive for Streptococcus pneumoniae. The patient and her siblings had not received any immunizations due to religious reasons. Family history was negative for intellectual disability, consanguinity or birth defects. Microarray revealed the presence of a pathogenic 5.33 Mb deletion of chromosome 1p36.31-p36.33 including 112 genes and transcripts and overlapping the critical region for 1p36 deletion syndrome. The deletion was confirmed by FISH. In addition, a 91.6 kb duplication on chromosome Xp22.31(8557747-8658296) was noted, encompassing exon 2 of KAL1, considered likely benign. Testing of parents is pending. She received a 2-week course of Ceftriaxone but was then lost to follow up. Four months later, the patient was readmitted with new fever and lethargy. Blood culture was positive for Hemophilus influenzae with CSF analysis suggestive of bacterial meningitis. Her condition improved with IV antibiotics and she was discharged. Four months later, she was re-admitted with new fever and convulsions. Blood cultures were negative. A physical examination revealed no abnormalities. White blood cell count and platelet count were in the normal range. Her peripheral smear was notable for Howell-Jolly bodies and abdominal ultrasound showed a 2.5 cm structure at the site of the spleen. Echocardiogram was normal. She was discharged on amoxicillin prophylaxis after completing a course of Ceftriaxone with recommendations for catch-up vaccinations. This is the first report of congenital functional asplenia in 1p36 deletion. The deletion does not include any known genes implicated in hypothyroxinemia, and is of a commonly reported size. Unmasking of a recessive allele could be considered as a potential mechanism, but no known recessive gene could be identified in the deletion interval. This report highlights the presence of a rare birth defect in a relatively well characterized microdeletion syndrome.

2648S
De novo deletion of 5q23.2-q31.1 in a boy with global developmental delay, contractures and dysmorphic features: a contiguous gene deletion syndrome involving morphogenesis and DNA repair. A. Guerin1, R. Gatti4, C. Brown5, M.S. Mayn5,6, M. Carter5,6. 1) Division of Medical Genetics, Department of Pediatrics, The Hospital for Sick Children, Kingston General Hospital, Kingston, Ontario, Canada; 2) Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada; 3) Program in Genetics and Genome Biology, SickKids Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada; 4) School of Medicine Department of Pathology and Laboratory Medicine, Los Angeles California.

Contiguous gene deletion syndromes represent a diagnostic dilemma in interpreting complex phenotypes. Genes that are deleted can have important roles in morphogenesis, which can result in malformations, but genes with ongoing expression, which are essential for cell survival, can also be affected. We present a patient with a de novo 8.4 Mb deletion of 5q23.2-q31.1, a rare deletion syndrome. The patient presented with congenital contractures, dysmorphic features, symmetric growth retardation and global developmental delay. The deletion encompassed 12 OMIM genes, including FB2N2, which is expressed in early morphogenesis. Deletions of FBN2 have been found in patients with Beals syndrome (MIM 121050); this is likely the cause of our patient’s congenital contractures; however, he did not have the other characteristic features (arachnodactyly, Marfanoid habitus). Our patient also had haploinsufficiency for the RAD50 gene, whose product is part of the MRN complex, required for repairing DNA damage. Complete RAD50 deficiency causes for progressive bone and muscle disease (OMIM #613078), characterized by microcephaly and growth retardation, “birdlike” facies and chromosome instability. Given the possibility of decreased production of key repair proteins, chromosome breakage and colony survival (radiosensitivity) assays were performed. Our studies showed that these studies were within normal limits. Western blot showed normal levels of ATM, NBS, and RAD50. Heterozygous missense variants in RAD50 may be low penetrance risk factors for breast cancer, so it remains uncertain if there is an association with this deletion. This case adds to the list of candidate syndromes that present with severe neurodevelopmental delay. This report represents a challenge in providing counseling and anticipatory guidance regarding microarray findings, especially when genes responsible can have variable clinical phenotype, and are responsible for DNA repair.
2649M

A case with mild phenotype of holoprosencephaly is caused by de novo hemizygosity for chromosome 2q14.1-q14.3 involving GLI2 gene. E. Kirtas1, E. Kirat1, A. Koparir1, E. Guzel1, E. Fenercioglu1, M. Seven1, H. Ulucan1, M. Ozlen2, G. Guven1. 1) Istanbul University, Medical School of Cerrahpasa, Department of Medical Genetics, Istanbul, Turkey; 2) Department of Pathology & Immunology Baylor College of Medicine, Houston, TX, 77030, USA.

GLI2 is mediated as an activator by central transcriptional factor on SHH signaling although its precise role on SHH signaling remains unclear. Defects in this gene are associated with holoprosencephaly (HPE) with extremely variable phenotypic penetrance or HPE-like phenotype, abnormal pituitary gland formation and/or function, and polydactyly. Here we present a case with HPE-like phenotype having a 2q14.1-q14.3 deletion. The case was a female, born to healthy and 1.5 degree consanguineous parents, 28-year-old mother and 42-year-old father. She was delivered at term with a birth weight of 3040 g (25-50p) and birth length of 47 cm (10-25p). She was referred to our department at the age of 2.5 years for mild neurodevelopment delay. Bilateral postaxial polydactyly, grade 3 vascularer reflux, and recurrent infections were treated short after the birth. In physical examination, her weight was 13 kg (%50p) and her height was 90 cm (%50p) while head circumference was 46 cm (<3p). Urinary and sacral USG were within normal ranges while atrophy of parenchyma in left kidney with 99mTc-DMSA SPECT scan detected. Cranial CT, MRI and elbow radiography showed no abnormality. Amino acid and metabolic screen were normal except from moderate high levels of lactate and ammonia. Chromosome analysis from the blood lymphocytes revealed 46,XX, ins (del)(2)(q14.2;q21).inv(9)(p11q13) mat karyotype. Her mother’s and father’s karyotype were 46,XX,inv(9)(p11q13) and 46,XX, respectively, indicating maternal origin. A detailed analysis of the constitutional chromosomal changes in proband was defined by array comparative genomic hybridization (aCGH) revealed a 5.6 mb heterozygous deletion spanning from 2q14.1 to 2q14.3 has been found. This region involves whole GLI2, which is thought to be responsible for this mild form of HPE. Our results on GLI2 deletion in this case confirms phenotypic heterogeneity and mild form of HPE.

2650T

Severe fetal phenotype of a dominant mesomelic dysplasia, associated with a 790 kb microduplication of HOXD gene cluster at 2q31.1, S. Odent1, C. Jeanne-Pasquier2, P. Balouet3, L. N. Leporrier1, S. Jaillard2, V. Jauffret2, E. Cheron1, M. Fradin1, C. Quelin1, L. Pasquier1, V. David1, C. Dubourg2. 1) Genetique Clinique, CHU de Rennes Hosp SUD, Rennes CDX 2, France; 2) Service d’Anatomie Pathologique, CHU de Caen, 14033 CAEN, France; 3) Service Maternite gynecologie, CH de Saint-Lo, 50009 Saint-Lo, France; 4) Service de Genetique, CHU de Caen, 14033 CAEN, France; 5) Service de cytogenetique, CHU de Rennes, Hospital Pontchaillou, Rennes, France; 6) Service de genetique moleculaire et genomique, CHU de Rennes, Hospital Pontchaillou, Rennes, France; 7) CNRS UMR 6290, IGDR, Universite Rennes 1, France.

Mesomelic dysplasias are a very heterogeneous group of skeletal disorders. Among them, the 2q31.1 microduplication syndrome and Mesomelic Kantaputra dysplasia (MIM 156232) have been associated with microduplication of a cluster of HOXD genes. Phenotype of early fetal mesomelic dysplasia was essentially described in Langer mesomelic dysplasia (MIM 249700) caused by homozygous deletion or mutation of the SHOX gene or dysplasia was essentially described in Kanto dysplasia (MIM 156232) have been associated with microduplication of a cluster of HOXD genes. Phenotype of a mesomelic dysplasia inherited from the father, who was 249700) caused by homozygous deletion or mutation of the SHOX gene or dysplasia was essentially described in Kanto dysplasia (MIM 156232) have been associated with microduplication of a cluster of HOXD genes. Phenotype of a mesomelic dysplasia inherited from the father, who was

2651S


Despite its importance to clinicians and pediatricians, there are relatively few studies currently which examine developmental trajectories or profiles of adaptively and maladaptively behavioral in children with genetic disorders that produce intellectual disability (ID) and/or autism (ASD). Recently, we undertook to study children with subtelomeric deletions that produce ID and/or ASD. We recruited 47 children, 26 females, 21 males, diagnosed with Wolf-Hirschhorn Syndrome (WHS), or one of three other subtelomeric deletions (invdup22p23, 2q37 or 11q23-25 (Jacobsen syndrome; JBS). Mean age for the sample at T1 was 11.1 years (<4.4). Initially, we assessed their cognitive abilities, adaptive behavior (DQ) and maladaptive behavior using standardized instruments at Time 1 (T1). Two years later, at Time 2 (T2), we were able to retest 31 children. We compared adaptive and maladaptive scores at T1, but also compared adaptive and maladaptive difference scores from T1 to T2 for the entire group, and by genetic disorder, DQ score at T1 or maladaptive score at T1. Results show an inverse correlation between DQ score and maladaptive score, but it was not statistically significant. However, results from ANOVA indicate statistically significant test-retest differences in maladaptive scores among the four syndromes (F=3.336; P=0.034). Post-hoc analysis showed 2q37 and JBS maladaptive difference scores were statistically significantly different (P<0.048). Children with WHS were most severely impacted in cognitive and adaptive behavior, but with relatively lower maladaptive behavior scores which tend to improve as children age. Curiously, children with JBS have the highest mean IQ but the worst maladaptive behavior and adaptive behavior scores. We employed a general linear model to analyze maladaptive behavior and found maladaptive scores were significantly related to both genetic disorder and changes in IQ scores from T1 to T2 (F=3.88; P<0.01). Aspects of maladaptive behavior, particularly as they relate to attention deficits and hyperactivity are currently being analyzed and will be discussed.

2652M

Mosaic 15q11-q13 maternal duplication without Autism. N. Urraca1, E. Pivnick2, K. McVicker1, R. Thibe1, S. Pasinotto1, H. Pedra2, L.T. Reiter1,2. 1) Neurology, UTHSC, Memphis, TN; 2) Pediatrics, UTHSC, Memphis, TN; 3) Neurology, Mass General Hospital, Boston, MA; 4) Genetics, UMC, Hackensack, NJ.

Most individuals with an interstitial 15q duplication share common deletion breakpoints with Prader-Willi/Angelman Syndrome (PWS/AS) and are the result of the reciprocal non-allic homologous recombination (NAHR) events that forms the PWS/AS deletions. Because the 15q11-q13 region has a maternal allele preference of NAHR, a parent of origin effect was expected in maternal duplication cases. However, the offspring with the duplication is de novo; however, there are few informative familial cases that will help us better understand the genes contributing to the ASD phenotype. The majority of duplications found in this region are maternal in origin, due to the presence of ASD. We report two families with mosaic inherited 15q duplications: family 1 with a Class II duplication and family 2 with a Class I duplication. Interestingly, in both families the mothers have a maternally derived duplication with no apparent ASD phenotype as would be expected in maternal duplication cases. However, the offspring with the duplication met criteria for ASD by ADOS/ADI-R testing. In family 1, the mother had learning disabilities and ADHD as a child, while the mother in the second family had school/learning issues as well. The Broad Autism Phenotype Questionnaire (BAPQ) and the Social Responsiveness Scale (SRS) did not reveal an ASD phenotype in the mother in family 1 and will be completed in the second mother. Fluorescence In Situ Hybridization in peripheral blood lymphocytes and skin fibroblasts showed mosaic 15q duplication for the first mother. Cases of paternal inher- ited 15q11-q13 have been reported with no obvious autism phenotype, but the child who inherits this same duplication maternally has ASD. The two families we will present came to our attention because of their affected offspring. We were surprised to find a maternal duplication in the mothers as previous studies have indicated a higher risk for an ASD phenotype in the child if the duplication is maternally derived. This finding is important as it expands the explanation of ASD. These two families confirmed our previous finding of ASD being fully penetrant in maternal 15q duplication cases and indicates that in any maternal case with no ASD phenotype a mosaic status needs to be ruled out.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
2653T
1) Gujarati Genetic Diagnostic Center (GenDiCe), Department of Zoology, University School of Sciences, Gujarat University, Ahmedabad-09, Gujarat, India; 2) Bawishi Fertility Institute, Near Adani CNG Station, Opp Municipal Garden, Ahmedabad-07, Gujarat, India.

“Y” chromosome microdeletions play an leading genetic cause of male infertility. Hence, the purpose of present study is to screen various “Y” chromosomal STS markers in infertile male patients. The blood samples were collected from infertile male patients (n=41). DNA was isolated by standard protocol and various “Y” chromosomal STS markers were analyzed. In our study 41 cases, attending the clinic, considered to have infertility were analyzed for semen, hormonal and microdeletion profiles along with other socio-economic factors including previous history. These data showed 3 out of 41 (8%) showed deletions at “Y” chromosome level (AZF regions) leading to infertility in relation to other factors. In the remaining patients (38) in our study the causes for infertility depend on other factors like hormonal imbalance etc. The data revealed that microdeletion is also one of the causes for inducing sterility in the male in addition to others. Thus, 8% in this western population, it may be contributory for infertility, although such infertility caused ranges 10-15% in other regions of India. In conclusion, “Y” chromosome microdeletions play an important role in male infertility, where our region constitutes 8% within our samples analyzed.

2654S
A microdeletion encompassing only three genes within the Potocki-Shaffer syndrome interval at 11p11.2 associated with intellectual disability and craniofacial anomalies. J.D. LAbonne1, J. Vogt2, L.C Layman1, H.G. Kim1, 1) Department of Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA; 2) West Midlands Regional Clinical Genetics Service, Birmingham Women’s Hospital, Birmingham, UK, B15 2TG.

Potocki-Shaffer syndrome (PSS) characterized by intellectual disability (ID), craniofacial anomalies (CFA), biparietal foramina, and multiple exostoses is a rare contiguous gene disorder caused by a minimal 2.1 Mb deletion at 11p11.2. EXT2 and ALX4 have been shown to be causative for multiple exostoses and parietal foramina, respectively. Haploinsufficiency of PHF21A was the underlying mechanism of ID and CFA in PSS since affected individuals with balanced translocations and deletions showed a reduction of PHF21A protein level. Furthermore, knock-down of PHF21A in zebrafish resulted in CFA and neuronal apoptosis. The death of neurons is a likely cause of ID in patients. We report a 5-year old boy with ID and CFA with a deletion including PHF21A, providing corroborating evidence for a role of PHF21A in these specific phenotypes. BlueGnome 8x60K ISCA design was used for the phenotypes of ID and CFA seen in this deletion patient.

2655M

Chromosome microarray analysis in patients with cleft lip/ cleft palate. P. Eydoux1, F. Kozak2, L. Ogilvie1, J. Pauwels2, B. Tsang1, L. Brown1. 1) Dept Pathology, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Genetics, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA.

Chromosome microarray analysis in patients with cleft lip/ cleft palate. P. Eydoux1, F. Kozak2, L. Ogilvie1, J. Pauwels2, L. Brown1. 1) Dept Pathology, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Genetics, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA.

Chromosome microarray analysis in patients with cleft lip/ cleft palate. P. Eydoux1, F. Kozak2, L. Ogilvie1, J. Pauwels2, L. Brown1. 1) Dept Pathology, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Genetics, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA.

Chromosome microarray analysis in patients with cleft lip/ cleft palate. P. Eydoux1, F. Kozak2, L. Ogilvie1, J. Pauwels2, L. Brown1. 1) Dept Pathology, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Genetics, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA.

Chromosome microarray analysis in patients with cleft lip/ cleft palate. P. Eydoux1, F. Kozak2, L. Ogilvie1, J. Pauwels2, L. Brown1. 1) Dept Pathology, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Genetics, Children’s & Women’s Hosp, Vancouver, BC, Canada; 2) Dept Obstetrics & Gynecology, Institute of Molecular Medicine and Genetics, Georgia Regents University, Augusta, GA 30912, USA.
2657S

The purpose of this abstract is to present a family with Xq27.3-q28 duplication syndrome, compare their findings to those of previously reported cases, and provide further delineation of this syndrome that has strong phenotypic overlap with Prader-Willi syndrome. Patient 1, a 7-year-old male who was born at term after an uncomplicated pregnancy. Development first became a concern to the family when he was not able to sit independently at 1 year of age. He continued to have delays in motor and language milestones. At 3 years of age, etiologic evaluation was initiated. Chromosomal micro-array identified duplication of Xq27.3-q28. Diagnosis at presentation to our clinic comprised; mild cognitive disability, strabismus, excessive hunger, obstructive sleep apnea, and hypogonadism. His physical exam is significant for short stature, truncal obesity, synophrys, bulbous nasal tip, small scrotum, cryptorchidism, and hypotonia. Patient 2, maternal cousin of Patient 1, is an 11-year-old male identified independently to have the same duplication. His medical problems and exam findings are congruent with Patient 1’s with the addition of fatty liver and Chiari malformation. In addition to microarray he has had a muscle biopsy and urine glycosaminoglycans which were normal. The mothers of both patients, as well as a third sister, all carry the dupXq27.3-q28. This is the third family reported with Xq27.3-q28 interstitial duplication. The patients described here show significant phenotypic overlap with cases of dupXq27.3-28 described in the literature with key features of hypotonia, gross motor delay, mild intellectual disability, hypogonadism, short stature and obesity. Deficiency of a number of genes in the duplicated region is associated with human disease; FMR1, IDS, MAMLD1, MTM1 and VMA21. FMR1 duplications have also been associated with intellectual disability. Other genes in this area are likely also dosage sensitive. Xq27.3-q28 interstitial duplication causes an X-linked recessive syndrome with a remarkably consistent phenotype, has very strong overlap with Prader-Willi syndrome (hypotonia, cognitive disabilities, short stature, obesity/polyyphagia) and should be considered in the differential for that syndrome. Any patient with phenotypic findings suggestive of Prader-Willi who has negative methylation testing, or who has family history suggestive of X-linked inheritance, should have testing for copy number variants in this region of chromosome X.

2659T

Axenfeld-Rieger syndrome (ARS) is a developmental disorder that associates eye abnormalities and multiple congenital malformations. An abnormal migration of neural crest cells results in dysgenesis of the anterior segment of the eye. Other manifestations are mild dysmorphic features, dental, cardiac and urogenital anomalies. ARS is inherited in an autosomal dominant manner with high penetrance of FOXC1 and PITX2 mutations. However, over 50% of cases remain of unknown cause. Herein, we report a case of ARS associated with epilepsy in a two-year-old boy. Clinical examination found some dysmorphic features and growth delay. Ophthalmologic examination revealed a bilateral corectopia, iris hypoplasia and a bilateral embryotoxon, consistent with ARS eye disorders. Whole genome microarray analysis using Nimblegen 135K array showed a 4.55 Mb de novo deletion on Xq25q26 bands (nt 112,009,675 to nt 116,558,399 bp - hg18). The deleted region encompasses 9 OMIM genes and its proximal boundary is located 230 Kb upstream of PITX2. In silico investigation identified the presence of regulatory sequences of PITX2 in this region. PCR analysis confirmed the deletion of these sequences. Further studies are carried out to assess the disruption of PITX2 expression. This report of ARS arising from deletion of PITX2 gene cis-regulatory elements sequences suggests that analysis of PITX2 upstream region could be helpful for ARS patients without PITX2 or FOXC2 mutations.

2660S
Fetal Skeletal Dysplasias in a Tertiary Care Centre: Radiology, Pathology, and Molecular Analysis of 112 cases. D. Chityal1, P. J. Healy2, U. Bokoya2, A. Mohan2, S. Keating4, J. T. Ford1, J. Frank1, R. Frank1, G. Tomlinson4, P. Glanc2. 1) Mount Sinai Hospital, Department of Obstetrics and Gynecology, The Prenatal Diagnosis and Medical Genetics Program, University of Toronto, Toronto, Ontario, Canada; 2) The Hospital for Sick Children, Division of Clinical and Metabolic Genetics, University of Toronto, Ontario, Canada; 3) Department of Medical Imaging, South Shore Regional Hospital, Bridgewater, NS, Canada; 4) Department of Obstetrics and Gynecology, University of Calgary, Calgary, AB, Canada; 5) Mount Sinai Hospital, Department of Laboratory Medicine and Pathobiology -Perinatal Pathology; 6) Mount Sinai Hospital, Department of Diagnostic Imaging; 7) University of Toronto, Institute of Health Policy, Management & Evaluation; 8) Sunnybrook Health Sciences Centre, Department of Medical Imaging.

Fetal skeletal dysplasias are a heterogeneous group of rare genetic disorders, affecting approximately 2.4-4.5/10,000 births. We performed a retrospective review of the perinatal autopsies conducted between the years 2002-2011. The study population consisted of fetuses diagnosed with skeletal dysplasia with subsequent termination, stillbirth and live-born who died shortly after birth. Of the 2,002 autopsies performed, 112 (5.6%) were diagnosed with skeletal dysplasia. These 112 cases encompassed 17 of the 40 groups of Nosology 2010. The purpose of this abstract is to present a family with Xq27.3-q28 deletion, using array CGH to assess the disruption of FOXC2 and PITX2. The deletions identified were Xq27.3-q28 deletions in Patient 1, a 7-year-old male who was born at term after an uncomplicated pregnancy. Development first became a concern to the family when he was not able to sit independently at 1 year of age. He continued to have delays in motor and language milestones. At 3 years of age, etiologic evaluation was initiated. Chromosomal micro-array identified duplication of Xq27.3-q28. Diagnosis at presentation to our clinic comprised mild cognitive disability, strabismus, excessive hunger, obstructive sleep apnea, and hypogonadism. His physical exam is significant for short stature, truncal obesity, synophrys, bulbous nasal tip, small scrotum, cryptorchidism, and hypotonia. Patient 2, maternal cousin of Patient 1, is an 11-year-old male identified independently to have the same duplication. His medical problems and exam findings are congruent with Patient 1’s with the addition of fatty liver and Chiari malformation. In addition to microarray he has had a muscle biopsy and urine glycosaminoglycans which were normal. The mothers of both patients, as well as a third sister, all carry the dupXq27.3-q28. This is the third family reported with Xq27.3-q28 interstitial duplication. The patients described here show significant phenotypic overlap with cases of dupXq27.3-28 described in the literature with key features of hypotonia, gross motor delay, mild intellectual disability, hypogonadism, short stature and obesity. Deficiency of a number of genes in the duplicated region is associated with human disease; FMR1, IDS, MAMLD1, MTM1 and VMA21. FMR1 duplications have also been associated with intellectual disability. Other genes in this area are likely also dosage sensitive. Xq27.3-q28 interstitial duplication causes an X-linked recessive syndrome with a remarkably consistent phenotype, has very strong overlap with Prader-Willi syndrome (hypotonia, cognitive disabilities, short stature, obesity/polyyphagia) and should be considered in the differential for that syndrome. Any patient with phenotypic findings suggestive of Prader-Willi who has negative methylation testing, or who has family history suggestive of X-linked inheritance, should have testing for copy number variants in this region of chromosome X.

2658M
Behavioral, Biochemical and Anthropometric Characteristics of patients with PWS. H. El-Bassyouni, E. M. Salisi, S. Khoulous, M. Shehab, W. Kandeef. 1) Clinical Genetics Department, National Research Centre, Cairo, Egypt; 2) Child Health Department, National Research Centre, Cairo, Egypt; 3) Immunogenetics Department, National Research Centre, Cairo, Egypt; 4) Cytogenetic Department, National Research Centre, Cairo, Egypt; 5) Biological Anthropology Department, National Research Centre, Cairo, Egypt.

Background: Prader-Willi syndrome (PWS) is a genetic disorder characterized by a recognizable pattern of physical findings with significant cognitive, neurologic, endocrine, and behavioral abnormalities. Objective: To study behavioral, cognitive, hormonal and anthropometric characteristics of children with Prader-Willi syndrome (PWS) compared with an age-and gender-matched control group. Subjects and methods: A case- control study of thirteen children and adolescents with PWS (mean age 7.69±4.44 years), and 14 age-and sex-matched non-PWS controls were enrolled. Measurement of anthropometric parameters and body fat percentage (BF%) screening of childhood dysphagia, assessment of intellectual function and estimation of plasma adiponectin and leptin levels were carried out in all studied subjects. Specific questionnaire for assessment of the behavioral phenotype of PWS was fulfilled by parents. Results: PWS subjects had significantly higher BF%, adiponectin and leptin levels compared to controls (38.8±5.66 vs. 20.26±3.92, P<0.01; 38.6±13.43ng/ml vs. 24.32±6.04ng/ml, P<0.01; 10.39±7.74ng/ml vs. 4±0.71 ng/ml, P<0.01 respectively). The majority of PWS subjects (84.6%) (11/13) had mild to moderate mental retardation. The most commonly registered behavior problems in PWS subjects were hyperphagia and cognitive rigidity. Inattention was reported in 76.9% (10/13) of PWS vs. 25% (3/14) of the controls. Conclusion: PWS subjects have substantial behavioral problems which are not associated with the degree of hyperphagia or intellectual disability. The distinct body fat distribution in PWS subjects is associated with significantly higher levels of adiponectin and leptin than their lean controls. The primary focus on management of PWS should be placed on weight control and behavior modification.
Osteogenesis imperfecta (OI) is a clinically and genetically heterogeneous disease characterized by fragility of bones and other soft tissue. OI type XI is a rare form of this disease, which is inherited in an autosomal recessive manner; mutations are described in the FKB10 gene. Epidermolysis bullosa (EB) is a group of inherited diseases of the skin manifested in formation of bubbles on the skin because of mechanical trauma. Mutations of simple type are frequently identified in genes of keratin 5 and 14 (KRT5, KRT14). Here is a description of the case of the combination of OI type XI and simple type of EB. The proband is a 4 years old girl from the State Medical University, Moscow, Russian Federation; 2) MV Vladimirsky University, Moscow, Russia. The child was admitted to the orphanage with a diagnosis of OI (fracture of the left humerus and of the right radius, multiple rib fractures in the medical history), the left hip dysplasia, EB of the simple type. Repeated cases of spontaneous fractures during the observation period happened. Examination: weight -12 kg, height - 99 cm, can't sit by her own, can't walk. Light blue sclera. Sweaty skin, multiple foci of de-and hyperpigmentation, blisters on the skin of the hands, elbow bends were observed. Feet hyperkeratosis. Pigion breast, thighs saber deformity, scoliosis, flat feet, hypermobility of joints of the hands, limitation of movement in the hip and knee joints, muscular hypotonia. There were no other singularities in other systems. Analysis of the literature showed that there was one case of combination of OI type XI and simple type of EB, where were identified homozygous mutations in FKB10 and KRT14 genes. We conducted the ArrayCGH to confirm the assumption of the similarity of our case with the previously described one and to narrow the candidate genes list for further genetic diagnosis. Results: stretches of homozygosity > 3,000,000 bp were identified in the proband. This value exceeds the population level 10 times that may indicate kinship parents. One of the longest stretches of homozygosity located on the long arm of chromosome 17, where the genes FKB10 and KRT14 are located. This result indirectly confirms our hypothesis about the possible mutations in the homozygous state in the genes responsible for the development of OI and BE. Sequencing of these genes in work.

MuSK - a new target for lethal fetal akinesia deformation sequence (FADS). M. Wilbe1, S. Ekvall2, N. Aneren1, M.-L. Bondeson3, 1) Immunology, Genetics and Pathology, IGP, Uppsala, Uppsala, Sweden; 2) Department of Women’s and Children’s Health, Uppsala University, Uppsala, Sweden.

Fetal akinesia is clinically and genetically heterogeneous disorders, with the common feature defined as reduced or loss of fetal movement. Several disease genes of fetal akinesia have been described. This includes genes involved in motor neuron development and survival, encoding components of the neuromuscular junction (NMJ), adult skeletal muscle proteins and fetal myoskeletal proteins. However, the genetic etiology of majority of cases with fetal akinesia is still unknown.

We report on a family with recurrent fetal loss, where the parents had five affected fetuses with fetal akinesia deformation sequence (FADS [MIM 208150]) and one healthy child. The fetuses displayed no fetal movements from the gestational age of 17 weeks, extended knee joints, flexed hips and elbows and clenched hands. There was polyhydramni and no visible fetal stomach.

Whole exome sequencing (WES) of one affected fetus and the parents were performed. A recessive inheritance model was assumed and data was filtered against dbSNP, ANNOVAR and ~900 exomes in our in-house database. Only one candidate homozygous variant was identified in the fetus, c.40dupA (p.Thr14Asnfs*9), located in the first exon of MuSK (muscle, skeletal, receptor tyrosine kinase). The variant was verified using Sanger sequencing. Segregation analysis in the family revealed homozygosity for all affected fetuses, while the variant was not present in the healthy child. The c.40dupA variant leads to a frameshift in MuSK predicting a premature stop codon, affecting the muscle-specific localization of protein. MuSK is an agrin-dependent receptor tyrosine kinase required for formation of the neuromuscular junction and misfolding mutations in this gene have previously been described in congenital myasthenic syndrome (CMS). Interestingly, MuSK is located in the same acetylcholine receptor pathway as several other genes reported to cause CMS and/or FADS (CHRNA1, CHRNB1, CHRNA4, CHRNA5, RAPSN, DOK7, CNTN1 and SYNET1). To our knowledge, this is the first time MuSK is reported to the spectrum of FADS and we propose that MuSK should be included in genetic analysis and prenatal screening for FADS.

2662T INCONTINIENTIA PIGMENTI: A Case Report Associated With Cleft Lip Palate in a Patient at Smile Operation Foundation, Bogotá - Colombia. m. montiel, l. brceno, j. martinez, a. patino, j. rincon. medicine, universidad de la sabana, chia, Colombia.

Introduction Incontinentia pigmenti is a neuroectodermal disease, with an autosomal X-linked transmission, also known as Bloch Sulzberger syndrome (1). Epidemiology The Bloch syndrome is a genetic disease X-linked dominant inheritance, with a reported incidence of 1 case per 50,000 live births (2). Genetic. They are two types of the disease: sporadic and genetic IP. The gene involved in the development of IP1 is located in Xp11 chromosome (3). On the other hand, mutations that occur in the NEMO gene (essential modulator of nuclear factor Kb) located on chromosome Xq28 , are responsible for IP2 , with a dominant X-linked inheritance (4). Histology of Ito Fig 3. Associated cleft palate, dental abnormalities and alopecia. The case was diagnosticated at Operation Smile Foundation in 2011. Fig 1. The child presents hypomelanosis of Ito in the back Fig 2. The abdomen has hypomelanosis of Ito Fig 3. Associated cleft palate, dental abnormalities and alopecia. Diagnosis for the diagnosis of incontinentia pigmenti, there are some criteria simplified by Landy et al . , but the management is based on an accurate physical examination which must be addressed to the search for skin disorders which are present in all patients, dental (absence of deciduous or permanent teeth, delayed in the eruption, deformations and conical crowns and hook) teeth, and eye CNS disorders (7). When the physician has the clinical suspicion, the histopathology analysis of the lesions is the method that identifies the type of injury and the stage at which the disease is found (7).
An interstitial microdeletion of 4q21 in a girl with pituitary insufficiency associated with empty sella, epilepsy, severe growth impairment, and profound intellectual disability.

E. Nishi1,2, K. Wakui3, M. Arakawa3, S. Hirabayashi4, Y. Fukushima1, T. Koshio1,5.

1) Dept Genetics, Nagano Children’s Hosp, Azumino City, Nagano, Japan; 2) Dept Medical Genetics, Shinsyu University School of Medicine, Matsumoto, Japan; 3) Dept Paediatric Neurology, Nagano Children’s Hosp, Azumino, Nagano, Japan. 

Microdeletion 4q21 syndrome (OMIM#613509) is characterized by severe intellectual disability, lack of speech, hypotonia, significant growth restriction, and distinctive facial features. The deletion sizes varied from 2.0 to 15.1 Mb, and the common features have been associated with a 1.37 Mb minimal critical region including five known genes: PRKG2, RASGEP1B, HNRNPDL, HNRPDL, and ENOPHT1 [Mitchell et al., 1981; Nowaczyn et al., 1997; Harada et al., 2002; Friedman et al., 2006; Bonnet et al., 2010; Dukes-Rimsky et al., 2011; Lipska et al., 2011; Tsang et al., 2012; Bhoj et al. 2013]. We report a girl with a de novo interstitial microdeletion of 4q21, who showed pituitary insufficiency associated with empty sella, epilepsy, severe growth impairment, and profound intellectual disability. The patient was the first child of a healthy 26-year-old mother and a healthy 30-year-old non-consanguineous father. She was born by cesarean-section at 38 weeks and 6 days of gestation. Her birth weight was 2206g (-2SD), length was 43.4cm (-2.4 SD), and OFC was 32.8 cm (-0.07 SD). She showed hypotonia and severe developmental delay. She raised her head at the age of 9 months, and walked unsupported in a wide gate at the age of 6 years. She also spoke no words, and profound intellectual disability. Brain MRI showed thinning of the pituitary gland with empty sella. At the age of 10 years, she had severe growth impairment with the weight of 20.5 kg (-1.8 SD), the height of 112 cm (3 SD), and the OFC of 49.1 cm (-2SD). She was found to have growth hormone deficiency and hypoplastic pituitary gland with empty sella, and growth hormone replacement therapy was started. G-banded karyotype was 46,XX,inv(12)[p13.13]. Microarray analysis revealed a de novo 2.16 Mb microdeletion at 4q21 ([arr[hg18] 4q21.22q21.23(29,297,619-30,447,117)x1 dn]). The deleted segment encompassed 15 OMIM genes (HNRNPDL, HNRPDL, SC5D, SEC31A, THAP9, LIN54, PLAC8, CO2, HPSE, HELQ, MRPS18C, FAM175A, and FAM175) that are involved in epigenetic regulation, neuronal development and differentiation, and immune response. This is the first patient with an interstitial deletion of 4q21 who showed pituitary insufficiency with empty sella.
2668T Clinical Implementation of Chromosome Microarray Analysis in Singapore, H. Law1, M. Brett2, M. Yong2, H. Yoon2, R. Roch2, M. Tan2, H. Ee2, E. Tan3, B. Cham1, J. Lim1, E. Tan1, I. Ny1, S. Jamuar1, A. Lai1. 1Pediatric Medicine, KK Women’s & Children’s Hospital, Singapore; 2KK Research Centre, KK Women’s & Children’s Hospital, Singapore; 3Department of Pathology and Laboratory Medicine, KK Women’s and Children’s Hospital, Singapore. Background: Chromosome microarray analysis (CMA) is recommended as the first-tier genetic test for children with intellectual disabilities, development delay, autism spectrum disorder and/or multiple congenital anomalies. Although CMA is readily available in USA, our lab only recently started offering it as a clinical test. Method: From May 2013 to May 2014, 100 patients were screened using the Agilent 4x180K CGH+SNP array. The indications for testing included developmental delay, intellectual disability, autism spectrum disorder and/or multiple congenital anomalies. Results: Copy number variants (CNVs) ranging in size from 10kb to 36.4 Mb were found in 38 patients (38%). Pathogenic and likely pathogenic CNVs were found in 18 (18%) patients. These included 14 deletions, 3 duplications and a patient with both a deletion and a duplication. Recurrent microdeletion and microduplication syndromes including the 1q21.1 microdeletion (2), Williams syndrome (2), Cat eye syndrome, Cri du Chat syndrome, Miller Diecker syndrome, 3q29 microdeletion, 15q24 microdeletion, and 1q43q44 syndrome were detected in our patients. CNVs of uncertain clinical significance were detected in 20 (20%) individuals: 11 were duplications and 9 were deletions. However, as Singapore is a self-pay health care system, parental testing could not be performed in all 20 cases and hence, significance of these variants could not be established conclusively. Conclusion: CMA is a powerful tool in identifying pathogenic chromosomal copy number alterations. However, in countries with self-pay health care systems like Singapore, cost plays an important role in successful implementation of a diagnostic clinical test. In addition, more data relating to the local population needs to be collected to enhance interpretation of the CMA results.

2669S Natural history and clinical management of patients with ASXL1 mutations and Bohring-Opitz Syndrome, including the first report of Wilms Tumor in two patients. B. Russell1, N. Kramer2, L. Biepecker1, J. Johnston3, W. Rhead2, A. Pickard2, A. Dobson2, L. Clarkson2, J. Graham1. 1Medical Genetics, Cincinnati Children’s Hospital and Medical Center, Cincinnati, OH; 2Medical Genetics Institute, Cedars Sinai Medical Center, Department of Pediatrics, Harbor-UCLA Medical Center, David Geffen School of Medicine at UCLA, Los Angeles, CA; 3National Human Genome Research Institute, National Institute of Health, Bethesda, MD. Introduction: Bohring-Opitz syndrome is a rare genetic condition (35 reported cases) characterized by distinct facial features (glabellar nevus flammeus, wide spaced eyes, depressed and wide nasal bridge, antverted nares, palatal anomalies, micrognathia, low-set posteriorly rotated ears), typical posture (elbow flexion with ulnar deviation and flexion of the wrists and metacarpophalangeal joints), variable microcephaly, severe intellectual disability, hyperthecosis, and feeding problems. Nine published patients with Bohring-Opitz syndrome have been reported as well as 2 with ASXL1. We report natural history and clinical management of five previously unpublished patients with mutations in ASXL1. Natural History: The 5 patients we discuss range from 2-12 years of age with complicated medical histories that included feeding issues, cyclic vomiting, respiratory infections, insomnia and Wilms tumor. Severity of illness improved with age after the first 1-2 years of life. Dysmorphic features such as nevus flammeus also faded with age. Severe myopia was present in all patients. They also had distinctive personalities (interactive, happy, and independent), alopecia areata, and sparse hair growth which were features not described in the literature. Clinical Management: Permanent feeding tubes due to gastric dysmotility, silent aspiration, chronic emesis and poor weight gain were required in 4 of the 5 patients. Two patients had cyclic vomiting that was managed with cyphe- trophamide, lorazepam, omeprazole, ondansetron and acetaminophen. Recurrent respiratory infections with components of reactive airway disease occurred in 3 patients and 1 patient required a tracheostomy. Insomnia was a significant challenge for 4 of the patients, 2 of which improved with treatment of their severe aspirations, and 3 patients had obstructive sleep apnea that improved with CPAP or mandibular distraction. With the occurrence of bilateral Wilms tumor in one of the patients, screening renal ultrasounds were recommended to the other patients. This led to the identification and treatment of bilateral Wilms tumor in a second patient. Given the known association between ASXL1 and myeloid malignancies along with the two patients presented here, consideration of Wilms tumor screening in patients with ASXL1 mutations will be discussed.

2670M Clinical and epidemiological study of orofacial clefts. S. RASKIN, J. SOUZA. Group for Advanced Molecular Investigation (NIMA), Graduate Program in Health Sciences, School of Medicine, Pontificia Universidade Católica do Paraná (PUCPR), Curitiba, Paraná, Brazil. Cleft lip with or without cleft palate (CL±P) or cleft palate (CP) are groups of malformations named orofacial clefts (OC), which are the second leading cause of birth defects. This study aimed to analyze clinical and epidemiological features of Brazilian patients with OC, studying cases treated in the reference center of the state of Paraná (PR). Methods 2,356 charts were reviewed and 1,838 were evaluated by the same clinical geneticist. Data were collected in the reference center, and compared with those of the Health Department of the state of Paraná. Clinical characteristics, presence of other anomalies, and birth prevalence were evaluated. 389 (21.2%) patients had CP, 437 (23.8%) had cleft lip (CL) and 1,012 (55%) had cleft lip and palate (CLP). Syndromic OC were identified in 15.3% of patients, 10.4% of patients with CL±P, and 33.9% of patients with CP. Common additional anomalies were: central nervous system, limbs, cardiovascular, and musculoskeletal defects. The number of syndromic cases was lower when clinical evaluation was performed by other medical specialists when compared to that of the clinical geneticist. Birth prevalence was 1,101.0 live births. Lack of notification with the national birth registry was observed in 49.9% of CL±P. The present data suggests a decrease of 18.52% in the prevalence of non-syndromic OC after folic acid fortification in Brazil. Conclusion: Better understanding of clinical and epidemiological aspects of OC is crucial to improve the understanding of pathogenesis, promote preventive strategies, and guide clinical care, including the presence of clinical geneticists in the multidisciplinary team for OC treatment.

2671T Lateral Meningoceles (Lehman Syndrome): A Rare Connective Tissue Disorder Craniofacial Dysmorphism. M. Carter1, S. Blaser2. 1) Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada; 2) Diagnostic Imaging, The Hospital for Sick Children, Toronto, ON, Canada. Multiple lateral meningeoceles are a rare anomaly. Lehman et al. (1977) first described a mother and daughter with similar craniofacial dysmorphisms, skeletal sclerosis, and multiple meningeoceles. Subsequently, eight more cases have been described (Philip et al., 1995; Gripp et al., 1997; Chen et al., 2005 and Correia-Sa et al., 2013). The causative gene for Lehman syndrome is not yet known. The patient described in this abstract is a two-year-old boy with Lehman syndrome, born to a non-consanguineous Caucasian couple. Prenatally, a complex cardiac malformation was diagnosed (tubular hypoplasia of the aortic arch with coarctation, large ventricular septal defect, hypoplastic aortic valve, and bilateral superior vena cavae). MRI of brain showed middle cranial fossa encephaloceles and Chiar I malformation, with cerebellar tonsils extending to the posterior arch of C1. After discharge at 2 months of age, parents noted bilateral flank swelling. An ultrasound of the abdomen showed lateral lumbar meningeoceles displacing the kidneys and causing the “swelling” due to deficient posterior wall of the bladder. Spine MRI showed enlarged spinal canal, and bilateral exten- sive lateral neural foramina meningeoceles. The baby was discharged with NG tube in situ for feeding, and eventually required gastrostomy due to dysphagia. He was diagnosed with moderate bilateral conductive hearing impairment; hearing aids were prescribed. He has left amblyopia. Cardiac status has been stable. Dysmorphic features include left posterior plagiocephaly and tall cranial vault, sparse hair, hypoplastic supraorbital ridges, epicranius, hypertelorism, ptosis, downsloped palpebral fissures, midface hypoplasia, small nares, long philtrum and thin upper vermilion, highly arched palate, bifid uvula, short upper lingual frenulum, asymmetric low set ears with short canals, significant microretroglossa, and bilateral single palmar creases. Neurological examination is significant for paucity of facial movement and diffuse hypotonia. Development at 13 months of age was significantly delayed. We present a child with Lehman syndrome, and review the cases reported to date. International collaboration will be useful for discovery of the gene(s) responsible for this complex condition.
2672S

We report here the case of a female child with multiple congenital abnormalities and intellectual disabilities(MCA/ID) and growth retardation whose dysmorphic features and anomalies overlap those observed in CHARGE syndrome. In this patient we identified a de novo EP300 splice mutation by whole exome sequencing. The patient was the firstborn of nonconsanguineous parents(37-year-old father and 32 year-old mother). She was born at 37 weeks gestation by caesarian section because of polyhydramnios. Her weight and length were 2.350 g (+1.4 SD) and 41 cm (-3.5 SD), and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallowing problems which required gastrostomy, strabismus, and scoliosis. Her facial features included a square outline, sparse hair, long eyelashes and severely malformed earlobe. She had moderate to severe intellectual impairments and her occipitofrontal circumference was 32.5 cm (-0.5 SD). She was noted to have a congenital heart disease, esophageal hiatal hernia, small cleft palate, choanal stenosis, deformed earlobes, and congenital hydronephrosis at birth. She underwent several surgical procedures for correction of a large ventricle septal defect and third degree vescicoureteral reflex. At three years of age, she showed severe growth retardation, complete deafness, swallow-
2675S

De novo heterozygous deletion involving NFIX in a Japanese subject with severe intellectual disability, postnatal growth delay and relative macrocephaly. D.T. Uehara1, S. Hayashi2, S. Mizuno3, J. Inazawa2. 1) Department of Molecular Cytogenetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan; 2) Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Japan.

In this report, we describe a Japanese patient with a 394 kb de novo deletion at 19p13.2 involving NFIX gene (OMIM *164005), detected by SNP arrays. Clinical features include severe ID, hypotonia, small intestinal atresia, low stature (-3.5 SD) and weight (-2.5 SD) along with a relative macrocephaly (+0.7 SD) at the age of seven years old, although her birth size was normal. NFIX belongs to the nuclear factor one (NFI) family, which encodes four transcription factors essential for normal development. Previous reports have identified NFIX heterozygous deletions and point mutations in a few patients whose common features were ID, overgrowth and macrocephaly. A subset of those reports were described in patients with Sotos-like or Marshall-Smith syndromes, two overgrowth disorders characterized by advanced bone age, and in the latter case, skeletal anomalies. Our patient is the first description of haploinsufficiency of NFIX associated with postnatal growth delay. A possible explanation for this case could come from two NFIX mouse models with marked differences in phenotypes: in one model, Nfix deficiency produced brain malformation and severe skeletal defects, while the other primarily showed defects in brain development. Therefore, this presents the second report of NFIX in development. Nevertheless, the contribution of the other genes (LYL1, TRMT1, NACC1, STX10, IER2 and CACNA1) encompassed by the rearrangement cannot be discarded, as well as other variants elsewhere.

2676M

Facial dysmorphism, skeletal abnormalities and central nervous system abnormalities in two sibs born to a consanguineous couple: A new autosomal recessive condition. L. Chad1,2, M. Thompson1, I. Miron1, P. Shannon4, S. Keating2, D. Chitayat1,2, 1) Department of Obstetrics and Gynecology, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada; 2) Department of Pediatrics, Division of Clinical & Metabolic Genetics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada; 3) Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada.

We report two female fetuses born to a consanguineous Sri-Lankan couple with facial dysmorphism, central nervous system and skeletal abnormalities. To our best knowledge this is a hitherto new autosomal recessive condition. The fetuses, both female, presented with thickened nuchal folds, echogenic bowel and kidneys, rocker-bottom feet, ventriculomegaly and intrauterine growth restriction. Detailed autopsies following termination of pregnancy at 23.4 and 22.3 weeks gestation respectively revealed short sloped forehead and hypertelorism with webbing of the neck, hydrocephalus with aqueduct stenosis as well as marked narrowing of the spinal canal and platyspondylidy with delayed ossification and flattened acetabular roofs, broad hands with brachydactyly and narrow wrists. Microarray analysis was normal on both. These findings likely represent a new genetic syndrome with most probably autosomal recessive mode of inheritance. Whole genome sequencing is being done to try and identify the causative gene.

2677T

Paraspinal neurofibromas in LEOPARD syndrome. E. Conboy1, R. Dhamija2, B. Babovic-Vuksanovic3. 1) Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN; 2) Department of Medical Genetics, Mayo Clinic, Rochester, MN.

LEOPARD Syndrome (LS) is an autosomal dominant disorder characterized by lentigines, EKG abnormalities, ocular hypertelorism, pujmonic stenosis, abnormal genitalia, growth retardation and deafness. There is significant clinical overlap between LS and other disorders that result from dysregulated Ras/mitogen-activated protein kinase (MAPK) pathway, including Neurofibromatosis type 1 (NF1) Noonan syndrome, Costello syndrome, Legius syndrome and cardio-facio-cutaneous syndrome. Except for NF1, other RASopathies are not known to be associated with development of neurogenic tumors. We describe two unrelated adult patients with clinical diagnosis of LS and massive paraspinal neurofibromas (dumbbell neurofibromas). Both patients were initially evaluated for NF1, but on the clinical exam they had lentigines, ocular hypertelorism, hearing loss, and positive family history lentigines. One of patients was found to have a heterozygous mutation ([T468M]) in the PTEN11. Dumbbell neurofibromas are an unusual complication of LEOPARD syndrome and may be an under-recognized manifestation of this disorder. We suggest surveillance for internal neurofibromas in patients with LS, since the risk for development of malignant peripheral nerve sheet tumors in these patients may be increased.

2678S

Transcriptional hallmarks of Neurofibromatosis type I in whole blood cells. G. Picco1,2, F. Natacci3, E. Trisolini4, D. Cantarella1, C. Cesaretti1, G. Mellon1, E. Riberi5, F. Dutto6, M. Ciniol Silengo5, S. Vannelli5, E. Medico1, G.B. Ferrero1. 1) Department of Oncology, University of Torino, Torino, Italy; 2) Candiolo Cancer Institute, FPO IRCCS, Candiolo, Italy; 3) Medical Genetics Unit, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milano, Italy; 4) Department of Medical Science, University of Torino, Torino, Italy; 5) Department of Pediatrics, University of Torino, Torino, Italy.

Neurofibromatosis type I (NF1), an inherited neurocutaneous disease that has a major impact on the nervous system, eye, skin, and bone, is one of the most common congenital non-chromosomal disorders affecting development and growth (1 in 3000 live births). This complex disorder is characterized by an extremely variable clinical presentation and Individuals with NF1 have a predisposition to benign and malignant tumor formation. Agressive neoplasms as malignant peripheral nerve sheath tumors (MPNST) and juvenile myelomonocytic leukemia (JMML) can be observed in a significant fraction of patients. NF1 is caused by a wide spectrum of mutations affecting the NF1 gene, resulting in loss of function of the gene product neurofibrin, a negative regulator of Ras. These mutations reduce inhibitory activity of the protein, promoting Ras/MAPK pathway signaling, which regulates cell proliferation and differentiation by ultimately controlling gene expression. To investigate the transcriptional consequences of the aberrant activation of the RAS signaling driven by NF1 molecular lesions, we performed Global mRNA Expression Profiling (GEP) in human blood cells, a target tissue of the syndrome. In details, we analyzed 43 samples from molecularly defined NF1 patients and 17 additional samples collected from age- and sex-matched controls. Total RNA extracted from whole blood (PAXgene RNA collection tubes) was processed for expression profiling on Illumina Beadarrays. Subsequently, GEP analysis allowed the identification of a transcriptional signature composed by approximately 100 genes differentially expressed between NF1 cases and control samples. Interestingly, the gene expression patterns highlight a subset of genes subdividing NF1 samples in distinct subgroups. This evidences suggest that the transcriptome of NF1 patients presents clear elements of heterogeneity, possibly reflecting the clinical variability typical of this syndrome. These data establish peripheral blood expression profiling as a powerful tool to appreciate perturbations driven by germine mutations affecting the NF1 gene.
Fibroblastosia ossificans progressiva (FOP): A case report.

Introduction: Fibroblastosia ossificans progressiva (FOP) (OMIM #135100) is a rare autosomal dominant disease with complete penetrance caused by heterozygous mutation in the ACVR1 gene (102576) on chromosome 2q23. 95% of FOP cases are sporadic. FOP involves progressive ossification of skeletal muscle, fascia, tendons, and ligaments. FOP has a prevalence of approximately 1 in 2 million worldwide. (Petrie et al. 2009). The treatment is mainly preventive including avoidance of trauma, passive physiotherapy and prevention of chest infections. (Dhamangaonkar et al. 2013). Objective: To present a FOP male patient. Case report: The proposi-
tus aged 21 year-old presented, at age of 15 years, intense pain on right dorsal and lumbar region, later developed a painful lump on right costal region and limitation movement of right arm. He was diagnosed with FOP. Physical examination at present: weight 64kg, height 1.72 m, sparse eye-
brows, cervical movement restriction, in posterior thorax presented three masses (2 x 2cm each), two on right scapula and one on the left scapula.

Osteogenesis imperfecta (OI) is a heterogeneous genetic disorder characterized by bone fragility and fractures. Patients with OI have clinical features that may range from mild symptoms to severe bone deformities and neonatal lethality. OI is classified into five different types, the most severe one being type I, and the mildest one being type V. The former patients are generally lethal within 150 days of life. The silencing of the referred mutations is considered to be autosomal recessive (AR) inheritance indicating a higher percentage of recessive types in our community compared to Western countries as a result of a higher consanguinity rate in our population (20-40%). For clinical purposes, we proposed a quantitative phenotype evaluation and management in osteogenesis imperfecta: Egyptian Experience. M.S. Aglan1, G.A. Otalily1, R. El-Hous-
insi1, M.S. Abder-Hamid2, V.L. Ruiz-Perez3, S. Temtamy1. 1) Clinical Genetics Department, Center of Scientific Excellence for Human Genetics, Human Genetics and Genome Research Division, National Research Centre, Cairo, Egypt; 2) Medical Molecular Genetics Department, Center of Scientific Excellence for Human Genetics, Human Genetics and Genome Research Division, Cairo, Egypt; 3) Centro de Investigaciones Biomedicas, Consejo Superior de Investigaciones Científaas Universidad Autónoma de Madrid, Madrid, Spain; 4) Centro de Investigación Bone/daica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III (ISCIll), Madrid, Spain.

Hajdu-Cheney syndrome is a rare, autosomal dominant syndrome characterized by skeletal dysplasia, characteristic craniofacial and dental features, acro-osteolysis and proportionate short stature. We describe, here, a 35-year-old male patient with typical clinical and radiological features of Hajdu-Cheney syndrome. His height was 164 cm. He had typical facial features including hypertelorism, bushy eyebrows, micrognathia, dental anomalies, low-set ears, short neck and short fingers. X-ray studies showed wormian bones in the skull, acro-osteolysis of distal phalanges, short bowed long bones. On echocardiography, minimal mitral and aortic regurgitation were observed. Otolaryngological examination revealed a conductive hearing loss. Regarding clinical findings, he was considered to have Hajdu-Cheney syndrome. Molecular analysis showed a heterozygous truncating c.6616 G-T (p.E2206X) mutation in the last exon of NOTCH2 gene. This Hajdu-Cheney case with a novel mutation is the first case whose molecular diagnosis was performed in Turkey and may help to establish phenotype-genotype correlation in the syndrome.
Expanding the diagnostic spectrum of terminal transverse limb defects: atypical mutations in ACVR1 result in a phenotype with elements of Adams Oliver Syndrome and Fibродysplasia Ossificans Progressive. R. Mendoza-Londono1, A. Al Maawal1,2, L. Dupuis3, F. Hyland1, C. Scafe2, T.A. Paton4, C.R Marshall2,5. 1) Division of Clinical and Metabolic Genetics, The Hospital for Sick Children and University of Toronto, Toronto, Canada; 2) Department of Genetics, Sultan Qaboos University Hospital, Sultan Qaboos University, Muscat, Oman; 3) Thermo Fisher Scientific, 200 Oyster Point Boulevard • South San Francisco • CA; 4) The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada; 5) Division of Molecular Genetics, Department of Pediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, ON, Canada.

Reduction limb defects have an estimated incidence of 1 in 1692 live births, with terminal transverse limb defects (TTLD) being the most common. Diagnosis is more common in patients with TTLD include: amniotic bands, teratogenic exposures, Adams Oliver Syndrome (AOS), Poland Anomaly, Moebius and Oromandibular-limb hypogenesis syndromes. Vascular disruption has been thought to play a major role in all of these disorders. We present an 11 year old boy of Italian descent with TTLD, hypotrichosis patchy alopecia suggestive of cutis aplasia, intellectual disability and brain malformations. We performed Proton AmpliSeq whole-exome sequencing with a trio design and prioritized variants in genes known to be associated with AOS (ARHGAP31, RBPJ, DOCK6 and EGOT). No pathogenic variants were identified in these genes. Further analysis of the sequence data revealed a de-novo variant in the ACVR1 gene (c.983G>A; p.G328E) that resides in a well-conserved glycine residue in the protein kinase domain of ACVR1. Mutations in ACVR1 have been previously associated with Fibrodysplasia Ossificans Progressiva (FOP), a syndrome characterized by heterotopic ossification of soft tissues. Most patients with FOP carry the same mutation in the glycine-serine rich domain (c.617G>A; p.R206H), have normal intelligence and great toe malformations. Only 3% of cases of FOP have atypical forms, and only 4 patients have been described with the G328E mutation. These patients present with TTLD, absent nails sparse scalp hair and cognitive impairment, supporting the causative role of this mutation. In addition the 5 other patients with atypical FOP and TTLD reported to date share mutations within the same region of the ACVR1 gene (c.983G>A; p.G328E). We analyzed the molecular pathways disrupted in AOS and postulate that the Cdc42/ Rac1 regulator (ARHGAP31 and DOCK6 genes) and the Notch pathways (RBPJ) converge in a common signaling route that is critical for vasculogenesis. Activating mutations in ACVR1 in endothelial cells have been shown to cause endothelial-to-mesenchymal cell fate transition, which could result in abnormal vascular development. Together, this data suggests that the original theory of vascular injury and impaired arterial angiogenesis in AOS and TTLD is still valid. With this report, we expand the diagnostic spectrum of TTLD to include atypical mutations of ACVR1 and suggest a unifying pathophysiological mechanism linking vascular disruption events to single gene mutations.

A novel homozygous mutation in FGFR3 causes tall stature, severe lateral bialateral deviation, scoliosis, hearing impairment, camptodactyly, and arachnodactyly. L. Chuchalin1,2,3, C.R Marshall4,5,6, S.A. Temtamy4,5. 1) Section of Medical Genetics, SUNY Upstate Medical University, Syracuse, NY; 2) Ambry Genetics, Aliso Viejo, CA; 3) Division of Pediatrics, SUNY Upstate Medical University, Syracuse, NY.

This female was large for gestational age at 35 weeks, delivered to a 27 year old primigravid Caucasian whose pregnancy was complicated by pre-eclampsia. Neonatal period was notable for hypoglycemia, apnea, bradycardia, hyperbilirubinemia, grade I intraventricular hemorrhage, subdural hematoa, laryngomalacia, hypotonia and feeding difficulties. She presented in the emergency department at 5 months of age with lethargy, emesis and MRI revealing progressive ventricular enlargement with cerebral atrophy; VP shunt was placed promptly. Erythrocyte of hydrocephaly was unknown. She has central apnea and hypothyroidism. Vitamin-K dependent clotting factor deficiency (VKCFD1) was diagnosed, and is now under good control with high dose vitamin K supplementation. At 18 months of age, she was diagnosed with a neuroblastoma. She has numerous minor dysmorphic features. At two years of age, the patient has global developmental delays and nystagmus. Karyotype and oligo-microarray were normal. Whole exome genomic analysis was undertaken. Mutations in 9 genes were identified (14 mutations), and these genes underwent medical review at Ambry Genetics. A de novo, frameshift mutation resulting in a truncated protein was identified in the CLTC gene. CLTC encodes clathrin heavy chain 1 (CHC1), one of the components of clathrin, which serves as the vesicle coat protein involved in intracellular trafficking and endocytosis. Mutations in CLTC have not been previously identified in human disease. However, CLTC is expressed at high levels in the brain, and it has been shown that inactivation of CHC1 in rat and fruit fly models prevents the recycling or release of vesicles in the synapse. Mutations in CLTC1, the second member of the Clathrin Heavy Chain family, have been associated with neurological disease: seizures, intellectual disability, autism and schizophrenia. Patients with a deletion of 17q23.1, including CLTC, have been reported with microcephaly, developmental delay, and growth retardation. A single patient with non-syndromic hydrocephalus has been reported with a duplication overlapping the CLTC gene. We propose that this mutation in the CLTC gene is the cause of neurological disease in this patient.
2688M
Two cases of lissencephaly with marked hydrocephalus caused by TUBA1A mutation. N. Ishihara1,2, S. Yokoi1,2, H. Yamamoto2, J. Natsume2, M. Tsutsui2, T. Ohye3, M. Kato3, S. Saitoh2, H. Kurashiki1, 1) Dept Pediatrics, Fujita Health Univ Sch Med, Toyoake, Japan; 2) Dept Pediatrics, Nagoya Univ Grad Sch Med, Nagoya, Japan; 3) Div Molecular Genetics, ICMS, Fujita Health Univ, Toyoake, Japan; 4) Dept Pediatrics, Yamagata Univ Faculty Med, Yamagata, Japan; 5) Dept Pediatrics and Neonatology, Nagoya City Univ Grad Sch Med Sci, Nagoya, Japan. 

Objective: TUBA1A encoding α-1a tubulin is one of the genes responsible for lissencephaly (LIS). Despite of wide spectrum of phenotypes associated with TUBA1A mutations, pathophysiology underlying the severity is still under discussion. Here we present two cases of LIS with marked hydrocephalus, the most severe phenotype of the spectrum, to clarify the function of mutated protein. Subjects and methods: Case 1 is a 2-year-old boy. Ventricile dilation was pointed out in fetal period. After birth, he was diagnosed as LIS. He had severe developmental delay with quadriplegia, but he had social smile and he could breathe and swallow by himself. At the age of 8 months, he began suffering from West syndrome, and visited our pediatric neurology office. Case 2 is a 3-year-old girl with marked hydrocephalus pointed out in fetal period. After birth, hydrocephalus rapidly progressed to become hydranencephaly, with cerebellar and brainstem hypoplasia. She did not have either spontaneous breathing or swallowing reflex, necessitating mechanical ventilation and gastrostomy for survival. Both patients were subject to exome sequencing for diagnosis. Results: Case 1 had c.74G>T (p.C25F) missense mutation in exon 2 of TUBA1A, and case 2 had c.190C>T (p.R64W) in exon 2 of TUBA1A. Both mutations were confirmed by Sanger sequencing. Mapping of the mutations on to 3D structure of microtubule interaction showed both C25 and R63 were not located at the contact surface of the tubulin interaction. Discussion: According to the spectrum of phenotypes associated with TUBA1A described by Kumar, case 1 is classified in group 4, most severe type of LIS, and case 2 is much more severe than TUBA1A showed both C25 and R63 were not located at the contact surface of the tubulin interaction. 

2689T
Comprehensive clinical characterization of VCP associated multisystem proteinopathy. V.E. Kimonis1, A. Surampalli1, M. Khare1, M. Wencel2, C. Nguyen3, S. Wigah3, S. Graft3, A. Wang3, S. Donkervoort1,4, M. Milla1, T. Mozafari1, V. Caiozzo2. 1) Pediatrics, UC Irvine School of Medicine, Irvine, CA; 2) Child Development Center, Pediatrics, University of California, Irvine, CA; 3) Department of Pediatric Exercise Medicine, University of California Irvine, CA; 4) ALS and Neurodegenerative Center, University of California, Irvine, CA; 5) Department of Orthopedics and Physiology & Biophysics, University of California, Irvine, CA; 6) National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD. 

VCP disease, also called multisystem proteinopathy is characterized by hereditary inclusion body myopathy with Paget's disease of bone and fronto-temporal dementia (FTD) and other less common varied phenotypes including amyotrophic lateral sclerosis, Parkinson's, and cardiomyopathy. TDP43 and ubiquitin positive inclusions are seen in affected tissues. We studied 41 individuals (24 affected/ 5 presymptomatic carriers/ 12 unaffected; M18/F23, mean ages 50.8, 44.9, 49.7 y, respectively) in 8 families harboring three missense mutations, R155H, R155C and R155P; with 300 population specific controls, two candidate pathological changes were found. Following in silico analysis and checking our variant was performed for one of the probands and we were able to identify a novel variant. This study represents the most comprehensive evaluation of individuals with VCP disease to date which will help establish baseline studies for future clinical trials in this underserved population.

2690S
Identification of a novel variant in TMEM67 gene responsible for JBTS6 by whole exome sequencing. A. Najmabadi1, M. Hosseini2, Z. Fattahi3, 1) University of California, San Diego, San Diego, CA; 2) Genetic Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

Joubert syndrome (JBTS) is a clinically and genetically heterogeneous disorder with autosomal recessive pattern of inheritance. The disorder is characterized by cerebellar hypoplasia, intellectual disability (ID), ataxia and oculomotor apraxia. Other clinical features include retinal degeneration, renal anomalies, hepatic fibrosis, and skeletal involvement. The hallmark of JBTS is a radiological pattern in magnetic resonance imaging (MRI), named “molar tooth sign”. In 2007 Baala et al. identified a new form of Joubert syndrome designated JBTS6 with a causative mutation in the TMEM67 gene. Until now only 5 mutations in this gene have been identified as responsible for JBTS6. Here, we are reporting an autosomal recessive Iranian family with novel mutations in the TMEM67 gene. This family has two-affected child, one whom has profound ID, seizure, strabismus, and renal failure while the other child only shows signs of ID. Exome sequencing was performed for one of the probands and we were able to identify a new variant of variants. Following in silico analysis and checking our variant with 300 population specific controls, two candidate pathological changes remained; of which only one of these changes co-segregated in the family. Upon MRI examination of the affected boy, molar tooth sign was observed. We can conclude the missense changes in TMEM67 gene are responsible for Joubert syndrome in our family. Underlying causes of ID remain unknown in many cases because of clinical and genetic heterogeneity; therefore exome sequencing is an effective and helpful technique in detection of de novo mutation in this type of disorders. This approach resulted in more precise genotype-phenotype correlation and clinical diagnosis.
Truncating mutation of NFIA causes a brain malformation and urinary tract defect. Y. Negishi1, A. Hatton2, K. Mizuno2, I. Hon3, N. Ando1, F. Miyaa, T. Tsunoda3, N. Okamoto4, M. Kato5, M. Yamashita6, Y. Kanemura7,8, K. Kosakia, S. Saitohb. 1) Department of Pediatrics and Neonatology, Nagoya City University Graduate School of Medical Science, Nagoya, Japan; 2) Department of Nephro-Urology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; 3) Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan; 4) Department of Medical Genetics, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; 5) Department of Pediatrics, Yamagata University School of Medicine, Yamagata, Japan; 6) Department of Neurosurgery, Takatsuki General Hospital, Osaka, Japan; 7) Center for Medical Genetics, Keio University School of Medicine, Tokyo, Japan.

Chromosome 1p32-p31 deletion syndrome (OMIM #163735) involving the NFIA gene is characterized by corpus callosum hypoplasia or absence, hydrocephalus or ventricular enlargement, and urinary tract defects. Here, we report the first case complicated by brain malformation and urinary tract defects observed with a truncating mutation in the NFIA gene. This male patient was born to healthy Japanese parents. Callosal agenesis had been suspected from the 28th gestational week. The boy was born by cesarean section on the 41st gestational week due to enlargement of the head circumference and post-term pregnancy. His birth weight was 3180g (+0.4SD), with a head circumference of 38.2cm (+3.3SD). Head MRI on the 3rd day of life revealed multicocular gliopendymal cysts, ventricular enlargement, and callosal agenesis. Regarding his developmental milestones, he started walking without support at 1 year and 3 months. He was observed speaking with mildly uplifted ear lobes. Brain MRI showed bilateral grade IV vesicouretical reflux. We performed whole exome sequencing on the proband and his parents, and identified a frameshift mutation (c.1093delC; p.P365fs) in the NFIA gene. This is the first report, to our knowledge, of a single nucleotide deletion (frameshift mutation) in the NFIA gene associated with brain malformation and urinary tract defects, confirming that the NFIA gene plays a fundamental role in development of brain as well as in urinary tract.

De novo 109 kb microdeletion of MED13L: report of a new patient with developmental delay, facial abnormalities and hypotonia. E.A. Ratnina1,2,3, C.E. Lawston1,2, S.D. Fiedler1, J.M. Joyce1, P.V. Thakor1, L.D. Cooley1,3,4, H.H. Ardinger2,4,1. 1) Pathology Dept, Children's Mercy Hospital, Kansas City, MO; 2) Pediatrics Dept, Children's Mercy Hospital, Kansas City, MO; 3) Pathology Dept, University of Missouri-Kansas City Medical School, Kansas City, MO; 4) Pediatrics Dept, University of Missouri-Kansas City Medical School, Kansas City, MO.

The MED13L gene is one of the subunits of the large mediator complex that functions as a transcriptional coactivator for most RNA polymerase II transcribed genes and encodes a primary transcript of ~319 kb in size. De novo MED13L is highly expressed in the brain, heart, skeletal muscle, kidney, placenta, and peripheral blood leukocytes. Missense mutations in MED13L have been linked to transient closure of the great arteries and non-syndromic intellectual disability. Deletions in MED13L have been recently proposed to result in a distinct syndromic phenotype consisting of motor and speech delay, moderate intellectual disability, hypotonic open-mouth appearance and facial dysmorphia. Only three cases with various size deletions (17 kb (exon 2), 41 kb (exons 6-20) and 115 kb (exons 3 and 4) in MED13L have been recently reported. In this study, we performed molecular and clinical characterization of a patient with an approximate 109 kb de novo deletion of exons 5 through 31 of MED13L detected by oligonucleotide array comparative genomic hybridization. The patient is the third child of healthy non-consanguineous parents born at 39 weeks gestation after a pregnancy complicated by polyhydramnios. At the age of 2 years, she had mild hypotonia, wide-based gait, drooling, motor and speech delay, and dysmorphic features that included triangular face with tall forehead, sparse scalp hair (hirsutism), prominent nose, short philtrum, and small mouth. She has normal intelligence and no cardiac anomalies. Our patient has moderate intellectual disability, distinctive facial dysmorphism and hypotonia. However, she lacks cardiac anomaly, a phenotypic feature reported in cases with MED13L nonsense mutations and two cases with deletions of exon 2, 3 and 4. This observation supports the hypothesis recently proposed by van Haelst et al. that the cardiac phenotype may have reduced penetrance. The expression of cardiac phenotype may also be dependent on the ratio of MED13L haploinsufficiency and the location within the MED13L gene. Our study thus adds to the increasing body of cases describing the existence of the novel MED13L haploinsufficiency syndrome.

Novel gene mutation in Schimmelpenning syndrome (nevus sebae- cous syndrome). Y. Kuroda1,2,5, I. Ohashi3, T. Niihori1, Y. Kant1, N. Okamoto1,2, N. Ando1, I. Hori1,3, Y. Kanemura1, N. Kanemura1, Y. Negishi1, A. Hatton2, K. Mizuno2, I. Hon3, N. Ando1, F. Miyaa, T. Tsunoda3, N. Okamoto4, M. Kato5, M. Yamashita6, Y. Kanemura7,8, K. Kosakia, S. Saitohb. 1) Pathology Dept, Children's Mercy Hospital, Kansas City, MO; 2) Pediatrics Dept, Children's Mercy Hospital, Kansas City, MO; 3) Division of Medical Genetics, Kanagawa Children's Medical Center, Yokohama, Kanagawa, Japan; 4) Department of Pediatrics, University of Tokyo, Tokyo, Japan; 5) Department of Medical Genetics, Osaka Medical Center and Research Institute for Maternal and Child Health, Izumi, Japan; 6) Department of Pediatrics, Tokohu University School of Medicine, Sendai, Japan; 7) Department of Epigenetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan.

We have identified a novel causative gene for Schimmelpenning syndrome. Schimmelpenning syndrome is a rare neurocutaneous disorder characterized by craniofacial nevus sebaeaeus with central nerve system, ocular, or skeletal abnormalities. Nevus sebaeaeus represents yellowish, waxy skin lesion and can undergo neoplastic degeneration in 24% in cases. Somatic mutations of HRAS and KRAS have been reported in patients of nevus sebaeaeus and Schimmelpenning syndrome. The proposita had nevus sebaeaeus at right frontal and vertex robe, angiona of the right postauricular skin, and bilateral preauricular tags. Ocular examination revealed retinal coloboma and macular hypoplasia of right eye. Head MRI revealed polymicrogyria and hemimeganencephaly of the right robe and bilateral cerebellopontine lipoma. She had also developmental delay and seizure. No mutation was detected by direct sequencing of exon 1 of HRAS in the nevus. Patient's sample from nevus was sequenced by MiSeq (Illumina Inc., San Diego, CA) by 121 bp pair-end reads, after the enrichment with TruSight Tumor. TruSight Tumor (Illumina Inc., San Diego, CA) was designed for targeted next-generation sequencing for 26 oncogenes and tumor suppressor genes, covering all coding regions, in total, 21 kb. Deep panel sequencing (average coverage of 158x6) detected novel mutation of RAS-MAPK pathway gene with the nevus, but not with the lymphocyte, indicating postzygotic mutation. The mutation has been reported as the recurrent mutation in other neurocutaneous disorder. RAS promotes cell growth through activation of the mitogen-activated protein kinase (MAPK) signal transduction pathway. Activating mutations in the RAS-MAPK signaling pathway play a pivotal role in cancer predisposition shown in epidermal nevus syndrome.
Neurofibromatosis type 1 and Optic Gliomas. E. Parkhurst, S. Abboy. Dept Genetics, Kaiser West Los Angeles, Los Angeles, CA.

Introduction: Neurofibromatosis type 1 (NF1) is a common autosomal dominant disorder that affects approximately 1 in 3,500 people. One of the most common tumors in children with NF1 is optic glioma with reported incidences ranging from 5-30%. Controversy surrounds the recommended frequency of eye exams and neuroimaging in asymptomatic individuals with NF1.

Methods: We queried the Southern California Kaiser Permanente electronic medical record database (Clarity) to find patients diagnosed with NF1 and seen in Ophthalmology to determine the incidence of optic glioma. Medical records showing a diagnosis of optic glioma were reviewed.

Results: 708 patients under 21- years-old have a diagnosis of NF1 in the Southern California Kaiser system. 347 of these patients were seen in the ophthalmology department and included in the study. Of the 347 with an ophthalmology exam, 30 (8.6%) had a diagnosis of optic glioma. Average age of diagnosis was 5 yrs; the youngest was diagnosed at age 18 months and the oldest was diagnosed at 12 yrs. 30% (9/30) of patients with optic glioma required treatment other than corrective lenses; two patients had surgery alone and seven had chemotherapy. Treated patients had an age of diagnosis ranging from 19 months to 10 yrs. The youngest treated patient was asymptomatic and had 10 weeks of chemotherapy before 2 yrs. 60% (18/30) of patients found to have an optic glioma on MRI presented with symptoms (vision loss, proptosis, precocious puberty, etc.) and 40% (12/30) were asymptomatic. 63% (19/30) of the gliomas were bilateral; 23% (7/30) right sided and 13% (4/30) left sided. 53% (16/30) of the gliomas involved the optic chiasm. The four patients diagnosed with precious puberty all had chiasmal gliomas and were treated with chemotherapy.

Discussion: Current standards of care for children with NF1 include annual ophthalmologic examination, especially for children <6 years of age. Baseline MRI to detect asymptomatic optic glioma is not warranted. Our findings show that most optic gliomas are asymptomatic and never require treatment. Children with vision changes or precocious puberty are most likely to require treatment such as surgery or chemotherapy, although very young asymptomatic children may have actionable gliomas as well.

Genetic heterogeneity in Van der Woude syndrome. P. Kumar1, A. Alif2, SK. Singh3, R. Raman1. 1) Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University; 2) Centre for Genetic Disorders, Banaras Hindu University, Varanasi; 3) G. S. Memorial Plastic Surgery Hospital and Trauma Center, Varanasi, India.

Van der Woude syndrome (VWS; OMIM 119300), an autosomal dominant disorder, is one of the most common syndromic forms of cleft lip and palate with lower lip pits as additional phenotype. IRF6 (OMIM 607199) is the foremost candidate gene in several populations. WDR65 (OMIM 614259), GRHL3 (OMIM 608317) and 17p11.2-11.1 have also been identified to be in linkage with VWS, showing locus heterogeneity. In earlier studies from India on several cases IRF6 is not found to be associated with VWS. Here we present results from genomic studies in two VWS families having 8 and 11 affected individuals, respectively. Initial analysis of IRF6 did not reveal any mutation in the total coding region, and no association with the polymorphism, rs642961, in the enhancer region. Genome-wide linkage study (GWLS) using Affymetrix 10K SNP microarray in one of the families resulted in 3 linked loci (chromosome #1: q31.3-q42.2, #16: q23.1-q24.4 and #18: q11.2-q21.2) which had a LOD score of 2.4. Since microsatellite marker-based linkage analysis from the 3 high LOD score regions did not narrow down the mapped region, NGS was performed for whole genome with paired-end read which obtained a novel mutation in NOL4 (OMIM 603577) that could affect its function by creating target site for a micro RNA. In the second VWS family, a haplotype involving an intron and two regulatory regions of IRF6 co-segregated with the affected family members. IRF6 haplinsufficiency being the major cause of VWS, its expression level was checked in blood of the affected and unaffected members of the family. Compared with the normal haplotype, IRF6 expression was 2.27 fold lower in the affected haplotype. In silico analysis discovered a putative repressive transcription factor that would bind to one of the variants in the risk haplotype. Thus we not only identify a novel candidate gene (NOL4) but find a novel haplotype in IRF6 as causal factor for VWS.

EXOMIC SEQUENCING AND MOLECULAR ANALYSIS OF IRF6 GENE IN PATIENTS WITH VAN DER WOUDE SYNDROME OR FAMILIAL HISTORY OF CLEFT LIP AND PALATE.

E. Prokudin1, L. Patino1, I. Briceno2,3, J. Martinez2, D. Mosquera2. 1) Bogota, Colombia; 2) Universidad de La Sabana; 3) Universidad Javeriana.

The advent of technologies for massively parallel sequencing has resulted in thousands of sequenced genomes. The Van der Woude syndrome (VWS [MIM 119300]) is a skull - facial malformation characterized by the association of cleft lip and palate with other ocular abnormalities including anophthalmia, anterior segment dysgenesis and cataracts. A number of causative variants have been identified in genes including SOX2, OTX2, CHX10, BMP4 and RAX. However, for the majority of patients the disease genes are still not known. Low penetrance and variable expressivity are among the factors contributing to the low detection rate in microphthalmia and coloboma patients. In this study we performed whole exome sequencing on 4 affected individuals from 2 generations in an Australian family with autosomal dominant microphthalmia and coloboma. Heterozygous variants shared by all four affected family members were selected for the analysis. The variants were filtered based on their population frequency, pathogenicity prediction and conservation scores. Subsequently we prioritised variants in previously investigated animal disease genes and pathways known to be involved in eye disease. The two candidate variants, a frameshift deletion, was in a gene in the WNT signaling pathway. The variant was confirmed by Sanger sequencing and further processed for in vitro analysis using a WNT reporter assay in HEK293 cells. Our results indicate a role for the WNT signaling pathway in the microphthalmia and coloboma phenotype in humans.
2698T


Fabry disease—a genetic disorder characterized by the accumulation of globotriaosylceramide in cell lysosomes resulting from an X-linked deficiency of α-galactosidase A activity—presents with multorgan manifestations. We report, mother and daughter with Fabry disease, the mother of 36-year-old is asymptomatic and daughter 1 year old with neurological, renal, cardiac and ophthalmological disorders. At 4 months of age begins with proteinuria and recurrent infection of urinary tract, such as kidney problems. Later are noted cardiac (left ventricular enlargement and leaky heart valves), gastrointestinal (abdominal cramps, frequent bowel movements shortly after eating and diarrhea) and neurological (transient ischemic attacks) disturbances. Was performed gene sequencing (GLA HPLC/Tandem MS lysosome-Gb3). The GLA gene was analysed by PCR and sequencing of the entire coding region and the highly conserved exon-intron splice junctions. Deep intrinsic mutations are not tested. The reference sequence of the GLA gene is: NM_000169.2. The biomarker lyso-Gb3 was measured by HPLC and tandem mass spectrometry. Was detected a previously unreported heterozygous variant in exon 6 of the GLA gene (c.966C>G p.P322R). It is located in a moderately conserved nucleotide and highly conserved amino acid position, with moderate physiochemical differences between the amino acids proline and arginine. Software analyses show inconsistent predictions: Mutation taster indicates this variant is probably damaging, whereas PolymapPhen, SIFT and Align-GVGD predict toleration. The concentration of the biomarker lyso-Gb3 was normal. Our findings therefore suggest that this defective gene in the heterozygote has resulted from a new mutation. Genetic counselling is recommended for our patient and other relevant family members, also in the heterozygote has resulted from a new mutation. Genetic counselling is recommended for our patient and other relevant family members.

2699S

Rasopathies and RAS/MAK pathway disorders: Genetic screening of a cohort of 37 Tunisian children. N. Abdelmoula1, R. Louati1, I. Trabelsi2, S. Kammoun3, M. Zenker4, T. Rebai1. 1) Lab Histology, Univ Medicine, Sfax, Tunisia; 2) Dept Cardiology, Hedi Chaker Hospital, Sfax, Tunisia; 3) Institute of Human Genetics, University Hospital Magdeburg, Magdeburg, Germany.

Introduction: Recent studies have shown that a group of genetic disorders baptized RASopathies result from genes mutations of the Ras/MAK pathway. These disorders include Noonan syndrome caused by mutations in PTPN11, SOS1, 2 mutations of Braf. No mutation of the new described gene; RIT1 was investigated in 10 patients. Mutations were not detectable in 21 patients. No mutation of the new described gene; RIT1 was investigated in 10 patients. Mutations were not detectable in 21 patients. In our study, we have investigated a cohort of 37 Tunisian children. Results: A total of 12 mutations have been retrieved: 8 mutations of PTPN11, 2 mutations of SOS1, 2 mutations of Braf. No mutation of the new described gene; RIT1 was detected for negative Noonan patients. Neurobehavioral and dysmorphic features in patients with confirmed RASopathies developmental disorders have been compared to negative patients. Conclusion: Our findings indicate that mutations promoting dysregulation of the RAS/MAK cascade mark an increased psychopathological risk and highlight that probably other new genes may be involved in Tunisian patients with mild, non-specific or atypical features. Using whole exome sequencing for molecular investigation of our negative patients will be a very powerful tool to confirm our hypothesis.
Antithrombin deficiency in a founder population: different genetic architectures for types 1 & 2. P. Saio1,4, M. Puurunen4, J. Corral4, S. Engelberth5, K. Javela3, M. Perola1,2,4,5. 1) Public Health Genomics, Natl Inst Health & Welfare, Helsinki, Finland; 2) Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland; 3) Hospital District of Helsinki and Uusimaa, Finland; 4) Hemostasis Laboratory, Finnish Red Cross Blood Service, Helsinki, Finland; 5) Centro Regional de Hemodonación, Servicio de Hematología y Oncología Médica, University of Murcia, Spain; 6) Estonian Genome Center, University of Tartu, Estonia.

Hereditary thrombophilia due to antithrombin deficiency (ATD) is a rare autosomal dominant disorder. It is caused by mutations in SERPINC1 coding for antithrombin, a central anticoagulant molecule in the blood. The affected individuals have an approximately 30-fold increase in risk for thrombotic complications such as deep vein thrombosis and pulmonary embolism. ATD is designated as type 1 when both antithrombin levels and its activity are reduced. In type 2 deficiency the level of antithrombin is normal but its activity is diminished. We characterized the genetic background of ATD in a sample containing the majority of patients (N=221) diagnosed with ATD in Finland using capillary sequencing, Sequenom genotyping, and multiplex ligation-dependent probe assays. The two ATD types have remarkably different genetic architectures in the Finnish population. Type 1 is dominated by allelic heterogeneity with 28 distinct mutations explaining the disease in the 36 families with type 1 ATD. Following from the requirement of decreased antithrombin levels for diagnosis with type 1 deficiency, these patients carried either frameshift-causing or exon-spanning indels, or point mutations affecting splice-sites or introducing premature stop codons. In stark contrast, only 5 distinct mutations were found in the 48 families with type 2 deficiency. A single mutation (p.Pro737Leu) explained almost 90% of the type 2 deficiency cases. The genetic structure of the Finnish population is a result of the small number of founding individuals, relative isolation, and slow population growth until a rapid expansion to the current size of 5.4 million inhabitants. This has reduced the allelic heterogeneity of most Mendelian diseases in Finland, where one or a few major disease-causing mutations typically explain most of the cases for a given disease. These major mutations were present in the population at the onset of the population growth and increased in number together with minor mutations at the step of ATD specific genetic structure. ATD is strikingly limited to only type 2 deficiency, possibly indicating historically stronger negative selection against alleles causing type 1 ATD. Given that both types are malignant and their phenotypic differences are not large, this may have selected for a strong reduction of the allelic diversity in Finland.

Evidence of germline mosaicism in Fibrodyplasia Ossificans Progressiva: Post discovery of the ACVR1 gene. M.B. Alcausin1,2, B.R. Miller3, R. Miller4, A. Braxton1,2. 1) Institute of Human Genetics, National Institutes of Health, University of the Philippines-Manila, Manila, Philippines; 2) Department of Pediatrics, Philippine General Hospital, Manila, Philippines; 3) Department of Orthopaedic Surgery, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, Pennsylvania, USA; 4) Center for Research in FOP and Related Disorders, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, Pennsylvania, USA; 5) Department of Medicine, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, Pennsylvania, USA; 6) Department of Genetics, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, Pennsylvania, USA.

Posters: Clinical Genetics and Dysmorphology 673

Multi-systemic Involvement in NGLY1-related disorder Caused by Two Novel Mutations. J. Heeley, M. Shinawi. Department of Pediatrics, Division of Genetics and Genomic Medicine, Washington University in St. Louis School of Medicine, St. Louis, MO.

Background: NGLY1-related disorder is a newly described autosomal recessive condition characterized by neurological (global developmental delay, severe hypotonia, movement disorder, seizures or EEG, acquired microcephaly, diminished reflexes and nerve conduction abnormalities), hepatic (neonatal jaundice, elevated liver enzymes, fibrosis, intrahepatic cytoplasmic inclusions), ophthalmological (alacrima/hypolacrimation, strabismus, chalazion, apraxia) findings and associated with dysmorphic features, constipation and scoliosis. The gene encodes an enzyme, α-N-glycanase 1 (NGLY1), involved in deglycosylation of glycoproteins, an essential step in endoplasmic reticulum-associated degradation pathway. It was hypothesized that loss of function of NGLY1 causes accumulation of misfolded glycoproteins in the cytoplasm. The condition has been described so far in 8 patients. Methods: Clinical and molecular characterization of a proband with novel mutations in the NGLY1 gene detected via whole exome sequencing. The findings in the proband are compared with previously reported cases. Results: The proband is a 13-year-old boy with profound hypotonia, non-ambulatory. He has severe scoliosis and osteopenia. He has acquired microcephaly after 10 months, involuntary movements, poor weight gain, muscle atrophy, absent reflexes, and seizures. He has had multiple procedures for lacrimal duct stenosis and strabismus and has intractable blepharitis. We also noted persistent hypochondroplasia, which was not previously described. Whole exome sequencing revealed two novel variants in the NGLY1 gene: a maternally inherited truncating mutation, c.347C>G (p.S116X), and paternally inherited splicing mutation, c.881+5G>T, predicted to abolish the splice donor site of exon 5. Neither change was previously reported in NHLBI. Conclusion: Mutations in the NGLY1 gene cause a severe, multisystemic but recognizable phenotype. This study along with previously reported cases suggests that targeted sequencing of the NGLY1 gene should be considered in patients with the typical combination of neurological, hepatic, and ophthalmologic findings. Our data reveal that osteopenia and hypochondroplasia can be part of the phenotypic spectrum of NGLY1-related disorder.
Is SMN2 related to severity in Spinal Muscular Atrophy?: a case report.

Introduction: Spinal muscular atrophy (SMA) is an autosomal recessive neuromuscular disease characterized by progressive degeneration and loss of the anterior horn cells in the spinal cord leading to progressive muscle weakness and atrophy. Incidence is 1 in 6,000 and carrier frequency is 1 in 35. It is caused in 96% of patients by homozygous absence of the survival motor neuron gene (SMN1). Copy number of SMN2 has been proposed to determine severity of the disease. SMA is classified by age of onset and severity of the disease. Case Presentation: This is a 15 years-old (y/o) male patient, who was born at first gestation from non-consanguineous parents. Mother was 18 y/o. Patient started to have frequent falls and trouble walking up and down stairs at 4 years. Progressive muscular weakness was evident first in legs and then in arms. After that, he started to have difficulty to solid swallowing. At 6 y/o, patient was not capable of deambulate anymore. He has had presented respiratory failure in several opportunities, requiring hospitalization. He has a 13 y/o brother with similar phenotype but with symptoms starting after 4 y/o. Physical examination revealed generalized muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced lumbar scoliosis and paravertebral muscle atrophy. MRI of the brain was normal. Muscle biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy. Spinal cord MRI evidenced left lumbar muscular hypotrophy, tongue twitching and knee retractions. Muscular biopsy showed muscular dystrophy.

Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2. Discussion: This individual is affected with SMA type II because onset was after age ten months and no copies of SMN1 were seen at molecular testing. It is known that SMN2 genes are capable to produce a protein identical to that of the SMN1 gene but at reduced capacity (10–20%). As a result, more than 3 copies of SMN2 are related to milder severity of the illness. Despite the anterior fact, muscular weakness in this patient was rapidly progressive and severity was significant. This case is also noteworthy because patient’s brother quantitative PCR showed SMN1 exon 7 and 8 homozygous deletion. This SMN1 homozygous deletion is typically detected by demonstrating the absence of exon 7 (~96% of SMA type II patients). This patient has 3 copies of SMN2.
2709M

Chromosome deletion 11q13.1 involving deletion of the CLCF1 gene in a female with features of Cold-Induced Sweating Syndrome. J.D. Weisfeld-Adams, K.E. Brown. Department of Pediatrics, University of Colorado, Denver, CO.

The CLCF1 gene (cardiotrophin-like cytokine factor 1) is known to be implicated in some cases of autosomal recessive Cold-Induced Sweating Syndrome (CISS). CISS, a designation used synonymously with "Crisponi syndrome", is characterized by paradoxical sweating response, high-arched palate, scoliosis, facial weakness, prominent ears, feeding difficulties in infancy, hyperthermia, and velopharyngeal insufficiency or hypernasal speech. CLCF1 is a member of the glycoprotein 130 cytokine family and encodes cardiotrophin-like cytokine factor 1 (CLCF1). CLCF1 partners with cytokine receptor-like factor 1 (CRLF1) to form the CRLF1/CLCF1 protein complex, with CLCF1 mutations accounting for significantly fewer historical CISS cases than CRLF1 mutations. The complex binds to ciliary neurotrophic factor receptor (CNTFR) and activates the Jak-STAT signaling cascade. CLCF1 is a potent neurotrophic factor, B-cell stimulatory agent and neuroendocrine modulator of pituitary corticotrophic function. Mutations in the CLCF1 gene may disrupt CNTFR signaling leading to disruption of sympathetic nervous system development, and resultant dysregulation of temperature homeostasis, manifesting as hyperthermia and pathologic sweating response. The CNTFR pathway is also implicated in motor neuron and skeletal development which may explain facial dysmorphology, facial weakness and skeletal abnormalities observed in some affected individuals. Here we report a 10 year old female with a history of frequent, unexplained fevers, lack of sweating with activity, feeding difficulties, short stature, scoliosis, subtle facial dysmorphism, and hypernasal speech. Chromosomal microarray analysis revealed a heterozygous de novo deletion at 11q13.3, of approximately 0.153 Mb in size and incorporating CLCF1. We believe this patient likely has a point mutation on the non-deleted CLCF1 allele. Research testing with sequencing of CLCF1 is currently underway. To our knowledge, this is the first CISS patient reported with a genomic rearrangement incorporating CLCF1.

2710T

Functional studies of EZH2 histone methyltransferase activity in Weaver syndrome. A.S.A. Cohen1,2, D. B. Yap3,4, J. Denny5, S. M.E. Lewis1,2,6, C. Chijiwa1,2,6, M.A. Ramos-Arroyo7, N. Tkachenko8, V. Milian9, M. Fradin10, C. J.D. Ross1,2,5, W. B. Dobyns11,12,13, D. D. Weaver14, S. Aparicio15, W. T. Gibson1,2, 1) Dept of Medical Genetics, University of British Columbia, Vancouver, BC, Canada; 2) Child and Family Research Institute, Vancouver, BC, Canada; 3) Dept of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada; 4) Dept of Molecular Oncology, British Columbia Cancer Research Centre, Vancouver, BC, Canada; 5) Centre for Molecular Medicine and Therapeutics, Vancouver, BC, Canada; 6) British Columbia Children’s and Women’s Health Center, Vancouver, BC, Canada; 7) Dept of Medical Genetics, Complexo Hospitalar de Navarra, Pamplona, Spain; 8) Medical Genetics Unit, Centro de Genética Médica Dr Jacinto Magalhães, Oporto Hospital Center, Porto, Portugal; 9) Instituto di Genetica Medica, Università Cattolica delSacro Cuore, Policlinico Universitario Agostino Gemelli, Roma, Italy; 10) PH Génétique Clinique, Hôpital SUD, Rennes, France; 11) Center for Integrative Brain Research, Seattle Children’s Hospital, Seattle, WA, USA; 12) Dept of Pediatrics, University of Washington, Seattle WA, USA; 13) Dept of Neurology, University of Washington, Seattle WA, USA; 14) Dept of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA.

Last year at ASHG (63rd annual meeting, 2013), I provided an update on our cohort, where 9 out of 35 individuals with Weaver-like phenotypes were found to have mutations in EZH2. Since then, we have identified a mutation in EZH2 in one more patient in our cohort (currently 45 patients), as well as 2 mutations in NSD1, the Sotos syndrome gene.

At the time of my presentation, I also showed results of an in vitro functional assay which supported the hypothesis that WS is caused by loss-of-function mutations in EZH2. These findings suggested that EZH2 inhibitors being developed for cancer treatment may not be as effective in rare cancers associated with WS. As such, a more detailed investigation of our WS mutants was required. Our most recent results suggest that WS mutants may be classified into different groups based on variable degrees of loss-of-function, which could explain some of the variability in the observed phenotype and also suggests a need for different therapeutic approaches based on the mutation identified.
2711S
Cantu syndrome: Delineation of cardiovascular abnormalities in six affected individuals evaluated at a research clinic. D. Grange1, B. Kozel1, G. Singh1, M. Levin1, P. Stein1, C. Nichols2 1) Dept Internal Medicine, Washington University; 2) Dept Pediatrics, Washington University; 3) Dept Cell Biology and Physiology, Center for the Investigation of Membrane Excitability Diseases, Washington University, St Louis, MO.

Cantu syndrome (CS) is associated with hypertrophicis, neonatal macrosomia, macrocephaly, craniofacial dysmorphic features, osteochondroplasia, edema and cardiovascular abnormalities including PDA, cardiomegaly, pericardial effusion and abnormal cerebral vasculature. CS is caused by heterozygotes activating mutations in either ABCC9 or KCNJ8 which encode the regulatory SUR2 and Kir6.1 subunits, respectively, of the ATP-sensitive potassium channels (KATP). KATP channels formed from SUR2 and Kir6.1 subunits are prominent in cardiovascular tissues, and KATP activation in vascular smooth muscle results in decreased vascular contractility. Six individuals with CS were evaluated in the first Cantu Syndrome Research Clinic, included 4 females and 2 males ranging in age from 12-46 years of age. All patients had heterozygous mutations in ABCC9. Participants had cardiovascular testing including ECG, echocardiograms, 24 hour Holter monitoring with heart rate variability (HRV) assessment and pulse wave velocity testing via Sphygmocor. Blood pressures were below the 50th percentile for age for all participants, although none had orthostatic changes with positional change. They had full pulses with normal peripheral perfusion. Echocardiograms revealed cardiomegaly with enlarged ventricles. Left ventricular diastolic volumes were +3 SD from the mean. Cardiac mass was +2.5 SD from the expected values for age, gender and body surface area. Left ventricular systolic strain, a measure of cardiac contractility, was decreased compared to age and sex matched normative values. Echocardiogram-derived cardiac output showed increased output when compared to controls. Pulse wave velocity was slow in the younger individuals, but average in the adults, suggesting a trend toward lax vasculature in childhood that becomes stiffer with age. Most participants showed peripheral edema. Although there was no clinical heart failure, overall, the findings were suggestive of a high output, hypertensive state with low peripheral vascular resistance in the absence of hypertrophic cardiomyopathy. 12-lead ECGs showed a repolarization abnormality with a persistent ST segment in inferior leads in 4/6 patients. The Holter analysis revealed autonomic dysregulation manifested by increased sympathetic activity and decreased vagally modulated HRV in all participants. The oldest participant had extreme bradycardia and HRV values. Cardiac function may deteriorate with age in patients with CS.

2711M
Exome sequencing reveals compound heterozygous mutations in ATP8B1 in a JAG1/NOTCH2 mutation-negative patient with clinically diagnosed Alagille syndrome. C.M. Grochowski1, R. Rajagopalan1, A.M. Falsey1, K.M. Loomes1, I.D Kranzler1,2, M. Devoto3, N.B. Spinner1,2 1) The Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3) University La Sapienza, Rome, Italy.

Exome sequencing is helping to revolutionize genetic screening, as it permits unbiased testing, demonstrating erroneous clinical diagnoses. Alagille syndrome (ALGS) is an autosomal dominant disorder characterized by bile duct paucity and the presence of distinct facial features, cardiac findings, musculoskeletal malformations, and ocular abnormalities. Prior to genetic testing, ALGS was diagnosed clinically in individuals with bile duct paucity and plus 3 of 4 secondary findings. Currently, 94% of ALGS patients with a clinically confirmed diagnosis have mutations detected in JAGGED1 (JAG1) and an additional 2% of patients have mutations in NOTCH2. We report a patient who presented with 4/5 features of ALGS including bile duct paucity, a cardiac murmur, posterior embryonytoxin and multiple skeletal features including a small forehead, triangular face), leading to a clinical diagnosis of the disorder. Genetic testing revealed no mutations in JAG1 or NOTCH2, warranting an expanded search for causal genes. Exome sequencing uncovered compound heterozygous mutations in the gene ATP8B1, previously reported to cause Progressive Familial Intrahepatic Cholestasis Type 1 (PFIC1), which has very limited overlapping clinical findings with ALGS. A splice-site mutation (c.3400+2T>C) inherited maternally and a frameshift mutation (c.1889_1890insTAAAC p.His630fs) inherited paternally were detected and validated through Sanger sequencing. A similar pathogenic variation has been previously reported in a patient with a clinical presentation similar to our patient, and a diagnosis of atypical PFIC1. We are investigating further clinical features within this patient to determine the consistency of a PFIC1 diagnosis versus an atypical PFIC1. This work helps to delineate the clinical spectrum of Alagille syndrome and more precisely characterize diagnostic criteria and avoid confounding diagnoses.

2713T
New dominant mutations in SF3B4 encoding an mRNA splicingosomal protein important in embryonic limb patterning underlie Rodrigue acrofacial dysostosis. M.D. Irving1,2, B. Dimitrov1, D. Chitayat1, J.J. Rodrigo1, M.W. Wessels3, M.A. Simpson4 1) Department of Clinical Genetics, Guy’s and St Thomas’ NHS Foundation Trust, London, London, United Kingdom; 2) Division of Medical and Molecular Genetics, King’s College, London, London, United Kingdom; 3) The Hospital for Sick Children, Toronto, Canada; 4) Department of Pathology, Hospital Universitario La Paz, Madrid, Spain; 5) Department of Clinical Genetics, Erasmus Medical Centre, Rotterdam, The Netherlands.

Acrofacial dysostosis syndrome of Rodrigue (RADS) (OMIM 201170) is a lethal condition comprising multiple congenital abnormalities. It is characterised by severe mandibular hypoplasia, upper limb phocomelia with oligodactyly, absent ibulae, cleft palate and microtia. Abnormal lung lobulation results in pulmonary hypoplasia and early neonatal demise. It was first reported by Rodriguez et al. who described three siblings with multiple congenital anomalies affecting the limbs, cardiovascular and neurological systems, and severe mandibular hypoplasia, resulting in lethal respiratory insufficiency. Thereafter five additional cases have been described in the literature expanding the clinical phenotype to include abnormal lung lobulation, absent ibulae, cleft palate and arhinencephaly. Given the occurrence in three siblings, RADS has been considered to be an autosomal recessive genetic condition. There are obvious similarities with RADS and other syndromes associated with severe mandibular hypoplasia, such as Nager syndrome (NS)(OMIM 1554400) and Miller syndrome (OMIM 263750). Whilst there are clinical features to distinguish these conditions from RADS, such as eyelid colobomata and accessory labella, there is clear phenotypic overlap, raising the possibility that RADS and NS are indeed allelic. Heterozygous loss-of-function mutations in SF3B4, encoding the mRNA spliceosomal protein Splicing factor 3B, subunit 4 that has a role in bone morphogenetic protein signalling in early embryonic limb patterning, have been shown to underlie NS. Recently a de novo mutation in SF3B4 exon 4 that is predicted to disrupt splicing was demonstrated in a patient with RADS. In this study, we obtained DNA samples for whole exome sequencing from four fetuses with RADS, including one originally described by Rodriguez et al., to investi-gate the molecular basis and to explore the concept that there is both phenotypic and molecular overlap between this and other acrofacial dysostosis conditions. Results demonstrated two heterozygous frameshift mutations of SF3B4 in three of the four cases. In addition, inheritance studies found one of the patients was mosaic. These data suggest that inactivating mutations might result in disease with both dominant inheritance and not recessive transmission, which will significantly influence genetic counseling. The clinical and molecular data will be presented, as well as a discussion of the role of SF3B4 mutations in causing this distinctive phenotype.
2715M
Disruption of HDAC8 gene due to partial duplication in a female with Cornelia de Lange syndrome diagnosed by SNP microarray. S. Ramanathan, R.D. Clark. Pediatric Genetics, Loma Linda University, San Bernardino, CA.

The clinical diagnosis of Cornelia de Lange syndrome (CdLS [MIM 122470]), a cohesinopathy, is evolving as its genetic heterogeneity is better understood. To date, mutations in 5 genes, 3 core cohesin subunits (SMC1A [MIM 300040], SMC3 [MIM 606062], RAD21 [MIM 606646]) and 2 cohesin-regulatory proteins (NIPBL [MIM 608667], HDAC8 [MIM 300269]) have been identified in CdLS. The distinctive features in classic CdLS include synophrys, arched eyebrows, long prominent philtrum, small, upturned nose and upper limb reduction defects. Significant growth and cognitive impairments are typical. Milder, subtle phenotypes have also been reported, with the variability likely due to inter- and intragenic mutation spectrum (Mannini L et al., 2013).

We report a 4-month old female with CdLS due to a partial duplication of the HDAC8 gene, who presented as a newborn with UIGR, tetralogy of Fallot, bilateral ptosis and asymmetric crying face due to right facial palsy. Her dysmorphic features did not raise suspicion for CdLS until SNP microarray analysis detected a de novo 210 kb duplication on chromosome X at Xq13.1 that disrupted the HDAC8 gene. An X-inactivation ratio of 100:0 in peripheral blood was highly skewed. At 2 months, she had a wide open anterior fontanelle, depressed supraorbital ridges with hooded appearance of the eyes, mild synophrys and long, prominent philtrum with thin upper lip. The 5th fingers were short with a single crease and clinodactyly. She had failure to thrive, dysphagia and G-tube dependency. She did not have hirsutism or arched eyebrows.

Mutations in HDAC8, most of which are missense, have been recently identified in CdLS. The facial features can be atypical, with some overlap with that of classic CdLS. Distinguishing features are delayed anterior fontanelle closure, hooding of the eyelids and a broader nose with no gross limb abnormalities (Deardorf MA et al., 2012, Kaiser FJ et al., 2014). Most patients are heterozygous females with marked X-inactivation skewing in peripheral blood, favoring the normal allele. To our knowledge, this is the first patient with CdLS caused by a partial duplication of the HDAC8 gene. The atypical phenotype associated with HDAC8, which was not recognized on clinical evaluation, illustrates the emerging phenotypic variability in CdLS.

2717S
Cervical myelopathy in a patient with metatropic dysplasia caused by a TRPV4 mutation. E. Zapata-Aldana1, L. Arnaud-Lopez1, E.L. Mellini-Sanchez1, C. Peña-Padilla1, J. Rivera-Vargas1, F.J. Martinez-Macias1, L. Bobadilla-Morales1,2,3, A. Corona-Rivera1,2,3, A. Superli-Furga1, J.R. Corona-Rivera1,2,3, T. Gambin1, T. Harel1, J.R. Lupski1,3. 1) Servicio de Genética, División de Pediatría, Hospital Civil de Guadalajara “Dr. Juan I. Menchaca”, Guadalajara, Jalisco, México; 2) University of Lausanne, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; 3) Instituto de Genética Humana “Dr. Enrique Corona Rivera”, CUCS, Universidad de Guadalajara, Jalisco, México.

Introduction. The name of metatropic dysplasia (MD[MIM #156530]) is derived from the Greek “metatropos,” (changing patterns), because initially it exhibited a long thorax with short limbs and later a short-trunked dwarfism. MD is an autosomal dominant disorder caused by mutations in the TRPV4 gene (MIM #156530), which encodes for a calcium channel that participates in the osmosensation of the chondrocytes, however, TRPV4 produces also other skeletal dysplasias and neurodegenerative disorders. Here, we present the clinical, radiological and molecular findings in a female patient with a classic form of MD, complicated by a cervical myelopathy. Clinical report. The proposita was born from healthy, and remotely consanguineous parents. She was born after an uncomplicated term pregnancy and a normal vaginal delivery. At birth, she weighed 3.5 Kg (50th percentile). Physical examination at 2.5 years of age showed a weight of 11 Kg (-2.9 SD), height of 81 cm (-3.4 SD), span of 87 cm. She had short trunk, severe kyphoscoliosis, hypermobile fingers, enlarged joints at knees and elbows, and hallux valgus. Radiographies demonstrated thoracolumbar scoliosis, platyspondyly, wide intervertebral space, lumbar hyperlordosis, halberd-shaped pelvis, dumbo-ball-shaped metaphyses of femurs, and wide-irregular metaphyses in fingers and toes. The spine MRI showed C1-C2 subluxation, cervical instability, and spinal cord compression. At the age of 5 years the proposita undergo thru a surgical procedure with cervical arthrodesis, unfortunately, she progressed to quadriparesis six months after the surgery. The molecular study of the TRPV4 confirmed the mutation P799L at the heterozygous state.

Conclusions. The clinical and radiographic features, in addition to the molecular study in our patient confirmed the diagnosis of MD. All of the cases with a P799L heterozygous mutation in the TRPV4 gene have been reported in MD1, in our case with a suitable genotype-phenotype correlation even considering the cervical myelopathy. Early detection with a careful neurological assessment, spine MR imaging, and appropriate surgical treatment should be carried out in all patients in whom MD is suspected, to prevent the observed neurological sequelae and improve its life quality.
CNS Involvement in OFD1 syndrome: A Clinical, Molecular, and Neuroimaging study. B. Franco1,2, M. Macca1, F. Imperati2, A. D’Amico2, P. Parent2, L. Pasquier5, V. Layet6, S. Lyonnet7, V. Stamboul-Darmency6, C. Thauvin Robinet1,2, E. del Giudice1, Oral-Facial-Digital Type I (OFD1) Collaborative Group, 1) TIGEM, Fondazione Telethon, Naples, Italy; 2) Department of Translational medical Sciences, Federico II University of Naples, Italy; 3) Department of Radiological Sciences, Neuroradiology Unit, Federico II University of Naples, Naples, Italy; 4) Service de Pédiatrie et de Génétique, CHU Brest, France; 5) Service de Génétique Médicale - Centre de Référence Maladies rares Labélisé « Anomalies de Développement et Syndromes Malformatifs » de l’Ouest, CHU Rennes, France; 6) Consultation de Génétique, Groupe Hospitalier du Havre, CH Le Havre, France; 7) Département de Génétique, Hôpital Necker-Enfants Malades, APHP, Paris, France; 8) Service de Pédiatrie, Hôpital d’Enfants, CHU Dijon, France; 9) Centre de Génétique et Centre de Référence Maladies rares Labélisé « Anomalies de Développement et Syndromes Malformatifs » de l’Est, Hôpital d’Enfants, CHU Dijon, France; 10) E4271 GAD « Génétique des Anomalies du développement », IFR 100 - Santé STIC, Université de Bourgogne, Dijon, France.

Background: Oral-facial-digital type 1 syndrome (OFD1; OMIM 311200) belongs to the expanding group of disorders ascribed to ciliary dysfunction. With the aim of contributing to the understanding of the role of primary cilia in the central nervous system (CNS), we performed a thorough characterization of CNS involvement observed in this disorder. Methods: A cohort of 117 molecularly diagnosed OFD1 type I patients was screened for the presence of neurological symptoms and/or cognitive/behavioral abnormalities on the basis of the available information supplied by the collaborating clinicians. Seventy-one cases showing CNS involvement were further investigated through neuroimaging studies and neuropsychological testing. Results: Seventeen patients were molecularly diagnosed in the course of this study and five of these represent new mutations never reported before. Among patients displaying neurological symptoms and/or cognitive/behavioral abnormalities, we identified brain structural anomalies in 88.7%, cognitive impairment in 68%, and associated neurological disorders and signs in 53% of cases. The most frequently observed brain structural anomalies included agenesis of the corpus callosum and neuronal migration/organisation disorders as well as intracerebral cysts, porencephaly and cerebellar malformations. Conclusions: Our results support recent published findings indicating that CNS involvement in this condition is found in more than 60% of cases. Our findings correlate well with the broad spectrum of developmental anomalies described in other ciliopathies. Interestingly, we also described specific neuropsychological aspects such as reduced ability in processing verbal information, slow thought process, difficulties in attention and concentration, and notably, long-term memory deficits which may indicate a specific role of OFD1 and/or primary cilia in higher brain functions.
2720S
Clinical-pathological features in a female infant with Pfeiffer syndrome type 3 negative for FGFR2 mutation. AK. SANDOVAL TALAMÁNTEZ1, C. PEÑA-PADILLA1, E.L. MELLIN-SANCHEZ1, J. RIVERA-VARGAS1, E. ZAPATA-ALDANA1, L. VIRAMONTES-AGUILAR1, J. TAVARES-MACIAS1, L. BOBADILLA-MORALES1, M. CUNNINGHAM2, S. PARK2, J.R. CORONA-RIVERA1. 1) Centro de Registro e Investigación sobre Anomalías Congénitas (CRIAC), Servicio de Genética y Unidad de Citogenética, Hospital Civil de Guadalajara Dr. Juan I. Menchaca. Guadalajara, Jalisco, México; 2) Craniofacial Center, Seattle Children's Hospital, Washington, USA.

Introduction. Pfeiffer syndrome (PS) is a genetic disorder characterized by craniosynostosis, broad thumbs and big toes, and variable syndactyly (OMIM #101600). Cohen classified three types of PS: Type 1, have a better prognosis, usually familiar and could have normal intelligence. Type 2, is lethal with cloverleaf skull, severe ocular proptosis, elbow ankylosis and visceral anomalies. Type 3, share the same features that type 2, without cloverleaf skull, and can have abdominal wall defects. All of the cases with PS types 2 and 3 are caused by different mutations in FGFR2 gene, 80% of them in exons 8 and 10. Here, we describe the clinical-pathological features found in a female infant with PS type 3. Clinical report. The proposita was the product of the first pregnancy of a healthy 21-year-old mother and a 25-year-old father, who were non consanguineous. Prenatal ultrasonography reported omphalocele and craniosynostosis. The mother has positive serology for parvovirus B19. Vaginal delivery was at the 40th week of gestation. Apgar scores were 5, 3, and 1 at 1, 5 and 10 minutes, respectively. Birth weight was 3200 g, length 50 cm, and occipitofrontal circumference 31.3 cm (<3rd percentile). Physical examination showed turri-brachycephaly, ocular proptosis, midfacial hypoplasia, down-slanting palpebral fissures, low-set ears, hemangioma on dorsum of the nose, prominent abdomen with prune-belly appearance, supraumbilical midline omphalocele; limitation of elbow extension, and square shaped first toes. He died at 30 min after birth. Autopsy was performed adding: brain microgyria, laryngotracheal stenosis, patent ductus arteriosus, omphalocele, hepatomegaly, partial intestinal malrotation with a subhepatic cecal appendix. Keratomyte with G-bands was normal 46,XY, at a 550 band-level resolution. Sequencing of the exons 3, 7, 8, 10, 11, 14, 16, 17 of the FGFR2 gene, was negative for deletions, low-set ears, hemangioma on dorsum of the nose, prominent abdomen with prune-belly appearance, supraumbilical midline omphalocele; limitation of elbow extension, abdominal wall defects, visceral anomalies, and early death supports the clinical diagnosis of PS type 3. Four previous patients with PS and prune belly have been reported –two with cloverleaf skull, however, molecular study was not performed in none of them. The negative result obtained for the search of common mutations in FGFR2 gene, probably implies a new or atypical mutation for the severe PS phenotype in our patient and consequently, further investigations are necessary.

2721M
Identification of structural alterations in the CX50 gene in patients with congenital cataracts. A.L. Araujo1, E. Figueiredo2, P.R.S. Cruz1, B.B. Souza1, C.E.L. Arieta2, M.B. Melo1. 1) Center of Molecular Biology and Genetic Engineering (CBMENG), University of Campinas, Campinas, São Paulo, Brazil; 2) Department of Ophthalmology, Faculty of Medical Sciences (FCM), University of Campinas, Campinas, São Paulo, Brazil.

Congenital cataract is the leading cause of reversible blindness in childhood, with a prevalence of one to five cases per 10,000 live births. The hereditary form represents about 50% of these cases. Most common mode of inheritance is autosomal dominant with high penetrance, although autosomal recessive and X-linked inheritance are also observed. Genetic alterations responsible for non-syndromic congenital cataracts lead to changes in lens proteins, such as crystallins, membrane transport proteins (connexins, aquaporin) and cytoskeletal proteins (BIFSP2). Connexins are structural proteins that compose an extensive network of intercellular channels of low resistance (gap junctions) present between the lens fibers. These proteins allow ions, metabolites and second messengers flow between cells, maintaining the balance of intercellular communication and ensuring the transparency of the lens. Connexins are encoded by CX46 and CX50 genes (among others), both previously linked to congenital cataract. In order to better understand the molecular basis of the disease, structural alterations in the CX50 gene were evaluated in patients with nuclear and lamellar bilateral congenital cataract with hereditary etiology by means of PCR and direct sequencing. Patients' recruitment and blood collection were approved by the Ethics Committee of FCM-UNICAMP (Campinas-SP, Brazil). Twenty seven patients from ten families were analyzed to date, sixteen affected and eleven unaffected. We observed eight point mutations, among which two were not yet described. There were alterations common to both affected and unaffected individuals (two), unique for affected individuals (five, including the novel ones) and unique for unaffected individuals (one) as well. Five alterations are located in the coding region, being three missense variations, including both novel ones, and two synonymous. All three missense variations were predicted to be probably damaging by SIFT, PolyPhen-2 and MutPred softwares. Noteworthy, the missense variation already described was observed exclusively in both affected members of a family of three individuals, being absent in one unaffected member. This might suggest a possible segregation with the disease, although the family size restricts the analysis power. The remaining members of the families are being analyzed to allow further conclusions about the alterations. Financial Support: Capes and FAPESP.

2722T
Female with a Complex Phenotype Associated with Variants in Two Neurodegenerative Genes Detected by Whole Exome Sequencing: Diagnostic and Counseling dilemma. M. Khalifa1, L. Naffaa2. 1) Medical Genetics and Genomics, Akron Children Hospital, Akron, Ohio, USA; 2) Lena Naffaa. Department Radiology, Akron Children’s Hospital, Akron, Ohio, USA.

We report on a female child with severe intellectual disability, aphasia, short stature, ataxia, and structural brain abnormalities. A brain MRI obtained in infancy showed hypomyelination involving the central periventricular white matter and thinning of the corpus callosum. Initial reporting of Whole Exome Sequencing (WES) identified three POLR3A 3 missense heterozygous variants: two, a maternal variant (c.1724A>T) and a paternal variant (1745G>A) on exon13, and a novel maternally inherited variant (c.346A>G) on exon 4. These variants are likely damaging. The patient’s clinical features, early MRI findings, and WES results supported the diagnosis of hypomyelinating leukodystrophy type 7, and the family was counseled accordingly. An updated WES report revealed a novel WDR45 deletorius frameshift mutation in Exon 9 (c.587-588del). This, with later brain MRIs that showed progressive iron accumulation in the globus pallidus and substantia nigra, raised the possibility of Neurodegeneration with Brain Iron Accumulation (NBIA). Both WDR45 and POLR3A are newly recognized genes; each is associated with a distinct neurodegenerative disease. WDR45 is an X-linked gene associated with a dominant form of NBIA, manifested by progressive neuropsychiatric abnormalities, dystonia, cognitive decline, spastic paraplegia, and iron deposition in the basal ganglia. POLR3A, on the other hand, is an autosomal gene causing a recessive form of a hypomyelination with leukodystrophy disease, also known as 4H syndrome. This syndrome is characterized by variable manifestations including congenital Hypomyelination with thinning of the corpus callosum, Hypodactyly and Hypogonadism. This patient’s complex clinical presentation and mixed brain neuroimaging findings might be attributed to the confounding effects of the expression of these two genes. Alternative, they may represent an expansion of the phenotype of either condition. In any event, these made the diagnosis and genetic counseling of the family challenging.
2723S
Whole exome sequencing allows the identification of a novel large deletion in PRPF31 in a family with autosomal dominant retinitis pigmentosa. B. Almoguera1, P. Fernandez-San Jose2, Y. Liu1, J. Liang1, R. Golhar1, M. March1, R. Pellegrino1, M. Corton2,3, F. Blanco-Kelly2,3, M. Lopez-Molina2,3, B. Garcia-Sandoval2,3, Y. Guo1, L. Tian, X. Liu1, L. Guan1, B. Keating1, X. Xu1, C. Ayuso1,3, H. Hakonarson1. 1) Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Department of Genetics and Genomics, IIS-Fundacion Jimenez Diaz, 28040, Madrid, Spain; 3) Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Spain; 4) BGI-Shenzhen, Shenzhen 518033, China; 5) Department of Ophthalmology, Fundacion Jimenez Diaz, 28040, Madrid, Spain.

With all types of Mendelian inheritance patterns described and an extreme phenotypic, clinical, genetic and allelic heterogeneity, diagnosis of patients with retinitis pigmentosa (RP) is a complex task. Next generation sequencing (NGS) is increasingly being applied to the diagnosis of highly heterogeneous monogenic diseases, such as RP, as a rapid and affordable alternative to the former methods. However, one of the main limitations of NGS is the detection of copy number variants (CNV), which can account for a significant number of RP cases. In this study, we identified a large deletion in PRPF31 using whole exome sequencing (WES) data in a family with autosomal dominant RP. Eight members of the family were enrolled in the study and two affected and two unaffected members were subjected to WES using Illumina Hiseq 2000 instruments. The Burrows-Wheeler Alignment tool, Genome Analysis Tool Kit and ANNOVAR were used for mapping, genomic variant detection and variant functional annotation, respectively. Single nucleotide variant (SNV) analysis was performed searching among all poten-
tially relevant genes for CNV in the affected family. Our exome Markov Model (XHMM) was used. CNV validation was performed using TaqMan predesigned probes (Life Technologies) and the Universal Probe Library (Roche). SNV analysis did not yield any candidate variant. A hetero-
ygous deletion in the genomic region chr:19:54600186-54629178, spanning the entire PRPF31 gene, was identified by XHMM in the two affected members of the family. Both CNV assays used confirmed that the entire coding sequence of PRPF31 was hemizygotically deleted in the two affecteds and not in the unaffecteds. With this study we demonstrate the feasibility of detecting CNVs in RP genes using WES, thus expanding the potential of this tool in the diagnosis of this disease.

2724M
Whole gene duplication and partial duplication and triplication of OPHN1. J.G. Pappas1, E. Ward1, P.R. Papenhausen2. 1) Dept Pediatri

Inactivating mutations and inactivating deletions in OPHN1 cause x-linked mental retardation with cerebellar hypoplasia, ataxia, seizures and distinctive facial appearance (Al-Owain M et al 2011). Duplication X(q12q13.3) containing OPHN1 along with other OMIM annotated genes has been associated with developmental delay, microcephaly, short stature and autism in males (Bedeschi, MF et al 2008, Kaya N et al 2012, Prontera P et al 2012). We present a 14 year old boy (case 1) with dup(X)(q12q13) that includes OPHN1 and AR and a 3 year old boy (case 2) with duplication and triplication in Xq12q13 only. Our cases are unrelated and presented with mild to moderate mental retardation. Case 1 was born full term male, small for gestational age with microcephaly. He started walking at 17 months. At 28 months, he was speaking a few words; his face was triangular and elongated with prominent forehead, hypotelorism and bulbous tip of nose. Blood from case 1 was analyzed by SNP microarray (Affymetrix 6.0) which showed a 983 KB interstitial duplication of X(q12q13): arrXq12q13(136,720,159-67,702,721)x2. This de novo copy number change was confirmed by the RP11-963N10 region specific BAC. Case 2 was tested with a higher resolution microarray (Affyme-
trix CytoScan 750 which revealed a maternally inherited duplication Xq12(67,295,132-67,404,398) and triplication Xq12q13(136,750,664-
67,908,008). The mother was phenotypically normal. The duplication in case 1 shares only OPHN1 with the previously reported duplications and it contributes to the association of the phenotype to OPHN1 only. Case 2 is unique because it lacks ataxia and seizures which are common in cases with disruption of OPHN1. Our cases have common facial features with cases reported with duplications, inactivational deletions and inactivating mutations in OPHN1 and this observation may help elucidating the role of this gene during development.

2725T
A Unique Family with Progressive Pseudorheumatoid Arthropathy of Childhood. A. Neogi1, V. Kimonis1, J. Soni1, E. Chao1. 1) Division of Genetics and Genomic Medicine, Department of Pediatrics, University of California, Irvine, CA; 2) Gandhi Lincoln Hospital, Deesa, Banas Kanta, Gujarat, India.

We report a family in India with five affected female siblings with an unusual manifestation of progressive pseudorheumatoid arthropathy of childhood. Progressive pseudorheumatoid arthropathy of childhood is an autosomal recessive condition that affects the bones and cartilage of joints with a childhood onset. Major symptoms include stiffness of joints, bony swellings of toes and fingers, weakened muscles, fatigue and bowed legs. The disorder is progressive and occurs due to mutations in the WISP3 gene encoding for a signaling protein that is essential for normal postnatal skeletal growth and cartilage homeostasis. In the affected girls, symptoms typically started around 3-5 years of age and lead to crouching followed by a squatting gait. They developed severe joint contractures of the elbows, knees, ankles and hips, with motor weakness and difficulty sitting and walking. The girls developed joint swellings. They did not have pain but have severe restriction of their mobility, which has been documented in a video clip. Skeletal X-Rays revealed enlargement and cystic malformations of the epiphyses and metaphyses and joint erosions. We performed exome sequencing through Ambry Genetics for the affected individuals. Two novel mutations in the WISP3 gene were reported after assessment. These c.172A>T, p.K58X and c.737_738del, p.L246LfxX32 mutations are considered most likely con-
tributory to the progressive arthropathy. The girls have an unaffected brother and two unaffected sisters and consanguinity was denied in the family.

Confirmatory testing is in progress in the parents in order to confirm phase matching. Our findings provide further support for the molecular etiology of the disorder in this family. Future research of this pathway could lead to promising therapies to modify this progressive disease.

2726S
Mutation in the EZH2 gene in a Brazilian family - Complex clinical findings. D.L. Polla1, N.M. Kokitsu-Nakata2, M.C.B. Silva1, I.C.C. Cardoso1, A. Richieri-Costa2, R. Pogue1. 1) Catholic University of Brasilia, Brasilia, Distrito Federal, Brazil; 2) Hospital for Rehabilitation of Craniofacial Anomalies - University of Sao Paulo, Bauru, Sao Paulo, Brazil.

Weaver syndrome (WS) is a multi-systemic pre- and post-natal overgrowth syndrome associated with variable intellectual disability. Most of the reported cases have been described as sporadic, however in a few instances vertical transmission has been documented indicating autosomal dominant inherit-
ance. Here we describe a family with 3 affected patients (mother, son and daughter), and an unusual clinical presentation within the family. There was a wide variation in clinical manifestations ranging from severe vestibule-
lar defects for the proposita to behavioral disturbance, delayed lan-
guage acquisition and learning difficulties in her brother. The original diag-
nostic hypothesis was Sotos syndrome. However, sequencing was carried out using an NGS (Illumina) skeletal dysplasia-specific panel which included the Sotos syndrome genes (NDS1 and NFIX) and the WS gene, EZH2. The affected individuals shared c.149A>G; p.50L>S variant. A hetero-
duplication of the variants and provide further support for the molecular etiology of this family. Future research of this pathway could lead to promising therapies to modify this progressive disease.
A Novel Mutation, p. (Lys1474*), in a Female adds Seizures and Ptosis to Clinical Findings in MED13L Haploinsufficiency Syndrome. M.M. Ali1, N. Smouei2, A.M. Slavotinek3, 1) Genetics, UCSF, San Francisco, CA; 2) GeneDx, Gaithersburg, Maryland, USA.

We describe a 12 year old female with a history of seizures and epileptiform activity on EEG, choreiform movements, moderate developmental delays, macrocephaly (OFC >97th centile), central obesity and dysmorphic findings, including a prominent and square forehead, left ptosis, epicanthic folds, a right ear pit, prominent lower jaw with underbite, thin and tapered fingers, cubitus valgus and an increased lumbar lordosis. An MRI showed T2 hyperintense foci in the periventricular and subcortical white matter with low white matter volume and a dysmorphic craniocervical junction. Whole exome sequencing demonstrated heterozygosity for a de novo, nonsense mutation in the Mediator Complex Subunit 13-like (MED13L) gene, p. (Lys1474*), that has not been reported previously and is predicted to cause loss of normal protein function. The MED13L protein is a subunit of the Mediator complex, a large complex of proteins that functions as a transcriptional coactivator for most of the RNA polymerase II-transcribed genes. MED13L is highly expressed in early development of the heart and brain. Heterozygous, missense mutations in MED13L were first described in patients with transposition of the great arteries, dextro-looped (dTGA) and one patient with a translocation disrupting MED13L between exons 1 and 2 had intellectual disability (ID) and dTGA. Subsequent reports have described mutations and translocations hypothesized to disrupt MED13L expression in patient with phenotypes including non-syndromic ID, hypotonia, congenital heart defects, facial anomalies and a ‘‘Noonan-like’’ phenotype with hypotonia and facial anomalies. MED13L dosage is important and mutations have been non-penetrant, heterozygous or homozygous; a de novo triplication of chromosome 12q24.2 including MED13L and MAP1LC3B2 has also been reported to result in a milder phenotype. Our case is a further example of MED13L haploinsufficiency syndrome and adds seizures, choreiform movements and ptosis as possible clinical findings in this condition.


The (acro)mesomelic skeletal dysplasias, classified as n 16 and 17 under the current skeletal dysplasias nosology, comprise a heterogeneous group of disorders, ranging from mildly affected Leri-Weill, Robinow dominants, Maroteaux type to more severe phenotypes such as Langer, Kantaputra, Niemeyer, and Savarirayan types. We report on a 14 year-old girl with disproportionate acromesomelic short stature and bilateral congenital glaucoma. Product of an uneventful pregnancy, she was born by cesarian section at term, small for gestational age, BW: 2320g, BL: 38.5cm (below 5th%ile), OFC: 34.5cm, Apgar score 4/9. Parents are nonconsanguineous, with a negative family history. Primary congenital glaucoma surgery was performed at 1mo. She evolved with normal developmental milestones and cognitive scores. Echocardiogram, abdominal US, brain CT scan, and chromosome analysis were all normal. At 12 y 3mo she presented with H: 98.5cm (below 5th%ile), and normal OFC. Mesomelic shortening of the upper and, especially, lower limbs, brachydactyly with hypoplastic nails and an increased distance between the 1st and the 2nd toes were observed. Radiological findings included a mild deformity of the forearm, hypoplastic middle and distal phalanx and carpal bone fusions, and a more prominent deformity of the lower limbs, with short and broad tibiae-fibulae, brachymetatarsalia with the 2nd metatarsal and the intermediate cuneiform bone. Small iliac wings with coxa vara were also found. The skeletal findings presented by our patient seem distinct from the previous described acromesomelic disorders, although the mesomelic involvement resembles the one described in Savarirayan type. Besides, the eye abnormality is rarely reported in skeletal dysplasias. The etiological genetic background for some of the acromesomelic dysplasias is still unknown. NGS of the exome for the case here described is underway, with preliminary results. This could soon unmask the causative gene, and whether this is a unique entity or part of the spectrum of a known genetic defect.

An autosomal recessive microcephaly syndrome with primordial growth failure and pigmentation changes is caused by mutations in the gene ANKLE2. R. Clark, C. Curry, W. Dobyns, J. Lupsko, H. Bellien, M. Wangler. 1) Division of Medical Genetics/Peds, Loma Linda University Medical Center, Loma Linda, CA; 2) UCSF/Geneic Medicine Central California, Fresno, CA; 3) Center for Integrative Brain Research, University of Washington, Seattle, WA; 4) Dept. of Human and Molecular Genetics, Baylor College of Medicine, Houston, TX.

We report a non-consanguineous Hispanic family with two children affected with severe congenital microcephaly, primordial growth failure, pigmentation changes and novel mutations in ANKLE2 identified through whole exome sequencing (WES). The proband, now aged 8, presented with a history of mildly low birth weight (2.67 kg) and severe microcephaly. His growth parameters have been severely reduced with weight at -4 SD, height at -6 SD and OFC at -9 SD. His development is at the <6 mos level. He has had unexplained anemia, glaucoma, seizures and feeding issues requiring a G tube. His examination is notable for hyperpigmented and hypopigmented macules over his entire body, a sloping forehead, ptosis, microenopsia and generalized spasticity. Brain MRI revealed a simplified gyral pattern, absence of the corpus callosum, thick cortex and small frontal horns of the lateral ventricles with mildly enlarged posterior horns. Normal evaluations included chromosomes in blood and skin, a microarray, and studies for the Nijmegen breakage syndrome and Fanconi Anemia. His affected sister had a similar phenotype and died within hours of birth of progressive cardiac failure without a pathological etiology. Her OFC at 37 wks GA was 26.5 cm and she had diffuse hyperpigmented and hypopigmented skin macules. A severe autosomal recessive microcephaly was suspected in this family and subsequently his severely decreased severity of the brain abnormality, this appeared to be a unique phenotype.

DNA from both siblings and parents was sent for WES, and no mutations were found in previously published microcephaly loci. Among the candidates that met Mendelian expectation in this family, four genes were conserved and expressed in the CNS. One of these genes had been flagged as a potential novel Mendelian disease gene from a forward mutagenesis screen of the Drosophila X-chromosome which identified 165 fly genes. A subsequence of the DNA variants in ANKLE2 are the best candidate for disease-causing mutations in this family because of the segregation of the phenotype and the consistent neurologic phenotype in mutant flies with small brain size, and loss of neurons and support cells of the fly sensory organs.

Background: The association of TWIST1 haploinsufficiency with Saethre-Chotzen syndrome (SCS [MIM 101400]; coronal synostosis with ptosis, small ears, and minor limb anomalies) is well known. Here we present evidence that a specific region of TWIST1 lead to a radically different craniofacial phenotype, possibly through a dominant-negative mechanism.

Methods: The proband was a 3-year old male with hypertelorism, a wide anterior fontanelle, upper eyelid colobomas, deficient bony orbits with pseudoprotosis, small low-set dysplastic cupped ears, syndactyly of fingers, bilateral talipes, bilateral undeveloped testes, imperforate anus and hydropenis. Parent/child trio-based exome sequencing revealed a de novo c.350A>T (p.Glu117Val) substitution in TWIST1. As the patient did not resemble typical SCS, the remaining exome was analysed but no other convincing pathogenic changes were identified. Presentation of the findings at the UK Dysmorphology group prompted TWIST1 sequencing of a female subject with similar, but more severe features including bilateral eyelid colobomas, hypertelorism, UK and choanal atresia, campodactyly and abnormal hair distribution. This revealed a different heterozygous mutation at the identical nucleotide (c.350A>G; p.Glu117Gly). Of note, mutations at TWIST1-Glu117 have never before been associated in report. This residue lies in the DNA binding domain, is highly conserved and makes specific contacts with bases involved in DNA binding. A heterozygous mutation at the equivalent residue (p.Glu29Lys) in the C.elegans homolog hih-8 shows dominant inheritance whereas null alleles are recessive

Discussion: On the basis of the similarity of the patients’ clinical features, and their dissimilarity to SCS, we propose that heterozygous substitutions at Glu117 of TWIST1 are associated with a unique phenotype. Whereas in SCS, haploinsufficiency of TWIST1 causes a change in homo/heterodimer balance in the coronal suture leading to craniosynostosis, in the cases presented here, a dominant-negative mechanism may act at an earlier stage of development to disrupt facial tissues derived from the cranial neural crest.


Background: Opitz GBBB syndrome is a well-described genetic syndrome characterized by hypertelorism, hypospadias, dysphagia, and congenital heart disease. Opitz GBBB syndrome has been linked to Xp22 (XLOS) and 22q11.2 deletion syndrome.

Methods: We have shown that ADOS is a clinical entity distinct from XLOS and from 22q11.2 deletion syndrome. Whole exome sequencing of chromosome 22 will likely identify the causative gene for this syndrome and gene discovery for ADOS will have a significant impact on counseling affected families about transmission risks and in their decision for reproductive options. Additionally, finding the responsible gene including functional studies will lead to new scientific insight into molecular pathways involved in normal and abnormal midline craniofacial, cardiac and urogenital development.


Center de Registro e Investigación sobre Anomalías Congénitas (CRIAC), Servicio de Genética y Unidad de Citogenética, Hospital Civil de Guadalajara “Dr. Juan I. Menchaca”, Guadalajara, Jalisco, Mexico; 2) Anatomia Patológica, Hospital Civil de Guadalajara “Dr. Juan I. Menchaca”, Guadalajara, Jalisco, Mexico; 3) Instituto de Genética Humana “Dr. Enrique Corona Rivera”, CUCS, Universidad de Guadalajara, Jalisco, Mexico.

Introduction. C syndrome (COS) MIM 211750 is an autosomal recessive (AR) entity manifested by trigonencephaly, mental retardation, congenital cardiac defects, facial dysmorphism, postaxial polydactyly, CNS and renal malformations. CS is caused by mutation in CD96, mapped on 3q13. Otherwise, renal-hepatic-pancreatic dysplasia (RHPD) MIM 208540 is a rare AR condition characterized by renal cystic dysplasia (RCD), hepatic fibrosis (HF), hepatic cyst (HC), and pancreatic cyst (PC) or dysplasia, due to mutations in NPHP3, whose locus was at 3q22.1. The concurrence of these two entities in a single patient has not been reported previously.

Here we present a male infant with clinical pathological findings of CS and RHPD, supporting a common pathogenesis as both entities have the same mode of inheritance and chromosome localization. Clinical report. The postnatal growth was the product of the third pregnancy of nonconsanguineous, young, and healthy parents. He was born after uncomplicated pregnancy and normal vaginal delivery at 38 weeks of gestation. Birth weight was 2340 g (below 10th centile); length was 48 cm (25th centile), and occipitofrontal circumference (OFC) was 34 cm (below 10th centile). Physical examination showed hypotonia, mental and motor development delay, delay growth, postaxial polydactyly, and brachydactyly of both hands, and syndactyly of 2-3 toes. The CT brain scan showed corpus callosum hypoplasia and colpocephaly. 3D CT skull reconstructions confirm craniosynostosis of the metopic suture, hypotelorism, small posterior fontanelle, bilateral talipes, constrictive aortic, delayed ossification, and bony orbits with pseudoprotosis. Autopsy results showed severe congenital HF, biliopancreatic dysplasia, chronic renal failure, and sepsis at 1 month of age. The patient died of renal failure and sepsis at 1 month of age. Autopsy results confirm a complex phenotype, resembling that of CS and RHPD.

Conclusions. The patient is of interest because of his unusual concurrence of CS and RHPD. Although RHPD has been occasionally reported in CS, other obligate features of RHPD such as PC or HF, are not parts of the CS. This is the first time that the concurrence of these two entities is observed in a single patient. The patient is characterized by the presence of multiple congenital anomalies, most of them not previously reported in CS.

The proband was a 3-year old male with hypertelorism, a wide anterior fontanelle, upper eyelid colobomas, deficient bony orbits with pseudoprotosis, small low-set dysplastic cupped ears, syndactyly of fingers, bilateral talipes, bilateral undeveloped testes, imperforate anus and hydropenis. Parent/child trio-based exome sequencing revealed a de novo c.350A>T (p.Glu117Val) substitution in TWIST1. As the patient did not resemble typical SCS, the remaining exome was analysed but no other convincing pathogenic changes were identified. Presentation of the findings at the UK Dysmorphology group prompted TWIST1 sequencing of a female subject with similar, but more severe features including bilateral eyelid colobomas, hypertelorism, UK and choanal atresia, campodactyly and abnormal hair distribution. This revealed a different heterozygous mutation at the identical nucleotide (c.350A>G; p.Glu117Gly). Of note, mutations at TWIST1-Glu117 have never before been associated in report. This residue lies in the DNA binding domain, is highly conserved and makes specific contacts with bases involved in DNA binding. A heterozygous mutation at the equivalent residue (p.Glu29Lys) in the C.elegans homolog hih-8 shows dominant inheritance whereas null alleles are recessive

Discussion: On the basis of the similarity of the patients’ clinical features, and their dissimilarity to SCS, we propose that heterozygous substitutions at Glu117 of TWIST1 are associated with a unique phenotype. Whereas in SCS, haploinsufficiency of TWIST1 causes a change in homo/heterodimer balance in the coronal suture leading to craniosynostosis, in the cases presented here, a dominant-negative mechanism may act at an earlier stage of development to disrupt facial tissues derived from the cranial neural crest.
Molecular genetic study of 75 patients with X-linked alpha-thalassemia and mental retardation (ATR-X) syndrome in Japan. H. Shimbo1, K. Kurosawa1, N. Okamoto2, S. Ninomiya2, T. Wada3, 1) Kanagawa Children’s Medical Center, Yokohama, Kanagawa, Japan; 2) Osaka Medical Center and Research Institute for Maternal and Child Health; 3) Department of Clinical Genetics, Kurashiki Central Hospital, Kurashiki, Japan; 4) Department of Medical Ethics, Kyoto University Graduate School of Medicine, Kyoto, Japan.

[Introduction] ATR-X syndrome (MIM#603040) is an X-linked intellectual disability syndrome caused by mutations in the ATRX gene. The syndrome is characterized by severe intellectual disability, dysmorphic faces, hypopia, genital and skeletal abnormalities, and the presence of o-talassemia. More than 200 patients have been diagnosed worldwide; more than 80 reside in Japan. We have been studying ATR-X syndrome for more than a decade and have diagnosed almost all of the Japanese patients molecular genetically. In order to delineate molecular and clinical characteristics, we reviewed identified mutations in the ATRX gene of 75 patients from 62 families in Japan.

[Methods] We analyzed genomic DNA and/or cDNA synthesized from RNA extracted from the leukocytes of patients. PCR was performed using 40 and 12 pairs of primers for genomic DNA and cDNA, respectively, followed by sequencing by Sanger method.

[Results] Most of the mutations are clustered in two functionally important regions: 35 out of 75 mutations (47%) reside in ADD (ATRX-DNMT3-DNMT3L) domain, 27 (36%) in helix-c domain, and the other 13 (17%) are outside of these domains. The types of mutations are as follows: 55 missense mutations, 2 nonsense mutations, 5 small deletions, 5 frame shifts, 3 intragenic deletions, 2 nucleotide exchanges in introns, 1 large insertion in an intron, 2 nucleotide exchanges in 5’- and 3’-UTR, and 1 large deletion. [Discussion] We did not find a clear correlation between genotype and phenotype. However, patients with typically severe clinical features tended to have mutations in ADD domain. We identified several cases with atypical clinical features. (1) A case with intractable seizures with an intragenic deletion of exon 1 of the 35 exons in the ATRX and exons 2-10 in the neighboring gene MAGT1, which would result in a premature translation stop codon. (2) Two siblings presented with clinically quite different severity with the intragenic deletion involving linker regions, 2 in the fourth transmembrane segments, and 1 in the C-terminal domain. The type of the initial seizures was variable including generalized tonic-clonic, atypical absence, partial, apneic attack, febrile convolution, and loss of tone and consciousness. Onset of seizures was during the neonatal period in 2 patients, and between 3 and 7 months of age in 5 patients. Brain MRI showed cerebellar and cerebral atrophy in 1 and 6 patients, respectively. Our data reveals that initially uncontrollable seizures by any drugs, but eventually one was seizure-free and 3 were controlled at the last examination. All patients showed developmental delay or regression in infancy, resulting in severe intellectual disability. None of these patients had either upper limb dystonia, dyskinetic limb movements and poor eye contact.

This may suggest the existence of a mechanism of re-initiation at the downstream ATG codon, or the possibility of other mechanisms.

Impact of Plexiform Neurofibromas on Adult Patients with Neurofibromatosis type 1. S. Stueber, R.J. Hopkin, E. Schorry, L. Martin, K. Wusik, Cincinnati Children’s Hospital, 3333 Burnett Ave, Cincinnati, OH.

Neurofibromatosis type 1 (NF1) is one of the most common autosomal dominant genetic conditions diagnosed today. Up to 50% of patients with NF1 will develop a specific type of neurofibroma tumor called plexiform neurofibromas (PNFs), which develop along the nerve endings or along multiple nerve systems. Though typically benign, these tumors can cause significant complications. Some of the more common morbidities reported to be associated with PNFs are pain, loss of function, disfigurement, risk of malignancy development, mortality, and risk of bleeding complications. We performed a retrospective chart review of 69 adult patients that investigated rates of surgery and development of malignancy in patients with NF1 and PNFs. Results demonstrated that 50.7% of patients required surgery, and patients with greater numbers of PNFs were more likely to have surgery (p = 0.02). PNFs located in the head/neck and paraspinal region had a higher frequency of surgeries compared to their counterparts, respectively.

Recent studies have shown that recessive mutations in the TBC1D24 gene cause early-onset epileptic encephalopathy and sensorineural hearing loss. K. Wirtitz1, D. Neubauer2, D. Pappalardo1, N. J. Furlan1, B. Gnidovec Stražišar3, 1) Institute of Medical Genetics, University Medical Centre, Ljubljana, Slovenia; 2) Department of Child, Adolescent & Developmental Neurology, University Children’s Hospital, Ljubljana, Slovenia; 3) Department of Neonatology, University Children’s Hospital, Ljubljana, Slovenia.

Recent studies have shown that recessive mutations in the TBC1D24 gene cause a variety of epilepsy syndromes (familial early onset myoclonic epilepsy, early infantile epileptic encephalopathy 16), nonsyndromic deafness and DOORS syndrome (Deafness, Onychodystrophy, Osteodystrophy, Mental retardation and Seizures). We report on two siblings with early-onset epileptic encephalopathy and deafness. The patients presented with clonic and myoclonic jerks within one hour after birth. The seizures were resistant to conventional antiepileptic treatments. They had severe neuropsychological impairments with axial hypotonia, upper limb dystonia, dyskinetic limb movements and poor eye contact.

The older sibling died at the age of two years due to respiratory failure following severe respiratory infection. Audiologic examination showed bilateral sensorineural hearing loss in both siblings. The clinical and radiological investigation showed hypoplasia or any fingers of both hands, which is a typical feature of normalist phenotype. Next-generation sequencing panel for epileptic encephalopathy revealed compound heterozygous mutations in the TBC1D24 gene: a novel missense mutation c.52A>G (p.Asp11Gly) in exon 2 and a frameshift mutation c.1008dupT (p.His336Glnfs*12) in exon 4. Interestingly, the frameshift mutation was previously reported as unique to DOORS syndrome, yet our siblings did not meet the DOORS syndrome criteria. This report supports previous observations that mutations in TBC1D24 cause diverse phenotypes. We recommend testing all infants with unclassified early onset epileptic encephalopathy and sensorineural hearing loss for TBC1D24 mutations.
Complex Genomic Presentation in the NICU. A. Khromykh1, B.D. Solomon1,2, D.L. Bodian1, R.K. Iyer1, R. Baveja2, S.W. Wong1, K.C. Huddleston1, E.Z. Klein1, D. Ascher3, J.G. Vockley2, J. Niederhuber1, 1) Inova Translational Medicine Institute, Falls Church, VA; 2) Fairfax Neonatal Associates, Falls Church, VA; 3) Inova Children's Hospital, Falls Church, VA.

Congenital anomalies are a leading cause of infant morbidity and mortality. The goal of our IRB-approved 'Impact of Genetic Disorders' study is to use trio-based whole genome sequencing (WGS) to elucidate the genetic and genomic factors that contribute to the etiologies of congenital anomalies, as well as to assess these methods in real-time clinical practice, evaluating feasibility and cost-benefit ratio vs. conventionally available clinical genetic tests. Analysis was done through our ongoing trio-based (mother, father, proband) study (n=80+) that focuses on the subset of the NICU patient population suspected of having an underlying etiological genetic component and who underwent a comprehensive but unrevealing clinical/genetic work-up. Our research utilizes trio-based WGS, RNA expression, methylation, and miRNA characterization. Clinical data are obtained from the electronic health records as well as parental input on nutrition, stress, behavior and environmental exposure via study-specific surveys. A variety of principal component, familial-based, pathway and genomic network analyses are utilized to identify the underlying genomic explanation of the observed clinical presentation. Our analyses are bolstered by our database of ~5,000 WGS (along with other biological data and clinical information) derived from multiple ongoing trio-based genomic studies. Data generated from an initial subset of 30 probands reveals that not only were we able to bioinformatically identify the causative genetic/genomic mutations in multiple probands in our cohort utilizing WGS including molecular evidence for novel findings as well as improved diagnostic yield. A significant portion of our cohort was due to the potential of WGS analysis to identify molecular etiologies for previously undiagnosed patients in an efficient and increasingly cost-effective way. The establishment of our analysis pipeline can allow effective, timely analysis of undiagnosed patients in an efficient and increasingly cost-effective way. Our study demonstrates the potential of WGS analysis to identify molecular etiologies for previously undiagnosed patients in an efficient and increasingly cost-effective way.

The establishment of our analysis pipeline can allow effective, timely analysis of undiagnosed patients in an efficient and increasingly cost-effective way. The establishment of our analysis pipeline can allow effective, timely analysis of undiagnosed patients in an efficient and increasingly cost-effective way.
Deep sequencing detects very low-grade somatic mosaicism in the unaffected mother of siblings with nemaline myopathy. E. Koshimizu1, S. Miyatake2, Y.K. Hayashi2, K. Kiyomi3, Y. Imai3, M. Shira3, N. Nakashima1, Y. Tsurusaki1, N. Miyake1, H. Saito1, K. Ogata3, I. Nishino3, N. Matsumoto1. 1) Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 2) Department of Neurophysiology, Tokyo Medical University, Tokyo, Japan; 3) Department of Pediatrics, Faculty of Medicine, University of Toyama, Toyama, Japan; 4) Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan; 5) Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan.

When an expected mutation in a particular disease-causing gene is not identified in a suspected carrier, it is usually assumed to be due to germline mosaicism. We report here very low-grade somatic mosaicism in ACTA1, in an unaffected mother of two siblings affected with a neonatal form of nemaline myopathy. The mosaicism was detected by deep resequencing using a next-generation sequencer. We identified a novel heterozygous mutation in ACTA1, c.448A>G (pThr150Ala), in the affected siblings. To explore the effect of this mutation, we mapped the mutation onto reported crystal structures. Thr150 is located near the polymerization/interaction interfaces between actin monomers and between actin and its interacting proteins. Thus, pThr150Ala may affect polymerization and/or the interactions of actin with other proteins. In this family, we expected autosomal dominant inheritance with either parent demonstrating germline or somatic mosaicism. Sanger sequencing identified no mutation. However, further deep resequencing of this mutation on a next-generation sequencer identified a very-low-grade somatic mosaicism. The total read depth of c.448A in ACTA1 was 131495x to 425933x. Very-low-grade somatic mosaicism was confirmed in the mother: 0.4%, 1.1%, and 8.3% in the salvia, blood leukocytes, and nails, respectively. We used allele-specific PCR to confirm the presence of the mutation in the mother. Both the wild-type and mutant alleles were amplified in the proband and the affected sister at a similar level. Both alleles could also be amplified in the mother, but the wild-type allele was amplified at a much greater level than the mutant allele. The wild-type allele was amplified in DNA derived from the mother (DNA derived from blood), and the mother (DNA derived from the nails, which showed the highest rate of mosaicism (8.3%). The melting curves of both alleles were calculated, and the mother’s calls were called normal. Our study demonstrates the possibility of very-low-grade somatic mosaicism in suspected carriers, rather than germline mosaicism.

Identification of RIT1 mutations in patients with RASopathies by clinical whole exome sequencing. P. Li1, F. Xia2, W. He3, L. Potocki2, F. Scriver1, P. Magopoulous1, D.M. Suznjev2, A.L. Beaudet1, R.A. Gibbs2,6, C.M. Eng1, Y. Yang1. 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.

RASopathies including Noonan syndrome (NS) are a group of autosomal dominant developmental syndromes characterized by short stature, distinctive facial features, and congenital heart defects. These disorders are caused by germline mutations that alter the activity of the RAS/MEK/ERK signaling pathway in the majority of causal genes. Over the past years, recently, potential-of-function mutations in RIT1, which encodes a small GTPase that belongs to the Ras supergene family of low molecular weight GTP-binding proteins, were reported in nine unrelated Japanese patients with NS. The same mutation was also reported previously in two unrelated probands with NS. Two of these two mutations are located within the evolutionarily conserved GTPase domain in RIT1. Additionally, we identified the recurrent c.104G>C (p.S35T) mutation in a third patient who has a clinical diagnosis of NS, demonstrating the c.104G>C (p.S35T) mutation as a common mutation in RIT1-related RASopathy. Our results expand the molecular and phenotypic spectrum of RASopathy caused by RIT1 mutations, and demonstrate the advantage of WES in diagnosing newly delineated genetic disorders.

Splicing mutation in IQSEC2 gene modulating the phenotype in three siblings with intellectual disability. I. Madrigal1,2, J. Rosell1, L. Rodriguez-Reveenga1,2, M. Alvarez-Mora1,2, O. Karberg3, D. Elurbe1,2, A. Bayes3, S. Sauer1, AC. Syvänen4, M. Mita5,1. 1) Hospital Clinic Barcelona and IDIBAPS, Barcelona, Spain; 2) Centre for Biomedical Research on Rare Diseases (CIBERER), Spain; 3) Genetics Section, Hospital Universitari Son Espases, Palma de Mallorca, Spain; 4) Department of Medical Genetics, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 5) Molecular Physiology of the Synapse Laboratory IIB, Sant Pau, Barcelona, Spain; 6) Max-Planck Institute for Molecular Genetics, Ihnestrasse 63-73 D-14195 Berlin Germany.

The IQSEC2 (IQ motif and Sec7 domain 2) gene is located on chromosome Xp11.22 and encodes a guanine nucleotide exchange factor for the ADP-ribosylation factor family of small GTPases. This gene is known to play a significant role in cytoskeletal organization, dendritic spine morphology and synaptic organisation. Mutations in IQSEC2 cause moderate to severe intellectual disability (ID) in affected males and a variable phenotype in females due to this gene escapes X chromosome inactivation. Here, we report on a novel splice-site mutation in IQSEC2 gene (c.3116_3194del) that cosegregates with the disease in a family diagnosed with a nonsyndromic form of X-linked ID. The mutation is a deletion of 2bp (c.3116–2delCA) that activates an intronic splice acceptor site resulting in 78 nucleotides deletion in exon 12 (c.3116, 3194del). The activation of this acceptor site abolishes 26 amino acids from the highly conserved PH domain of IQSEC2 and creates a premature stop codon 36 amino acids later in exon 13. Two separate PCR products were obtained when IQSEC2 mRNA from exons 9 to 13 was analyzed. Interestingly, the percentage of aberrant splicing in each patient correlates with the severity of their clinical features. As we are concerned, we are describing the first splicing mutation affecting IQSEC2 gene that, moreover, modulates the severity of phenotype in these patients. Acknowledgements: This study was supported as a transnational access project at the European Sequencing and Genotyping Infrastructure (ESGI) which has received funding from the Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 262055. Exome sequencing was performed by the SNP&SEQ Technology Platform (www.sequencing.se), Department of Medical Sciences, Science for Life Laboratory, Uppsala University, a national infrastructure supported by the Swedish Research Council (VR-RFI) and the Knut and Alice Wallenberg Foundation. We also thank AAGU from the Autonomous Catalan Government (2009-SGR1337) and Fundación Agrupación Mutua (Premio Discapacidad 2012). The CIBER of Enfermedades Raras is an initiative of the ISCIII.
As the study expands in scope it has the potential to identify mutations in the treatment of a previously reported family with an autosomal recessive disorder. Successful identification of mutations assuming a recessive mode of inheritance. We have found putative mutations in the ANTXR1 gene, which encodes anthrax toxin receptor 1 has been reported in the progression of the clinical findings of GAPO syndrome in the first family consisting of two affected female siblings who presented with gait ataxia at 12 and 18 months of age. MIR revealed vermicular hypoplasia and increased spaces between the folia. WES identified compound heterozygous inheritance of two novel mutations (non-synonymous and splice site) in the aarF domain containing kinase 3 gene. Mutations in ANTXR1 cause GAPO syndrome and AMN, J. Jacobsen, S. Robertson, C. Wilson, B. Swan, E. Giammuzina, J. Taylor, R. Hill, D. Love, K. Lahner, R. Snell. 1. Centre for Brain Research and School of Biological Sciences, The University of Auckland, New Zealand; 2. Dunedin School of Medicine, The University of Otago, New Zealand; 3. Adult and Paediatric National Metabolic Service, Auckland City Hospital, New Zealand; 4. Genetic Health Service New Zealand, Auckland City Hospital, New Zealand; 5. Department of Neurology, Auckland City Hospital, New Zealand; 6. Diagnostic Genetics, LabPLUS, Auckland City Hospital, New Zealand.

New Zealand has a unique and diverse population for genetic studies. We are currently investigating the genetic underpinnings of genetically unconfirmed neurodevelopmental disorders in the New Zealand population. In a pilot study, using an exome sequencing approach, we have sequenced two affected individuals from each of four families and have searched for mutations assuming a recessive mode of inheritance. We have found putative mutations in ANTXR1 in the first family consisting of two affected female siblings who presented with gait ataxia at 12 and 18 months of age. MIR revealed vermicular hypoplasia and increased spaces between the folia. WES identified compound heterozygous inheritance of two novel mutations (non-synonymous and splice site) in the aarF domain containing kinase 3 gene. Mutations in ANTXR1 cause GAPO syndrome and AMN.
2747S

Mutations and clinical outcomes of copper-histidine therapy in Menkes disease patients. G. Kim1, J. Kim2, J. Choi3, B. Lee3,4, H. Yoo1,2.
1) Medical Genetics Center, Asan Med Ctr, Seoul, South Korea; 2) Dept. Pediatrics, Asan Med Ctr, Seoul, South Korea.

Menkes disease is a very rare X-linked copper metabolism disorder caused by mutations in the ATP7A gene. With the advent of subcutaneous copper-histidine therapy, the early diagnosis of Menkes disease becomes of utmost importance for patients' prognosis. The clinical characteristics of 12 Korean patients with Menkes disease (11 males and 1 female from 11 unrelated families) were studied along with the mutation spectrum. Only 2 males were diagnosed in the neonatal period, and the other male patients were diagnosed at age 4.3 ± 1.9 months. The presenting signs included depigmented kinky hair, neurologic deficits, and hypotonia. Serum copper and ceruloplasmin levels were markedly decreased. Intracranial vessels were dilated with tortuosity and accompanied by regional cerebral infarctions, even at an early age. Notably, a female patient who presented with developmental delay, was diagnosed at age 18 months by characteristic MRA findings, biochemical profiles, and genetic analysis. CAG repeats in the AR gene at Xq12 were investigated to evaluate the X-inactivation pattern, and the patient revealed a 76% inactivated normal X-chromosome. A total of 11 mutations in ATP7A were identified, including five novel mutations: p.L825fs1, c.1219-930del2626+48del (exons 3-12 deletion), c.3744_4123+576del (exons 19-21 deletion), p.N886fs¹, and p.H1086PfsX3. Most mutations were truncated (except 1 nonsense mutation), including 3 frameshift, 2 nonsense, 3 large deletion, and 2 splice-site variants. The age at commencement of copper-histidine treatment was 7.3 ± 7.5 (0.5 - 27) months. Despite treatment, seven patients died before age 5 years, and the remaining patients were severely retarded in neurodevelopment. The poor outcomes of our patients might be related to delayed therapy as well as truncated nature of ATP7A mutations.

2748M

The first two AUTS2 mutations on the nucleotide level causing AUTS2 syndrome. G. Beunders1, S.A. Munnik de2, J. Motton3, P. Vasudevan4, E. Voorhoeve1, A.J. Grafton5, W.M. Nillesen6, H. Meijers-Heijboer7, H.G. Yntema8, E.A. Sistemas9. 1) clinical Gen, VU Med Ctr, Amsterdam, Netherlands; 2) Department of Human Genetics, Radboud university medical centre, Nijmegen, The Netherlands; 3) Department of Clinical Genetics, Birmingham Women's Hospital, Edgbaston, Birmingham, UK; 4) Department of Clinical Genetics, University Hospitals of Leicester, Leicester, UK.

Background: We recently described a new ID syndrome, AUTS2 syndrome which is characterised by low birth weight, feeding difficulties, intellectual disability, microcephaly, and mild dysmorphic features. All cases thus far were caused by chromosomal rearrangements. Mutations at the base pair level disrupting AUTS2 have not yet been described.

Methods: Here we present the full clinical description of the first two cases with AUTS2 syndrome caused by mutations at the nucleotide level, found by diagnostic exome sequencing.

Results: The phenotypic features of both cases include: intellectual disability, microcephaly, feeding difficulties, dysmorphic features and mild contractures. Both cases have AUTS2 mutations in exon 7 (one nonsense mutation and one two-basepair deletion), both causing a premature stop of the full length transcript and without an predicted effect on the shorter 3’ transcript starting in exon 9 that is also expressed in human brain.

Conclusions: The similarities between the phenotypes of these two cases and formerly described patients with AUTS2 syndrome, further confirm that AUTS2 syndrome is a single gene disorder with a recognizable phenotype that can be caused by haploinsufficiency of the full length transcript of AUTS2.

2749T

Novel mutations and clinical outcomes of copper-histidine therapy in Menkes disease patients. G. Kim1, J. Kim2, J. Choi3, B. Lee3,4, H. Yoo1,2.
1) Medical Genetics Center, Asan Med Ctr, Seoul, South Korea; 2) Dept. Pediatrics, Asan Med Ctr, Seoul, South Korea.

Menkes disease is a very rare X-linked copper metabolism disorder caused by mutations in the ATP7A gene. With the advent of subcutaneous copper-histidine therapy, the early diagnosis of Menkes disease becomes of utmost importance for patients’ prognosis. The clinical characteristics of 12 Korean patients with Menkes disease (11 males and 1 female from 11 unrelated families) were studied along with the mutation spectrum. Only 2 males were diagnosed in the neonatal period, and the other male patients were diagnosed at age 4.3 ± 1.9 months. The presenting signs included depigmented kinky hair, neurologic deficits, and hypotonia. Serum copper and ceruloplasmin levels were markedly decreased. Intracranial vessels were dilated with tortuosity and accompanied by regional cerebral infarctions, even at an early age. Notably, a female patient who presented with developmental delay, was diagnosed at age 18 months by characteristic MRA findings, biochemical profiles, and genetic analysis. CAG repeats in the AR gene at Xq12 were investigated to evaluate the X-inactivation pattern, and the patient revealed a 76% inactivated normal X-chromosome. A total of 11 mutations in ATP7A were identified, including five novel mutations: p.L825fs1, c.1219-930del2626+48del (exons 3-12 deletion), c.3744_4123+576del (exons 19-21 deletion), p.N886fs¹, and p.H1086PfsX3. Most mutations were truncated (except 1 nonsense mutation), including 3 frameshift, 2 nonsense, 3 large deletion, and 2 splice-site variants. The age at commencement of copper-histidine treatment was 7.3 ± 7.5 (0.5 - 27) months. Despite treatment, seven patients died before age 5 years, and the remaining patients were severely retarded in neurodevelopment. The poor outcomes of our patients might be related to delayed therapy as well as truncated nature of ATP7A mutations.

2750S

Genetic analysis of an atypical case of Branchio-Oto-Renal (BOR) Syndrome. R. Birkenhager, T. Jakob, E. Prera, S. Arndt, W. Maier, R. Laszig, A. Aschendorff. Department of Otorhinolaryngology and Head and Neck Surgery, University Medical Center Freiburg, Kilianstrasse 5, D-79106 Freiburg, Germany.

Introduction: Branchio-oto-renal (BOR) (MIM 113650) and Branchio-otic (BO) (MIM 602588) syndromes are one of the most common forms of autosomal dominant syndrome hearing loss. BOR/BO syndromes are genetically heterogeneous and caused by mutations in EYA1, SIX1 and SIX5 genes. The phenotypes of BOR syndrome are highly variable, with common clinical features including hearing impairment (HI), malformations of the pinnae, the presence of branchial fistulae, and various renal abnormalities, congenital heart defects without renal abnormalities. In this study, clinical and genetic analysis were performed in a family with BOR syndrome focusing on auditory characterization and rehabilitation. Methods: Hearing testing BERA/Electrocochleography was only done in the Patient other family members were not tested because they have no hearing impairment or clinical abnormalities. Radiological a high-resolution CT scan was made. Genetic analysis was performed for EYA1, SIX1, and SIX5 genes of the affected individual and first degree relatives using direct sequencing of the coding exon and intron transitions of the genes, including deletion analysis. The modality and outcome of auditory rehabilitation were evaluated. The genotype-phenotype correlation was investigated. Results: One novel SIX1 mutation, a small deletion (c.376_378delAG, p.Glu126del) was identified in the family with BOR syndrome. EYA1 and SIX5 mutations were not detected in the present study. No renal and branchial defects were observed in the patients. Cochlear implantation performed in one patient resulted in significant hearing improvement. Conclusions: Successful outcome can be expected with cochlear implantation in patients with BOR syndrome who cannot benefit from hearing aids. The novel SIX1 mutations may add to the genotypic and phenotypic spectrum of BOR syndrome in the European population.
Mutations in ERF gene as a new genetic cause of craniosynostosis - enabling parents and clinicians to understand why a child is affected.

A. Chaudhry1, 2, P. Sabatini1, 2, L. Harr1, P. Ray3, C. Forrest1, 2, S. Bowdin2, 2. 1) Division of Clinical and Metabolic Genetics. Department of Paediatrics. The Hospital for Sick Children, 525 University Avenue, Toronto, ON, M5G 2L3; 2) University of Toronto, Toronto, ON; 3) Molecular Genetics Division, Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto; 4) Division of Plastic & Reconstructive Surgery, Centre for Craniofacial Care & Research, The Hospital for Sick Children.

Craniosynostosis is a genetically heterogeneous condition resulting from a prematurely fused suture or sutures. In recent years, mutations have been identified as a cause of craniosynostosis including FGFR1, FGFR2, FGFR3, TWIST, EFN1, FAM20C and LMX1B. Twigg et al. reported that for the first time in the literature, 3 patients with ERF-related craniosynostosis phenotype presented with a new and rare phenotype. The cases presented included microcephaly, hypertelorism, proptosis, prominent nose, bilateral flat nasal bridge, depressed nasal bridge and retrognathia in addition to the abnormal head shape as a result of craniosynostosis. These patients did not have developmental delays other than speech delay. At birth, the newborns exhibited multiple-suture synostosis, craniofacial dysmorphism (hypertelorism, shortening and/or vertical displacement of the nose, prominent orbits and forehead), Chiari malformation, language delay and behavioral issues. These patients presented after infancy and did not need repeated cranial surgeries. ERF encodes an inhibitory ETS transcription factor directly bound by ERK2 (refs. 2-7). The help of mouse models, the authors were able to show that reduced dosage of the ERF gene causes complex craniosynostosis in humans. In our prospective study, we sequenced the coding regions of the ERF gene in 43 patients who tested negative for mutations in FGFR1, FGFR2, FGFR3 and TWIST as well as by cytogenetics testing. In our cohort of patients with multi-suture or sagittal suture synostosis, we identified heterozygous pathogenic mutations in the ERF gene (c.374delA, c.1154delGC, c.1154delT, c.374delA, c.979delG and c.979delC). We sequenced the coding regions of the ERF gene in three of our patients (60%) presented later in infancy. Subject A presented at 6 and half years of age with headaches and visual loss secondary to oxycephaly. Z/S (40%) of our patients had other abnormalities on head imaging such as Chiari I malformation. These patients had an impressive dysmorphic phenotype including multiple-suture synostosis, hypotelorism, proptosis, depressed nasal bridge and retrognathia in addition to the abnormal head shape as a result of craniosynostosis. These patients did not have developmental delays other than speech delays in infancy. They did not need repeated cranial surgeries in general. In conclusion, ERF-related craniosynostosis should be suspected in patients who present later in infancy with multiple suture or sagittal synostosis and who have other cranial malformations like Chiari I malformation. For the purposes of counseling, it is important to note that most of these individuals do not have developmental delays, but few may have speech delay. In general, they do not need repeated cranial surgeries.

Hypotrichosis-Lymphedema-Telangiectasia (HLTS) is an ultra-rare, congenital condition that is characterized by blood and lymph vessel dysfunction, as well as hair follicle defects (OMIM #607823). Mutations in the transcription factor MITF are associated with Waardenburg syndrome (WS). In this study, we report a rare family composed of previously undescribed cases with the phenotype of HLTS. In this cohort, our study reported a rare family composed of previously undescribed cases with the phenotype of HLTS. In this cohort, we identified heterozygous pathogenic mutations in the MITF gene (MIM 136540), including two novel mutations (c.481C>T) that results in a premature STOP codon (p.Q161*). The mutation was found in the largest affected female family member. The mother and daughter of the patient had more severe symptoms compared to the father. For the first time, we are able to show that reduced dosage of the MITF gene may have implications for the development of the retina and the cardiovascular system. Moreover, we validated the MITF-wt protein and MITF-V307G protein in yeast and Drosophila melanogaster models. The MITF-wt protein is able to activate all of the above mentioned promoters, including the MITF-K307N mutant also activated all of the above mentioned promoters, while MITF- V307G mutant was unable to activate any of the tested promoters. Current studies are focused on how mutant MITF molecules may interact with the wild type protein, with each other, and with proteins important in ocular and other tissues. These results describe a novel syndrome caused by biallelic mutations in MITF (COLObotamato Microphthalmia, Macrocephaly, Albinism & Deafness (COMMAD syndrome) and a novel “activating” mutation in MITF gene.

A novel missense mutation of ryanodine receptor 1 (RYR1) in a Japanese idiopathic hyper CK-emia family. K. Sano1, 2, S. Miura1, 2, T. Fujisawa2, R. Fujikawa2, A. Yorita1, K. Noda1, H. Kida1, K. Azuma1, S. Kaieda1, K. Yamamoto1, T. Taniwaki1, Y. Fukumaki1, H. Shibata1. 1) Department of Respiratory, Neurology and Rheumatology, Department of Medicine, Kurume University School of Medicine, Kurume, Fukuoka; 2) Division of Genomics, Medical Research Institute, Kansai University Medical School, Osaka, Japan. Objectives: It is known that RYR1 mutations cause various clinical phenotypes, such as malignant hyperthermia or central core disease. In our study, we identified a novel missense mutation of RYR1 in a Japanese family with idiopathic hyper CK-emia.

A 4-year-old, female patient with anomalies in hair growth, subcutaneous herniation and hypotonia was referred for interdisciplinary evaluation. The patient had some growth retardation and slight delay in speech development. At age 13 years, a 1,014 base pair deletion in intron 15 of the RYR1 gene was identified by Sanger sequencing method. The patient was negative for mutations in the genes encoding the proteins of the triadin/furin (TDP-43/FTLD) family, such as TDP-43. A multiplex ligation-dependent probe amplification analysis was performed using a panel of 351 probes. The patient was negative for deletions and duplications of the genes of this family. The disease followed an autosomal dominant inheritance pattern. At age 13 years, a 1,014 base pair deletion in intron 15 of the RYR1 gene was identified by Sanger sequencing method. The patient was negative for mutations in the genes encoding the proteins of the triadin/furin (TDP-43/FTLD) family, such as TDP-43. A multiplex ligation-dependent probe amplification analysis was performed using a panel of 351 probes. The patient was negative for deletions and duplications of the genes of this family. The disease followed an autosomal dominant inheritance pattern.

A 13 year old, female patient with anomalies in hair growth, subcutaneous herniation and hypotonia was referred for interdisciplinary evaluation. The patient had some growth retardation and slight delay in speech development. At age 13 years, a 1,014 base pair deletion in intron 15 of the RYR1 gene was identified by Sanger sequencing method. The patient was negative for mutations in the genes encoding the proteins of the triadin/furin (TDP-43/FTLD) family, such as TDP-43. A multiplex ligation-dependent probe amplification analysis was performed using a panel of 351 probes. The patient was negative for deletions and duplications of the genes of this family. The disease followed an autosomal dominant inheritance pattern.
2755T
A novel case of Epidermolysis Bullosa with Pyloric Atresia due to homozygous mutation of c.600delC in integrin 4 gene: Clinical, Immunohistological and Molecular Diagnosis. S. Yılmaz1, I. Mungan Akın2, C. Chiaverini3, A. Charlesworth3, D. Buyukkayhan3, J.P. Lacour4, I. Akalin4. 1) Medical Genetics, Istanbul Medeniyet University, Department of medical Genetics, Istanbul, Turkey; 2) Neonatology, Istanbul Medeniyet University, Istanbul, Turkey; 3) Neonatology, Goztepe Training and Research Hospital, İstanbul Medeniyet University, Istanbul, Turkey; 4) Reference Centre for EBH, University Hospital of Nice, France.

Epidermolysis Bullosa (EB) with pyloric atresia (EB-PA) is a hemidesmosomal subtype of EB characterized by blister formation following minor trauma and pyloric atresia with/without renal disorders. The main pathology in EB-PA is basement membrane disruption due to abnormal expression of integrin and/or plectin proteins that are responsible for adherence of epidermis to lower layers in dermoepidermal junction. The genetic etiology of hemidesmosomal EB is homozygous or compound heterozygous mutations of PLEC, ITGA6 or ITGB4 genes. The proband was a male newborn with the findings of bilateral hydronephrosis, pyloric atresia and large aplasia cutis on four extremities. The parents were 2nd degree cousins and they had another newborn baby with the same clinical findings who died at 1st week. The patient was clinically diagnosed as EB-PA. Immunohistology of the biopsy specimen showed ITGB4 deficiency. Molecular analysis revealed a homozygous c.600delC mutation in exon 7 of ITGB4 gene. To our knowledge there is only one previous case of EB-PA with homozygous c.600delC mutation reported in the literature.

2756S
Increased Susceptibility to Attention Deficit Hyperactivity Disorder risk in Marfan Syndrome and Other Connective Tissue Disorders. A. Hall, M. Tucker, L. Escobar. Medical Genetics and Neurodevelopmental Center, St. Vincent Hospital, Indianapolis, IN.

Marfan syndrome is an autosomal dominant connective tissue disorder that involves the cardiovascular, ocular and skeletal systems and is caused by mutations in FBN1. Based on clinical observations, an increased susceptibility to difficulties with sustained visual attention and visuoconstructive skills may also be present in Marfan syndrome. We report a population of patients clinically diagnosed with a connective tissue disorder, including Marfan syndrome, seen in our medical genetics and neurodevelopmental clinic who had a high prevalence of attention deficit hyperactivity disorder (ADHD). Patients were referred to our clinic for genetic evaluation of various connective tissue disorders, including Marfan syndrome, between 2009-2014. Six out of 20 or 30% of patients with a clinical diagnosis of Marfan syndrome had comorbid ADHD with 3 out of 12 or 25% of patients with a molecular confirmation of Marfan syndrome having a clinical diagnosis of ADHD. In addition, 10 out of 32 or 34% of patients with a clinical diagnosis of a connective tissue disorder were also clinically diagnosed with ADHD. Our clinic observations suggest a possible link between connective tissue disorders, including Marfan syndrome, and an increased susceptibility to ADHD as has been previously reported in the literature. This possible link should be further investigated with increased patient populations.

2757M
A novel mutation in OPN1MW in a brazilian patient with x-linked retinal cone dystrophy type 5. A.CV. Castro1, A.T. Rassi2, A.N. Rocha2, L.FOB. Chaves1, L.SM. Mendonca1, T. Oliveira1, M.P. Avila1, I.M.M. Silva1, J. Chiang2, L.A.R. Gabrielli1. 1) Ophthalmic Genetics Department, Federal University of Goias, UFG, Goiânia, Goiânia, Brazil; 2) Molecular Diagnostic Laboratory, Casey Eye Institute, Portland, Oregon, USA.

Purpose: To report a novel mutation in the OPN1MW gene in a patient with x-linked retinal cone dystrophy type 5 presenting severephotophobia and tritanopia. Methods: Eye examination was performed. Best corrected visual acuity (BCVA) was done (Snellen chart). Color vision was measured with Farnsworth D-15 Dichromatous Color Blindness Test. Retinography, (TRC50DX Topcon) and optical coherence tomography-OCT (Spectralis Heidelberg Engineering Inc.) were done. Full-field electroretinogram (FF-ERG) was performed (Roland Consult RETIport Science 6.12.5). Direct testing for mutations in the GUCA1A gene was performed by PCR amplification and bidirectional DNA sequencing of all coding exons and exon/intron boundaries in a CLIA certified laboratory. Additionally, whole exome sequencing (WES) was performed. The results were analyzed and compared to their NCBI reference sequences. The bioinformatic tool Mutation Taster was used to predict the effect of the mutation on the protein. Protein domains were identified with the bioinformatic tool SMART. Results: Examination revealed a 65-year-old brazilian male with severe photophobia, obligating the patient to use dark lenses sunglasses all the time and pilocarpine eyedrops in order to obtain some relief. BCVA was 20/20 at the right eye (OD) and 20/40 at the left eye (OS). The Farnsworth 15 T est showed tritanopia. Retinography revealed moderate arteriolar vasoconstriction, severe venular vasoconstriction, retinal pigment epithelium atrophy, peripapillary atrophy, and retinal temporal atrophy with dark pigment clumping. The macula was relatively spared. The OCT showed a decreased foveal thickness on both eyes: 174μm in OD and 205μm in OS. The FF-ERG showed an important decrease in cones response (30%) and, in contrast, a discreet decrease in rods response (90%) on both eyes. The GUCA1A sequencing didn’t reveal any mutations. WES identified the novel missense mutation c.538G>T:p.Ala180Ser on the OPN1MW gene. This novel mutation was predicted as disease causing by Mutation Taster. It affects the 4th transmembrane domain of the OPN1MW protein according to SMART. Conclusions: Herein we’ve described a novel mutation in OPN1MW (c.538G>T:p.Ala180Ser) in a patient with severe photophobia and tritanopia.
2758T

Cerebrofaciothoracic dysplasia: a case report with molecular search for TMCO1 mutation. J. Rivera,1 P. Campeau,2 E.L. Mellin-Sánchez,2 F.H. Martínez-Miclia,1 L. Bodea-Chibucas,3,4 A. Corona-Rivera1,2,3,4, J.R. Corona-Rivera1,2,3,4, C.B. Chaves,1,2,3,4, A. Chintanapoke,1 J.R. Pascual-Castroviejo1,2,3,4. 1) Department of Pediatrics and Adolescent Medicine, Queen Mary Hospital, Li Ka Shing Faculty of Medicine, The Chinese University of Hong Kong, HKSAR, China; 2) Department of Obstetrics & Gynecology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, HKSAR, China; 3) Prenatal Diagnostics and Counselling Department, Tsan Yuk Hospital, HKSAR, China; 4) Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajara, México.

Cystic fibrosis (CF) is an autosomal recessive disorder in Caucasians, yet less than 20 Chinese patients with molecularly confirmed CF have been reported in both English and Chinese biomedical journals. Our department is the only center that offers sweat test in Hong Kong, a city of 7.3 million in southern China. We estimate that the incidence of CF is around 1 in 300,000 live births in this country. In Hong Kong. Here, we report 6 unrelated Chinese patients with CF with detailed description of their clinical manifestations. With a molecular approach combining NGS, Sanger sequencing and MLPA, we screened the entire protein coding regions of CFTR and selected deep intronic sites for known disease causing mutations and disease haplotypes. A new cystic fibrosis mutation of 4 patients was confirmed, and interestingly 3 patients were found to carry a missense mutation, p.I1023R (CFTR: NM_000492.3; c.T3068G), reported previously in 2 Taiwanese siblings with CF. All 3 patients inherited the mutation from either parent. Bioinformatic tools predict I1023R is possibly damaging to the protein. The clinical manifestations of patients carrying this recurrent mutation are typical CF features including Pseudomonas aeruginosa pneumonia, bronchiectasis and meconium ileus. From our search in literature, the CFTR mutation database, p.I1023R has not been reported in Chinese CF patients of other ethnicities. Furthermore, it is not found in our own exome database (~200 exomes), ESP6500 and 1000GP, suggesting a low allele frequency. Linkage analysis is currently underway to study whether I1023R is an inherited mutation in Hong Kong Chinese population. Our results confirm the hypothesis that I1023R is a rare but recurrent disease-causing CFTR mutation important in Chinese CF patients. p.I1023R is not included in the CFTR panel recommended by the ACMG. This finding has implications in the design of mutation panels/ analysis of NGS for molecular diagnosis of CF in Chinese population.

2760M

Identification of a novel ERCC8 mutation in a 10 year old brazilian female with Cockayne Syndrome type 1. A.A. Rocha,1 A.C. Jordao,1 M.M. Silva,1 A.C. Castro,1 L. Lando,1 J. Chiang,1 M.P. Avila,1 L. Gabriel,2 A.A.N. Rocha1,2, J. Chiang1,2, M.P. Avila1, L. Gabriel1,2, I. Chaves1,2 and E.A. Pascual-Castroviejo1,3. 1) Ophthalmic Genetics Department, Federal University of Goias, UFG, Goiania, Goias, Brazil; 2) Molecular Diagnostic Laboratory, Casey Eye Institute, Portland, Oregon, USA.

Purpose: To present a case of a female patient with Cockayne Syndrome type I presenting two mutations on the ERCC8 gene, one being an already known mutation and the other a novel missense mutation. Methods: Clinical examination was performed. Direct testing for mutations in the ERCC6 and ERCC8 gene was performed by PCR amplification and bidirectional DNA sequencing of all coding exons and exon/intron boundaries in a CLIA certified laboratory. The results were analyzed and compared to the NCBI reference sequences NM_000124.3 (ERCC6) and NM_000082.3 (ERCC8). The bioinformatic tools PolyPhen-2 and Mutation Taster were used to predict the effect of the novel mutation on the resulting protein. Protein domain identification was done with the bioinformatic tool SMART. Results: Clinical examination revealed a female patient born with 38 weeks of gestation by elective cesarean delivery, without any maternal or obstetric complications. At birth she presented APGAR (1st and 5th minutes) 8 and 9 respectively, weight of 2850g, height of 48cm, head circumference of 33.5 cm and chest girth of 32cm. She was developing properly until 3 months at which point failure of growth associated with developmental deterioration, microcephaly with a final head circumference of 45 cm (08 years old), mental retardation, ataxia, tremors, cutaneous photosensitivity, thin and dry hair. The ocular findings were sunken eyeballs, strabismus (exotropia), pigmentation of the optic disc and diffuse and faint yellow dots between the retinal vessels. Radiographs showed platybasia, wedged vertebræ, and multiple thoracic hemivertebrae. MRI of the brain exhibited cerebellar atrophy, posterior fossa cyst, corpus callosum hypoplasia, and cavum septum pellucidum. On fluorescein angiography the fluorescein angiography highlighted these eye fundus alterations, with discreet mottling of the posterior pole. Farnsworth land, Oregon, USA.

Purpose: To present a case report of a female patient with Cockayne Syndrome type I presenting two mutations on the ERCC8 gene, one being an already known mutation and the other a novel missense mutation. Protein domain identification was done with the bioinformatic tool SMART. Results: Clinical examination revealed a female patient born with 38 weeks of gestation by elective cesarean delivery, without any maternal or obstetric complications. At birth she presented APGAR (1st and 5th minutes) 8 and 9 respectively, weight of 2850g, height of 48cm, head circumference of 33.5 cm and chest girth of 32cm. She was developing properly until 3 months at which point failure of growth associated with developmental deterioration, microcephaly with a final head circumference of 45 cm (08 years old), mental retardation, ataxia, tremors, cutaneous photosensitivity, thin and dry hair. The ocular findings were sunken eyeballs, strabismus (exotropia), pigmentation of the optic disc and diffuse and faint yellow dots between the retinal vessels. Radiographs showed platybasia, wedged vertebræ, and multiple thoracic hemivertebrae. MRI of the brain exhibited cerebellar atrophy, posterior fossa cyst, corpus callosum hypoplasia, and cavum septum pellucidum. On fluorescein angiography the fluorescein angiography highlighted these eye fundus alterations, with discreet mottling of the posterior pole. Farnsworth land, Oregon, USA.
CRANIOFACIAL SYNDROMES AND GENETIC VARIABILITY IN A PEDIATRIC HOSPITAL IN MEXICO. W. San-Martín-Briekie, J.M. Apacino-Rodríguez, L. Vázquez-Olarte, L. Aparicio, A. Perel García, S. Sahelatien-Mendoza, 1) Maxilofacial Surgery; 2) Genetics, Hosp para el Nino Poblano, Puebla; 3) Estomatology, Benemérita Universidad Autónoma de Puebla; 4) Biotechnology, Universidad Metropolitana de México. It has been observed that a 60 to 70% of congenital malformations, there is not a definite cause. Within the causes that are known to exist: alterations chromosome 3-5%, 20% genetic mutations, environmental agents, radiation 1%, 2-3% infections, metabolic disorders maternal 1-3%, drugs and agents chemical 2-3%. All these alterations are the cause of the development and growth disorders affecting craniofacial structures Pinto (1979); Alfaro et al. (1994); Gorlin (1985); Wittkop (1975); Slavkin (1996) is achieved through knowledge of embryology, genetics and histology of these structures. The clinical factors present in these alterations in the development and growth of the oral cavity, maxillary and various soft tissues. Wittkop in 1975 and Gorlin in 1983 stressed that in certain craniofacial diseases; genetic and hereditary factors may be decisive or just contribute to the emergence of a specific disease. Most of craniofacial malformations are of unknown etiology, and as a result, the classification is mainly based on features of form and structure. There are many types of anomalies relating to the shape, number and structure and it has a hereditary origin. The nature of the abnormalities depends mainly by genetic factors. The frequency which these problems may arise depends on the form of inheritance and other laws of probability. Some anomalies may occur as the only apparent hereditary alteration. Others are presented as part of a much more complex genetic problem. Specifically congenital agenesis of structures that occur in the only one parental group. Agenesis and malformations can be part of a syndrome and be related to alterations in other ectodermal tissues such as hair, skin and mucus membranes, assessed as syndromes in this study: Moebius, Goldenhar, Cockayne, Retts, Rondebosch, Cri-du-Chat, VGPRD, and Down, Klinefelter and Turner. In Mexico was established in 1978 the "registration and surveillance epidemiological of the malformations craniofacial external" (RYVEMCE), generating preventive information programmes targeting the populations and are found in dbSNP. The identified variants were queried per individual, 95% of them have been previously seen in the human population and are found in dbSNP. The 37 year old male propostus was referred to our clinic due to chest pain, dyspnea and shortness of breath. The main findings of this propostus included severe myopia and mild marfanoid facial features. Echocardiography and computerized tomography showed chronic type I aortic dissection, ascending aortic aneurysm (90 mm) and severe aortic valve insufficiency. His family history was significant for sudden death of his father and his older sister, and similar phenotypes for three generation in the family. The patient’s 41 year old brother’s main findings included chest abnormalities (pectus carinatum), myopia-astigmatism, and lens subluxation. Echocardiography and computerized tomography showed also chronic type I aortic dissection, ascending aortic aneurysm (70 mm) and severe aortic valve insufficiency. The two brothers underwent immediate Bental Procedure. The FBN1 sequencing of the proband, brother and available family members revealed the presence of a pathogenic variant (c.7828G>C, p.Glu2610Gln). None of the unaffected family members did reveal this variant. Because of its complete segragation and due to the fact that Glu2610 residue is part of the conserved DINE motif found at the beginning of each cEBG domain of FBN1, it is considered a pathogenic variant. This variant has been reported before in the FBN1-UMD database by personal communications, but no description of the phenotype is given. The present family demonstrated variable clinical expression of the Marfan phenotype. For their study, malformations were divided as genetic or congenital abnormalities observed in this study into five groups according to the structural defect of the same as well as the genetic, congenital, and multifactorial cause.

Combination Biotin Responsive Encephalopathy and Hemiplegic Migraine Disorder presenting as Autism and episodic limb dysfuction/seizures in a 10 year old girl. P. Benke. Genetic Division, Joe Dimaggio Children’s Hosp, Hollywood, FL. A 10 year old girl presented to Genetic Clinic with findings of autism. She had all the features of autism, including developmental delay/late language and a paucity of verbal expression, poor understanding, poor social skills, poor eye contact and oppositional behaviors. Some improvement in her symptoms had been accomplished with a low gluten, dairy free diet, but she was still far behind an age appropriate level. Physical examination did not yield any clues, but when asked whether or not her hair and nails grow, her mother said that they did not. The mother was asked to observe hair and nails. When it was confirmed that they did not grow, she was instructed to start 10 mg/day biotin, which resulted in good growth of both. In addition, there was immediate improvement in school performance. The child began to talk better and her memory dramatically improved. It was brought to have had partial seizures from age 2-3 yrs, with changes in her EEG. Antiepileptic medicines and the change in diet decreased, but did not completely eliminate the episodes, during which one or another limb would lose function, and she developed a concurrent inability to talk or communicate. She would recover after a few hours time. She improved and was free of the latter symptoms when she vacationed in Colorado, and took Diamox. She was then treated for presumptive hemiplegic migraine with Diamox, and dramatically improved. Her EEG normalized, and she began to communicate. There is a history of congenital anomalies between mother and advanced artistic skills. Her school work improved further, with normal reading and math skills, and a gain in IQ scores. An exon genome scan (GeneDx) demonstrated a new mutation (p.T364M, c.1091 C>T ) in the ATP1A2 associated Hemiplegic Migraine gene, not found in either parent, but did not show a mutation in a biotin dependent gene. This study illustrates that 2 genetically unrelated disorders that together can lead to a diagnosis of autism, and symptoms can improve dramatically when specific gene appropriate measures are employed.
2766M
New cases of patients with developmental delay and incidental findings of chromothripsis. O. Caluero1, J. Chemos2, N.J. Leonard3, C.M. Klotz2, B. Argoiropoulos2. 1) Dept Med Gen, University Hospital, Edmonton, AB, Canada; 2) Dept Med Gen, Alberta Children’s Hospital, Calgary, AB, Canada.

Chromothripsis refers to a phenomenon of a local apparent shattering event of chromosomes and subsequent rearrangement, observed in 2-3% of cancers (Stephens et al., 2011). It has also been identified in the germline of a subset of individuals with congenital developmental disorders (Kloosterman et al., 2011). While the understanding of the complex chromosomal rearrangement (CCR) resulting from chromothripsis has started to be unveiled (Kloosterman et al., 2012), the relation to developmental problems and long term consequences for the patient’s health are unknown. We are presenting two paediatric patients with incidental findings of chromothripsis, their clinical phenotypes and potential correlations with their genotype known to date. PATIENT 1 was evaluated at 7 years of age for short stature (H -3SD), relative macrocephaly (OFC 75-90%), non-dysmorphic features, mild global developmental delay, hypothyroidism, breath holding spells with normal EEG, and a query paracentric inversion 13q21-q31. A microarray global developmental delay, hypothyroidism, breath holding spells with no, -3SD), relative macrocephaly (OFC 75-90%), non-dysmorphic features, mild global developmental delay, hypothyroidism, breath holding spells with normal EEG, and a query paracentric inversion 13q21-q31. A microarray

2768S
Clinical Aspects associated with Syndromic forms of Orofacial Clefts in Colombia. I. Briceno1,2, j. carrillo3, a. collins2, i. arias2. 1) Bioscience, U Universidad de La Sabana, Chia, Colombia; 2) University of Southampton. UK; 3) universidad Javeriana. Bogota, Colombia.

Abstract Objectives: 1) to present descriptive epidemiology of Orofacial Clefts (OFCs) and; 1) to determine the association of syndromic forms with antenatal high-risk conditions, preterm birth, and comorbidities, in a group of patients from Operation Smile Colombia. Design: A cross-sectional study was conducted. Frequencies of cleft type, associated congenital anomalies, syndromic, non-syndromic and multiple malformation forms, and distribution of OFCs according to gender and affected-side were determined. Odds ratios were calculated as measures of association between syndromic forms and antenatal high risk conditions, preterm birth and comorbidities. Setting: Operation Smile Colombia. Participants: A total of three hundred and eleven patients with OFCs treated in a 12-13 month period. Results: The most frequent type of OFC was cleft lip and palate (CLP), CLP was more frequent in males, whereas cleft palate (CP) occurred more often in females. The most common cases occurred as non-syndromic forms. Aarskog-Scott syndrome (ASS) showed the highest frequency. Hypertensive disorders in pregnancy, developmental dysplasia of the hip, central nervous diseases and respiratory failure showed significant statistical associations (p < 0.05) with syndromic forms. Conclusions: These data provide an epidemiological reference for OFCs in Colombia. Novel associations between syndromic forms and clinical variables are determined. In order investigate causality relationships between these variables further studies must be carried out.

2769M
Distribution of the AKT1 p.Glu17Lys mutation in a patient with Proteus syndrome. M.J. Lindthurst, M. Douce, H.M. Bloomhardt, M.R. Younck1, K. Moroz2, L.G. Bieseker1. 1) MGBMG, NHGRI/NIH, Bethesda, MD; 2) Department of Pathology and Laboratory Medicine, Tulane University School of Medicine, New Orleans, LA.

Proteus syndrome (PS) is a rare disorder that occurs sporadically and is characterized by progressive, disproportionate, segmental overgrowth that can affect any tissue in the body. It is caused by a post-zygotic activating mutation, c.49G>A, p.Glu17Lys in AKT1, resulting in individuals who have both mutation positive and mutation negative cells. Because the mutation occurs in a somatic cell, each individual has a unique constellation of manifestations making this disorder extremely heterogeneous. Recently, we had the opportunity to assess the mutation level in multiple tissues due to the unfortunate death of a 21-year old woman with PS. We sampled 25 tissues that were visibly affected and 13 tissues that were apparently normal. In most cases, the specimens were subdivided and multiple DNA extractions were performed. The AKT1 p.Glu17Lys mutation was measured using a custom PCR-based RFLP assay and the levels were averaged for each specimen. Of the 13 unaffected samples, six had detectable levels of the PS mutation (2 - 29%) although four of the mutation positive samples had no significant histopathologic abnormalities. Of the seven unaffected samples that had no mutation, only three were histologically normal. The mutation level in the affected samples ranged from 3 to 35% and all but the kidneys with mutation levels of 15 and 19% had abnormal histopathology. Interestingly, the highest mutation level was in a bone sample from the second digit of the right foot while the lowest levels were found in the soft tissue surrounding that bone, as well as in an omental fat nodule. The ovaries, vagina, breast, as well as several bones and pieces of skin were mutation positive, while the spleen, liver and lungs were mutation negative. Samples from a uterine leiomyoma, splenic hemangiomia, intracranial osteoma and an intracranial meningioma, were all mutation positive. In addition, a papillary carcinoma was found in the right lobe of the thyroid that was positive for both the AKT1 PS and BRAF p.Val600Glul mutations whereas uninvolved thyroid was mutation negative. This case has given us a rare opportunity to correlate mutation burden with histopathology in many tissues typically unfeasible to sample.
Phenotypic characterization of Microtia in Bogota, Colombia. L.P. Barragan Osono1, M. Garcia1, I. Zarante1, D.V. Luquetti2, 1) Instituto de Genetica Humana, Pontificia Universidad Javeriana, Bogota, Colombia; 2) Secretaria Distrital de Salud, Bogota, Colombia; 3) Seattle Children's Craniofacial Center, University of Washington, Seattle, US.

Introduction: Microtia is a congenital anomaly of the shape and size of the ear with a wide range of defects which is often associated with hearing loss in 80-90% of cases, and can occur as an isolated condition. The prevalence in South America is considered to be higher, especially in the Andean population ranging from 8 to 18 per 10,000 new-borns. The pathogenetic mechanisms are still uncertain and are supported by the environmental role as well as genetic causes of microtia. The aim of this study is to characterize the phenotype and risk factors of a population with isolated microtia in Bogota. Methodology: All potential participants with isolated microtia who were registered in the Bogota congenital malformations surveillance program (BCMSP) database, as well as the patients who assisted to otolaryngology visit or were born in any of following up hospitals, were invited to participate in this study. Each patient was cisted to a single clinical visit during which they were asked to undergo a physical examination, photographs taken and provide blood specimen for later DNA extraction. In addition, the medical history, results of hearing tests and ear computed tomography were collected. Participant characteristics were analyzed using descriptive statistics. Results: 27 patients with isolated microtia were recruited during September 2012 to July 2013. Microtia occurs most frequently in male patients (74%), children of multiparous mothers, 74% had unilateral microtia in whom right ear was mostly affected (80%) and 85% of them had atresia of the external auditory canal regardless the microtia grade (mild to severe or profound) that was not correlated with the type of microtia. All individuals had prenatal ultrasounds with a mean of 4 per patient, however no prenatal diagnosis was made. 63% of the mothers had acute infections during pregnancy. Maternal exposure to physical factors (e.g.: insecticides, organophosphates) was presented in 37% mostly during the first trimester. Discussion: A higher percentage of factors risk factors exposition was observed as: acute illness and physical risk factors exposures. Also the severity grade and middle ear involvement was greater than previously reported in the literature. These findings could suggest a possible environmental factor association in our population, but is necessary to increase our sample to calculate and extrapolate our results.
Phenotypic diversity in patients diagnosed with VACTERL association.

M. Husain 1,2, M. Wencel 2, B. Lemieux 1,2, V. Kimonis 3, B. Solomon 4, 1) Department of Pediatrics, UC Irvine, School of Medicine, Irvine, CA; 2) Division of Genetics and Genomic Medicine, Department of Pediatrics, UC Irvine Medical Center, Irvine, CA; 3) National Human Genome Research Institute, NIH, Bethesda, MD; 4) Inova Translational Medicine Institute, Inova Health System, Falls Church, VA.

The combination of vertebral, anal, cardiac, tracheo-esophageal, renal and limb anomalies termed VACTERL association is a clinical descriptor and a diagnosis of exclusion for a specific set of phenotypic manifestations observed to co-occur non-randomly and more frequently than expected by chance. Because VACTERL is clinically heterogeneous with an elusive etiology, we investigated the frequency and variety of these anomalies, other co-occurring manifestations and underlying causes of defects in embryogenesis or genetic mutations that elucidate VACTERL’s etiology. Unique cases in our cohort include the first Trisomy 18 male with an associated Dandy-Walker syndrome exhibiting the VACTERL phenotype, another with a 498.59 kb microdeletion in the 16p11.2 region, a 215 kb duplication in the 3p25.2 region with no known clinical phenotypes, and a genotypic female with dysmorphic facies and ambiguous genitalia. These are four among the 36 patients described in our study. Phenotypically, the most common clinical feature was vertebral anomalies, ranging from vertebral body/sacral malformations to spinal cord/rib anomalies, reported in 30 (83.3%) patients. 29 (80.6%) had cardiac defects ranging from ASD/VSDs to dextrocardia and atrial septum aneurysms. 24 (66.7%) had tracheo-esophageal anomalies ranging from TEF with esophageal atresia to Sandler syndrome. 23 (62.9%) had renal anomalies ranging from agenesis/duplex kidneys to multicystic dysplastic kidneys. 21 (58.3%) had anal anomalies ranging from imperforate anus/anos/anal atresia to a variety of GI/GU fistulas and atresias. 19 (52.3%) had limb anomalies encompassing both upper and lower extremities. Lastly, 14 (38.9%) also had urogenital anomalies. Indubitably, VACTERL is more complicated than previously thought. Although we found no evidence of teratogenic exposure of the mother during pregnancy or other contributing factors, our findings are promising nonetheless. The 16p11.2 microdeletions notably identified in patients with mental retardation, autism and learning/speech problems, have also been associated with malformations commonly found in VACTERL. Additionally, the 3p25.2 duplication with no previously known clinical phenotypes can now be scrutinized further, along with other emerging microarrays abnormalities, to identify candidate genes responsible for VACTERL association. We hope these findings enlarge upon the current understanding of VACTERL and guide research aimed at exploring its etiologies.

Brooke-Spiegler syndrome: A rare association of thichoepithelioma, cylindroma and spiradenoma. Report of a familial mexican case. N.O. Davalos1,2, M.E. Sanchez-Castellanos2, I.M. Salazar-Davalos1, M.A. Aceves-Aceves1, J.O. Higareda-Gonzalez4, M. Hernandez-Torres4, S.A. Ramirez Garcia1, M. Van-Dick1, I.P. Davalos1, D. Garcia-Cruz1, C. Garcia-Silva2, 1) Instituto de Genetica, UCIS, Biolmol, Universidad de Guadalajara, Guadalajara, Guadalajara, Mexico; 2) Instituto Dermatologico de Jalisco, SSJ, Zapopan, Mexico; 3) Universidad de la Sierra Sur, Oaxaca Mexico; 4) Hospital General de Occidente, Secretaria de Salud Jalisco, Guadalajara, Mexico.

INTRODUCTION Brooke H (1892) and Spiegler E (1899), independently described an epithelioma adenoides cysticum and skin endotelioma, as distinct entities. Brooke-Spiegler syndrome (BSS, OMIM # 605041) characterized by benign adnexal neoplasia. The predominating tumor such as thichoepithelioma, cylindroma and spiradenoma appear in late childhood and early adolescence. BSS is an autosomal dominant entity. The tumors located on head and neck, and increase throughout life with overlapping clinical features. Clinically thichoepithelioma showed flesh-colored papules specially on nasolabial folds; cylindromatosis lesions presents with multiple erythematous nodules arise on the scalp and is associated with alopecia; and blue-colored-painful lesions on breast and back suggested spiradenomas. The pathogenesis in BSS considered as a defect in the differentiation of folliculo-sebaceous-apocrine unit. Mutations have been identified in CYLD tumor suppressor gene, mapped to chomosome 16q12-q13. CASE REPORT: Case I: The proposita 9 years-old female. Normal psychomotor development. Between 5-6 years-old displayed flesh-colored papules on nose. The physical examination showed clinical and histopathological features indicating thichoepithelioma, with keratinizing cystic spaces. Case 2:42 years-old female showed since childhood, multiple flesh-colored-pink papules on her scalp, face and trunk. The lesions were asymptomatic. Three biopsies for the patients were reviewed, the first one on right preauricular region revealed thichoepithelioma, the second one on scalp conclusive cylindroma, and the third on lumbar region conclusive ecrine spiradenoma. Patients are being given this cryotherapy treatment and cosmetic surgery with good results. DISCUSSION: BSS an autosomal dominant disease. Female are commonly affected, with a female to male ratio of 3:1. BSS include multiple skin appendage tumors such cylindroma (OMIM# 132700), thichoepithelioma (OMIM# 601606) and spiradenoma, that share a common genetic basis. It may be related to other skin disorders such, basal cell adenomas, basal cell carcinomas. Histopathology studies are necessary to diagnosis. The treatment included excision, demabrasion, carbon dioxide laser, cryotherapy and in some cases radiotherapy. We present a Mexican family, the mother and daughter affected, presenting the classical clinical features of the disease. Molecular studies are needed to understand the genetic bases of the disease.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Disruption of the otoconic niche signaling in craniosynostosis: primary cilium and prostaglandin pathways crosstalk. W. Lattanzi1,2, M. Barba1, C. Cicione1, L. Massimo1, G. Di Taranto1, F. Frassanito1, F. Pignotti1, M. Baranzini1, M. Caltarelli2, F. Michetti1,2, C. Di Rococo2, G. Tamburini3, C. Bernardini4, 1) INSTITUTE OF ANATOMY AND CELL BIOLOGY, Università Cattolica S. Cuore, Rome, Rome, Italy; 2) Latium Musculoskeletal Tissue Bank, Rome, Rome, Italy; 3) Pediatric Neurosurgery Unit, Università Cattolica S. Cuore, Rome, Rome, Italy.

Nonsyndromic craniosynostosis (NSC) is a highly prevalent craniofacial malformation, with a strong and heterogeneous genetic background. A possible involvement of genes involved in the primary cilium signaling has been proposed in recent studies, proposing brand new candidate genes, with variable genotype/phenotype correlations. Through microarray genome-wide expression profiling and in silico bioinformatics analysis we have shown an altered expression of cilium-associated genes, and of genes involved in prostaglandin metabolism, in caival tissues and cells of NSC patients. We have developed and in vitro model for the functional assays aimed at demonstrating the functional role of the selected molecular pathways in the aberrant osteogenic process occurring at the site of premature suture closure. The expression on cilia-related genes (BBS9, GLI3, SMO) and of osteospecific genes (BMP2, RUNX2, OSX, OC), was significantly upregulated in cells isolated from fused sutures (syn-cells) compared to cells isolated from matched normal sutures (control cells), and increased upon 5 days of osteogenic induction. Confocal microscopy showed that syn-cells produced less primary cilia compared to control cells; BBS9 expression was spread throughout the cytoplasm in syn-cells, while appeared organized in polarized structures surrounding the cilium basal body in control cells. Upon BBS9 silencing in syn-cells, the expression of osteo-specific transcription factors (RUNX2 and OSX), and of SMO (key molecule of the hedgehog pathway), were significantly down-regulated. The osteogenic potential of BBS9 silenced syn-cells decreased and reverted to the physiological behavior observed in controls. Also, syn-cells displayed reduced HPGD expression, leading to reduced catabolic degradation of the osteoinductive prostaglandin E2. Interestingly, PGE2 administration induced differential effects in cells overexpressing BBS9, suggesting an unpredicted crosstalk between the primary cilium and prostaglandin metabolism. Our data may provide an in-depth insight into the events occurring in the calvarial osteogenic niche and leading to premature suture closure, which may be useful to: 1. clarify unclear etiopathogenic processes, 2. assist prioritization of new genomic variants, 3. pave the way to further translational outcomes.

Genotype-phenotype correlation in 12 patients with Oculoauriculovertebral Spectrum. S. Bragaglino1, M.E.S Colovati2, R.S. Guilherme3, A.G. Dantas1, C.A. Kim3, M.I. Melanaro1, A.J. Souza1, 1) Universidade Federal de São Paulo, SÃO PAULO, Brazil; 2) Genetics Unit, Instituto da Criança, Universidade de São Paulo, Brazil.

Oculoauriculovertebral spectrum (OAVS [MIM 164210]) is a clinically heterogeneous condition among the syndromes of first and second branchial arches related to embryonic craniofacial development. Most cases are sporadic with probably multifactorial inheritance, but rare familial cases and several chromosomal abnormalities have been associated with OAVS. The minimal criteria according to Tasse et al. [2005] are unilateral or bilateral asymmetric ear anomalies and hemifacial microsomia. Ocular defects, conductive and/or sensorineural hearing loss, vertebral malformations, and more rarely, cardiac, renal and cerebral malformations with intellectual impairment may appear. The identification of the OAVS genes/genomic position is important to the better understanding of the molecular pathways involved in craniofacial development and clarify clinicogenetics of the syndrome. Considering a sample of 62 patients that met Tasse et al. [2005] OAVS criteria (Eur J Med Genet 48:397-411), we identified whole-exome sequencing and pathogenic copy number variations (CNV). Genomic arrays identified: 7 deletions (4p16.3p15.33, 4q13.3q21.1, 8q13.3, 10q26.2q26.3, 16p13.3, 22q11.21, Xp22.33) and 6 duplications (2q32.1, 4p16.1.1, 16p13.11, 17q11.2, Xp22.33). We have identified the genes contained within these regions that are involved in this process and performed a genotype phenotype correlation in our patients. Even though some deleted and duplicated regions found in this study were relevant to the phenotype OAVS as BAPX1 (hemifacial microsomia HM [MIM 164210]; HMX1 [Oculoauricular syndrome OAS [MIM 612109]); EVAT1 (Otocephaly Otocephaly Hemifacial Microsomia LIKE 1 YPEL1 [MIM 608082]) and ERK1 (MITOGEN-ACTIVATED PROTEIN KINASE 1; MAPK1 [MIM 176948]), larger deletions and duplications found in our patients included many genes described in OMIM and genomic regions that are involved in the development of the 1st and 2nd branchial arches. These genes were investigated and correlated by reverse dysmorphology. Financial support: FAPESP, Brazil (2013/04623-2).

An autosomal recessive PGAP3 novel mutation was identified in patients with severe intellectual disability, dysmorphism and hyperphosphatemia from 2 unrelated families using whole-exome sequencing. V. Adir, A. Shalata, K. Golinker, E. Shahak, M. Mahroum, S. Tzur, Z.U. Borowicz. Molecular Genetics, Bnai Zion Medical Ctr, Haifa, Israel.

Whole-exome sequencing (WES) is very efficient method, that can be used as a diagnostic tool for identifying the molecular basis of genetic syndromes that are challenging to diagnose. In this study we report on hereditary intellectual disability and dysmorphism in two large Arab families with consanguineous marriage in northern Israel. WES revealed that the patients from the two unrelated families carry a novel homozygous mutation in the PGAP3 gene; c.H17.r.37,829,35T>C; c.b.845 A>G; p.D282G. According to previous studies, the PGAP3 gene encodes a product which is a post-Glycospophosphatidylinositol (PGI) attachment to protein factor 3, a protein that is involved in GPI-anchored maturation. The aspartic acid in position 282 is highly conserved from worms to humans, it is likely that this change has critical impact on the PGAP3 function, and ability to interact with other proteins and harm the GPI-anchor maturation process. GPI-anchor is a glycolipid structure that is added to the C-terminus of many proteins and is anchor their attachment to the cell membrane. Thus mutation in proteins that are involved in this process can affect signal transduction, neuronal development and can cause many pathologies including a wide spectrum of intellectual disabilities (Chesbrough et al). Recently Howard et al described four different mutations in the PGAP3 gene that causes a subtype of hyperphosphatasia with intellectual disabilities. The four patients we analyzed had normal MRI and have characteristic phenotypic features, they suffer from microcephaly, severe intellectual disabilities, low head circumference and high body weight. They also have coarse face features which includes narrow foreheads, long palpebral fissures, cleft palates, short noses with a wide nasal tip, full cheeks, tent shape wide mouths and large ear lobes. The phenotype of these patients is similar to the one described previously by Howard et al. Since we found the same mutation in four patients, in two unrelated families, we wanted to develop an easy and affordable test that will allow us to screen the relevant population from the same ethnicity. We developed an easy and affordable test that will allow us to screen the relevant population from the same ethnicity. We developed an easy and affordable test that will allow us to screen the relevant population from the same ethnicity. Whole exome sequencing demonstrated compound heterozygosity for p.(P643L) and p.(G741R) variants in the SLC12A3 gene; both of these mutations have been reported in Gitelman syndrome (Cruz et al., 2001, Am J Hum Genet 69:491-9). For p.(P643L) patients a males developed a PCR RFLP test that would enable easy analysis of the c.845 A>G mutation in the PGAP3 gene. The mutation was not found in 40 control chromosomes from the same ethnicity and geographical region. We suspected that there is a blood relation between the this two families and that they are probably connected via a third family from the same area.

Significant Secondary Findings of Exome Sequencing in Minor Anomalies with Autism Spectrum disorder. A. Alsdah. Benioph Children's Hospital, University of California at San Francisco, San Francisco, CA.

Significant Secondary Findings of Exome Sequencing in Minor Anomalies with Autism Spectrum disorder A. Alsdah, J.T.C. Shieh. Division of Medical Genetics, Department of Pediatrics, Institute of Human Genetics, Benioff Children's Hospital, University of California at San Francisco, San Francisco, CA. The results of exome sequencing in patients with minor anomalies with autism spectrum is unclear. Secondary or incidental findings on exome testing may also yield unexpected results, however findings of clinical consequences may go beyond current ACMG incidental findings recommendations. We present a diagnostic example of exome sequencing performed as part of developmental delay/autism genetic work up after normal microarray and metabolic testing at the UCSF Genetics Exome Clinic. A dedicated pre-test counseling visit was performed and incidental findings were discussed. The male patient avoided eye contact and had sloped shoulders and vague knee pains. Whole exome sequencing demonstrated compound heterozygosity for p.(P643L) and p.(G741R) variants in the SLC12A3 gene; both of these mutations have been reported in Gitelman syndrome (Cruz et al., 2001, Am J Hum Genet 69:491-9). Gitelman syndrome is an autosomal recessive disorder characterized by hypokalemia, metabolic alkalosis, hypomagnesemia, hypercalcuria, and hyperreninemia and typically presents in adults. As mutations in SLC12A3 affect the thiazide-sensitive Na/Ci cotransporter, we assessed electrolytes and found low potassium and abnormally low urine calcium confirming an early diagnosis of Gitelman syndrome. Our results demonstrate that exome sequencing can detect significant and treatable secondary findings as in this patient. He was started on potassium supplements to correct his electrolyte abnormalities. He also had intense paroxysmal bradycardia and arrhythmias, which could be fatal. Autism is not the expected presentation for this renal disease, however growth retardation and developmental delays have been reported in few patients with Gitelman syndrome (Skalova et al., 2002, Am J Med Genet 108:99). This case study highlights the importance of secondary findings in diagnostic clinical practice, and future studies will reveal potential effects on clinical outcomes.
2780S

Introduction: Autism spectrum disorders (ASD) are neurodevelopmental conditions characterized by language abnormalities, impaired social function and repetitive behaviors. Prevalence among 8-15 year-old (y/o) children is 1% and intellectual disability (ID) occurs in 70% of them. Together affect 3-5% of population. Genetic causes are identified in 10-20% of ASD and 25-50% of ID. Case presentation: This is an 8 y/o male patient who was born by uncomplicated pregnancy; presented with developmental delay, first full term gestation from non-consanguineous parents. Mother was 21 y/o. Birth weight was 3.4 kg. Patient had normal development until 2 y/o when started to have hampered verbal communication. Significant family history was denied. Clinical examination revealed a severe language problem but no other abnormal findings. Hearing tests, electroencephalography and MRI of the brain showed no abnormalities. Whole genome sequence analysis on DNA extracted from patient and parents’ blood was performed. Heterozygous variants of unknown clinical significance in UBE3B and GRIN2B genes were detected. Detection: GRIN2B variant was inherited from mother (autosomal dominant). For UBE3B variant (autosomal recessive), a second allele was not detected. Deletions or duplications cannot be ruled out because both genes are located in chromosomes 12. UBE3B codifies for an E3 ubiquitin-protein ligase that plays an important role on neurodevelopment. The variant c.421T>G - p.D141A was identified in the exon 6, which is known to be a preserved exon. Other UBE3B variants have been associated with blepharophimosis/ptosis-intellectual disability syndrome (BPID), in the patient’s brother (autosomal dominant, consistent with ASD and ID; BPID). Therefore, we propose a new clinical significance for this variant. Additionally, the variant c.3823A>C - p.T1278P was identified in the exon 13 of GRIN2B. This gene codifies for glutamic N-methyl-d-aspartate (NMDA) receptor 2B subunit. Other GRIN2B variants have been related to ASD, ID, psychiatric disorders and problems in retina, skin and testis. Both genes are related to glutamic acid signaling pathway which is important in neurodevelopment, therefore we propose an association between these variants and ASD/IDNSID. Conclusions: Heterozygous variants in GRIN2B and UBE3B detected in patient with autism and NSID. Further studies are necessary to uncover the whole phenotypic spectrum of disruption in these genes.

2781M
1) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 2) Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada; 3) Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto, Ontario, Canada; 4) Division of Hospital, Department of Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada; 5) Division of Plastic and Reconstructive Surgery, Hospital for Sick Children; 6) Department of Surgery, University of Toronto, Toronto, Ontario, Canada; 7) Department of Medical Imaging, Mount Sinai Hospital, Toronto, Ontario, Canada; 8) Division of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, Ontario, Canada.

Congenital limb reduction defects (LRDs) comprise an etiologically heterogeneous group of conditions caused by amniotic bands, early chorionic villus sampling, single-gene disorders, chromosome abnormalities, and teratogen exposure. Moreover, a substantial number of cases occur without an obvious cause. Vascular disruption events may be responsible for some of these cases, especially the terminal transverse reduction defects. Some studies have found an association between maternal thrombophilia and congenital LRDs while other studies have not confirmed this association. We further investigated this association through a review of all prenatally identified LRDs at a major tertiary care center in Toronto, Canada over a 12-year period. Research ethics board approval was obtained at all applicable institutions. Using a meta-analysis, our results showed a higher prevalence of inherited thrombophilia (specifically, Factors V Leiden, prothrombin G20210A, and heterozygosity for both mutations) among women with pregnancies affected with an LRD when compared to the general population ($\chi^2(3) = 54.63, p < .01$). Our work was strengthened by the inclusion of affected family members and the use of strict criteria to try to eliminate including cases of LRDs with an unrelated etiology. The LRDs in our sample were identified at mean age of 19.7 weeks gestation ($SD = 2.8$). We found an excess of left-sided defects among terminal transverse but not longitudinal reductions; additionally, all cases of thrombophilia occurred in the terminal transverse group. Our results support the hypothesis that maternal thrombophilia is associated with fetal LRDs. This supports the use of appropriate maternal thrombophilia screening (i.e., screening that is not affected by pregnancy, such as molecular genetic testing or anti-phospholipid assays) in pregnancies with an identified LRD. Additionally, our findings emphasize the importance of careful examination of the extremities, including digits, during the routine 18-20 week anatomy ultrasound in women with a known thrombophilia or history of thrombosis.

2782T
Whole Exome Sequencing of Moyamoya Disease. S. Jang1, S. Lee1,2, J. Chae3, B. Lim3, J. Kim1,4,5.
1) Biomedical Sciences, Seoul National University, Seoul, South Korea; 2) Genomic Medicine Institute (GMI), Medical Research Center, Seoul National University, Seoul, South Korea; 3) Department of Pediatrics, Seoul National University College of Medicine, Seoul National University Children’s Hospital, Seoul, South Korea.

Moyamoya disease (MMD) is a cerebrovascular condition distinguished by bilateral stenosis and occlusion in the internal carotid arteries, causing development of an abnormal collateral network of vessels for compensatory circulation. Previous studies have reported RNF213 as a major susceptibility gene for MMD. In this study, we suspected there may be another element that could cause or have an effect on MMD since there are patients with MMD who do not carry any variant in RNF213. To discover additional factors that may cause MMD, we performed exome sequencing on 15 individuals (11 MMD patients and 4 of whom are the parents of 2 of these patients). We narrowed down the variants to SNPs in exonic region that were frequent in MMD patients, but not in general population. Our study discovered 6 nonsynonymous variants: EML6, SYNE1, CHRNA3, CDK5RAP2, ZEB1 and SIPA1L3: all of these variants were shown twice in our patient samples. All 6 genes are candidate genes for causing MMD and may explain why RNF213 negative patients still display MMD symptoms. Further study is needed for verifying the specific pathways these genes are involved in to be asserted of these results.

2783S

Neuregulins are signaling proteins that mediate cell-cell interactions. They play a critical role in the normal growth of organ systems and are essential for the embryologic development of the central nervous system. NRG3, encoding neuregulin3, is located at 10q23.1 and is expressed mainly in brain, testis, skin and muscle. Extracellular cleavage product of NRG3 binds to ERBB4 resulting in ligand-mediated tyrosine phosphorylation. A newborn with holoprosencephaly, monoventericillar, cerebellar agenesis and severe cutis laxa was diagnosed with a novel developmental syndrome. Skin biopsy of the patient showed normal elastin and collagen staining but an increased thickness of the epidermis with excessive wrinkling, decreased hair growth and decreased pigmentation. Microarray CGH analysis of the patient’s DNA revealed a 34.7kb deletion in intron 5 of the NRG3 gene. The deletion was confirmed by qPCR as present in the proband but not in the apparently normal mother. Transgenic mice that overexpress NRG3 in skin are hairless and have thick, pale and wrinkled skin as seen in cutis laxa. Patient fibroblasts showed increased cell growth and proliferation suggesting overexpression of cell signaling. The deleted region of the intron 5 of NRG3 contains binding site for a strong transcription repressor. We initiated a reporter assay to evaluate the functional consequences of the intronic NRG3 deletion. Similar to the NRG3 transgenic mice, immunohistochemistry in cell culture confirmed accelerated cell proliferation, compared to control cells, without increased apoptosis. We defined a new syndrome of holoprosencephaly, cerebellar anomalies and cutis laxa. Based on the presence of the strong repressor site in the intron 5 of NRG3, the overexpression phenotype in our patient, and the finding of cutis laxa in the overexpression mice model of NRG3 we hypothesize the pathogenic role of NRG3 in this complex developmental skin and brain anomalies.
2784M

The 16p11.2 duplication syndrome is characterized by a spectrum of clinical findings including autistic behavior, developmental delay, intellectual disability, cranio-facial dysmorphism, microcephaly and epilepsy. Severity and expressivity of clinical symptoms can vary based on duplication size and other yet unknown factors with no evidence of gain or loss of specific gene function. The estimated prevalence of the 16p11.2 duplication syndrome is 3:10,000 individuals and as high as 1% in patients with autism spectrum disorder. The syndrome is a result of genomic structural change likely due to the 16p11.2 locus of about 600 kb (minimal critical duplication region) being situated between flanking segmental duplications having >99 percent sequence identity and regions of non-allelic homologous recombination (NAHR) events. Here we report a five year old female patient with a 3.6 Mb duplication including the 16p11.2 locus obtained by using Illumina’s Infinium SNP Array platform. Genes in the duplicated region that may contribute to the patient’s phenotype include PRR2, PKC, DYT10, EKD1, BRIC2, ICCA, FUS, TLS, ALS6 and ETM4. The patient presented with developmental delay involving motor, speech and language, learning difficulties and microcephaly (<3rd percentile). There is no definitive history of seizures and/or autistic spectrum disorders. She was born as the first child of healthy, non-consanguineous parents of mixed ancestry. Diagnostic work-up including EEG, brain MRI/MRS and basic metabolic testing were within normal limits. We report this case to serve as awareness of 16p11.2 duplication syndrome and its complexities as indicated in current literature. Variation in and outside the 16p11.2 locus may or may not be considered part of the syndrome.

2785T
Maternal UPD(16) with IUGR, transient neonatal hypoglycemia and cholestasis. H. Lesmana, R. Hopkin. Human Genetics, Cincinnati Children’s Hosp Med Center, Cincinnati, OH.

Uniparental disomy (UPD) is defined as the presence of homologous chromosomes from one parent. Maternal UPD of chromosome 16 is the most common reported UPD in the literature other than UPD(15) and has previously been reported to cause intrauterine growth retardation (IUGR) and a variety of congenital malformations. However there is no specific pattern of features documented in these cases suggesting for imprinted genes in this chromosome. Most cases are reportedly associated with confined placental mosaicism of trisomy 16 and some degree of placental insufficiency contributing to IUGR. Neonatal cholestasis has not been previously reported in maternal UPD(16). We report a new case of maternal UPD (16) which was initially suspected due to the presence of long contiguous stretch of homozygosity in chromosome 16 by SNP microarray. Maternal UPD(16) was subsequently confirmed through UPD testing using 12 microsatellite markers spanning chromosomes 16 in the proband, the mother and the father. This test revealed both maternal segmental heterodisomy and isodisomy on chromosome 16. Main clinical features include severe IUGR, transient neonatal hypoglycemia and cholestasis. Metabolic and molecular genetics testing for neonatal cholestasis and other inherited liver disorders failed to reveal any pathologic findings. We discuss reported cases of maternal UPD(16) and consider whether our patient’s features may be due to disordered imprinting or unmasking of an autosomal recessive condition.
2786S
Falling serum estradiol levels prior to human chorionic gonadotropin on follicle growth and pregnancy outcomes in in vitro fertilization cycles. X. Bao, JW. Xu, YP. Sun. The first affiliated Hospital of Zhengzhou University, Zhengzhou,China, Henan, China.
Background: Whether the falling of oestradiol on the day of hCG administration could influence in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI) pregnancy outcomes remains unknown. The aim of the study was to analyse the relationship between the falling estradiol levels prior to hCG administration and follicle rupture before the oocyte retrieval procedure and outcomes on IVF /ICSI treatment. Material and Method: A retrospective study of the database in the First Affiliated Hospital of Zhengzhou University (Zhengzhou, China) to identify 813 IVF /ICSI stimulation cycles whose E2 levels fell prior to the day of hCG administration during IVF/ICSI treatment. Patients were assigned to three groups according to serum E2 drops on the day of hCG administration compare to the day before as follows: group 1 with <10% drop in serum E2 concentration; group 2 with 10-20% drop; group 3 with >30% drop; control group with E2 levels continued to rise until the day of hCG. Results: Follicle ruptured rates (6.1%, 9.1%, 13.5% vs 3.3%, P=0.004) prior to the oocyte retrieval procedure was highly significant among the three groups. The cancellation rates (21.3%, 24.4%, 28.1% and 17.4%, P=0.018) was significantly higher with the increasing serum E2 drops on the day of hCG administration compare to the control group. Implantation rates (33.0%, 30.9%, 34.6% vs 33.4%, P=0.826) and pregnancy rates (50.0%, 45.6%, 50.4%, 51.0%, P=0.732) were no statistical significant in all study and control groups respectively. Conclusions: The falling estradiol levels prior to hCG administration might have a potential negative effect on follicular growth and lead to higher cancellation rates, but the pregnancy rates for embryo transfer cycles are not compromised.

2787S
The LH gene mutation and controlled ovarian hyperstimulation. M.R. Ranjouri, R. Davar,M.D. 2, M.H. SheikhM,M.D. Ph.D. 3, 1Department of Genetics and Molecular Medicine, Zanjan university of medical science, Zanjan, Iran; 2) Yazd Research and Clinical Center for Infertility, Shahid Sadoughi University of Medical Science,Yazd,Iran.
Background: One of the most difficult problems in vitro fertilization (IVF) treatment is the variability in the response to controlled ovarian hyperstimulation (COH) which ranging from poor to high, leading to IVF failure or complications related to ovarian hyperstimulation syndrome (OHSS). Objective: To evaluate the correlation between LH G1502A polymorphisms in exon 3 of the LH gene and ovarian response to COH. Materials and methods: A total of 220 women treated with a long protocol for ovarian stimulation were studied. Three genotypes of GG, GA and AA were detected by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis. Results: The most frequent genotype was GA (55.5%) whereas 44.5% of patients showed GG genotype and there was no patient with AA genotype. In total, 34 patients were poor responder, 154 were normal responder and 12 were hyper responder. In total 54.6% of normal responder, 61.8% of poor responders and 50% of hyper responders showed GA genotype. Discussion: Our results did not establish a significant relationship between this polymorphism and the ovarian response. Therefore it is still very difficult to use the genotype of patients for prediction of the ovarian response to stimulation.

2788S
Match Study of Sperm Relative Factors on the IVF Outcome. M. Zhang, J. Xu, Y. Sun. The first affiliated hospital of Zhengzhou University, Zhengzhou, Henan, China.
Objective: To investigate the effect of sperm relative parameters on the outcome of fertilization, in order to direct the selection of fertilization method. Methods: Collect 2 601 first Controlled ovulation IVF cycles from January 2012 to June 2013, match female factors(age, weight, cause of infertility, base FSH and LH, endometrial thickness on hCG day and the number of acquiring ovum ) and get 457 pairs. Analyze the sperm relative factors on different outcomes. Results: Sperm normal morphology rate in pregnant cycles was significantly higher than those of non-pregnancy cycles(P<0.05), but qualified rate were same; sperm density, activity and survival rate were same in different groups(P>0.05). Conclusions: IVF cycles with different outcomes own different sperm normal morphology rate, which has higher reference values then other sperm factors.

2789S
Prenatal diagnosis of Apert Syndrome: fetal brain phenotype on imaging. Z. Stark1, 2, , R. Palma-Dias3, A. Edwards4, 5, J.M. Sach2, 3, 4, 6, 7, G. Whiteley7, A.M. Fink7, 8, 9, 10, 11, 1 Victoria Clinical Genetics Services, Murdoch Children’s Research Institute, Melbourne, Australia; 2 Fetal Medicine Unit, Royal Women's Hospital, Melbourne, Australia; 3 Maternal Fetal Medicine, Sunshine Hospital, Western Health, Melbourne, Australia; 4 Pauline G. Gandel Imaging Centre, Royal Women's Hospital, Melbourne, Australia; 5 Pregnancy Research Center, Department of Obstetrics and Gynaecology, University of Melbourne, Australia; 6 Fetal Diagnostic Unit, Monash Medical Centre, Melbourne, Australia; 7) The Ritchie Center, Monash Institute of Medical Research, Melbourne, Australia; 8) Northwest Academic Centre, University of Melbourne, Australia; 9) Department of Radiology, Monash Medical Center, Melbourne Australia; 10) Department of Radiology, University of Melbourne, Australia; 11) Medical Imaging Department, Royal Children’s Hospital, Melbourne, Australia.
Objectives: The diagnosis of Apert syndrome relies on the identification of the classic triad of complex syndactyly of the hands and feet, craniosynostosis and midface hypoplasia. Diagnosis in the prenatal setting can be difficult as the onset of craniosynostosis is highly variable. We describe the brain imaging findings in 6 fetuses affected by Apert syndrome, and highlight the utility of detailed neuroimaging as an adjunct to diagnosis. Methods: Retrospective review of ultrasound and MRI brain imaging obtained in 6 fetuses with a diagnosis of Apert syndrome. Results: Five fetuses had attenuation of the septal leaflets, and two had callosal corpuscule dysgenesis. All six had temporal lobe expansion and overconvolution. The temporal lobe abnormalities preceded the development of cranial deformity in two fetuses. Conclusion: Overconvolution of the temporal lobe cerebral mantle can be detected antenatally and is particularly conspicuous in the fetus when the normal brain is still relatively smooth (approximately 24 to 28 weeks of gestation). Detailed fetal brain imaging by neurosonography or MRI is capable of demonstrating this temporal lobe malformation, and can contribute to diagnostic certainty in cases of suspected Apert syndrome.

2790S
Fibroblastosis ossificans progressiva: bilateral hallux valgus on ultrasound as clue for the first prenatal diagnosis for this condition —case report. C. Mattei1, I. Thiffault1, , J. Dubé2, A-M. Laberge1, E. Lemyre1, 1) Genetics, Ste-Justine Hospital, Montreal, Quebec, Canada, 3175 ch de la Côte-Ste-Catherine, Montréal, QC H3T 1C5, tel: (514) 345-4931, fax: (514) 345-4781; 2) Obstetric & Gynecology, Ste-Justine Hospital, Montreal, Quebec, Canada, 3175 ch de la Côte-Ste-Catherine, Montréal, QC H3T 1C5, tel: (514) 345-4931, fax: (514) 345-4781; 3) Center for Pediatric Genomic Medicine, Children’s Mercy Hospital, Kansas City, MO, USA,61408.
Fibroblastosis ossificans progressiva (FOP: MIM: 135100) is a very rare autosomal dominant condition with an incidence of 1:1,000,000. The clinical picture is characterized by congenital malformations of the great toes and progressive disabling heterotopic ossification. The genetic explanation is an activating mutation in the ACVR1, a BMP type I receptor, which promotes mesenchymal differentiation, part of the transforming growth factor superfamily. We report here how the presence of fetal bilateral hallux valgus, seen on second-trimester ultrasound, prompted clinicians to consider the diagnosis of FOP in the absence of family history. This classical foot malformation is present in all FOP patients. We present two cases of this French-Canadian pregnant woman was referred to our center for investigation and genetic counseling after a second trimester ultrasound identified a right ectopic kidney and a mega cisterna magna. A repeat ultrasound at 23 weeks in our center confirmed a right crossed ectopic kidney fused to the inferior pole of the left kidney, with good cortico-medullary differentiation. Bilateral hallux valgus was also observed. The cisterna magna was normal and all the other structures, including bones, were normal. Fetal growth was normal. Postmortem external examination showed antenatal fetal demise. The karyotype was normal (46,XY), as well as array CGH. A tridimensional ultrasound was performed to better characterize the foot and showed a hallux valgus deformation with plantar deviation of the first phalanx, and the absence of the second phalanx. The fetus was found to have the recurrent c.617G>A; p.R206H mutation in ACVR1 gene, confirming our diagnosis of FOP. Post-moratal external examination showed normal growth parameters, some minor dysmorphic features but no other malformations. Bilateral hallux valgus was present and the first metatarsophalangeal articulation was rigid. Fetal X-rays showed important bone malformation in hallux and plantar deviation. This report highlights that malformed great toes on prenatal ultrasound, especially hallux valgus, should raise the suspicion of FOP even when other atypical anomalies are found. Establishing the diagnosis allows for appropriate genetic counseling and informed decision-making for the family.
Posters: Prenatal, Perinatal, and Reproductive Genetics

2791S
Campomelic dysplasia: Prenatal Ultrasound and Autopsy Findings in Early Pregnancy. K. Chong1, N. Martin1, A. Tofte2, S. Keating2. 1) Dept OB/Gyn, The Prenatal Diagnosis and Medical Genetics Program, Mount Sinai Hospital, Toronto, ON, Canada; 2) Dept of Medical Imaging, Mount Sinai Hospital, Toronto, ON, Canada; 3) Dept of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON, Canada.

Campomelic dysplasia (CD) is a severe skeletal dysplasia characterized by cleft palate, tracheomalacia with respiratory compromise, bowing of the long bones and ambiguous genitalia. CD is an autosomal dominant genetic condition with few published case reports of gonadal mosaicism. Fetal ultrasound findings of CD have mostly commonly been reported in the second trimester with bent femurs, normal upper limbs and hypoplastic/absent scapula, however, some reports of cystic hygroma with lower limb anomalies have on fetal autopsy revealed CD. Translocations, small rearrangements and mutations involving the SOX9 gene on chromosome 17q have been identified as the cause of CD. We report a couple with 2 early pregnancies affected with cystic hygroma +/- other anomalies in the first trimester. The pregnancies were terminated by dilatation and curettage, however, careful pathology examination and X rays combined with molecular testing for SOX9 revealed a diagnosis of CD. Follow-up testing and examination of parents confirmed gonadal mosaicism in a symptomatic parent. This case highlights the importance of fetal pathology, even in early surgical samples, and the importance of clinical examination of parents in rare cases of skeletal dysplasia recurrence.

2792S

Introduction: The Harmony™ Prenatal Test uses an assay method, Digital Analysis of Selected Regions (DANSR™), for analysis of chromosomes 13, 18, 21, X and Y as well as other chromosomes to measure fetal fraction. Products from the DANSR assay are then analyzed with the Fetal fraction Optimized Risk of Trisomy Evaluation (FORTE™) algorithm to assess patient-specific risk of trisomy. Alternative tests, such as those using massively parallel shotgun sequencing, use a Z-statistic or Normalized Chromosome Value (NCV) to discriminate between normal and abnormal chromosomal counts without accounting for fetal fraction. Methods: A general pregnancy population cohort of 15,841 women between 10.0 to 14.3 weeks gestation were followed to pregnancy outcome. Z-statistics were computed using previously described standard Z-test of proportions and compared to FORTE risk scores generated by the Harmony test. A Z-statistic of ≥ 3 was considered ‘Positive’ for trisomy and a FORTE risk score of ≥ 1% was ‘High Risk’ for trisomy. Results: The cohort included 38 cases of trisomy 21 (T21) and 10 cases of trisomy 18 (T18). A subset of participants (n = 11,185) had trisomy 13 (T13) analysis.

PPV: Positive Predictive Value

<table>
<thead>
<tr>
<th>Method</th>
<th>T21</th>
<th>T18</th>
<th>T13</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORTE (%)</td>
<td>38/38 (100)</td>
<td>9/10 (90)</td>
<td>1/1 (100)</td>
</tr>
<tr>
<td>Z ≥ 3</td>
<td>38/38 (100)</td>
<td>3/4 (75)</td>
<td>3/4 (75)</td>
</tr>
<tr>
<td>FORTE (%)</td>
<td>38/47 (81)</td>
<td>9/10 (90)</td>
<td>1/1 (100)</td>
</tr>
<tr>
<td>Z ≥ 3</td>
<td>38/72 (53)</td>
<td>8/10 (80)</td>
<td>4/4 (100)</td>
</tr>
</tbody>
</table>

The cumulative test discordant rate for non-trisomies with FORTE was 0.08% compared to 1.07% with the Z-statistic. Conclusion: Analysis using FORTE to incorporate fetal fraction in cfDNA testing yields a >10 fold reduction in discordant results compared to the Z-statistic. This will become increasingly relevant as cfDNA testing gets more broadly used in the general pregnancy population.

2793S
Maternal subchromosomal abnormality identified through noninvasive prenatal testing (NIPT). C. Setterfield, T. Boomer, J. Newell, P. Santiago-Munoz2, N. Teed1, J. Saldivar1, N. Dhariaiya1, 1) Sequenom Laboratories, San Diego, CA; 2) Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas, TX.

Introduction: Noninvasive prenatal testing (NIPT) uses circulating cell free DNA (cfDNA) for the evaluation of fetal chromosomal abnormalities. Studies indicate that about 10-15% of the cfDNA in maternal plasma is derived from the pregnancy, and thought to be placental in origin. The remaining cfDNA is contributed by the mother. Thus, evaluation of cfDNA to identify chromosomal abnormalities has the potential to identify both fetal and maternal abnormalities. Here we report a case of a maternal deletion identified through a combination of NIPT and clinical history. Methods: Maternal blood samples submitted to Sequenom Laboratories for MaterniT21™ PLUS testing were subjected to DNA extraction, library preparation, and whole genome massively parallel sequencing as previously described. Sequencing data were analyzed using a novel algorithm to detect trisomies and other subchromosomal events. Results: A 20 year old G3P1010 had NIPT secondary to positive serum screening. NIPT results were negative, but ultrasound identified multiple anomalies. Karyotyping performed on amniocentesis sample subsequently revealed a ring chromosome 18 [46,XY,r(18)(p11.3q23)]. The patient elected pregnancy termination. Maternal karyotype was normal (46,XX). The patient conceived again and NIPT was again negative; no anomalies identified by 20 weeks' gestation. The history of the ring 18 chromosome was provided to the laboratory, and based on this information, detailed bin data was reviewed for chromosome 18, revealing an apparent 0.75 Mb deletion at 18q22.1-q22.1, likely maternal in origin. Maternal microarray analysis confirmed a 773 Kb interstitial deletion with the following result: arr 18q22.1 (63,884,429-64,657,305)x1. The patient declined invasive testing for the current pregnancy (ongoing). Familial studies are in progress to characterize the segregation of the chromosomal abnormality.

Conclusion: Clinicians should be cognizant that maternal subchromosomal abnormalities may be identified by NIPT using massively parallel sequencing. Accurate clinical information provided to the laboratory is critical for interpretation of test results. Additionally, standard karyotyping and in many cases microarray analysis do not have sufficient resolution to confirm subchromosomal events detected by NIPT. In these cases, microarray studies are superior to standard karyotype and FISH testing for confirmation of the suspected abnormality.

2794S
Thrombophilic Mutations for Recurrent Miscarriage in Iranian women with or without Thrombophilia. H. Mirtavvos-Mahyari1,2, B. Poopak2. 1) National Institute of Tuberculosis and Lung Disease, Tehran, Iran; 2) Department of Hematology and Laboratory Blood Bank, Tehran Medical Branch, Islamic Azad University, Tehran, Iran, PhD; 3) Department of Medical Genetic, Tehran University, Tehran,Iran, PhD Student.

Background: Inherited thrombophilia may increase susceptibility of fetal loss. This study was performed to investigate the frequency of some factor's mutations in ethnic Iranian patients . Methods: This epidemiologic case study performed with 609 patients whom were referred to Payyand Medical and Specialty Laboratory in Tehran, Iran, from 2011 through 2012. Peripheral blood samples of the patients with abnormal thrombosis were analyzed for mutations in Factor V Leiden, FII , MTHFR and PAI-1 . Mutation analysis was accomplished by PCR-reverse dot blot. Results:The mean age of patients was 29.58 years. The results are shown in Table 1. Table1.Distribution of inherited thrombophilia

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Homozygous</th>
<th>Heterozygous</th>
<th>Wild type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor V Leiden (n=170)</td>
<td>0</td>
<td>17</td>
<td>159</td>
</tr>
<tr>
<td>MTHFR (n=195)</td>
<td>21</td>
<td>90</td>
<td>84</td>
</tr>
<tr>
<td>FII (n=102)</td>
<td>0</td>
<td>6</td>
<td>96</td>
</tr>
<tr>
<td>PAI-1 (n=136)</td>
<td>23</td>
<td>78</td>
<td>35</td>
</tr>
</tbody>
</table>

Discussion: MTHFR and PAI-1 polymorphisms were identified in more than fifty percent and factor V leiden mutation is higher than some other countries . Therefore, testing and providing genetic consultation may be needed in Iran.
Validation of a taxonomy of genetic conditions for pre-conception genetic carrier testing. T.L. Kaufman1, N. Neill2, C. McMullen1, M.C. Leo3, J. Reiss4, B. Wilfond4, J. Davis5, M. Gilmore6, P. Himes6, F. Lynch3, K.A.B. Goddard1. 1) Center for Health Research, Kaiser Permanente Northwest, Portland, OR; 2) Kaiser Permanente Northwest, Portland, OR; 3) Decision Research, Eugene, OR; 4) Truman Katz Center for Pediatric Bioethics, Seattle Children’s Hospital, Seattle, WA.

Background: We are studying the use of genome sequencing in the context of pre-conception genetic carrier testing for women and their partners. The study team created a taxonomy of four categories of results that participants could choose from to categorizing their result. We then compared their perceptions of and reactions to these conditions found. We created five versions of the survey, each with four conditions. Participants rated their reactions to each condition on a seven-point Likert scale, including the condition on a child and the family, predictability and controllability of the condition, and how well they would change their plans for pregnancy. Enrollment was limited to the first approximately 40 respondents per version of the survey. Results: We assigned the 20 conditions into the taxonomy categories. We then determined if survey participants were able to correctly place the 20 conditions into the four categories. Based on the results of factor and cluster analyses, the conditions were grouped according to ratings of hopeful feelings, controllability, predictability, visibility, severity, and negative feelings. Overall, survey responses supported our five category taxonomy of genetic conditions. Conclusions: Survey respondents made meaningful distinctions between the genetic conditions presented to them. While the categories were derived from experts and revised with focus group input, the results provide empirical support for the category distinctions. Future research utilizing the same methodology with a wider variety of conditions could be beneficial.


Introduction Y-chromosomal microdeletions are the second most frequent genetic cause of male infertility. As stated in the guidelines of the European Academy of Andrology (EAA) for molecular diagnosis of Y-chromosomal microdeletions, this kind of analysis should be performed in patients with azoospermia or severe oligospermia with sperm concentrations < 2x106/ml. In these patients is also convenient analyze other clinical parameters such as hormone levels, testicular volume, varicocele, infections etc, although these parameters do not have any predictive value. Here we report a lab case of an azospermic patient candidate to a testicular sperm extraction (TESE) for performing an assisted reproductive technology. Methods According to EAA guidelines, we first performed the multiplex PCR amplification of genomic DNA with the Ampli-Y Chromosome and Ampli-Y Chromosome Extension kits (Diachern - Italy). As a confirmatory test we also utilized the Deyser AZF kit (Deyser Ab - Sweden). The execution of karyotype was performed in an external laboratory. Conclusions The Y-chromosomal microdeletions analysis performed showed a lack of AZF b and c regions of Y-chromosome, confirmed by both kit utilization for Y microdeletions diagnosis. The patient had also high levels of FSH and LH, low level of inhibin b and normal levels of prolactin and testosterone hormones. Moreover after the execution of karyotype test, the patient was found to be carrier of a severe form of chromosomal mosaicism (45X;99; 46XY[1]) that confirmed Y microdeletion diagnosis. In the overall these results suggested clinicians to do not perform TESE since, as reported in literature, the chance for testicular sperm retrieval in these cases is virtually zero. Our case confirm the need of a good laboratory practice to support clinical decisions.

First Experiences with Non invasive Genetic Testing in Switzerland. J. Esslinger1, S. Hotz1, L. Risch2, U. Wiedemann2. 1) Institute for family research and counseling, University of Fribourg, Fribourg, Switzerland; 2) Labormedizinisches Zentrum Dr. Risch, Liebefeld BE, Switzerland.

Introduction In July 2012 the NIPD Praena-Test of Lifecodexx, was implemented in Switzerland. NIPD-testing received high media coverage. Initially, a lack of clarity regarding the target patient-community, the length of reporting time and the indication for this test existed. Referral was made via local gynaecologist, hospital centre or private person. We evaluated data to clarify whether a significant “patient pattern” exists for this new test. Material and methods A total of N=378 cases were included in this analysis, representing all NIPD-tests requested in the laboratory Dr. Risch, as a typical big private laboratory in Switzerland, from the beginning in August 2012 until the reference date December 31st 2013. Results As expected due to location, most patients were Swiss (92,1%), wherefrom 72,1% from the German-speaking part of Switzerland, 27,6% from the French-speaking part, 0,3% from the Italian-speaking part and foreign patients were from Liechtenstein, France and Italy or other European countries. Patients were between 18 and 46 years old, with a mean of 36 years (SD=4,4), and a most prevailing age of 39 years. The mean pregnancy week (PW) was 14 (SD=2,4), with a min. of 8 and a max. of 27, PW of 19 and max. PW of 29 showed main indications for testing: maternal age >35 years (46%), conspicuous first trimester screening (1TT) (14%), combination 1TT/age (13%) and wish (13%). Further combinations such as age/wish, age/previous trisomy as well as familiar burden have each been indicated in 2% of patients. Other indications were the case in 1% or less of women. Two thirds (67%) of patients were referred from local gynaecologists, 30% from hospital centres, 3% were self-registered patients. We found no significant difference in indication between gynaecologists and hospital centre. The median delay for a clinical result was 16 days (SD=4,8), with a min. of 6 and a max. of 53 days (incl. repetition). In 4% of the cases, the first blood sampling did not provide a valid result. One case showed a trisomy 13, one case a trisomy 18 and 4 cases a trisomy 21. The typical NIPD-patient after 18 months of test-availability was Swiss, 39 years old, with a gestational age of 16 weeks and maternal age as indication. After initially long reporting time, a statistically significant decrease in reporting time was achieved throughout the examination period during months 10-18 (M=15, SD=3,4) vs. months 1-9 (M=17, SD=5,4; p<0.001). The majority of patients (86%) receive their result with in 15 days.
2799S

Carrier screening for recessive disorders through exome sequencing.

P. Makrythanasis1, A. Massouras1, S.E. Antounarakis1,2,3, 1) University of Geneva, Geneva, Switzerland; 2) University Hospitals of Geneva, Geneva, Switzerland; 3) IGE3 Institute of Genetics and Genomics of Geneva, Geneva, Switzerland.

Detection of carrier status for certain recessive Mendelian disorders is a well-accepted health-care practice in several countries since the 70s and aims at the prevention of frequent severe monogenic disorders. Exome sequencing in combination with the increasing knowledge of human pathogenic variation provide the possibility to perform a carrier screening for all the known recessive Mendelian disorders. Such a screening would allow for a much more informative genetic counseling and may alter the total prevalence of the known recessive disorders. In order to test this hypothesis we have used exome sequencing data from 104 individuals of European origin and have identified the total number of likely pathogenic variants in the >1600 recessive disorders for which the responsible gene is known. The mean value was 18.2 variants per individual. Consequently we have randomly paired these exomes in order to create 5356 fictive couples. 33.14% of these couples have at least one gene for which both members are heterozygous for a likely pathogenic variant. These preliminary results exhibit an upper estimate of at risk couples but more precise knowledge and definition of the pathogenic potential of each variant will render the carrier detection more accurate and make it a potent tool for family planning.

2800S

IL-10 Promoter polymorphism (592C/A) in women with recurrent miscarriages in Punjabi population (India). A. Kaur, N. Sudhir, B. Badarud doza. Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Cytokines, expressed by several cell types and tissues, are diverse family of soluble proteins which play a fundamental role in differentiating the nature of immunologic response. Interleukin -10 (IL-10) a type of cytokine has a dual immunological role as it act as either immunostimulatory or immunosuppressive. IL-10 is fundamental in promoting normal pregnancy outcomes as it is involved in placental angiogenesis and inhibits the secretion of inflammatory cytokines. The promoter region polymorphisms in IL-10 gene have reported to interfere in its regulation. IL-10 down regulation may affect the placental angiogenesis affecting fetal growth and might end in miscarriage. In the present study IL-10 592C/A promoter polymorphism was investigated in women with recurrent miscarriages and in women with at least one live birth and no history of miscarriages. For this, a total of 182 samples (92 cases and 90 controls) were genotyped using PCR-RFLP method. In the present study, 56.5% of women with recurrent miscarriages has homozygous wildtype genotype (CC), while 4.76% and 2.04% of women were having heterozygous and homozygous mutant genotype, respectively. In case of control group 52.2% showed CC genotype, 40% had CA genotype and homozygous mutant (AA) was found in 7.8% of women. Three genetic models such as dominant, codominant (additive) and recessive were used to analyze the data, none of the models have been found significant except recessive model which showed positive association (OR: 1.45;0.53-3.98,P=0.373). no significant difference were also observed in either genotype (P=0.542) or allele (P=0.907). The present case control study did not show any significant association with recurrent miscarriages. Therefore, no overall significant association was found between IL-10 592C/A and recurrent miscarriages in Punjabi population. (India).

2801S

Tay-Sachs carrier screening by enzyme and molecular analyses in the New York City Black population. G.A. Lazarn1, E. Spiegel2, K. Berent- sen1, K. Brennan1, N. Mehta1, H. Haque1, R. Hawthorn3, 1) Counsyl, South San Francisco, CA; 2) Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, NY.

Introduction: Carrier screening for Tay-Sachs disease (TSD [MIM 272800]) analyzes the HEXA gene or HEXA enzymatic activity. The former can be more costly or may not detect all carriers. Enzymatic assay approximates phenotype but is subject to misinterpretation (e.g., indeterminate results). Enzymatic assay is often suggested as the optimal detection method in non-Ashkenazi Jewish populations. However, this protocol has recently been considered for routine, pan-ethnic TSD carrier screening with the advent of expanded carrier screening (ECS). Columbia University implemented universal ECS in 2013. Results of DNA and enzymatic analyses for TSD carrier status in the Black population are reported. Methods: 123 individuals reporting Black ancestry underwent ECS, including HEXA mutation screening (9 mutations) and lymphocyte HEXA activity levels (WBC HEXA %). Mutation screening was performed at Counsyl (South San Francisco, CA) and most enzyme assays were performed at Mount Sinai Genetic Testing Laboratory (New York, NY); one enzyme assay was performed at Mayo Medical Laboratory (Rochester, MN). Individuals with carrier or indeterminate enzyme levels and negative targeted mutation analysis were subsequently offered next-generation sequencing (NGS) of exons 1-14, also at Counsyl. Results: Of 123 individuals screened by enzyme, n=16 (13.0%) were deemed carriers and n=15 (12.2%) had inconclusive results. Mean WBC HEXA % were: 58% in all individuals, 45% in positive results, 52% in inconclusive results and 81% in negative results. All of these 31 individuals were heterozygous for a 9-mutation. NGS was completed on 4 individuals with positive results and 4 with inconclusive results. One likely deleterious variant, c.1510C>T, was found in 1 individual. One individual had no variation from the reference sequence. In the remaining 26, 5 known benign variants (total of 6 unique variants detected 34 times) were identified. Conclusions: Using enzyme analysis in this Black population, we found an unexpectedly high positive or inconclusive rate. Such results necessitate molecular testing, and we have initially found most individuals to have benign mutations. We will continue to collect data characterizing non-negative enzyme results, which may reveal unrealized deleterious variants or pseudodeficiency alleles. The data may suggest recalibration of enzyme reference ranges or an alternative routine screening protocol for certain populations.

2802S

Prospective exome sequencing of consanguineous couples: First steps. H. Meijers-Heijboer1, M. Teewu, MD, O. Waisfisz, PhD, P.J.G. Zwijsen, MD, PhD, E.A. Sleijfer, PhD, J. Weiss, PhD, L. Henneman, PhD, L.P. ten Kate, MD, PhD, M.C. Cornel, MD, PhD, VU University Medical Center, Amsterdam, Netherlands.

With 10% of people worldwide in a consanguineous relationship or having consanguineous parents, consanguinity is one of the most frequent risk factors for congenital disorders. In theory, prospective exome sequencing of consanguineous couples could identify couples who are carriers for autosomal recessive diseases, and empower such couples to make informed reproductive decisions. This study aimed to test the feasibility of this approach at the University Medical Center, New York, NY. Carrier screening for Tay-Sachs disease (TSD [MIM 272800]) in non-Ashkenazi Jewish people. However, this protocol has only recently been validated in the Black population, we found an unexpectedly high positive or inconclusive rate. Such results necessitate molecular testing, and we have initially found most individuals to have benign mutations. We will continue to collect data characterizing non-negative enzyme results, which may reveal unrealized deleterious variants or pseudodeficiency alleles. The data may suggest recalibration of enzyme reference ranges or an alternative routine screening protocol for certain populations.

Testing Laboratory (New York, NY); one enzyme assay was performed at Counsyl. Results: Of 123 individuals screened by enzyme, n=16 (13.0%) were deemed carriers and n=15 (12.2%) had inconclusive results. Mean WBC HEXA % were: 58% in all individuals, 45% in positive results, 52% in inconclusive results and 81% in negative results. All of these 31 individuals were heterozygous for a 9-mutation. NGS was completed on 4 individuals with positive results and 4 with inconclusive results. One likely deleterious variant, c.1510C>T, was found in 1 individual. One individual had no variation from the reference sequence. In the remaining 26, 5 known benign variants (total of 6 unique variants detected 34 times) were identified. Conclusions: Using enzyme analysis in this Black population, we found an unexpectedly high positive or inconclusive rate. Such results necessitate molecular testing, and we have initially found most individuals to have benign mutations. We will continue to collect data characterizing non-negative enzyme results, which may reveal unrealized deleterious variants or pseudodeficiency alleles. The data may suggest recalibration of enzyme reference ranges or an alternative routine screening protocol for certain populations.

Carrier testing (CT) is unique among diagnostic tests in that disease can only be observed in persons other than those being tested. As a result, clinical validation of many rare disease-causing mutations may be difficult if not impossible. Nevertheless, CT can be applied only to validated mutations in the detection of recessive disease risk prior to conception. CT is a particularly poor indicator of reproductive risk for anonymous donors to sperm and egg banks. We have developed a computational system that simulates the creation of haploid gametes from a person’s genomic or exomic sequence information. The Monte Carlo-generated “virtual gametes” from two prospective genetic parents are combined to form a pool of diploid “virtual progeny” (VP) genomes. Each VP genome is analyzed individually for the likelihood that it will induce a disease phenotype based on clinical annotations and the combined biochemical properties of the products of the two copies of each gene. Our isolated variant scoring method shows a sensitivity of 95% on defined clinical variants. We applied our analysis to 440 well-characterized recessive disease-associated genes in virtual matings of 2,500 exomes from phase 3 1,000 genomes data. The detected risk of recessive disease is over 10-fold greater than predicted by an ideal all-encompassing carrier testing protocol.

**2080S**

**Exome chip evaluation of genetic variants for association with uterine fibroids.** M.J. Bray, T.L. Edwards, K.E. Hartmann, D.R. Velez. 1) Vanderbilt University, Nashville, TN; 2) University of Pennsylvania, Philadelphia, PA; 3) New York University Medical Center, New York, NY; 4) Massachusetts General Hospital, Boston, MA; 5) Cedars-Sinai Medical Center, Los Angeles, CA.

Uterine fibroids (UFs) affect up to 77% of women by menopause, account for $10 billion in yearly medical costs, and have a significant impact on minority women. Although UFs are heritable, genetic causes are poorly understood. The first genome-wide association study (GWAS) of UFs was performed in 2011 in a Japanese population. However, to date few large-scale genome-wide association studies have been performed in UFS populations. The objective of this study is to conduct a whole exome association study of UF risk in European Americans (EA) and African Americans (AA). This is a case-control genetic association study of UFs defined using pelvic imaging data obtained from an electronic medical record biorepository (BioVU DNA Repository). Logistic regression adjusted for ancestry and age was used in 728 EA and AA DNA samples (EA: 246 cases, 243 controls; AA: 121 cases, 118 controls) to evaluate SNPs for association, stratified by race. Meta-analyses of results with fixed effect models were performed to obtain combined evidence for associations across racial groups. Our strongest association within EAs was within t-complex 11, tests-specific like 1 (TCP11L1, rs1064005, OR = 1.79, 95% CI = 1.36-2.32, p = 3.94×10^-5) and within multiple SNPs within the HLA region (smallest p = 9.28×10^-10). It is of note that the HLA region showed the strongest evidence for association compared to GWAS arrays. Among AAs strong associations were observed within a confident (HPC4, OR = 0.46, 95% CI = 0.31-0.68, p = 8.73×10^-5). Meta-analysis across EA and AA further strengthened the associations observed in EA for all SNPs with p<0.05. (1) Meta-analysis across all races). Prior gene expression studies of TCP11L1 have shown increased expression in tumor tissues compared to normal tissues across several different cancers, and variants within HLA genes have been associated with multiple SNPs within the TCP11L1 gene. The objective of this study is to evaluate the relationship between gene variants and UF risk in prior studies. These pilot data suggest correlation of variants in exonic regions increase risk for UF in both EA and AA populations. However, further validation of our study findings is needed to confirm our results. We are currently genotyping >4,000 samples for GWAS and exome arrays to further evaluate the relationship between gene variants and UF risk.
Expression of Aurora Kinase C splice variants in human oocytes and cumulus cells. J.E. Fellmeth1, K. Schnider1, N.R. Treff2. 1) Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, New Jersey, United States of America; 2) Division of Reproductive Endocrinology, Department of Obstetrics, Gynecology and Reproductive Science, Robert Wood Johnson Medical School, Rutgers University, New Brunswick; Reproductive Medicine Associates of New Jersey, Morristown, New Jersey.

Objective: Accurate chromosome segregation during meiosis I (MI) is essential for generating eggs with the proper chromosome complement. This segregation is highly error-prone in humans leading to miscarriage or offspring with developmental disorders. The Aurora family of protein kinases is a well-established regulator of chromosome segregation during mitosis and meiosis. Aurora kinases A and B (AURKA and AURKB) are expressed in nearly every cell type, whereas Aurora kinase C (AURKC) is only highly expressed in gametes. AURKC was identified in sperm and missense mutations in AURKC are linked to formation of polyploid sperm and infertility.

Studies in mouse oocytes indicate that AURKC is also required to regulate MI chromosome segregation in females. To begin to determine if AURKC function is conserved between mouse and human, we used quantitative real-time PCR to detect the presence of AURKC message in single human oocytes. Design: Observational Materials and Methods: To determine which variants are present in human oocytes, we designed Taqman probes that are specific for detecting each variant, 1, 2, and 3. To compare the expression levels between meiotic and mitotic cells, we compared the expression of AURKC in oocytes to that of its expression in oocyte-matched cumulus cell samples. Sperm samples known to express all 3 variants were used as positive controls. Results: We found that AURKC expression levels varied in oocytes and cumulus cells, and that expression levels were greater in the oocyte than sperm, where variants were originally found. Oocytes express all three variants, however, variant 1 is expressed approximately 10 fold greater than variants 2 and 3. Surprisingly, cumulus cells also express AURKC, although this expression was restricted to variant 1. Similar to the single oocytes, AURKC expression levels varied amongst the cumulus cell samples. Conclusion: In sum, like sperm, human oocytes and cumulus cells express AURKC suggesting that missense mutations may also affect female MI and thereby compromise fertility. Our work may involve characterizing whether varying expression levels amongst oocytes correlate with aneuploidy development.

Support: ASRM Research Grant.

Understanding the genetics of spermatogenic failure by resequencing the sex chromosomes of infertile men. R. George1, J. Hughes1, L. Brown1, L. Lin2, D. Koboldt2, K. Meltz-Steinberg2, R. Fulton2, R. Wilson2, R. Oates2, S. Silber2, S. Repping2, D. Page2, J. Hughes1. 1) Whitehead Institute, Cambridge, MA; 2) The Genome Institute at Washington University, St. Louis, MO; 3) Urology Department, Boston University, Boston, MA; 4) Infertility Center of St. Louis, St. Louis, MO; 5) Academic Medical Center, Amsterdam, The Netherlands; 6) Biology Department, MIT, Cambridge, MA; 7) HHMI.

The sex chromosomes harbor a large number of genes involved in spermatogenesis and are hemizygous in males. For these reasons, mutations that cause spermatogenic failure—the production of very few or no sperm—are disproportionately likely to be found on the X and Y chromosomes. While several genetic causes of spermatogenic failure, such as microdeletions of the Y chromosome and large-scale cytogenetic abnormalities (e.g. Klinefelter's syndrome; XXY), have been identified, they only account for 20-30% of cases and the majority of genetic causes remain unknown. To identify additional factors causing spermatogenic failure, we sequenced the coding and conserved non-coding regions of the X and Y chromosomes in 300 men with nonobstructive azoospermia and 300 controls. Our targeted regions include 838 and 53 protein coding genes from the X and Y chromosomes respectively, along with 22 Mb of non-coding sequence that contain ncRNAs and putative regulatory regions. We will describe the initial analysis of these sequences, and the identification of putatively causal coding mutations and copy number variants.


Robertsonian translocations are among the most common balanced structural rearrangements in humans involving two (often non homologous and rarely homologous) acrocentric chromosomes (13, 14, 15, 21 and 22). Nevertheless, non-Robertsonian translocation involving these chromosomes is a rare event and only few cases were reported. Here, we report a familial non-Robertsonian translocation involving chromosomes 15 and 21. The index case is a 38-years-old man for whom chromosomal investigation was carried out to explore a male infertility of 3 years related to azoospermic profile at the semen level, a first unsuccessful attempt of sperm retrieval using TESE and a second diagnostic biopsy of testis showing a sertoli cell only syndrome. Cytogenetic analysis carried out using RHG banding, disclose the presence of a balanced translocation between the long arm of chromosome 15 and the long arm of chromosome 21: (15;21)(q12;q21). At the genetic counseling, familial history showed that the patient had consanguineous parents, an infertile paternal uncle and two hypofertile brothers with only one son for each of them (as well as another brother and 2 sisters who were fertile). One of hypofertile brother's patient had a history of recurrent pregnancy losses (RPL) (n=3) and a teratozoospermia with a normal count of spermatozoa (88x106/ml). His karyotype revealed at the occasion of familial genetic counseling, was abnormal with the same translocation (t(15;21)(q12;q21)). Non-Robertsonian translocations, involving chromosomes 15 and 21 are uncommon and were described only three times at literature. However, these translocations involved other break-points and were reported for parents of malformed children harboring partial trisomy 15 [t(15;21)(q13;q21.3)][mat], partial trisomy 21 [t(15;21)(q26;q22.1)][pat], and Prader willi syndrome [t(15;21)(q15;q22)][pat]. This new translocation is different by its breakpoints and its association with male infertility and RPL. Further studies need to investigate the susceptibility locus and DNA methylation of recurrent miscarriage, most commonly between 8-12 gestational weeks. The majority genetic factors of early pregnancy losses is attributed to aneuploid embryos. However, the etiology of approximately 40% of early abortions remains unknown. To investigate additional factors causing spontaneous abortions, we screened uniparental disomies (UPD) in cytogenetically normal diploid spontaneous abortion couples.

Materials and methods: We screened 165 couples who came for assisted reproduction at Reproductive medical center of The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China. And all the patients have undergone more than 2 times miscarriage. The karyotype of these couples was 46,XX and 46,XY detected at literature. However, these translocations involved other break-points and were reported for parents of malformed children harboring partial trisomy 15 [t(15;21)(q13;q21.3)][mat], partial trisomy 21 [t(15;21)(q26;q22.1)][pat], and Prader willi syndrome [t(15;21)(q15;q22)][pat]. This new translocation is different by its breakpoints and its association with male infertility and RPL. Further studies need to investigate the susceptibility locus and DNA methylation of recurrent miscarriage.
2811S  
**The correlation between Y chromosome micro-deletions and recurrent pregnancy loss.**  
Z. Sarraf1, R. Mirfakhraie1, M. Salimi1, G. Modabber2  
1) Shahid Beheshti University of Medical Sciences, Tehran, Iran; 2) Tadbir Fan Azma Co., Tehran, Iran.

Introduction: Abortion is defined as preterm labor with no chance of live birth. It is considered as the end of pregnancy before the 20th week of gestation, or a fetus birth weighing less than 500g. Recurrent pregnancy loss, which involves 5% of couples, means 2 or more consecutive pregnancy losses. The studies made so far on the matter have been mainly focused on female causes however, among men, evaluations have only been made on karyotype and chromosomal abnormalities. There are 3 azoo-sper-mia factor regions on Yq11, called AZFa, AZFb and AZFc, of which AZFc is more than others prone to deletions. In AZFc, there are a few kinds of partial deletions including gr/gr, b1/b3 and b2/b3 of which the most common sub-deletions among infertile persons are gr/gr and b2/b3. It was only in recent years that some studies were conducted to realize the relation between Y chromosome micro-deletions and recurrent pregnancy losses, in which different results were achieved. Among them the relation between partial micro-deletions of the Y chromosome and recurrent pregnancy losses, has been less investigated. Materials and methods: In this study 87 men of couples with idiopathic RPL and 50 healthy fertile men, who had at least one healthy child were examined. After DNA extraction from peripheral blood by salting out and phenol-chloroform methods, multiplex PCR was used to investigate the existence of gr/gr and b2/b3 micro-deletions. Results: No significant difference in the frequency of gr/gr micro-deletions were observed between cases (5 patients, 5/7%) and the controls (1 person, 2%) (p Value = 0.262). There was no significant relationship in the frequency of b2/b3 micro-deletions between the cases (2 persons, 2/3%) and the controls too (1 person 2%) (p Value=0/699). Conclusion: gr/gr and b2/b3 deletions have no correlation with recurrent pregnancy losses among Iranian couples. Key words: Recurrent pregnancy loss, Y chromosome, gr/gr deletion, b2/b3 deletion.

2812S  
**Predictive value of sperm count and motility in the assessment of sperm morphology in infertile men.**  
Faculty of Medicine, Sfax, Tunisia.

Introduction: Sperm morphology has an important impact on the success of fertilization. However, it is known that morphology assessment is subjective and highly variable between laboratories and technicians. We thus aimed to determine whether the value of sperm count or motility in routine semen analysis would be predictive of normal morphology. **Patients and Methods**: Data from semen analyses performed in 763 patients undergoing routine fertility evaluation in our laboratory were collected. 689 (90,3%) of them had abnormal morphology. Of the 74 specimens having normal morphology, 61 (8%) had normal count and motility. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were obtained from 2 x 2 contingency tables of sperm count and motility against morphology. Fisher’s exact test was used for the statistic analysis. **Results**: A significantly greater proportion of normal morphology samples exhibited a normal count (96% vs 75,6%; p<0,001) and normal motility (86,3% vs 59,3%; p<0,001). A normal count predicted concurrent normal morphology with a NPV of 96% and a sensitivity of 98,2%. Normal motility predicted normal morphology with a NPV of 86,3% and a sensitivity of 96,4%. A significantly greater proportion of normal morphology samples exhibited either normal count or motility (83,6% vs 53%; p<0,001) and both normal count and motility (83,6% vs 53%; p<0,001). The presence of either normal count or normal motility predicted concurrent normal morphology with a NPV of 98,5% and a sensitivity of 96,4%. **Conclusion**: We have found that normal morphology is more likely in the presence of both normal count and motility compared with the presence of normal count, normal motility, or either normal count or motility. This category includes only 8% of specimens in our study population. So, a complete assessment of sperm morphology remains necessary in order to have all the important keys of male infertility diagnosis.

2813S  
**Nef effects GnRH migration and secretion in mouse puberty and fertility.**  
EK. Ko, SD. Quaynor, RC. Cameron, LP. Chonich, LC. Layman.  
Institute of Molecular Medicine & Genetics, Georgia Regents University, Augusta, GA.

The hypothalamic-pituitary-gonadal (HPG) axis controlled by gonadotropin releasing hormone (GnRH) plays a crucial role in normal puberty and fertility. Proper specification, migration, and regulation of GnRH neuron are prerequisites for normal GnRH function. When there is impairment of GnRH action, hypogonadotropic hypogonadism results. These patients present with delayed puberty, low sex steroids, and low or inappropriately normal levels of gonadotropins. Sense of smell may be normal, as in normosmic hypogonadotropic hypogonadism (nHH), or impaired as with Kalmann syndrome (KS). Mutations in the NELF gene have been identified in human nHH/KS patients, but the mechanism of how NELF mutations impair puberty is unknown. To address this question, we have generated Nelf knockout (Nelf -/-) mice, and have shown that females manifest delayed puberty, while both males and females have reduced litter sizes. GnRH neurons occupy a smaller region of the brain in adult female Nelf -/- mice compared with Nelf +/- mice suggesting altered migration. We hypothesized that pubertal aged mice (30 days) will have impaired GnRH neuron migration and manifest hypogonadotropic hypogonadism. Serum follicle stimulating hormone (FSH) and luteinizing hormone (LH) were drawn on all animals as were serum testosterone (males) and estradiol (females). Brains from 6 Nelf -/- and 6 Nelf +/- mice of each sex (24 total at 30 days of age) were frozen and 10 um-thick coronal sections collected anterior to the optic chiasm through the posterior cerebrum were processed for immunofluorescence microscopy using antibodies directed against GnRH. Immunolabeled sections were examined at 20x using LSC microscopy and GnRH neuron number determined and compared with the number of GnRH neurons in age paired Nelf +/+ mice. Our preliminary findings suggest that in Nelf -/- mice have reduced GnRH neuron numbers at all ages tested. Elevation of pubertal mice for GnRH neuron number and distribution is ongoing, and could indicate changes that may explain the altered patterns of GnRHS neurons observed in adults.

2814S  
**High throughput sequencing of short sequence tags (STS) uncovers novel Y chromosome deletions associated with non-obstructive azoospermia.**  
1) BGI, Shenzhen, Guangdong, China; 2) Department of Biology, University of Copenhagen, Copenhagen, Denmark; 3) Peking University Shenzhen Hospital, Shenzhen, Guangdong, China; 4) 2nd people’s hospital, Shenzhen, Guangdong, China.

Male infertility, which affects approximately one in ten men among population, is closely associated with Y chromosome deletions, especially in AZOOSPERMIA FACTOR(AZFa) region. Traditional PCR based deletion detection method traces a couple of STS markers located in palindromic regions, to check the existence of long range deletions in AZFa, AZFb and AZFc region. This low resolution technique limits our exploration on extensive and novel Y chromosome deletions associated with male infertility. Here we present a novel methodology by sequencing thousands of short sequence tags(STS) distributed across the entire male specific Y chromosome(MSY) in next generation sequencing platform to identify Y deletions with significantly increased sensitivity and resolution. By exquisite algorithm and validation the accuracy of our method was proved. We applied this approach in 766 Chinese men with non-obstructive azoospermia(NOA) and 683 ethic matched healthy individuals, and identified 481 and 98 STS deleted in NOA and control group respectively, which enormously expanded the previous picture. Overall, we found NOA patient tended to carry significantly more yet rarer deletions. We also discovered a couple of novel deletions in Y chromosome significantly impact functions of spermatogenic genes thus very likely directly resulted in the incidences. Haplogroup O2* seemed to be associated with NOA in Chinese population. In summary, our work reflected a new high-resolution portrait of deletion in Y chromosome, both in NOA patients and normal population.
Expression of hsa-miR-34b, hsa-miR-181c, hsa-miR-449b, hsa-miR-517c and hsa-miR-605 in FFPE testicular tissues of infertile men with different impairments of spermatogenesis. D. Plaseska-Karanfilska, K. Popovska-Jankovic, P. Noveski, V. Filipovski, K. Kubelka. 1) RGCEB “Georgi D. Efremov”, Macedonian Academy of Sciences and Arts, Skopje, Macedonia; 2) Clinical Hospital “Acibadem Sistina”, Skopje, Macedonia.

Using microarray analysis we have previously detected 32 differentially expressed miRNAs in FFPE testicular tissues of infertile men with hypospermatogenesis (Abstract presented at the 5th Florence-Utah Symposium on the Genetics of male infertility, September 2013, Florence, Italy). Among these, hsa-miR-34b, hsa-miR-449b and hsa-miR-517c were one of the most significantly down regulated miRNAs, while hsa-miR-181c and hsa-miR-605 were up-regulated only in the patients with hypospermatogenesis and AZFc deletion. In this study the expression of these five miRNAs was studied using qRT-PCR in a total of 74 infertile men with different impairments of spermatogenesis. Based on the histopathological examination and molecular analysis the studied men were divided in the following groups: normal spermatogenesis (n=18), hypospermatogenesis (n=27), hypospermatogenesis combined with azoospermia (n=28), and azoospermia (n=11). Total RNA was extracted from FFPE testicular tissues using commercial DNA/RNA FFPE tissue kit (Qiagen, Hilden, Germany). RNA quality and purity were determined using the NanoDrop Spectrophotometer and Agilent Bioanalyzer 2100. MiRNA quantitation was performed by stem-loop RTPCR followed by Taq-Man PCR analysis using TaqMan MicroRNA Reverse Transcription Kit, TaqMan Universal PCR Master Mix and 7 TaqMan MicroRNA Assays (miR-34b, miR-449b, miR-517c, miR-181c, miR-605, as well as RNU44 and RNU6b as control genes) (Life Technologies, Carlsbad, CA, USA). The expression of each miRNA relative to RNU6b was determined using the ∆∆Ct method. Our results showed that hsa-miR-449b and hsa-miR-34b were down-regulated in all studied groups of infertile men; hsa-miR-517c was also down-regulated in all groups, except in the MA group. Hsa-miR-605 and hsa-miR-181c were up-regulated in all, except the hypospermatogenesis group, although the later with a fold change below 2. In conclusion, the qRT-PCR results of this study are consistent with the microarray data. They, furthermore, suggest that the expression pattern of these miRNAs in patients with hypospermatogenesis who also have AZFc deletions is more similar to that of the men with more severe impairments of spermatogenesis, such as SCOS than to the hypospermatogenic patients without AZFc deletion.

Novel mutations in spermatogenesis genes in azoospermic and severely oligospermic men. K.A. Fakhro1, A. Robay1, J. Rodriguez-Flores2, A. Al-Shakiki1, H. Mqdad1, C. Abi Khalil2, M. Atra2, H. El-Bardissi2, S. Sait2, J. Mezey2, R.G. Crystal1,2. 1) Genetic Medicine, Weill Cornell Medical College - Qatar, Doha, Qatar; 2) Genetic Medicine, Weill Cornell Medical College - New York, NY. Primary male infertility - the inability of a healthy male to achieve pregnancy in a fertile female - affects up to 7% of couples worldwide. In some cases, infertility is a result of physical or hormonal causes, often treatable by surgical or pharmacological intervention. However, in cases with no clear environmental cause, a primary genetic defect may be suspected, prompting routine genetic evaluation. In these cases, the infertile men may present with any combination of impaired sperm motility (asthenospermia), abnormal morphology (teratospermia) and/or reduced count (oligozoospermia), and in the most severe instances, a complete absence of sperm (azoospermia). At the Urology Department at Hamad Medical Corporation in Qatar, over 80% of these patients are classified as having idiopathic infertility, where cytogenetic testing is negative in addition to environmental causes being ruled out. Of these patients, about 10% have a strong family history of infertility (at least one concordant brother), further implicating a genetic basis. We have collected DNA from 18 such families, each comprising a core set of at least two infertile brothers and one fertile brother, plus other available family members. We use exome sequencing to identify rare, severe protein-altering variants segregating with disease in these families and have found novel candidate mutations in all 12 families sequenced to date. In 2 families, we identified novel deleterious mutations in CFTR and DNAI2, genes in which mutations have previously been reported to cause male infertility. In the remaining 10 families, we identified genes not previously implicated in human infertility, including TEX1, TULP3, MED14, SPO11, PDZD7, SPATA9, SPATA21, NOTCH4, ODF3, EEAT1 and SYNO2. All mutations were novel or very rare in all public databases, including >800 ethically similar controls sequenced on the same platform. Notably, for the 5 latter genes, we observed the same mutations in 14 additional individuals from a cohort of 128 men with sporadic infertility. Preliminary examination of these genes supports involvement in spermatogenesis, with expression in the testis, and potential roles in the development and/or survival of spermatogenic cells.

HLA-G gene polymorphisms in Mexican women with recurrent abortions. A. Porras1,2, A. Lazcano1, T. Da Silva2, C. Juarez1,4, J. Juarez1,2, F. Perea1,2, J. Garcia1,2. 1) DIVISION DE GENETICA, UNIVERIDAD DE GUADALAJARA, GUADALAJARA, JALISCO, Mexico; 2) DIVISION DE GENETICA, CENTRO DE INVESTIGACION BIOMEDICA DE OCCIDENTE, INSTITUTO MEXICANO DEL SEGuro SOCIAL, GUADALAJARA, JALISCO, MEXICO; 3) SERVICIO DE GINECOOBSTETRICIA, HOSPITAL GENERAL 180, INSTITUTO MEXICANO DEL SEGURO SOCIAL, TLAJOMULCO DE ZUNIGA, JALISCO, MEXICO; 4) DIVISION DE MEDICINA MOLECULAR, CENTRO DE INVESTIGACION BIOMEDICA DE OCCIDENTE, INSTITUTO MEXICANO DEL SEGURO SOCIAL, GUADALAJARA, JALISCO, MEXICO; 5) DIVISION DE GENETICA, UNIVIVERSIDAD DE GUADALAJARA, GUADALAJARA, JALISCO, Mexico.

The human leukocyte antigen (HLA)-G is an important immunomodulatory molecule for the maintenance of maternal-fetal relationships because it contributes to the acceptance of semi-allogeneic fetuses. Some single nucleotide polymorphisms (SNPs) in the noncoding regions of the HLA-G gene may influence cellular levels of HLA-G, contributing to pregnancy complications such as preeclampsia or recurrent spontaneous abortions (RSA). In this study, we analyzed the 725C>G (rs12333334), .201G>A (rs12333333) and 14-bp deletion/insertion (14-bp del/ins) (rs66554220) polymorphisms in the HLA-G gene by polymerase chain reaction amplification sequence-specific oligonucleotide probing (PCR-SSOP) and polymerase chain reaction amplification (PCR), respectively in 58 RSA women (≥ 2 miscarriages) without identifiable risk factors and 56 unrelated fertile women (≥ 2 live births). The ages of individuals in both groups ranged from 18 to 42 years. We found no significant differences in the genotype distributions that were analyzed between the RSA women and the fertile women. In addition, the combinations of SNPs were not in linkage disequilibrium (r2<0.1140). Therefore, this study suggests that these polymorphisms in the HLA-G gene are in linkage equilibrium and do not influence the risk of RSA in Mexican women. Key words: Human leukocyte antigen, recurrent spontaneous abortion.
2818S
Performance and Limitations of Sequenced-Based Cell-Free DNA Aneuploidy Screening: Experience of a Tertiary Referral Center and Importance of Confirmatory Follow-Up Studies. Y. Liu, W. Neufeld-Kaiser. Dept Pathology, Univ Washington, Seattle, WA.
Non-invasive prenatal screening (NIPT) using cell free DNA in maternal serum has had a major influence on diagnosis of fetal aneuploidies in the clinical setting. Much higher sensitivity and specificity than maternal serum screening was reported in recently published validation studies. Follow-up diagnostic cytogenetic testing and/or ultrasound evaluation can provide the positive predictive value (PPV) and negative predictive value (NPV) of NIPT that are the more relevant metrics in the clinical setting. We performed a retrospective chart review for 56 abnormal cfDNA screen results from March 2012 through December 2013 in high risk pregnancies. Of 648 consecutive patients referred for non-invasive prenatal screening and genetic evaluation for chromosomal aneuploidy, 5 patients (0.8%) were referred either because of inconclusive results or repeated test failures. There were 586 true negative NIPT results and one false negative NIPT finding, resulting in 99.8% negative predictive value. The false negative NIPT case was trisomy 22 mosaicism diagnosed by amniocentesis and microarray analysis. For a variety of aneuploidies, there were 45 true positive NIPT results and 12 false positive NIPT results, resulting in 78.6% Positive Predictive Value. Therefore, NIPT had a sensitivity of 97.8% and a specificity of 98% in this clinical setting. Among the 12 false positive results by NIPT, there were 3/33 trisomy 21, 4/12 trisomy 18, 1/2 trisomy 13, 2/5 monosomy X, and 2/5 XXX. Although NIPT has been shown to be the best non-invasive prenatal screen test to date, it is not a diagnostic test as shown by the false positive rate of 21.4%, an indicator of its diagnostic performance. A number of case reports have been published showing that discordant results can be due to a vanished twin, confined placental mosaicism, low-level maternal mosaicism, complex fetal chromosomal imbalances, and statistical false abnormal. Selected cases with some of these interesting findings will be discussed. Our experience shows that clinicians must recognize the limitations of cfDNA as a screening test. Diagnostic testing via fetal chromosome analysis should always be offered following abnormal NIPT results. In addition, the correlation between the maternal age and gestational age and NIPT test results will be discussed.

2819S
Referral of patients for pre-implantation genetic diagnosis: a survey of clinicians. K. Barlow-Stewart, A. Morrow, S. Seetho, B. Meiser, J. Fleming, J. Karatas. 1) Northern Clinical School, University of Sydney, Sydney, NSW, Australia; 2) Perinatal Research Group, Kolling Institute of Medical Research, Sydney, NSW Australia; 3) Psychosocial Research Group, Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia; 4) Centre for Genetics Education, Royal North Shore Hospital, Sydney, NSW Australia; 5) IVF Australia, Sydney, NSW, Australia.
BACKGROUND: Pre-implantation genetic diagnosis (PGD) is an assisted reproductive technique, in which embryos are tested for specific genetic abnormalities to enable the selection of those unaffected for implantation and pregnancy. This provides an alternative to prenatal diagnosis and potential pregnancy termination for couples at risk of transmitting a genetic condition to their children. There is recent Australian evidence to suggest that women who are not informed by their obstetricians about the availability of PGD feel disempowered and distressed about not having been provided this option. This study aimed to explore obstetrician knowledge regarding PGD and to identify potential barriers to referral for PGD. METHODS: An online questionnaire was distributed to fellows and accredited trainees of the Royal Australian and New Zealand College of Obstetricians and Gynaecologists. The questionnaire assessed respondents’ knowledge of PGD and barriers to referral. RESULTS: Of 372 practicing obstetricians who responded, 315 were fellows and 57 trainees. Obstetricians’ perceptions of their patients’ financial status and ability to access PGD services were identified as the principal reasons that justify PGD and the principal players that should be involved in the regulation of PGD. Methodology: Qualitative Research Design — Online questionnaire containing 34 questions (13 open, 11 semi-open, 10 closed-ended) has been completed by 15/30 obstetricians, 15/30 geneticists and 17/30 genetic counselors engaged in activities related to prenatal diagnosis in the province of Quebec, Data analysis: General inductive and thematic analysis. Results: Overall, the participants agree that the use of PGD for multifactorial illnesses or for social reasons should be prohibited. For them, PGD should be restricted to fatal or incapacitating illnesses. However, professionals lack consensus on certain diseases or syndromes. Regarding regulation, the majority of participants want PGD to be regulated while allowing them the flexibility to adjust their clinical management or medical counselling to particular cases. While they consider that the patients should be involved in the decision-making process, they do not believe that patients should decide everything. Conclusion: This project has made it possible to develop novel knowledge through the comparison of the perceptions of geneticists, obstetricians and genetic counsellors in Quebec. In light of these results, it is necessary to find new approaches to understand the concepts of serious genetic diseases and the sharing of responsibility, to build guidelines that are more adapted to social and professional contexts.

2821S
Combining accurate preimplantation comprehensive chromosome screening with micro-deletion and microduplication detection (SGD) in single gene disorders (diagnosed with microarray analysis) can increase the yield of PGD results. SGD results can be obtained in 4 hours by a different technology (single gene disorder (SGD) or insertion/deletion (indel) diagnosis from the same biopsy has presented many challenges. Current methods either fail to detect all origins of aneuploidy, require excessive workup times, involve long PGD turn-around times, or are expensive to perform. A new method was developed which combines each of these limitations using quantitative real-time (q)PCR. Workups involved identifying informative SNPs in the parents using SNP arrays and assigning the markers using qPCR on family members. Blastocyst biopsies underwent targeted multiplex PCR of several loci followed by qPCR for SGD, indel, contamination, and CCS analyses. For insertions, a linkage only approach was used. For deletions, multiple informative markers within the deletion were used along with linked markers outside the deletion. SGD results were compared for consistency and reliability to conventional diagnoses which were obtained from a reference lab. Additional linked informative markers were evaluated to resolve discordant cases. The typical workup involved approximately 4 weeks to complete. 17 SGD cases, including autosomal and X-linked, recessive and dominant, received and processed 150 embryos for analysis. 150 (100%) gave a 4-hour diagnosis by SNP qPCR, while 139 (92.7%) gave a result by conventional STR analysis. Concordance in diagnosis was 97%, with 4 discrepancies. Subsequent analysis of additional markers confirmed the SNP qPCR diagnosis in all cases. The use of SNPs rather than STRs greatly reduces the number of markers nearer the mutation, enhancing the ability to avoid recombination-based misinterpretation or a failure to make a diagnosis in many cases. 14 indel cases were processed including a total of 157 embryos, all of which are liveborns and 12 of which have transferred embryos. 8 had undergone transfer from which 5 have delivered, 1 has an ongoing pregnancy, and 1 did not become pregnant. In 2 cases, DNA from the newborns was available for analysis and confirmed the PGD and CCS diagnoses. This new approach to combine CCS and SGD or indel PGD involves a short and inexpensive workup and the ability to reliably and rapidly produce accurate PGD results in parallel with CCS from the same biopsy.
2822S  
Preimplantation genetic risk reduction (PGR) - a new concept in the era of microarray CGH and exome sequencing.  
G. Altarese,\textsuperscript{1} R. Been,\textsuperscript{2} T. Eldar-Geva,\textsuperscript{3} G. Lazer-Orboko,\textsuperscript{1} E. Levy Lahav,\textsuperscript{1} P. Renbaum,\textsuperscript{1} Zohar PGD Lab, Medical Genetics Institute, Shaare Zedek Medical Center, Jerusalem, Israel; 2 In vitro Fertility Unit, Shaare Zedek Medical Center, Jerusalem, Israel.  
While the ESHRE guidelines of best practice standardizes the genetic and WF approach prior to and during the PGD cycle, new technologies such as Chromosomal Microarray (CMA) and exome or genome sequencing reveal variants of unknown significance (VOUS). These variants raise challenges in the decision making process in respect to the indications calling for preimplantation genetic diagnosis. Results: We present two couples that came to our PGD unit at the Shaare Zedek Medical Center requesting preimplantation genetic diagnosis for VOUS. In the first couple, a healthy male, carrier of the variant V51D in the MLH1 gene requested PGD as risk reduction for Lynch syndrome. His family history included his father who died of colon cancer (no DNA sample), paternal grandmother died of gastric cancer (no DNA sample), a paternal uncle with colon polypos carries the familial variant and another paternal uncle who also underwent polypectomy, but is not a carrier of this variant. While the MLH1 V51D has not been reported as pathogenic it is predicted by PolyPhen to be probably damaging (score 0.96), by SIFT to be damaging (score 0.9) and by SNPs@Go to be a polymorphism. Despite the uncertainty of the pathogenicity of this variant as the cause of multiple cancers in this family, the couple, who do not require IVF for infertility, requested PGD to possibly reduce the risk of cancer. The second couple came to our PGD unit with one child diagnosed with PDD. Results from CMA revealed a 570,000 bp microduplication on chr-X:p.22.3, and a 637,000 bp microdeletion on chr-17q11.2-q21.31 in the affected child. The X linked microduplication was maternal inherited and also present in the couple’s healthy daughter, while the microdeletion was “de novo”. Both microduplications and microdeletions in these regions have been reported to be associated with intellectual disability and congenital malformations. The couple has requested PGD for both findings. Conclusions: Chromosomal microarrays and exome findings of uncertain significance pose dilemmas for both couples and geneticists in regard to counseling and performing PGD. As CMA and exome are becoming standard tests, couples are requesting PGD for VOUS to reduce their risk of possible genetic disease. These technologies demand urgent discussion and guidelines for preimplantation risk reduction.

2824S  
A Simple and Streamlined Next-Generation Sequencing-based approach to Preimplantation Genetic Screening.  
M. Umbarger, J. Gole, A. Gore, G. Porreca. Good Start Genetics, Cambridge, MA.  
Preimplantation genetic screening (PGS) is used to assess the chromosome copy number of embryos. Although increasing evidence indicates that euploid embryo transfer increases and decreases implantation and miscarriage rates, respectively, PGS adoption has been limited at least in part due to the high cost associated with traditional PGS approaches. However, increased use of trophectoderm biopsy followed by vitrification and subsequent frozen embryo transfer, coupled with streamlined workflows employing next-generation DNA sequencing (NGS), are poised to enable broader PGS adoption.  
We have developed and implemented an automated PCR-based method that amplifies regions from each chromosome and simultaneously attaches the sequencing adapters and sample-specific barcodes necessary for multiplexed NGS. 12 pg DNA purified from cell lines (~2 diploid cells) or lysate derived from 2-cell isolated cultured lymphocytes served as template for the PCR reactions. The products were sequenced to generate count data for each chromosome, and this data was subsequently used to infer chromosome copy number.  
A total of 37 true positive aneuploid chromosome calls were made across the DNA from 21 aneuploid cell lines. The method generally 789 common diploid chromosome calls, 2 incorrect aneuploid (false positive) chromosome calls, and zero incorrect diploid (false negative) chromosome calls. Both incorrect aneuploid calls were in samples containing other aneuploid chromosomes, thus yielding perfect sample level specificity and perfect chromosome-type sensitivity. Ancestral detected include trisomies 8, 9, 13, 18, 20, 21, 22, 16+21, 2+21, monosomy X, tetrasomy X, XXX, and disomy Y. The technique also detected trisomy 21 and XXX when lymphocytes were used as a template. Collectively, our results indicate that we have developed a simple and accurate NGS-based PGS approach.

2825S  
A Simple and Streamlined Next-Generation Sequencing-based approach to Preimplantation Genetic Screening.  
M. Umbarger, J. Gole, A. Gore, G. Porreca. Good Start Genetics, Cambridge, MA.  
Preimplantation genetic screening (PGS) is used to assess the chromosome copy number of embryos. Although increasing evidence indicates that euploid embryo transfer increases and decreases implantation and miscarriage rates, respectively, PGS adoption has been limited at least in part due to the high cost associated with traditional PGS approaches. However, increased use of trophectoderm biopsy followed by vitrification and subsequent frozen embryo transfer, coupled with streamlined workflows employing next-generation DNA sequencing (NGS), are poised to enable broader PGS adoption.  
We have developed and implemented an automated PCR-based method that amplifies regions from each chromosome and simultaneously attaches the sequencing adapters and sample-specific barcodes necessary for multiplexed NGS. 12 pg DNA purified from cell lines (~2 diploid cells) or lysate derived from 2-cell isolated cultured lymphocytes served as template for the PCR reactions. The products were sequenced to generate count data for each chromosome, and this data was subsequently used to infer chromosome copy number.  
A total of 37 true positive aneuploid chromosome calls were made across the DNA from 21 aneuploid cell lines. The method generally 789 common diploid chromosome calls, 2 incorrect aneuploid (false positive) chromosome calls, and zero incorrect diploid (false negative) chromosome calls. Both incorrect aneuploid calls were in samples containing other aneuploid chromosomes, thus yielding perfect sample level specificity and perfect chromosome-type sensitivity. Ancestral detected include trisomies 8, 9, 13, 18, 20, 21, 22, 16+21, 2+21, monosomy X, tetrasomy X, XXX, and disomy Y. The technique also detected trisomy 21 and XXX when lymphocytes were used as a template. Collectively, our results indicate that we have developed a simple and accurate NGS-based PGS approach.

2822S  
The frequency, type and classification of chromosome errors differs at the pre-implantation stage from that observed during pregnancy.  
M. Schweitz, S. McReynolds, W.B. Schoolcraft, M.G. Katz-Jaffe. FLC, Lone Tree, CO.  
Women at 40 years of age have a 50% reduction in fecundity compared to women a decade younger, primarily due to progressive oocyte depletion and an increase in meiotic errors resulting in aneuploidy. This well documented that oocyte aneuploidy is a major contributor to spontaneous miscarriage. The aim of this study was to evaluate the frequency, type and classification of chromosome errors in human pre-implantation embryos relative to maternal age. Infertility patients consented, under IRB approval, to an in vitro fertilization cycle with comprehensive chromosome screening (CCS). Embryos were cultured to the blastocyst stage for a trophectoderm (TE) biopsy (n=15,750 blastocysts). Biopsied TE cells were examined for chromosome number using quantitative PCR (RMA-NJ). Statistical analysis involved Chi square test with significance at P<0.05. Blastocyst CCS results revealed all 22 autosomes and both sex chromosomes were involved in chromosome gains and losses. Aneuploidy was identified in 50.2% (7,905) of these human blastocysts, with close to even numbers of trisomies (51.1%) and monosomies (48.9%) that are not typically observed in pregnancy. Predictably, for younger infertility patients (<38 years) a lower incidence of aneuploidy was observed (35.9%). This percentage was significantly less than for women of advanced maternal age (AMA; ≥38 = 63.1%; P<0.0001). Interestingly, specific chromosome errors were associated with AMA, chromosomes 14, 15, 16, 17, 18, 20, 21 & 22 were more likely (average 2.3 fold) to be aneuploid in human blastocysts from AMA infertility patients (P<0.05). In addition, the classification of these chromosome errors was significantly different relative to maternal age. Blastocysts from young infertility patients showed a significant increase in errors involving the large metacentric and submetacentric chromosomes (1-5), which are not typically observed in clinical pregnancy losses (<38=18.5% vs. ≥38=13.3%; P<0.01). The frequency and classification of chromosome errors in pre-implantation human blastocysts were significantly different in relation to maternal age. In addition, the incidence of aneuploidy in pre-implantation human blastocysts has a different profile to aneuploidy observed in first trimester losses and maternal age (AMA) ≥38 = 63.1%; P<0.0001). Interest-ingly, specific chromosome errors were associated with AMA, chromosomes 14, 15, 16, 17, 18, 20, 21 & 22 were more likely (average 2.3 fold) to be aneuploid in human blastocysts from AMA infertility patients (P<0.05). In addition, the classification of these chromosome errors was significantly different relative to maternal age. Blastocysts from young infertility patients showed a significant increase in errors involving the large metacentric and submetacentric chromosomes (1-5), which are not typically observed in clinical pregnancy losses (<38=18.5% vs. ≥38=13.3%; P<0.01). The fre-
2826S
The Israeli experience of the first 300 Panorama™ tests that use 19,488 single nucleotide polymorphisms (SNPs) followed by high-throughput sequencing for common trisomies risk assessment. H.N. Baris Feldman1, Z. Weiner2, I. Solt3, M. Shoahat4, D.M. Behar5. 1) The Genetics Institute, Rambam Health Care Campus, Haifa, Israel; 2) OBGYN department, Rambam Health Campus, Haifa, Israel; 3) The Recanati Genetic Institute, Rabin Medical Center, Petach Tikva, Israel.

Background Cell free DNA (cfDNA) has emerged over the last year as an alternative for amniocentesis for diagnosis of the common aneuploidies looking at trisomy 21, 13, 18, sex chromosomes and triploidy. Methods We present our experience of the first 300 Panorama™ tests sent from Israel. This method is based on massively multiplexed PCR amplification of cfDNA isolated from maternal plasma, targeting 19,488 SNPs, followed by high-throughput sequencing. The fetal fraction is determined. The SNP pattern of maternal DNA (from buffy coat) is compared to the SNP pattern of free DNA from maternal plasma, which contains maternal and fetal DNA. Paternal genomic samples, when available, were included in the analysis; in the absence of a paternal sample, the algorithm considers population allele frequencies. Combining the maximum likelihood ratio with a priori risk generates a risk score. Results The results of the first 300 sequential tests performed in Israel were analyzed. Fifteen samples necessitated redraw, two samples failed analysis. Four samples yielded high risk scores: two cases for trisomy 21, one for Klinefelter syndrome (KS) (47,XXY) and one for trisomy 18. Confirmation of both trisomy 21 and one KS were done by CVS or amniocentesis. The mother of suspected trisomy 18 was not interested in invasive testing in view of normal ultrasound scans and delivered a healthy baby. Karyotype was not done so mosaic state was not excluded. There were no false positives for non-invasive detection. Discussion We applied our previously established Panorama™ test as a fetal-specific epigenetic marker for noninvasive detection of trisomy 18 in maternal plasma. This study assessed the accuracy of prenatal screening for trisomy 18 by a combination of U-maspin concentration and fetal NT thickness in the first trimester of pregnancy. We recommend confirmation of the diagnosis for high risk scores pregnancies using invasive tests.

2827S
First trimester trisomy 18 screening using fetal epigenetic marker and nuchal translucency. D.E. Lee1, S.Y. Kim1, S.Y. Park1, J.W. Kim1, D.J. Kim1, D.W. Kwak1, H.M. Ryu1,2. 1) Laboratory of Medical Genetics, Cheil General Hospital and Women's Healthcare Center, Kwandong University College of Medicine, Seoul, Korea; 2) Department of Obstetrics and Gynecology, Cheil General Hospital and Women's Healthcare Center, Kwandong University College of Medicine, Seoul, Korea.

Objective: The association between fetal trisomy 18 and increased nuchal translucency (NT) (≥3.5mm) in pregnancy is well-established. Recently, the placental-derived maspin (U-maspin) gene was applied as a fetal-specific epigenetic marker for noninvasive detection of trisomy 18 in maternal plasma. This study assessed the accuracy of prenatal screening for trisomy 18 by a combination of U-maspin concentration and fetal NT thickness in the first trimester of pregnancy. Methods: A nested case-control study was conducted using maternal plasma samples collected from 65 pregnant women carrying 11 trisomy 18 and 54 normal fetuses. Fetal NT measurement was performed in the first trimester of pregnancy. Results: U-maspin concentrations were significantly elevated in women with trisomy 18 fetuses compared with controls (5.9 vs 2.0 mm; P<0.001). The sensitivities of U-maspin concentration and NT thickness for prenatal screening of fetal trisomy 18 were 90.9% and 90.9%, respectively, with a specificity of 98.1%.

Discussion: The combination of U-maspin concentration and NT thickness is highly efficient for prenatal screening of fetal trisomy 18 in the first trimester of pregnancy.

2828S
Prenatal array CGH and follow up of fetuses with increased nuchal translucency: results from VUMc. K.E. Stuurman1,2, S.L. Bhoia1, M.A.J. Engels1, E.J. Meijers-Heijboer1. 1) Clinical Genetics, VU University Medical Center, Amsterdam, Netherlands; 2) Medical Genetics, University of British Columbia, Vancouver BC, Canada, Women's Hospital & Health Centre Site.

Introduction: Increased nuchal translucency (NT) (≥3.5mm) in pregnancy is associated with chromosomal aneuploidy, genetic syndromes and structural defects. Prenatal array CGH for all fetal anomalies increases the percentage of genetic abnormalities with 2-12% compared to standard karyotyping. We assessed the outcome of prenatal array CGH in a consecutive series of fetuses with increased NT only seen in our clinic. We also followed up this cohort postnatally with respect to clinical outcome. Methods and Materials: Included were all women pregnant between January 2011 and August 2013 with a fetus with increased NT, but without additional abnormalities seen on ultrasound. We offered array CGH (Agilent 180k oligoarray) subsequent to QF-PCR and karyotyping when enough fetal material was present. Results: In total 155 fetuses had increased NT only. 124 had invasive prenatal testing (80%), of which 65 had normal QF-PCR results and karyotype (52%). Of these 65 fetuses, 37 received subsequent prenatal array CGH (60%). Two had an abnormal result (11 Mb duplication on chromosome 10q25.1-q26 and 8 Mb duplication on chromosome 2p25(5)% and these pregnancies were terminated. In total 48 of 65 fetuses, with and without prenatal array CGH, were followed up in further pregnancy and postnatally (75%). Of these 48 fetuses, five had termination of pregnancy (10%); three because of hydrops foetalis (one was diagnosed with RyR1-related congenital myopathy postnatally), two because of low trust in a good outcome. Two resulted in IUFD (4%) and one had postnatal death (2%). Four syndromes and structural defects were diagnosed postnatally (Noonan syndrome, Beckwith-Wiedemann syndrome, RyR1-related congenital myopathy and aniridia(8%), and three had undiagnosed psychomotor retardation (6%). Conclusio: Prenatal array in fetuses with increased NT only increases abnormal findings with a huge negative predictive value showing that the vast majority of fetuses with increased NT is normal. However, more than 25% of the fetuses with increased NT are abnormal. The abnormalities seen postnatally could not have been picked up with array CGH.

2829S

Background: Noninvasive prenatal testing (NIPT) that will allow genetic testing of a fetus within the first trimester of pregnancy by isolating cell-free fetal DNA (cfDNA) in the mother's plasma raises a range of ethical and legal issues. The goal of this study is to provide an Islamic ethical framework for health care providers and government agencies providing NIPT. Methods: We refer to our previous experience in medical genetics, screening tests and ethics in combination with the Islamic “Sharia/ha” (Figh) principles and authoritative “consensus edicts” “Fatwas” of Islamic scholars, literature review and our publications. We developed a set of best practices for the provision of NIPT within an Islamic framework. Results: Applying the Islamic “Sharia/ha” principle “The basic concept in useful matters is permissiveness” which indicates that everything is lawful, as long as it is useful to people, our principal recommendations include promotion of NIPT to high risk pregnant women for the prevention of fetal aneuploidy and in certain cases at high risk of single gene disorders, with the amendment of current informed consent procedures to include attention to the noninvasive nature of this new testing and the potential for a broader range of results earlier in the pregnancy. However, the need for confirmatory testing by amniocentesis must be discussed carefully with the pregnant woman, as abortion in Islam should be done before 120 days of gestation in severe malformations or devastating disorders which are incompatible with life. Conclusions: Pregnant women at increased risk of aneuploidy or other genetic conditions may be offered cfDNA testing. Its performance in low-risk women and women with multiple gestations is unclear. Such test should be regulated by government agencies. Since limited professional guidance is available clinician performing the test should adopt responsible best practices in the provision of the test within an ethical framework that continues to uphold the patient’s need to receive information with sensitivity to the options of individuals and families dealing with choices and necessities within the laws, norms and traditions of their society.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
2830S Susceptibility loci for neurodevelopmental disorders -prenatal genetic counseling and psychological impact. K.E.M. Diderich1, L.C.P. Govaerts3, J. Verhagen-Visser3, S.L. van der Steen3, M. Joosten1, M.F.C.M. Knapen1,2,3, F.A.T. De Vries1, D. Van Opstal1, M.I. Srebniak2, S.R. Riedijk2, R.J.H. Galjaard1. 1) Clinical Genetics, Erasmus MC, Rotterdam, Netherlands; 2) Dept. of Obstetrics and Prenatal Medicine, Erasmus Medical Centre, Rotterdam, the Netherlands; 3) Foundation of Prenatal Screening South-west Netherlands, the Netherlands.

Objectives: Since 2012 genomic SNP array is used as a first-tier prenatal cytogenetic test for all indications in our laboratory. Next to causative and unexpected diagnoses, array may also detect susceptibility loci (SL) for neurodevelopmental disorders, with a yet unquantifiable risk for the fetus. While some may argue that we should not disclose such uncertain results, potentially leading to unnecessary anxiety, the fact that an abnormal phenotype has been described in association with these SL, may be a reason to disclose SL. We classify SL as pathogenic. With this study we evaluated the effect of reporting and counseling SL and explored the psychological impact of hearing about a SL during pregnancy. Methods: To assess the psychological impact of releasing SL 9 couples with a child that was diagnosed with a SL during pregnancy were approached for an interview. From these, 8 women and 4 men were interviewed by phone 3-18 months after the prenatal diagnosis. During their pregnancy, they all received pre-test counseling. The post-test genetic counseling concentrated on the phenotype of the particular SL, its incidence in the normal and affected population and the difference between postnatal and prenatal ascertainment. Targeted parental array testing was offered. Extensive US examination was offered when the SL was associated with physical abnormalities. Results: Hearing about the SL during pregnancy was initially shocking to 8 and concerning for 4 parents. However, at the time of the interview, eleven parents had no worries anymore about the SL while one mother considered it a stigma. All parents considered their child healthy without concern about the SL. They stressed the importance of pre- and posttest counseling and expressed their wish for a choice regarding SL disclosure during pre-test counseling. Eleven parents indicated they wished to learn about a SL again in a next pregnancy. One mother did not know whether she wanted to learn about a SL again.

Conclusions: Although our group is small, it provides a preliminary insight into the psychological impact of releasing SL. Further studies have to be carried out in order to investigate the diagnostic results, postpartumplacental studies, and pregnancy management considerations. Case 1: NIPT performed at 14.6 weeks gestation indicated trisomy 13 in a sample sent for advanced maternal age. Amniocentesis and blood studies at birth were normal. The fetus was small for gestational age in late pregnancy. Post-delivery, chromosome 13 FISH testing revealed three out of four placental sections as mosaic for trisomy 13: 81% (162/200), 34% (67/200), and 76% (152/200). Case 2: NIPT performed at 10 weeks gestation indicated trisomy 13 in a sample sent for advanced maternal age. FISH on chorionic villus sampling (CVS) revealed trisomy 13 in 34/200 cells (17%), though cultured CVS results showed a normal karyotype. Amniocentesis results were normal by FISH (45 cells) and karyotype (37 cells). The pregnancy is ongoing with normal fetal growth. Postnatal studies are pending. Case 3: NIPT performed at 18 weeks gestation indicated trisomy 18 in a case referred for positive biochemical screening (1/11 risk for Down syndrome). Amniocentesis and blood studies at birth were normal. Post-delivery FISH testing on cultured placental tissue revealed 3.2% (8/250 cells) mosaicism for trisomy 18. Conclusions: These cases describe three pregnancies with positive NIPT tests, negative amniocentesis results, and confirmed CPM. Postpartum analysis of two placentas and CVS analysis of a third demonstrate a wide spectrum of discordance between NIPT and invasive diagnostic results, highlighting the importance of counseling and psychological support for parents in these cases.

2831S Placenta whispers: Discordant noninvasive prenatal testing (NIPT) results and the role that confined placental mosaicism (CPM) plays. T. Boomer1, S. Cherry2, J. Miles1, C. Pitrolo4, J. Wardrop1, J. Jesiolowski1, N. Teed1, C. Settler1, N. Dhariaiya1, T. Monroe2, J.-S. Saldivar1. 1) Sequenom Laboratories, San Diego, CA; 2) Central Dupage Hospital Maternal Fetal Medicine, Winfield, IL; 3) Medical University of South Carolina, Charleston, SC; 4) Capitol Health Maternal Fetal Medicine, Pennington, NJ; 5) Sequenom Laboratories, Morrisville, NC.

Background: Noninvasive prenatal testing (NIPT) screening for aneuploidy relies on the presence of circulating cell-free DNA believed to be largely placental in origin. The genetic material in fetal and placental tissue matches in most pregnancies. However, discordance between these tissues can occur due to post zygotic non-disjunction or trisomy rescue, causing uneven distribution of cells between fetus and placenta. A comprehensive study of three confirmed cases of confined placental mosaicism (CPM) identified by positive NIPT are highlighted, detailing invasive diagnostic results, postpartumplacental studies, and pregnancy management considerations. Case 1: NIPT performed at 14.6 weeks gestation indicated trisomy 13 in a sample sent for advanced maternal age. Amniocentesis and blood studies at birth were normal. The fetus was small for gestational age in late pregnancy. Post-delivery, chromosome 13 FISH testing revealed three out of four placental sections as mosaic for trisomy 13: 81% (162/200), 34% (67/200), and 76% (152/200). Case 2: NIPT performed at 10 weeks gestation indicated trisomy 13 in a sample sent for advanced maternal age. FISH on chorionic villus sampling (CVS) revealed trisomy 13 in 34/200 cells (17%), though cultured CVS results showed a normal karyotype. Amniocentesis results were normal by FISH (45 cells) and karyotype (37 cells). The pregnancy is ongoing with normal fetal growth. Postnatal studies are pending. Case 3: NIPT performed at 18 weeks gestation indicated trisomy 18 in a case referred for positive biochemical screening (1/11 risk for Down syndrome). Amniocentesis and blood studies at birth were normal. Post-delivery FISH testing on cultured placental tissue revealed 3.2% (8/250 cells) mosaicism for trisomy 18. Conclusions: These cases describe three pregnancies with positive NIPT results, negative amniocentesis results, and confirmed CPM. Postpartum analysis of two placentas and CVS analysis of a third demonstrate a wide spectrum of discordance between NIPT and invasive diagnostic results, highlighting the importance of counseling and psychological support for parents in these cases.

2832S Transcriptome expression analysis of amniotic fluid cell-free fetal RNA according to gestational weeks in Korean women. Y. Jung, S. Shim, S. Shim, S. Sung, J. Park, D. Cha. CHA University, Seoul, Rep of Korea, Seoul, South Korea.

Objective: To characterize the transcriptome expression patterns and biological pathways in amniotic fluid cell-free fetal RNA according to gestational weeks in Korean women. Method: The discarded amniotic fluids of Korean women in terms of gestational weeks were prospectively collected from euploid fetuses for this study. Total RNA was extracted from cell-free amniotic fluid supernatant and hybridized to Affymetrix GeneChip Human Arrays. Differentially expressed transcripts among 16-18 gestational weeks, 25-26 weeks and 37-38 weeks amniotic fluids were obtained by Welch’s t-test. Hierarchical clustering analyses were performed to visualize overall expression characteristics of all samples used for the study. The biological functions of selected genes were analyzed using various online Gene Ontology databases. Results: There were 2902, 6035 and 6289 genes significantly expressed in early second, late second and third trimesters, respectively. Hierarchical clustering showed differential transcriptome expression pattern according to gestational weeks. Comparing with the early second trimester, fetus related specific genes were more highly enriched in late second and third trimesters (471 vs 913 and 949, respectively). Gene expression analysis showed enrichment of brain transcripts in the late second trimester as compared with early second and third trimesters. Only 33 genes were differentially expressed in the first trimesters and pathway analysis revealed that those are related to cell proliferation and apoptosis pathway. The transcripts that are related to fetal brain were enriched in late second trimester. Conclusion: This study provides information regarding gene expression changes during normal fetal development in Korean women. Brain specific transcripts are enriched in the late second trimester rather than early second trimester. This data suggests that brain development occurs in late second trimester.
Our results demonstrate that amniotic fluid cffRNA is reflected in real-time the collection of the sample. Most studies of the fetus have obtained from development of the fetus have been limited by practical problems, such as heart, liver, kidney, lung, eye, etc. The gene expression studies for the brain and the 60 genes were confirmed expressed in various tissues, such line. Also, the 63 genes in amniotic fluid supernatant were expressed in specific transcripts through a search of pathway analysis software and Med- these genes, we were identified 123 genes that have been known to tissue- expressed in amniotic fluid supernatant compared with cultured cell. Among transcripts in the cultured cell. We found 403 genes that were differentially expressed in amniotic fluid supernatant and 15,418 GeneChip® PrimeView™ Human Gene Expression Array. The data were

BIO-AMP-2 complete medium (Biological Industries Ltd., Haemek, Israel) Cleanup kit (Qiagen, Germany). The collected amniocytes were cultured in er’s instructions. The RNA was purified with the RNeasy® MinElute® DNase digestion step to remove genomic DNA according to the manufactur- 

supernatant. The total 10 AF samples were collected each supernatants oropharynx, lungs, gastrointestinal tract, skin, and urinary system and amni- 

Center, Kangnam CHA Hosp, CHA University, Seoul, South Korea. Amniotic fluid is the only body fluid in direct contact with the the fetal 

phenotype may lead to serial dilemmas during the period of pregnancy and prenatal decision-making. Considerable discrepancy between genotype and results of ultrasonographic examination were effective and essential for have no small effect on fetal phenotype and survivability. In this case, we assume that genetic imbalance due to isochromosome 20q are likely to 

cytos cell in infancy showed non-mosaic normal karyotype different from 

distinctive abnormalities. Postnatal chromosome analysis of the lympho- 

term. At 40 weeks of gestation, female infant was delivered with no 

the presence of intermediate and small premutation alleles may provide some early developmental advantage. For larger alleles, some full mutation embryos may be non-viable from loss of the X chromosome carrying the full mutation. This has already been observed in an increased risk for mosaic Turner syndrome in females (Dobkin et al., AJMG, 2009). Early loss of an X in males and females may lead to skewed distributions of the normal versus the fragile X chromosome.

Methods for Isolation and Enrichment of circulating cell-free fetal DNA (ccfDNA) from maternal plasma, for Non-Invasive Prenatal Tests (NIPT), such as the MaterniT21™ PLUS Laboratory Developed Test. G. DeSantis, L. Chamberlain, A. Kulkarni, D. Wong,erveral collection and processing conditions along with several methods for DNA extraction, including both manual and automated plat- forms, were evaluated. The fetal DNA, total DNA and fetal fraction for each method as well as purity for downstream assays and applications (including fetal copy and fetal fraction analysis, library preparation, and next generation sequencing) are compared. Methods explored to enhance fetal fraction include: processing changes, centrifugation protocols, digestion with endo- 

nucleases, characterization of micro-particle compartments, and bead based strategies. Fetal fraction assessments are made by various approaches including qPCR, methylation sensitive restriction digestion, mass array analy- sis and sequencing based fetal fraction determinations. These genomic-technological approaches exploited the size distribution profile of isolated circulating cell free DNA (ccfDNA) and fetal fraction. Distribution of, and enrichment of, various DNA fragment sizes are discussed as well as mechanisms for enrichment. Micro-particle fractions are correlated to fetal and maternal DNA distribution to provide insight into enrichment approaches explored. Fetal mutations and fetal chromosomal abnormalities can be detected by molecu- lar analysis of circulating cell free fetal DNA (ccfDNA) extracted from maternal plasma and classification may be enhanced by fetal fraction enrich- ment. The collection and processing methods facilitate efficient extraction of circulating cell-free fetal DNA (ccfDNA) from maternal plasma. The enhanced understanding of fetal fraction enrichment approaches and com- partmentalization will lead to enhanced performance and decreased costs for NIPT.

Prenatal diagnosis of mosaic isochromosome 20q detected in amnio- centesis. S. Ito1, T. Kuchikata2, H. Yoshishita3. 1) Division of Nursing, Tokyo Metropolitan Children’s Medical Center, Tokyo, Japan; 2) Division of Medical Genetics, Tokyo Metropolitan Children’s Medical Center, Tokyo, Japan. Mosaicism of abnormal karyotypes in amniocentesis is observed approxi- mately 3%, which considered to be either true mosaicism or artifact of cultured amniocytes. When chromosome mosaicism was encountered, the definitive prenatal cytogenetic diagnosis depends on the type of fetal chro- mosome aberration. However, we often left in a conflict situation and precise information should be offered in prenatal genetic counseling. We describe a case of 34-year-old woman, anxious about the abnormal result of maternal serum screening and underwent the amniocentesis in 17 weeks of gestation. The result of chromosomal analysis was 46,XX,(20)(q10)[4]/46,XX[11]rate of mosaicism:19%), which derived from three independent cultures. Whether to continue the pregnancy or not, she was referred for genetic counseling in 19 weeks of gestation. She was informed that the most of fetuses prenatally diagnosed mosaic isochromosome 20q represented nor- mal phenotype features if not apparent abnormal echo findings, but part of them show several malformations and development delay after birth in the previ- ous reports. Decision must be made before 22 weeks of gestation in Japan, and it was impossible to perform further subsequent genetic analysis. After the session, she attended the fetal ultrasonographic examination in 20 weeks of gestation and no major malformation was noted, and decided to carry it to term. At 40 weeks of gestation, female infant was delivered with no distinctive abnormalities. Postnatal chromosome analysis of the lympho- 
cells in infancy showed non-mosaic normal karyotype different from the one in amniocentesis. All studies included in amniocytes cultured in BIO-AMP-2 complete medium (Biological Industries Ltd., Israel) and 

mRNA and mRNA originated from amniocytes in amniotic fluid using GeneChip® PrimeView™ Human Gene Expression Array. The data were

S. S. SHIM1, S. H. SHIM1, Y. W. Jung1, S. R. Sung2. 1) Dept OB/GYN, Kangnam CHA Hosp, CHA University, Seoul, South Korea; 2) Genetic Research Center, Kangnam CHA Hosp, CHA University, Seoul, South Korea. Amniotic fluid is the only body fluid in direct contact with the the fetal oophoraxym, lungs, gastrointestinal tract, skin, and urinary system and amni- 

Differences of transcriptional profiling analyses between cell free mRNA and mRNA originated from amniocytes in amniotic fluid using GeneChip® PrimeView™ Human Gene Expression Array. D.H. Cha1, S.S. SHIM1, S.H. SHIM1, Y.W. Jung1, S.R. Sung2. 1) Dept OB/GYN, Kangnam CHA Hosp, CHA University, Seoul, South Korea; 2) Genetic Research Center, Kangnam CHA Hosp, CHA University, Seoul, South Korea. Amniotic fluid is the only body fluid in direct contact with the the fetal oophoraxym, lungs, gastrointestinal tract, skin, and urinary system and amni- 

Differences of transcriptional profiling analyses between cell free mRNA and mRNA originated from amniocytes in amniotic fluid using GeneChip® PrimeView™ Human Gene Expression Array. D.H. Cha1, S.S. SHIM1, S.H. SHIM1, Y.W. Jung1, S.R. Sung2. 1) Dept OB/GYN, Kangnam CHA Hosp, CHA University, Seoul, South Korea; 2) Genetic Research Center, Kangnam CHA Hosp, CHA University, Seoul, South Korea. Amniotic fluid is the only body fluid in direct contact with the the fetal oophoraxym, lungs, gastrointestinal tract, skin, and urinary system and amni-
Phocomelia in Thrombocytopenia-absent radius (TAR) Syndrome due to compound heterozygosity for a 1q21.1 microdeletion and a RBM8A hypomorphic allele. Report of two cases. R. Jobling, S. Unger, P. Shannon, A. Tor, S. Keating, D. Chilavat 1,2. 1) Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Division of Molecular Pediatrics, Maternité, Clinique Infantile, Lausanne, Switzerland; 3) Mount Sinai Hospital, Department of Laboratory Medicine and Pathobiology; 4) Department of Diagnostic Imaging University of Toronto, Toronto, Ontario, Canada; 5) The Prenatal Diagnosis and Medical Genetics Program, University of Toronto, Toronto, Ontario, Canada.

Thrombocytopenia-absent radius (TAR) syndrome is a rare autosomal recessive disorder characterized by megakaryocytic thrombocytopenia and longitudinal limb deficiencies mostly affecting the radial ray. Most patients are compound heterozygotes for a 200 kb interstitial microdeletion in 1q21.1 and a hypomorphic allele in RBM8A, mapping in the deleted segment. We report two cases with TAR syndrome detected prenatally with phocomelia and involvement of the lower limbs. Case 1: The couple had two pregnancies affected with phocomelia and Roberts Syndrome (RS) was first suspected. In their second offspring the fetus had all four limbs affected with upper limb phacomelia, short femurs and contractures with pterygia of the lower limbs' joints. Case 2: During the couple's fourth pregnancy, routine anatomy ultrasound revealed upper limb phocomelia. The couple elected to terminate the pregnancy. Fetal autopsy revealed upper limb phocomelia, single palmar creases, adducted thumbs, bilateral 5th digit clinodactyly and partial syndactyly of finger 2-5 and lower limbs' contractures with pterygia. Both fetuses had a deletion at 1q21.1 encompassing the RBM8A gene inherited from one parent and a point mutation in the RBM8A gene inherited from the other. Fetal ultrasound finding of phocomelia leads to a wide differential diagnosis including maternal thalidomide exposure, RS, Raas-Rothschild Syndrome, Fuhrmann Syndrome among others, VACTERL association and. An important addition to this differential diagnosis is TAR syndrome.


Background: The array technique is able to detect at least 5% more cytogenetic aberrations than conventional karyotyping in cases of fetal ultrasound abnormalities. Its added value for other indications has also been shown. However, with the increase of resolution the number of pathogenic findings not related to the indication will extend as well. The chance of finding these so-called unexpected diagnoses (UD), which potentially may involve late-onset treatable disorders, is one of the reasons that the use of this technique in prenatal diagnosis is still controversial. Methods: Since 2009 we have performed ~2500 prenatal SNP-arrays, initially in pregnancies with ultrasound anomalies, and since 2012 as a first-tier test for all prenatal cytogenetic indications. We retrospectively investigated the prevalence, nature, counseling and pregnancy outcome of UD amongst these cases. Results: In approximately 1 in every 200 prenatal SNP-arrays a submicroscopic UD was encountered (n=12). This figure does not include susceptibility loci for mainly neurodevelopmental disorders which can be considered to be UD already but which we regard as a separate category of pathogenic findings (Srebniak et al., 2013). The risk of a UD is independent of indication. In most cases (9/12) the UD involved an early-onset disease. In the cases of a severe untreatable disorder (e.g. Angelman syndrome) the UD helped the couples in making a decision about the course of their pregnancy. Prenatal awareness of an early-onset treatable disease (e.g. Leri-Weill dyschondrostosis) may be considered beneficial for the newborn regarding therapy and follow up. The only late-onset untreatable disorder that we found in 3/12 cases in our cohort, was hereditary neuropathy with liability to pressure palsy (HNPP [MIM162500]), which is generally regarded as a milder disease. None of the dreaded CNVs, such as BRCA1-deletions, were found. Conclusion: Due to the added value of detecting extra clinically relevant submicroscopic chromosome aberrations without unmasking any severe late-onset untreatable disease in our cohort of ~2500 prenatal cases, we argue that SNP-array should be the first-tier test for prenatal cytogenetic studies in all indications.
2840S
Prenatal array comparative genomic hybridization (aCGH) in fetuses with structural cardiac anomalies in a medium-sized Canadian Prenatal Genetics Clinic.

J. Lazaro1, D. Fruman1, J. Lauzon1, F. Bernier1,2,4, J. Chernos1,9, R. Simrose6, M.A. Thomas1,2,4. 1) Department of Medical Genetics, University of Calgary, Alberta Children’s Hospital, Calgary, Alberta, Canada; 2) Department of Pediatrics, University of Calgary, Alberta Children’s Hospital, Calgary, Alberta, Canada; 3) Alberta Children’s Hospital Research Institute for Child and Maternal Health; 5) Cytogenetic Laboratory, Alberta Children’s Hospital, Calgary, Alberta, Canada; 6) Department of Genetics and Gynecology, University of Calgary, Calgary, Alberta, Canada.

OBJECTIVES: To determine the yield of aCGH over karyotype in cases of prenatal structural cardiac anomalies in a medium-sized Prenatal Genetics Clinic.

METHODS: 21 cases referred to the Calgary Prenatal Genetics clinic with a structural fetal cardiac anomaly, either isolated or with additional fetal anomalies, were prospectively recruited between December 2011 and March 2013. Cases with a normal karyotype and FISH for deletion 22q11.2 had an array CGH performed at Signature Genomics using the Prenatal Chip OS.

RESULTS: Of the 19 cases who met inclusion criteria, 12 presented with isolated fetal cardiac anomalies and 7 with additional fetal anomalies. The cardiac and extra-cardiac anomalies were varied. Array CGH was successfully completed on 16 cases. One case had a large 8p deletion that was also seen on karyotype and included the GATA4 gene, which has been associated with congenital heart disease. Three cases provided more information on the nature of the karyotype than aCGH, however, were not clearly the etiology for the cardiac anomaly. Two cases had an inherited CNV, including one with a duplication of 16p11.2 and another with a deletion of 15q11.2. One case had the incidental finding of being a carrier of cystinosis, a recessive disease not associated with cardiac anomalies.

CONCLUSIONS: Recent studies into prenatal aCGH have grouped fetal anomalies together but there is a reported need to better define the diagnostic yield of array CGH in individual fetal anomalies, such as cardiac anomalies, with studies that assess the full cohort of cases with that anomaly. (deWit et al., 2014). This study prospectively recruited all cases with a fetal cardiac anomaly who had an aCGH performed. We demonstrated a lower than expected diagnostic yield for pathogenic findings, after the exclusion of deletions/duplication of 22q11.2 deletion. There was a higher than expected detection (12.5%) of variants of unknown significance that could be defined as potentially clinically relevant CNVs. These CNVs are associated with variable penetrance or expressivity for neurocognitive issues and are challenging to counsel in the prenatal genetics clinic.

2842S

Non-invasive prenatal testing (NIPT) using next generation sequencing (NGS) is currently used for detection of fetal aneuploidies in cell-free fetal DNA present in the mothers’ blood. NGS however remains a relatively new technique with not all laboratory processes being automated, hence incidental sample swaps cannot be excluded. For other NGS applications such as exome sequencing, tests are developed that compare exome data and independently derived SNP data and, therefore, identifying potential sample swaps. However such an approach is not feasible for NIPT due to low sequencing depth. Thus, a reliable sample tracking method for NIPT is currently lacking. We have developed a sample tracking method for NIPT that overcomes the low coverage generally used. Based on a manuscript by Quail et al (2014), we generated PCR products of the bacteriophage PhiX174, including unique 11-mer barcodes. These barcoded amplicons were spiked into DNA isolated from human plasma, followed by library preparation and sequencing. As proof of principle, three different spike-in concentrations (0.1%, 0.5%, and 1%) were tested for three different barcodes. Following sequencing, reads that mapped against PhiX174 (BWA 0.5.9, allowing one mismatch per read) were extracted and inspected for the presence of any of 384 described barcodes. The correct spiked barcode could be unambiguously identified in all samples. A concentration of 1% spiked DNA thereby reflected the best signal to noise ratio (124:1). On average, 24×10^3 PhiX174 reads were assigned to the correct barcode and less than 0.2×10^3 reads were misassigned (2×10^2) to other barcodes. Three different barcodes were used in 3 cases to confirm the results (mean 11×10^5 mappable reads). In a second experiment, 6PhiX174 amplicons were directly spiked into human plasma. Again all spiked amplicons could be correctly identified within each sample, for all three different concentrations tested, with an even lower mean signal to noise ratio (3612:1) for the 1% concentration. Consequently, barcoded PhiX174 amplicons can be used for sample tracking in NIPT. The PhiX174 sequences do not map to the human genome, do not interfere with the NIPT data and can be spiked directly into human plasma, thereby permitting the independent analysis and library preparation. If these barcodes also remain stable when spiked directly into human blood, the addition of a barcode directly to each NIPT blood sample can be introduced into the diagnostic workflow to facilitate sample tracking and detect sample swaps.

2843S
Maternal copy number variants contribute to the burden of false positive prenatal aneuploidy test results. M.W. Snyder1, L.E. Simmons3, J.O. Mudd1,2, K. Han9, J. Hensons10, T. Maza1, J. Shendure1, H. Garmmili1,2,3. 1) Genome Sciences, University of Washington, Seattle, WA; 2) Obstetrics & Gynecology, University of Washington, Seattle, WA; 3) Division of Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, WA.

Prospective studies of noninvasive prenatal testing (NIPT) for aneuploidy by analysis of circulating cell-free DNA have demonstrated high sensitivity and specificity in both high- and low-risk cohorts. However, the overall low incidence of aneuploidy continues to limit the positive predictive value of NIPT. Few studies for general screening use. To date, the full spectrum of causes of false positive tests is not well understood. From a series of four pregnancies with discordant prenatal tests, we investigated one potential cause: maternal copy number gains. Here, we demonstrate the potential for maternally carried duplications on chromosome 18 to cause false positive results in prenatal testing of cell-free DNA. Using published metrics from one NIPT methodology, we estimated the minimum size of the maternal duplication required to give rise to false positive test results in NIPT frameworks based on exome sequencing, tests are developed that compare exome data and independently derived SNP data, thereby identifying potential sample swaps. For other NGS applications such as exome sequencing, tests are developed that compare exome data and independently derived SNP data, thereby identifying potential sample swaps. For other NGS applications such as exome sequencing, tests are developed that compare exome data and independently derived SNP data, thereby identifying potential sample swaps. For other NGS applications such as exome sequencing, tests are developed that compare exome data and independently derived SNP data, thereby identifying potential sample swaps. For other NGS applications such as exome sequencing, tests are developed that compare exome data and independently derived SNP data, thereby identifying potential sample swaps.


The Illumina HiSeq™ 2500 (Illumina®, San Diego, California) is capable of sequencing in two run modes - one for collecting large amounts of data and the other for acquiring data rapidly. This study demonstrates equivalent sequencing performance of a noninvasive prenatal test, the MaterniT21™ PLUS laboratory-developed test, using both the high output and rapid run modes of the instrument. The two run modes of the instrument were compared with the same set of sequencing libraries in a paired fashion. Five library plates, comprised of 385 euploids, 44 trisomy 21 samples, 8 trisomy 18 samples, and 3 trisomy 13 samples, were analyzed on the same HiSeq 2500 using the high output and rapid run modes with 12-plex multiplexing. In a second experiment focusing on the clustering method using rapid run mode, 48 paired libraries were sequenced with on-board cluster generation in 24-plex or following template hybridization on the Illumina cBot in 12-plex. All sequencing libraries were generated from circulating cell-free (ccf) DNA following extraction from 4 mL plasma aliquots. All flow cells were clustered on the Illumina cBot or using on-board cluster generation, and then sequenced on the HiSeq 2500. Sequencing reads were de-multiplexed, aligned to the human genome (hg19) with Bowtie 2, and chromosomal representations were calculated. Although the HiSeq 2500 rapid run mode generates fewer passed filter reads due to a reduced imaging area per flow cell lane, sequencing performance between high output and rapid run mode is remarkably similar. Representations for chromosomes 21, 18, 13, X and Y are highly correlated between the two modes. No statistically significant differences in z-scores were found for the three aneuploidy types and the discriminative distance between euploids and aneuploids is similar in either run mode. Trisomy 21 sensitivity/specificity was > 99.9% / 99.9% for the high output mode and > 99.9% / 99.7% for the rapid run mode. The sequencing metrics for onboard cluster generation and for cBot template hybridization using the rapid run mode were also found to be comparable. Therefore, classification performance of the MaterniT21™ PLUS test was shown to be equivalent using both high output and rapid run modes on the HiSeq 2500. Comparable results were also found in the rapid run mode using either on-board cluster generation or cBot template hybridization.


Non-invasive prenatal testing (NIPT) for fetal trisomy detection already revealed that there is a small chance of a false positive and false negative result. This is partly due to the fact that the fetal DNA present in the cell free maternal plasma fraction is derived from the cytotrophoblast of chorionic villi (CV), which is not always representative for the fetus. Since we already learned in the eighties from cytogenetic investigations in short-term-cultured CV (STC-villi) in which cells derived from the cytotrophoblast are studied. For this reason, accurate cytogenetic studies in CV should be done by using both STC-villi as well as long-term cultured CV (LTC-villi), the latter investigating cells derived from the mesenchymal core of CV which has the same embryonic origin as the fetus itself. We calculated the risk for a false negative trisomy 13, 18 and 21 NIPT result of a biological nature based on our experience with CV. Methods: All cases of fetal trisomy 13, 18 and 21 among ~6000 CV samples that were cytogenetically investigated (STC- and LTC-villi) in our centre between January 2000 and December 2011, were retrospectively studied for the presence of a normal karyotype or mosaicism < 50% in STC-villi. The main indications for CV sampling in these trisomic cases were fetal ultrasound abnormalities, increased nuchal translucency, advanced maternal age > 36 years, and/or abnormal first trimester screening (>1:200). Results: 404 (6,7%) cases of trisomies 13, 18 and 21 were found amongst ~6000 samples. Of these 404 cases, 15 (3.7%) had a normal (N=9) or <50% mosaic (N=6) karyotype in STC-villi and therefore would potentially be missed if NIPT was performed in these cases. It involved 6 cases of trisomy 21 and 9 cases of trisomy 18. Conclusion: Apart from technical reasons (large BMI, low fetal fraction etc) that may explain false negative NIPT results, in 2 to 3/1000 NIPT samples of patients at high risk, a trisomy 18 or 21 will be missed due to the biological phenomenon of absence of the chromosome aberration in the cytotrophoblast. It is important that patients opting for NIPT are informed about these figures so that an informed choice between invasive an non-invasive testing can be made.
2847S
Prenatal detection of fetal aneuploidy on the Ion Torrent Proton platform. T. Zwiebelhofer, P. Whitley, K. Roy, M. Saha, T. Burcham, D. van den Boom, M. Ehrlich. Sequenom Laboratories, San Diego, CA. Noninvasive prenatal testing (NIPT) for fetal aneuploidy detection via massively parallel sequencing has been successfully implemented in a number of high throughput clinical laboratories. Automation and parallelization of the complex workflow has reduced turnaround time and labor while maintaining high sensitivity and specificity. As these developments have improved workflow issues, the availability of new sequencing platforms on the market has introduced additional flexibility in implementation. Platform flexibility should encourage competitive pricing, foster innovation and ultimately improve patient satisfaction. Here, we examine the performance of an NIPT fetal aneuploidy test on the Ion Torrent Proton platform. We examined the performance of the MaterniT21™ assay using the Ion Torrent™ Proton Sequencer (Life Technologies™, San Diego California). One hundred and fifty-four patient samples, including sixteen from women carrying a known trisomy 21 fetus, as determined by fetal karyotyping, were analyzed. Libraries were prepared and sequenced according to manufacturer’s recommendations. Sequenced reads were aligned, filtered for quality and normalized for GC bias. Robust statistics were then applied to identify positive samples with a z-score greater than 3. All patient samples were correctly identified according to their karyotype results. The total number of aligned reads and uniformity of genome coverage was sufficient to generate the necessary discriminatory power to indicate aneuploidy status. Fetal aneuploidy status was correctly determined for 154/154 pregnant females, including 16 carrying a T21 fetus. Though the current Proton workflow requires more labor than is optimal for a production environment, significant improvements in that respect are anticipated in the launch of the Ion Chef template preparation system. Sequencing time was brief at <3 hours and data analysis consistent with standard platforms. In summary, the performance of the MaterniT21™ assay on the Ion Torrent Proton platform in this limited study suggests the possibility of its suitability for implementation in a clinical environment.

2848S
Utilization of a SNP array in prenatal diagnosis of Ellis van Creveld syndrome in a consanguineous couple; a case report. B. Suskin1, K. Erskine1, J. Gepp1,2, K. Uhrich1, P. Dar1,2. 1) Obstetrics & Gynaecology and Women’s Health, Albert Einstein College of Medicine; 2) Montefiore Medical Center, Bronx NY. Background: Osteochondrodysplasias and dysostoses include more than 350 disorders. As a group, their pattern of inheritance can be dominant, recessive, X-linked, or secondary to imprinting, somatic mosaicism or teratogens. Individually, they are relatively rare leading to challenges in definitive prenatal diagnosis. Case: 25 year old G2P1 who had a 3 week femur length lag on an 18 weeks dating scan presented for further detailed aneuploidy survey at 20 weeks. History was remarkable for gestational diabetes in her prior pregnancy and consanguinity; the patient and her husband are first cousins. The scan revealed rhizo-mesomelia, polydactyly, syndactyly, mildly shortened ribs, hypoplastic left ventricle, aortic coarctation and ventricular septal defect. Chest circumference was above the 10th percentile. After extensive genetic counseling, the couple opted to undergo amniocentesis with microarray analysis with the understanding the test may not have a diagnostic yield. The array includes 20 regions of interest identified for EVC and ECV2. These are relatively rare leading to challenges in definitive prenatal diagnosis. The analysis was performed on a total of 100 patient samples, including sixteen from women carrying a known trisomy 21 fetus, as determined by fetal karyotyping, were analyzed. Libraries were prepared and sequenced according to manufacturer’s recommendations. Sequenced reads were aligned, filtered for quality and normalized for GC bias. Robust statistics were then applied to identify positive samples with a z-score greater than 3. All patient samples were correctly identified according to their karyotype results. The total number of aligned reads and uniformity of genome coverage was sufficient to generate the necessary discriminatory power to indicate aneuploidy status. Fetal aneuploidy status was correctly determined for 154/154 pregnant females, including 16 carrying a T21 fetus. Though the current Proton workflow requires more labor than is optimal for a production environment, significant improvements in that respect are anticipated in the launch of the Ion Chef template preparation system. Sequencing time was brief at <3 hours and data analysis consistent with standard platforms. In summary, the performance of the MaterniT21™ assay on the Ion Torrent Proton platform in this limited study suggests the possibility of its suitability for implementation in a clinical environment.

2849S
Uniparental origin GWAS of human gestational age implicates genes involved in angiogenesis. J. Bacea1, V. Sengpieler2, R. Myhre3, G. Zhang4, L. Magi5, K. Ehrlich6, J. Jacobsson3, L. Mulder2, J. Fluen2, C. Mihai2, M. Jin2, N. S. L. Ellis1, L. D hommel7, D. Grohmann8, F. Thomsen9,10, O. Glendinning8,11, L. Muglia2,3, R. Myhre4, G. Zhang1, J. Gebb1,2, S. Klugman1, P. Dar1,2. 1) Obstetrics & Gynaecology and Women’s Health, Albert Einstein College of Medicine; 2) Montefiore Medical Center, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, United States of America; 3) Mathematical Sciences, Chalmers University of Technology, SE-41296, Gothenburg, Sweden. Background: More than 10% of the pregnancies World-wide end up in preterm delivery (PTD) - the major cause of neonatal death and morbidty. More than 30% of the variation in gestational age can be attributed to genetic factors. Our knowledge of such factors is very limited. Hypothesis: Genetic polymorphisms (SNPs) affecting gestational age are located in genomic regions that experience an uniparental silencing (imprinting) and thus are more likely to be detected by genome-wide association study (GWAS) in separated haplotypes of the fetal genome. Methods: We used data from 991 mother-child pairs from the Norwegian Mother and Child cohort. Genotypes were phased using family structure data and reference haplotypes from 1000 Genomes Project. SNPs at maternal and paternally derived haplotypes were tested for association with gestational age (in days), while adjusting for child’s sex, maternal age and parity. Results: None of the 525577 tested SNPs showed a genome-wide significance (p=2x10^-6). The top SNPs from paternal-haploid analysis were located in the introns of angiogenesis-related genes ZBTB16, ANG, COL6A1. Angiogenesis is an essential process during placental vascularisation and can influence nutrient supply to the fetus, thus affecting gestational age. ANG gene encodes angiogenin which is induced both by hypoxia and inflammation, and has been shown to be a marker for PTD. ZBTB16 is known to regulate expression of collagen, metalloproteinases, secretion of interleukins and interferon gamma, all of which are known to be involved in PTD. Moreover, ZBTB16 and ANG are involved in imprinting-related studies. The top SNPs from maternal-haploid analysis were located in the intron of the gene ABCC8, which is involved in insulin (also imprinting-related) secretion. Conclusion: Our findings are consistent with the intergenic-conflict hypothesis, which predicts that any developmental gene-environmental interaction(s) have been identified as etiological contributors. Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s). Using genetic predisposition, maternal exposure to physiopathological condition(s), or gene-environmental interaction(s) have been identified as etiological contributors(s).
2851S
Association of candidate gene single nucleotide polymorphisms with the clinical subtypes of preterm birth (PTB). L.G. Gimenez1,2, A.M. Momany2, J.A. Gil1, T.A. Poletta1,4, T. Buschi1, B. Cornia1,2, V. Cosentino2, C. Salem1, H. Krupitzki1, E. Castilla2,4, E. Gadow1, J.C. Murray1, J.S. Lopez Camel1o1,2,4. 1) Dirección de Investigación, CEMIC (Centro de Educación Médica e Investigaciones Clínicas), Buenos Aires, Argentina; 2) ECLAMC (Instituto de Latido fetal de México (Centro de Généticas)) at CEMIC, Buenos Aires, Argentina; 3) Maternidad Nuestra Señora de la Merced, Tucumán, Argentina; 4) INAGEMP (Instituto Nacional de Genética Médica Populacional) at CEMIC, Buenos Aires, Argentina; 5) Department of Pediatrics, University of Iowa, Iowa City, USA.

Background. Preterm birth (PTB) is the leading cause of perinatal morbidity and mortality worldwide. The etiology of PTB is multi-factorial, heterogeneous, and there is strong evidence of genetic susceptibility. The aim of this work was to investigate the association between 24 single nucleotide polymorphism (SNPs) and the different clinical subtypes of preterm birth: spontaneous (PTB-I), premature rupture of membrane (PTB-PROM) and medically indicated (PTB-M). These SNPs have been previously studied in a heterogeneous group of PTBs. Methods. The sample included 674 triads (proband, mother, father) recruited at the Nuestra Señora de la Merced Maternity Hospital in Tucumán, Argentina. Of these triads, 233 had probands from PTB-I, 241 had probands from PTB-PROM and 200 had probands from PTB-M. We studied 24 SNPs in 18 candidate genes. Genotyping was performed using Applied Biosystems Taqman probes and the Fluidigm genotyping platform. SNPs were chosen based on previous reports of significant association with preterm birth. Data were analyzed using the Transmission Disequilibrium Test (TDT) (Spelman et al. 1993). The p values (<0.05) were corrected for multiple comparisons. Preliminary results. We found a significant association (P<0.05) between a SNP in COL4A3 (rs10178458), a SNP in POL1 (rs2723865), and a SNP in CRHR1 (rs4458044) with PTB-I. A SNP in F3 (rs610277) showed significant association (P<0.05) with PTB-PROM and PTB-M. A SNP in KCN3 (rs885319) showed significant association (P<0.05) with PTB-M. Conclusions. This study suggests different genetic influences between the different clinical subtypes. These findings may have implications in understanding the pathophysiology of clinical subtypes of preterm birth.

2852S

Introduction: An important medical challenge is to mitigate the public health impact of preterm birth, both in its cost and complications. More than 30% of extremely premature infants are the products of pregnancies that have been complicated by infection. The fetal unit includes the fetus, the fetal membranes (chorion and amnion) and the placenta. Chorioamnionitis, or infection of the fetal membranes can lead to preterm birth via rupture of the membranes. Although chorioamnionitis is associated with infants of any gestational age, it is predominantly found in the preterm infant, and its incidence increases as the gestational age decreases. The first line of defense against pathogens is the innate immune system, and the portal of entry is the Toll-like receptor (TLR). TLRs are pattern-recognition receptors that recognize pathogen-associated molecular patterns. Activation of TLRs initiates an inflammation cascade involving downstream signaling moieties and cytokines. Genetic variation in these pathways is associated with a variety of disease processes, including chronic and deregulated inflammation. TLR2 and TLR4 are involved in the inflammatory events associated with labor. Purpose: We tested the hypothesis that genetic alterations of the fetal innate immune response are associated with an increased risk of prematurity. Approach: We evaluated three TLRs, 1, 2 and 4, using single nucleotide polymorphisms (SNPs). We utilized a unique and valuable set of newborn DNA samples derived from residual newborn screening specimens from over 3000 Wisconsin infants who represent the full spectrum of gestational ages (23-42 weeks). A univariate analysis was used to determine whether the TLR1 genotype is a predictor for gestational age and/or birth weight. Results: We found strong significant correlations between TLR2 or TLR4 SNPs and preterm birth. There was a correlation between the TLR1 SNP rs4986791 (N248S) genotype in both the birth weight and gestational age analyses when compared to full-term populations (37-42 weeks gestational age). There was a significant association between TLR1 SNP (N248S) and preterm birth in the non-Hispanic Wisconsin black population (P<1×10^-6). Conclusions: Our data suggest that TLR1 alterations may be associated with an increased risk of preterm birth in black infants in Wisconsin.

2853S
Whole Exome Sequencing of Hispanic Infants Reveal Novel Pathways Implicated In Spontaneous Preterm Birth. M.K. Veerapen1,2, E. Rampersaud1, L. Pelaez1, J.E. Potter2, S. Wu3, M.M. Rodriguez4, O.A. Bodamer1,2,3, 1) Dr John T. Macdonald Department of Human Genetics; 2) Department of Biochemistry and Molecular Biology; 3) John P. Hussman Institute of Human Genomics; 4) Department of Pathology; 5) Department of Obstetrics and Gynecology; 6) Department of Pediatrics, Division of Neonatology, University of Miami, FL.

Approximately 12% of all infants in the United States are born prematurely with considerable variation based on ancestry, social status and access to health care. Preterm birth (PTB) is multifactorial with an overall heritability of up to 37% and higher variance in spontaneous PTB (SPTB). Despite an increased risk for SPTB in Hispanic women, few studies have identified genetic factors associated with SPTB in this population. Consequently, much of the heritability remains unexplained. Whole exome sequencing (WES) has the potential to close this knowledge gap through the identification of additional genetic factors. The use of WES on archival tissues is novel, facilitating the study of well characterized cohorts from existing biorepositories. We proposed to utilize a cohort of retrospectively collected umbilical cords from SPTB Hispanic infants for the identification of genetic risk factors to SPTB by WES. Twenty-three Hispanic infants following SPTB were retrospectively identified through an umbilical cord collection at the University of Miami/Jackson Memorial Hospital. gDNA was isolated for ACE-capture WES (Personalis Inc., Menlo Park, CA). Individual exomes were filtered for variants with a genotype quality (GQ)>90 and analyzed using the likelihood ratio test (LRT) against the 1000 Genomes Project as background (VAAST v2.0.4). Significant shared genes (FDR<0.05) and VAAST scores were used for pathway analysis with MetaCore (v6.18, Fluid 65505). Mean gestational age was 34.1±1.72 weeks; birth weight was 2.36±0.46 kg; and APGAR scores at 10 min were 8.8±0.59 gDNA (4.6±1.43 g) were successfully extracted from archival umbilical cords for WES: average read depth was 29.5× and 13.1× reads for PTB-I, 241 had probands from PTB-PROM and 200 had probands from SPTB-M. We studied 24 SNPs in 18 candidate genes. Genotyping was performed using Applied Biosystems Taqman probes and the Fluidigm genotyping platform. SNPs were chosen based on previous reports of significant association with preterm birth. Data were analyzed using the Transmission Disequilibrium Test (TDT) (Spelman et al. 1993). The p values (<0.05) were corrected for multiple comparisons. Preliminary results. We found a significant association (P<0.05) between a SNP in COL4A3 (rs10178458), a SNP in POL1 (rs2723865), and a SNP in CRHR1 (rs4458044) with PTB-I. A SNP in F3 (rs610277) showed significant association (P<0.05) with PTB-PROM and PTB-M. A SNP in KCN3 (rs885319) showed significant association (P<0.05) with PTB-M. Conclusions. This study suggests different genetic influences between the different clinical subtypes. These findings may have implications in understanding the pathophysiology of clinical subtypes of preterm birth.

2854S

Noninvasive prenatal testing based on massively parallel sequencing (MPS) of circulating cell free DNA (ccfDNA) from pregnant plasma offers a powerful tool for detecting fetal chromosomal aneuploidies and other copy number variations; however, copy neutral structural rearrangements have proven challenging. We aimed to detect and characterize a balanced fetal translocation. Our algorithm identified the known specific translocation event by sequencing ccfDNA from maternal plasma. Simulations were used to develop an algorithm which leverages base increment changes in mapping characteristics of ccfDNA to identify paired end reads potentially harboring structural rearrangements. We then applied this methodology after performing high-coverase, 100bp paired-end sequencing of ccfDNA isolated from the plasma of a 38-year-old pregnant donor carrying a fetus with a balanced translocation. Our algorithm identified the known specific translocation event by sequencing ccfDNA from maternal plasma. Simulations were used to develop an algorithm which leverages base increment changes in mapping characteristics of ccfDNA to identify paired end reads potentially harboring structural rearrangements. We then applied this methodology after performing high-coverase, 100bp paired-end sequencing of ccfDNA isolated from the plasma of a 38-year-old pregnant donor carrying a fetus with a balanced translocation. Our algorithm identified the known specific translocation event by sequencing ccfDNA from maternal plasma. Simulations were used to develop an algorithm which leverages base increment changes in mapping characteristics of ccfDNA to identify paired end reads potentially harboring structural rearrangements. We then applied this methodology after performing high-coverase, 100bp paired-end sequencing of ccfDNA isolated from the plasma of a 38-year-old pregnant donor carrying a fetus with a balanced translocation. Our algorithm identified the known specific translocation event by sequencing ccfDNA from maternal plasma. Simulations were used to develop an algorithm which leverages base increment changes in mapping characteristics of ccfDNA to identify paired end reads potentially harboring structural rearrangements. We then applied this methodology after performing high-coverase, 100bp paired-end sequencing of ccfDNA isolated from the plasma of a 38-year-old pregnant donor carrying a fetus with a balanced translocation. Our algorithm identified the known specific translocation event by sequencing ccfDNA from maternal plasma.
2855S
Low folate levels and MTHFR polymorphism C677T in case-mothers of children with neural tube defects and control-mothers of Pakistani origin. A Case- control study. N. Nauman1, S. Jalali1, S. Shami1, S. Rafii2, G. Groje2, A. Hilger2, M. Draakken2, M. Nötgen2, M. Ludwig2, H. Reutter2. 1) Pathology, Rawalpindi Medical College, Rawalpindi, Pakistan; 2) Institute of Human Genetics, Bonn, Germany Department of Neonatology, Bonn, Germany; 3) Department of Pathology, Holy Family Hospital, Rawalpindi, Pakistan.

Introduction: Neural tube defects (NTDs) are congenital malformations of central nervous system which result from a failure of the neural tube to close during the fourth week of embryogenesis. Low folate levels are associated with increased risk of neural tube defects. Genetic studies examining the gene coding for the folate metabolizing methylenetetrahydrofolate reductase (MTHFR) enzyme suggest that the functional 677C/T thermolabile polymorphism contributes a genetic risk to NTDs. Previous studies suggest that not only the fetal genotype but also the maternal genotype might have an impact on the fetal development. The deleterious effects of this mutation can be overcome if folate levels are adequate. Here we carried out a case-control study of case-mothers and control-mothers of Pakistani origin. Patients and Methods: We examined 109 case-mothers of children born with NTDs and of 100 control-mothers without history of NTDs in their offspring. Case-mothers and control mothers were of Pakistani origin. Red blood cell folate and serum folate were determined. Genotype and folate level comparisons were carried out using chi-square analysis. Results: In the case mothers mean RBC folate and serum folate were highly significantly low as compared to control mothers (p<0.0001). Genotype analysis of case-mothers and control-mothers revealed 11 (10.09%) and 2 (2%) respectively to be homozygous for the MTHFR 677T allele. There were 32 (29.35%) case-mothers and 26 (26%) control-mothers to be heterozygous for this polymorphism. Genotype comparison by chi-square analysis of case-mothers and control-mothers was significant (p=0.0393). Conclusion: We found the homozygous MTHFR 677TT genotype to be more frequent among case-mothers compared to control-mothers. Mean RBC and serum folate levels were significantly low in case mothers as compared to control mothers. Combined analysis of these results with the information folate supplementation during the perinatal period of case-mothers and control-mothers is warranted to further elucidate the role of the MTHFR genotype and folate supplementation per se in the risk of NTDs in pregnancies of women of Pakistani origin. Key words: Case- control study, MTHFR polymorphism, folate levels, neural tube defects.

2856S

The purpose of this study was to assess the value of appropriate management of fetal tissue by obstetricians, pathologists and dysmorphologists.management of fetal tissue in order to provide adequate counseling, pregnancy management and prenatal diagnostic procedures, and management of fetal demise material. All patients were referred to our center for evaluation of fetal demise and to provide genetic counseling, pregnancy management, prenatal diagnostic procedures, and management of fetal demise material. We present data on our experience during the last five years of the state wide Perinatal Loss Evaluation Program in our institution. All patients were referred to our center for evaluation of fetal demise with the majority being second or third trimester losses. We either examined patients directly or through photographic material obtained by previously trained staff. We requested cord blood or tissue material for cytogenetic analysis and further history was obtained by our genetic counselor. Cases were evaluated by a pediatric pathologist, a dysmorphologist, and maternal fetal medicine specialist. The etiology of the demise in each case was suggested after reviewing medical records, autopsy, dysmorphologic exam and cytogenetic analysis.

Preliminary results indicated a likely identifiable cause for the demise in 88% of case (N=129). Fetal (33%), multiple risk factors (15%), placental (8%), infectious (8%), cervical (6%), umbilical cord (6%). We briefly review here the diagnoses seen, particularly those attributed to fetal abnormalities with likely or proven genetic factors. Our data demonstrate the importance of evaluating fetal demise material in order to provide appropriate information to families in terms of cause and recurrence of their loss. We review the value of appropriate management of fetal tissue by pathology and dysmorphology as well as the value of photographic evaluation post mortem when products of conception are not available.

2857S
Molecular and histopathological findings in placentas of newborns with Down Syndrome. R. García-Robles1, J. Martínez1, P. Ayala-Ramírez2, M. Olaya1, M. Bermúdez1. 1) Instituto de Investigación en Nutrición, Genética y Metabolismo. Universidad El Bosque, Bogotá, Columbia; 2) Instituto de Genética Humana. Pontificia Universidad Javeriana, Bogotá, Colombia; 3) Departamento de Patología. Hospital Universitario San Ignacio - Pontificia Universidad Javeriana, Bogotá, Colombia.

Down Syndrome is the most frequent chromosomal abnormality with a worldwide incidence of 1/600-700 newborns. Down Syndrome is the main cause of mental retardation and is caused by a trisomy of chromosome 21 involving a specific region named “Down Syndrome critical region”. The locus of the CBS gene is on this region and this gene codes for the enzyme Cystathionine β Synthase that participates in homocysteine metabolism. It has been reported overexpression of this gene in some cells and tissues of individuals with Down Syndrome and it is possible that this abnormal expression could be involved in some phenotypical findings of the disease. However, it has not been described this finding in placentas of affected individuals. The aim of this study was to assess the expression levels of mRNA of protein of Cystathionine β Synthase and histopathological changes in placentas of individuals affected with Down Syndrome and a control group of healthy and normal newborns. Thereby, we studied six placentas of newborns with Down Syndrome and 16 placentas of control group. We evaluated mRNA by real time PCR and protein by immunohistochemistry. An expert pathologist was blinded to diagnosis and evaluated the placentas. We found increased expression of Cystathionine β Synthase mRNA levels (p=0.0465) and protein levels (p=0.0090) in placentas of case group compared to control group. The histopathological changes showed statistically significant differences between groups in the following findings: irregular villous outlines (p=0.0007) and trophoblastic inclusions (p=0.0037). These results suggest overexpression of Cystathionine β Synthase in placentas of individuals with Down Syndrome, possibly causing alteration in homocysteine metabolism leading to mental retardation. Our findings in placentas were associated with trisomy 21. Finally, these molecular and pathological findings could help to identify individuals with Down Syndrome when the placenta is available to be assessed.

2858S
Rising Whole Body Counts of 125Cs (WBC) in Pregnant Women and Persistent Elevated Rates of Neural Tube Defects (NTD) and Microcephaly/Microphthalmia (M/M) in a Chornobyl Impacted Region of Rivne (R) Province in Ukraine. W. Wartelec1, L. Yevtushok2, N. Zymak-Zakumia2, Z. Sosyunik2, S. Lapchenko4, I. Kuznietsov2, A. Korblein1. 1) OMNI-Net USA; 2) Rivne Diagnostic Center, Rivne, Ukraine; 3) Medical Genetics Center, Khmelnytsky, Ukraine; 4) OMNI-Net Ukraine; 5) Eastern European University, Lutsk, Ukraine; 6) OMNI-Net Germany.

Purpose: A population-based surveillance of congenital anomalies in R demonstrated elevated rates of NTD and M/M. Two prevalent teratogenic risks in R are alcohol and 125Cs. Analyses indicated that alcohol is an unlikely primary cause. This study concerns WBC of ambulatory patients and pregnant women. Methods: Study of 44,438 WBC patterns recorded from adult ambulatory patients (2001-2013) inclusive of 6,425 pregnant women from non-contaminated counties and 5,090 pregnant women from contaminated counties. Results: WBC were statistically significantly higher among pregnant women from contaminated counties than those from non-contaminated counties (2001-2013). Population-based rates of NTD and M/M in polluted regions of R are persistently elevated over 28 years respectively. The corresponding values among women from non-contaminated counties were 13 and 29 per 10,000 live births respectively in the same period. Conclusions: Persistent elevated rates of NTD and M/M in a Chornobyl Impacted Region of Rivne, Ukraine. Further studies are needed to understand the mechanisms underlying the higher than expected rates of NTDs and M/M in R. The higher than expected rates of NTDs and M/M in R may be due to thalidomide intake in R in the early 1970s, although the evidence for this exposure is limited.

2859S
Posters: Prenatal, Perinatal, and Reproductive Genetics
2859S
Assessing the causal relationship between maternal height and birth outcomes: A Mendelian randomization analysis. G. Zhang 1, J. Bacelli 2, C. Lengyel 2, B. Jacobsson 2, L. Muglia 3, 1) Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; 2) Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; 3) Center for Prevention of Preterm Birth, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH; 4) Department of Obstetrics and Gynaecology, Sahlgrenska Academy, Sahlgrenska University Hospital/Ostra, Gothenburg, Sweden; 5) Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway.

Background: Epidemiologic studies have shown that maternal height is significantly associated with gestational age and various birth outcomes, such as birth length and birth weight. However, it is unclear whether these associations truly reflect causal effects of maternal height. Alternatively, as an indicator of maternal nutrition status, maternal height could be indirectly associated with those birth outcomes. In this study we investigated the causal influence of maternal height on birth outcomes by Mendelian randomization using height genetic score as an instrumental variable.

Methods: We used data of 3,486 mother/infant pairs from three birth cohorts collected from three northern European countries (Finland, Denmark and Norway). Genome-wide SNP genotype data were generated by SNP arrays and imputation. We constructed height genetic score using 171 SNPs to estimate the causal effects of maternal height on birth outcomes.

Results: In observational analysis, maternal height was significantly associated with gestational age (p=1.4x10^{-7}), birth length (p=3.26x10^{-12}) and birth weight (p=2x10^{-18}). The constructed height genetic score in mothers was significantly associated with maternal height (p<2x10^{-16}) and could explain more than 8% of the observed variance in maternal height. In Mendelian randomization analysis, height genetic score in mothers was significantly associated with infant’s birth length (p=0.0015) and birth weight (p=1.49x10^{-5}) and the estimated causal effects were directionally concordant with the observed associations (i.e. higher maternal height was associated with longer birth length and higher birth weight). The association between maternal height genetic score and gestational age was not significant (p=0.71).

Conclusion: Our results highly suggest that maternal height, as an important component of prenatal environment, causally affects both birth length and birth weight; however, whether maternal height causally impacts gestational age, or reflects nutritional or other influences on overall maternal health, requires further evidence.

2860S
Maternal and placental genome-wide and candidate gene association studies of placental abruption. M. Denis 1, D.A. Enquobahrie 2,3, M.G. Tadesse 2, B. Gelaye 2, S.E. Sanchez 1,2, M. Salazar 2, C.V. Ananth 1,2, D.A. Enquobahrie 1, 1) Department of Epidemiology, Harvard School of Public Health, Boston, MA; 2) Center for Perinatal Studies, Swedish Medical Center, Seattle, WA; 3) Department of Epidemiology, University of Washington School of Public Health, Seattle, WA; 4) Department of Mathematics and Statistics, Georgetown University, Washington, DC; 5) Sección de Post Grado, Facultad de Medicina Humana, Universidad San Martín de Porres, Lima, Peru; 6) A.C. PROESA, Lima, Peru; 7) Department of Obstetrics and Gynecology, San Marcos University, Lima, Peru; 8) Department of Obstetrics and Gynecology, College of Physicians and Surgeons, Columbia University Medical Center, New York, NY; 9) Department of Epidemiology, Joseph L. Mailman School of Public Health, Columbia University, New York, NY.

Introduction: Available evidence supports the role of genetics in the pathogenesis of placental abruption (PA), the premature separation of the placenta from the uterus, an important cause of perinatal mortality. Associations of maternal genetic variants with risk of PA have been demonstrated. However, PA-related variations in the placental genome and interactions between maternal and placental genomes on PA risk have not been investigated.

Methods: Maternal blood and placental samples collected from participants of the Peruvian Abruptio Placentae Epidemiology study were genotyped using the Illumina Cardio-Metabochip platform. A total of 118,782 genome-wide SNPs and 333 SNPs in 32 candidate genes from mitochondrial biogenesis and oxidative phosphorylation pathways were examined in placental DNA from 280 PA cases and 244 controls. For interaction analyses in maternal and placental genomes, 325 SNPs from 32 candidate genes were assessed among 222 PA cases and 198 controls. Univariate and penalized logistic regression models were fit to estimate odds ratios. We also examined associations of combinations of SNPs with PA risk using weighted genetic risk scores (WGRS) and haplotype-based analyses. A multinomial model was used to investigate maternal and placental genome interactions. Functions of genes represented by significant SNPs were further examined using pathway analyses.

Results: Overall, in placental genome-wide and candidate gene analyses, no SNP was significant after false discovery rate correction. The top SNPs using the WGRS approach were rs1434201 (CTNNA2), rs149270 (TNFRSF1A) and rs13055470 (ZNRF3) (nominal p-values: 1.11e-05 to 5.00e-05). Genes participating in cell cycle, growth and proliferation were overrepresented by the top 200 SNPs of the GWAS. Subjects in the highest quartiles for WGRS (9 SNPs selected by lasso regression) had a 5.5-fold higher risk (95%CI 3.2-9.7) of PA compared with those in the lowest quartile. A haplotype block in PPARG was associated with a 50% (95% CI: 0.29-0.88) lower risk of PA compared to the referent haplotype. We found a strong interaction between maternal and placental genomes on PA risk for two SNPs in PPARG (chr3:12313450 and chr3:12412978 with p-values 8.62E-08 and 7.89E-10).

Discussion: Variations in the placental genome and interactions between maternal-placental genetic variations can contribute to PA risk. Larger studies in this area can help advance our understanding of PA pathogenesis.
Comprehensive genotype phenotype correlations reveal NLRP7 role in regulating the balance between embryonic tissue differentiation and trophoblastic proliferation. NMP Nguyen1, L. Zhang1, P. Sauder2, C. Dery1, K. Rahim1, R. Reddy1, J. Arseneau1, A. Cheung1, U. Surti1, L. Hofner2, M. Seoud2, G. Zaatar2, R. Bagga2, R. Srivasan2, M. Brequert2, W. Buckett2, P. Coulth2, A. Ap3, R. Slim1. 1) Human Genetics, McGill University, Montreal, Quebec, Canada; 2) Service de Gynécologie Oncologique, Hôpital Notre-Dame, Montréal, Quebec, Canada; 3) Department of Pathology, University of Hong Kong, Queen Mary Hospital, Hong Kong, China; 4) University of Pittsburgh, Pittsburgh, Pennsylvania, USA; 5) American University of Beirut, Beirut, Lebanon; 6) Post Graduate Institute of Medical Education and Research, PGIMER, Chandigarh, India; 7) Endocrinologie et Génétique de la Reproduction et du Développement, Clamart, France.

Hydatidiform mole is a human pregnancy with excessive trophoblastic proliferation and abnormal embryonic development. NLRP7, a nucleotide oligomerization receptor, pyrin containing 7, is a major gene responsible for recurrent hydatidiform moles. Recurrent HMs are caused by recessive mutations, mostly in NLRP7, and have diploid biparental genomes. To better understand the effect of NLRP7 on the mechanism leading to hydatidiform moles, we used a combination of five different approaches, flow cytometry, fluorescent in situ hybridization, microsatellite DNA genotyping, and p57KIP2 and Ki-67 immunohistochemistry to comprehensively characterize 103 hydatidiform moles from patients with or without NLRP7 mutations and variants. We confirm that most molar tissues from patients with two NLRP7 defective alleles are diploid biparental with a single cellular population, importantly, our data demonstrate significant genotype-phenotype correlation between the nature of the mutations and the histological and cytogenetic features of the hydatidiform moles. Protein-truncating mutations were associated with more severe molar phenotype such as excessive trophoblastic proliferation and absence of embryonic tissues while missense mutations were associated with milder phenotype such as partial molar pregnancy while missense mutations were associated with milder phenotype such as partial molar pregnancy. Statistical analysis using a series of filtering steps, we identified rare, damaging, biologically relevant variants exhibiting co-segregation with affection. Potentially causative variants were screened in available family members, additional families, and a replicate group of over 570 cases with HG and over 425 unaffacted controls using Sanger Sequencing, Fluidigm, and Taqman. Our data suggest that NLRP7, depending on the severity of its mutations, regulates the time at which embryonic development arrests.

Kidney disease genes linked to Hyperemesis Gravidarum. M. Schoenberg-Fejo1, K. MacGibbon2, J. Sinseheimer3, P.L. Reddy4, P. Pajukanta5, K. Tabsh1. 1) Department of Obstetrics and Gynecology, David Geffen School of Medicine, University of California, Los Angeles, CA; 2) Hyperemesis Education and Research Foundation, Leesburg, Virginia; 3) Department of Biostatistics, Biomathematics, & Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA; 4) Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA.

Hyperemesis Gravidarum (HG), severe nausea and vomiting of pregnancy, often results in dehydration and undernutrition and is the 2nd leading cause of hospitalization in pregnancy after preterm birth. Current treatments are largely ineffective and the cause is unknown. Evidence for a genetic component comes primarily from a classic twin study and reports of familial aggregation. This study was undertaken to elucidate the etiology of HG by identifying predisposing genes. HG reduces reproductive fitness, which makes whole exome sequencing to find rare variants a reasonable approach. We performed exome sequencing of 18 individuals from 5 families with HG. Using a series of filtering steps, we identified rare, damaging, biologically relevant variants exhibiting co-segregation with affection. Potentially causative variants were screened in available family members, additional families, and a replicate group of over 570 cases with HG and over 425 unaffected controls using Sanger Sequencing, Fluidigm, and Taqman. Three of five families carry biologically related rare damaging variants in genes involved in kidney function, PKHD1, PKD1, and LAMA5. None of over 425 controls carried these variants. Thus, a common pathway may be responsible for the majority of cases of familial HG. This novel discovery of rare variants relating to HG susceptibility is the first step in understanding the biology of HG; may be relevant to understanding genetic susceptibility to nausea and vomiting in general; and provides a new path for development of more effective therapies.
2864S
The genetic basis of preeclampsia in a high altitude indigenous Andean population. C.R. Gignoux1, K. Sandoval Mendoza1, G.L. Wojcik1, A. Moreno Estudillo1, E. Ortiz-Tello1, A.F. Adams1, C. Eng1, S. Huntsman1, V. Villanueva Dávalos1, J. Manzaneda1, F. Manzaneda2, M. Hurtado2, V. Villagómez3, E. González Burchard4, C. Gallo5, L. Enrique Lencinas5, J.C. Baker1, C.D. Bustamante1. 1) Department of Genetics, Stanford University, Stanford, CA; 2) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA; 3) Dept of Obstetrics and Gynecology, Hospital Regional “Manuel Núñez Butrón”, Puno, Peru; 4) Universidad Nacional del Altiplano, Puno, Peru; 5) Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia.

Preeclampsia is the largest killer of pregnant women worldwide. It is a disorder characterized by the sudden onset of maternal hypertension, edema, and proteinuria, with a global prevalence of 3-8% of all pregnancies. The condition in particular has elevated rates in Hispanic/Latino women. Preeclampsia also poses a risk to newborns, particularly in areas of the world where resources are limited. The heterogeneity of the disorder has posed a challenge in our understanding of the molecular basis of the preeclampsia. However, the trait is known to be exacerbated in high altitude regions such as the Andes. In order to investigate the genetic basis of preeclampsia and its relationship to altitude we have begun a collaboration with researchers and clinicians in Puno, Peru, a city nearly 4,000 meters above sea level. The local population is primarily of Quechuan and Aymaran descent. Based on population genetic analysis then we expect these groups to have inherited a unique pattern of genetic risk factors from other populations previously studied. We have collected data on almost 500 individuals in affected trios (including maternal, paternal, and cord blood) and population controls, with sections of placenta alongside pertinent medical records. The trio design allows for measuring the contributions of maternal, paternal and proband effect on the expectant mother. We have generated genome-wide data at over 800,000 sites across the genome with the Affymetrix Axiom LAT array on all samples to date. Most individuals have very high levels of Native American ancestry (median value >98% as measured by ADMIXTURE). Using trio-aware genome-wide association analyses we identify several top candidates, including a group of SNPs in LD over a cluster of blood clotting factor genes on 13q34 (PROZ, F7, F10), with a relatively high odds ratio for preeclampsia (OR 3.4, 95% CI 2.0-5.7, p<4×10⁻⁶). Although this region has not been previously identified in genome-wide studies, preeclampsia is known to cause changes in clotting levels and blood thinners can be used as preventative treatment. PROZ in particular has been extensively studied as a biomarker in multiple pregnancy disorders. This finding of a novel association in a relevant biological pathway and other genetic discoveries in this study demonstrates the increased power of relevant study design and underscores the importance of more genetic analysis in diverse, under-studied populations worldwide.

2865S
Molecular and pathological abnormalities in placentas of pregnancies complicated by preeclampsia. P.A. Ayala-Ramírez1, R. García-Robles2, M. Olaya3, J.L. Rodríguez2, T. Buitrago2, 1) Instituto de Genética Humana, Pontificia Universidad Javeriana, Bogotá, Bogotá D.C., Colombia; 2) Instituto de Investigación en Nutrición, Genética y Metabolismo, Universidad El Bosque, Bogotá D.C., Colombia; 3) Departamento de Patología, Pontificia Universidad Javeriana-Hospital Universitario San Ignacio, Bogotá D.C., Colombia.

Preeclampsia (PE) is a hypertensive disorder of pregnancy of unknown etiology with impact in maternal and fetal health. Its clinical manifestations occur in the second half of pregnancy with a world frequency of 2-8% and in developing countries occurs in 10%. In Colombia, PE causes 42% of maternal mortality. It is proposed that an alteration in placental homeostasis may be involved in the pathophysiology of this disease. The aim of this study was to compare levels of mRNA and protein of Tissue Factor (F3) and Thrombomodulin (THBD) as well as histopathological findings in placentas with PE. We studied 18 term placentas of patients with diagnosis of PE and 18 term placentas from pregnancies without any complication and healthy newborn. It conducted an evaluation by an expert blinded pathologist of the presence of thrombosis, fibrin deposits, decidual arteriopathy, necrotic villi, leukocyte stasis, hyperplasia of syncytiotrophoblast and edema. The assessment of mRNA and protein levels of F3 and THBD was performed by Real time PCR and ELISA, respectively. The results showed statistically significant differences between cases and controls for these variables: decidual arteriopathy (p=0.027), necrotic villi (p=0.001) and hyperplasia of syncytiotrophoblast (p=0.0017). In addition, it found statistically significant differences in levels of F3 mRNA (p=0.011) and protein (p=0.0001) as well as in mRNA THBD (p=0.0001) and protein (p=0.0001) levels, with an increase in case group. We found an abnormal expression of F3 and THBD with increased levels of protein and mRNA in placenta of pregnancies complicated by PE, suggesting a possible role of these molecules in the pathophysiology of this disease. Also, these results agree with the existence of alterations in hemostatic mechanisms and histopathological changes in placenta of pregnancies affected by PE.
2866S
A ‘conditional-ON’ mouse model of Fibrosyplasia Ossificans Progressiva (FOP), A. N. Economides1,2, L. Huang2, L. Xie2, K. Namuru2, K. Feeley2, T. Persaud2, P. Yang2, V. Idone2, A. Lee2, P. Yu2, C. Schoeneweiss2, S. J. Hatself, A. J. Murphy. 1) Regeneron Genetics Center, Regeneron Pharmaceuticals, Tarrytown, NY; 2) Regeneron Pharmaceuticals., Tarrytown, NY; 3) National Center for Advancing Translational Sciences, NIH, Rockville, MD; 4) Brigham and Women’s Hospital, Boston, MA.

FOP (MIM 135100) is an autosomal dominant disorder characterized by early onset, episodic and progressive ossification of skeletal muscle and associated connective tissue (Pignolo et al., 2011). FOP is driven by mutations in the intracellular domain of ACVR1 (also known as ALK2), with the great majority altering Arginine 206 to Histidine (R206H). In order to enable the development of therapeutic approaches for FOP, as well as mechanistic studies of the disease process, we have engineered a Cre-regulated ‘conditional-ON’ allele of ACVR1[R206H] in the mouse - Acvr12372loxP[Cre]. This allele places in antiparallel orientation with the wild type human exon 5. Body-wide activation of the FOP allele in Acvr12372loxP[Cre]+/+ Rosa26Sor[CreERT2] adult mice resulted in progressive ossification resembling the human disease. Heterotopic ossification was spontaneous, and did not require experimentally induced inflammation. The FOP phenotype was rational treatments for a currently untreatable disease.

2866T
Retinal vascular lesions associated with mutations in Col4a1. M. Alawi, D. Gould. Department of Ophthalmology, University of California, San Francisco, CA. Type IV collagen alpha 1 (COL4A1) is a major component of almost all basement membranes and mutations in COL4A1 cause multisystem disorders in humans and mice. Notably, patients and mice harboring mutations in COL4A1 exhibit ocular defects including anterior segment dysgenesis, optic nerve hyperplasia and retinal vascular tortuosity. Here we present our findings studying the retinal phenotype of mice carrying a dominant-negative Col4a1 mutation (Col4a1ex41Δ+/−). We examined retinas of Col4a1ex41Δ+/−; C57BL/6J;129S1/SvImJ mice and Col4a1ex41Δ/− littermate controls at different ages from postnatal day (P) 21 to P75 in vivo using fluorescein angiography (FA), fundoscopy and optical coherence tomography. Histology, immunohistochemistry and quantitative PCR analyses complemented our analyses.

FA showed fully penetrant retinal vascular tortuosity and abnormal ramification. Retinal examinations revealed serous chorioretinopathy, retinal hemorrhages, fibrosis or signs of pathogenic angiogenesis with anastomosis of the choroid and the retinal vasculature in approximately 80% of all Col4a1ex41Δ/− eyes at various ages. Retinal hemorrhages and angiogenesis were observed as early as P21. We assessed expression levels of candidate genes involved in angiogenesis and found increased vascular endothelial growth factor (VEGF) expression in Col4a1ex41Δ+/− retinas compared to Col4a1+/− retinas. Our findings suggest that patients carrying mutations in COL4A1 may be at risk for sudden vision loss resulting from retinal vascular insults. What is more, retinal exams may help to identify individuals that are at greatly increased risk of sudden retinal hemorrhages. Elevated VEGF expression may be a primary cause or consequence of vascular lesions that then leads to further retinal damage. Currently VEGF is a major target for preservation of vision in patients with age-related macular degeneration. This suggests that anti-VEGF therapy may also be an effective strategy to prevent pathogenic vascular defects that result from COL4A1 mutations.

2869S
Inflammatory demyelination in a duplication mouse model of Pelizaeus-Merzbacher Disease. G. M. Hobson1,2, L. Sakowski2,2, K. Clark2,2, K. Sperle1. 1) Nemours Biomedical Research, A I duPont Hosp Children, Wilmington, DE; 2) Dept Biological Sciences, University of Delaware, Newark, DE. Pelizaeus-Merzbacher disease (PMD [MIM 312080]) is an X-linked leuкоatrophiа affecting the myelin sheath in the central nervous system (CNS). It is caused by alterations in the proteolipid protein 1 gene (PLP1 [MIM 300041]) that encodes the most abundant protein in CNS myelin. Our lab has developed the Plp1dup mouse model that is a genetically faithful model of the most common cause of PMD, a genomic duplication. Previous characterization on 129/B6 mixed background mice has shown disorganized myelin in white matter tracts, abnormal mRNA levels, and impaired gait. Further characterization on an inbred C57Bl/6 strain using the Noldus CatWalk has shown that two major phases of the gait cycle, stand and swing, are affected. Plp1dup mice also demonstrated a decreased latency to fall in accelerating trials on the Stoylling Rotarod, indicating a progressive effect on motor coordination and balance. Light microscopy on brain sections of Plp1dup mice showed degenerating nerve fibers and abnormal myelin organization with evidence of gliosis and activated microglia, suggesting an inflammatory response. Analysis of markers of gliosis and inflammation in brain mRNA showed upregulation of several key inflammatory markers, including TNF-α, and GFAP. Global transcriptome profiling revealed significant upregulation of 45 genes, many of which are involved in an inflammatory response. Further characterization of the Plp1dup mouse model will allow us to better understand the underlying mechanisms for PMD and develop rational treatments for a currently untreatable disease.
observed. Quantitative mRNA expression of the gene Myod, Myf5 and Myog, centrally nucleated fibers were observed, continuing in d21 and d30. In SJL d5, and signals of regeneration starting at d5. After d10, large areas of cal analysis showed small areas with foci of degeneration starting at d3 and was analyzed after 3, 5, 10, 15, 21 and 30 days. In control animals, histological and multipolar spindles. The defects in cell division and cilia likely account for the disorganized proliferative zone of the epiphysis growth plate because chondrocytes are not polarized properly and undergo increased apoptosis. This evidence underlies the role of primordial dwarfism in humans with two different mutations of POCA1 (2, 3). Male infertility is caused by arrested spermatogenesis at the pachytene stage and progressive germ cell loss. SpERMATOSpermogoniatal stem cell transplantation studies reveal that POCA1 is essential for normal function of both Sertoli cells and germ cells. We hypothesize that POCA1 regulates the progression of primordial germ cells in a number of genes associated with centrosome function are now known to cause skeletal dysplasias and primordial dwarfs. This mouse model will provide information about the global functions of POCA1, clarify how centrosome dysfunctions lead to defects in skeletal growth and development, and serve as a means to uncover potential therapies. 1. Cha et al., Skeletal dysplasia and male infertility focus on mouse chromosome 9. Genomics 83:951-960, 2004. 2. Shaheen et al. POCA1 truncation mutations cause SI choledochal cysts in humans characterized by immunoreactivity. Am J Hum Gen 91:330-336, 2012. 3. Sarig et al. Short stature, onychodysplasia, facial dysmorphism, and hypotrichosis syndrome is caused by a POCA1 mutation. Am J Hum Genet 91:337-342, 2012.

Using electroporation as a model of degeneration/regeneration to investigate the regenerative potential in neuromuscular disorders (NMD). M. Vainzof, C.F. Almeida, R. Ishiba, A. Martins-Bach, A.L.F. Santos, L. Nogueira. Human Genome Research Center, University of Sao Paulo, Sao Paulo, Brazil.

After injury, the healthy muscle is able to repair and grow new fibers. However, in neuromuscular diseases, the process of regeneration is not completely efficient and its constant activation leads to exhaustion of this ability. Understanding all the steps involved in muscle degeneration and regeneration is of major importance for a better comprehension of the genetic muscle diseases. Mouse models for NMD are good molecular models, but not always reflect the histopathological and phenotypic characteristics associated to the mechanisms of the respective disease. Protocols for in vivo gene therapies have used electrotransfer to optimize transfection efficiency. However, most of them have reported histological evidence of muscle degeneration and complete regeneration within 15 days. Therefore, we are proposing the use of this model of very fast muscle degeneration and regeneration processes to characterize different patterns of muscle regeneration in NMD. We are studying electroporated lower limb of normal mice and dystrogen deficient SJL mouse model of myostatin deficiency (myostatin -/-). Dys trophy type 2B, with the main objective to study the pattern of alterations along time, both through histological analysis and quantitative mRNA expression of genes involved in the degeneration/regeneration pathways. The animals were anesthetized and received 8 pulses of 100 V, duration of 20 milliseconds and an interval of 0.5 second at the cals. The group of animals was analyzed after 3, 5, 10, 15, 21 and 30 days. In control animals, histological analysis showed small areas with foci of degeneration starting at d3 and d5, and signals of regeneration starting at d5. After d10, large areas of centrally nucleated fibers were observed. The expression of genes related to the degeneration phase (Myod, Myf5 and Myog, showing multipolar spindles) and the regeneration phase (Myg1) was observed. In the control model, the degeneration was more intense in d3 and d5, and was also observed in d10. In d21 and d30, large areas of cen tronuclearized fibers were observed. Quantitative mRNA expression of the gene Myod, Myf5 and Myog, showing multipolar spindles and the regeneration-related gene Myg1, was observed only after d10. This new system of acutely induced degeneration and regeneration in mouse models is useful to study and quantify alterations in satellite cells and regeneration impairment in some mildly affected myopathic model, in which the pattern of degeneration and the pattern of regeneration is different from the pattern observed in human diseases. FAPESP-CEPID, CNPq-INCT, CAPES-COFECUB, FINEP.

2872S

Definitive Implication of Innate Immunity in the Pathogenesis of Sclerosis. E. Gerber, J. Beckett, D. Huo, H. Dietz. 1) Institute for Genetic Medicine, Johns Hopkins University, Baltimore, MD; 2) Department of Comparative Medicine, Johns Hopkins University, Baltimore, MD; 3) Howard Hughes Medical Institute, Chevy Chase, MD.

The most common and severe form of scleroderma (systemic sclerosis; SSC) shows sporadic onset of skin and visceral fibrosis in association with autoantibody production and other markers of autoinflammation. In contrast, stiff skin syndrome (SSS) shows congenital dermal fibrosis caused by heterozygous mutations in the domain of fibrillin-1 that mediates cell-cell and cell-matrix interactions. We previously showed that knock-in mice models of SSS recapitulate dense dermal fibrosis and show all of the autoinflammatory manifestations of SSC including antinuclear and anti-topoisomerase-1 antibodies, Th2 and Th17 skewing and B cell activation in the skin in association with recruitment and activation of plasmacytoid dendritic cells (pDCs), a cell type that performs a surveillance function for pathogens and responds via IL6 and IFNγ production upon presentation of RNA or DNA to TLR7/9 receptors, respectively. All therapeutic interventions that prevent fibrosis in SSS mice, including agents that downregulate tintin expression on pDCs and TGFβ and ERK inhibitors, associate with prevention of pDC recruitment and activation. This prompted the question regarding the relative contribution of innate vs. adaptive immunity in disease pathogenesis, an unaddressed question in SSC as well, and the mechanism of cross-talk between the matrix and pDCs. Using a transgenic allele that expresses diphtheria toxin receptor specifically on pDCs, we now show that postnatal pDC depletion is sufficient to prevent both skin fibrosis and disease (S. Rappoport, no T or B cells or antibodies) had no effect on disease expression. These data unequivocally point to a predominant, if not exclusive, contribution of innate immunity. In other autoimmune diseases such as lupus and psoriasis, abnormal protein aggregates present nucleic acids to TLRs, with the naturally occurring antimicrobial peptide LL37 serving as a critical intermediate. We wondered whether the macroaggregate of fibrillin-1 might be initiating similar events in SSS (and perhaps SSC). In keeping with this hypothesis, introduction of dominant-negative TGFβRII or treatment with a synthetic TGFβ-9 blocking peptide abrogated fibrosis in SSS mice. These data offer many novel therapeutic strategies in scleroderm and perhaps other more common presentations of fibrosis.

2873M

Human-Mouse: Disease Connection, new pathway to discovery. J.T. Epigg, C.L. Smith, P. Frost, K. Stone, L.E. Corbani, J. Campbell, K.L. Forhofer, J.E. Richardson, S.M. Bello. Jackson Laboratory, Bar Harbor, ME. The Human-Mouse: Disease Connection (www.diseasemodel.org) is a translational tool allowing simultaneous access to human-mouse genomic, phenotypic, and genetic disease information. Researchers can explore phenotypes and disease relationships, identify candidate genes, and evaluate therapeutic targets and biomarkers for human diseases. Within the framework of the Human-Mouse:Disease Connection portal, web display, links are provided to supporting mouse model publications and to repositories supplying mouse resources. Three primary approaches give users flexibility to search from a human or mouse perspective, using (1) orthologous groups of indicated genes, (2) genome location(s) from either species, and (3) mouse phenotype or human disease terms. Data also can be uploaded from VCF files. Thus, exploration can begin with a single gene or set of genes, a region for QTL, multiple deletion regions, or using phenotype/disease searches, such as “Crouzon Syndrome”, “neurofibromatosis”, or “cardiomyopathy”. All search methods initially return an interactive grid that presents a visual overview of results facilitating comparison of phenotypes and diseases across multiple genes, phenotypes, and diseases. Within the grid features color cues reflecting depth of annotated human and mouse data, and grid cells are active links leading to more detailed information, including availability of mouse models from repositories worldwide. Alternate web page displays with gene and disease-focused information are one click away. Integration of mouse and human data related to gene homologs, genotypes, mutations, phenotypes, and diseases is ongoing in the Mouse Genome Informatics resource (MGI, www.informatics.jax.org), the primary data source for the Human-Mouse:Disease Connection portal. These data include over 1300 OMIM-defined human genetic diseases with at least one mouse model or mutant. Each mouse model contains unique phenotypes for more than 51,600 unique genotypes. Data (to be released in June 2014) from the systematic screens of knockout mutations analyzed in the International Mouse Phenotype Consortium (IMPC) project will further augment phenotypic spectra for mouse models and human diseases. We will discuss the challenges of human and mouse data integration and new developing capabilities of the Human-Mouse: Disease Connection portal to ease translation between human and mouse data and to promote best matching of human disease and mouse model resource data. Supported by NIH grant HG000330.
Social and maternal behaviours are affected by a mutation in Gtf2ird2 in Williams-Beuren Syndrome mouse model system. N. SHARMIN1,2, M. Tassabehji1, R. Hager1. 1) Faculty of Life Sciences, University of Manchester, Manchester, Lancashire, United Kingdom; 2) Genetic Medicine, Faculty of Medical and Human Studies, St Mary’s Hospital, University of Manchester, Manchester, Lancashire, United Kingdom.

Human social behaviour is affected by both environmental and genetic factors. The human genetic disorder Williams-Beuren Syndrome (WBS), is characterised by neurological deficits that impact social and cognitive behaviour. Our hypothesis is that deletion of the Gtf2ird2 gene, from the Gtf2i family of transcription factors, affects social and neuropsychological development and consequently affects maternal behaviour during early development. By phenotyping maternal behaviour in a single-gene knockout mouse model of WBS, we show, for the first time, that deletion of Gtf2ird2 influences key maternal and social behaviours in homozygous and heterozygous animals, affecting offspring development. In addition, Gtf2ird2 heterozygous mothers exhibit high rates of filial cannibalism compared to either knockout or wild type mothers. These results show that this single gene mutation has an important role in the hyper-social behavioural symptom seen in WBS patients. The results are in line with existing literature on Gtf2i gene family.

Canine developmental disorder maps to the critical region of human 22q11.2 deletion syndrome. M. Hyönen1,2, A. Lappalainen3, H. Lohi1,2. 1) Molecular Neurology, Research Programs Unit and Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland; 2) Folkhälsoinstitutet Institute of Genetics, Helsinki, Finland; 3) Equine and Small Animal Medicine, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland.

Dogs have emerged as clinically and genetically relevant large animal models for human inherited disorders. We present here a genetic study of a congenital canine syndrome characterized by typical craniofacial features, patellar subluxation, microphthalmia and mental disturbances. We have performed genome wide association study to map the disease to canine chromosome 26. The associated locus overlaps human DiGeorge syndrome region characterized by a 22q11.2 deletion. Our ongoing targeted resequencing efforts aim to discover the causative mutation in the critical region. Comparative phenotypic analyses reveal similarities in canine and human phenotypes and genetic characterization of the canine locus may assist in the genetic dissection and understanding of the DiGeorge syndrome in both species. This study also enables the development of a genetic test for breeding purposes in dogs.

A transgenic zebrafish model for facioscapulohumeral dystrophy. A. Lek1,2, H. Mitsuhashi1,2, F. Rahimov1, C. Mosimann1,2, L. Zon1,2, L. Kunzel1,2. 1) Genetics, Boston Children’s Hospital, Boston, MA; 2) Pediatrics and Genetics, Harvard Medical School, MA.

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most common forms of muscular dystrophy, characterized by asymmetric weakness of the facial, shoulder and upper arm muscles, accompanied by hearing loss and retinal vasculopathy. Although the causative gene for FSHD remains controversial, the primate specific retrogene, DUX4, is a leading candidate. There is currently no mammalian model that recapitulates the human FSHD pathology. Previously, our lab has shown that misexpression of very low levels of human DUX4 in zebrafish development recapitulates the phenotypes seen in FSHD patients. It was demonstrated that microinjection of human DUX4 mRNA into zebrafish eggs caused asymmetric abnormalities of the eyes and ears, and disorganization of fin and trunk muscles. Using a tamoxifen-controlled CreER12loxP system, we have now generated a transgenic DUX4 line that successfully reproduces the mosaic, low-level expression of DUX4. We show that activating DUX4 expression during development results in a degenerative muscle phenotype by day 7 post-fertilization. This stable line will enable us to control when, where and how much DUX4 is expressed to best model FSHD pathogenesis in zebrafish, allowing for better functional studies.
2879M

Rapid Approaches in Functional Validation of Candidate Disease Genes Identified from Structural Rearrangements and Next-Generation Sequencing. R. Greenlees1,2, S. Yousoo1,2, E. Semina1, R.V. Jamieson1,2, C.Y. Chow, M.F. Wolfner, A.G. Clark. Dept Molec Biol & Gen, Cornell Univ, Ithaca, NY.

Retinitis pigmentosa (RP) is a hereditary disease characterized by progressive loss of vision due to degeneration of rods and cones in the retina. Autosomal dominant retinitis pigmentosa (ADRP) makes up 30-40% of all RP cases. Dominant mutations in the rhodopsin gene (RHO) comprise 25% of all ADRP cases and represent the most common cause of RP. Dominant mutations in the Drosophila melanogaster ortholog of RHO, Rh1 (or ninE), provide an important model for dissecting the pathophysiology of ADRP. The Rh1G69D mutation in Drosophila closely resembles many human mutations in RHO. This mutation results in a misfolded Rh1 protein that is retained in the endoplasmic reticulum (ER), and induces the ER stress response, leading to apoptotic cell death and retinal degeneration. Previous studies demonstrated that mutations in genes in the ER stress and apoptosis pathways can alter the phenotypic presentation of Rh1G69D. However, these previous genetic screens relied entirely on loss-of-function (LOF) mutations. In the human population, it is unlikely that severe LOF mutations contribute appreciably to variability in ADRP phenotypes. We take advantage of natural genetic variation in Drosophila to identify dominant modifiers of Rh1G69D induced retinal degeneration. We crossed the Rh1G69D mutation into ~200 strains from the Drosophila Genetic Reference Panel (DGRP). The DGRP is a collection of wild-derived Drosophila strains that harbor genetic polymorphisms present in a natural population. To assess the effect of DGRP backgrounds in modulating the phenotypic impact of Rh1G69D, we measured eye size to quantify the extent of degeneration and apoptosis. We found that eye size varied by more than ten phenotypic standard deviations, presenting a more than two-fold difference in eye size due to background modifiers. This dramatic phenotype allowed us to perform an association study to identify natural genetic polymorphisms that enhance or suppress the Rh1G69D induced retinal degeneration. We used functional studies to validate the novel candidate genes found through this approach. The candidate polymorphisms and genes nominated from this study may more accurately identify potential human modifiers of ADRP, as they are drawn from existing natural variation. These results have important implications for identifying modifiers of the ER stress response and identifying loci that modify human ADRP.

2881S

AIPL1 mutation in Persian cats defines a new model for Leber's Congenital Amnesia. B. Gandolfi1, H. Alhaddad2, R.A. Grahn3, D.J. Maggs4, H-C. Rah5, N.C. Pedersen6, L.A. Lyons1. 1) Department of Veterinary Medicine & Surgery, University of Missouri - Columbia, Columbia, MO; 2) College of Science, Kuwait University, Safat, Kuwait; 3) Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California - Davis, Davis, CA USA; 4) Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California - Davis, Davis, CA USA; 5) College of Medicine, Chungbuk National University, Chongju, Chungbuk Province, South Korea; 6) Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California - Davis, Davis, CA USA.

Leber's Congenital Amnesia (LCA) is a rare inherited eye disease that has onset at birth or early in life. LCA Type 4 (LCA4, MIM: 604393) is caused by mutations in the gene, aryl-hydrocarbon-interacting-protein-like 1 (AIPL1) and is considered the severe end of the spectrum for LCA. Newborns with AIPL1 mutations generally have severely impaired vision or blindness, nystagmus, and an abnormal or flat electroretinogram. A variety of mutations in AIPL1 have been identified to cause LCA in humans, presenting with varying ages of onset and degree of cone - rod dystrophies. Persian cats have been demonstrated to have an autosomal recessive progressive retinal atrophy (PRA) similar to LCA. Onset of photoreceptor loss is approximately 5 weeks of age with severe loss by 16 weeks of age. The cats often have divergent strabismus but do not suffer from bilateral ectasia with central thinning of the cornea as the cats age. A series of association studies, including linkage analysis and transmission distortion tests suggested that the casual gene for Persian PRA was localized to cat chromosome E1.

Aim: To search for new genetic variants in AIPL1.

Methods: Pedigree analysis and whole genome sequencing to identify potential causal AIPL1 variants. We sequenced 63985 and 426243960 nucleotides from 175 samples from 64 unrelated Persian pedigrees. Our approach was to screen for potential causal AIPL1 variants in a diverse group of unrelated Persian pedigrees.

Results: We failed to identify a causative mutation in AIPL1. Further analyses identified a pigmentary retinopathy in a single kindred that could not be explained by AIPL1 variants. This form of retinopathy is common in Persians, with a prevalence of approximately 1.6% in the extended pedigree of Persians with PRA. Two mouse models of AIPL1 deficiency include the Aipl1-hypomorphic (h/h) mouse (with reduced Aipl1 levels and a relatively slow degeneration), and the Aipl1-null mouse (with no functional Aipl1 and a very rapid retinal degeneration). Gene therapies using the sc-Y733F-AAV2/8 viral vector have shown rescue vision loss in the null mouse model. Cats have a longer life span than mice and can be evaluated for repeated and longer term gene and stem cell therapy trials. The cat can now become an efficient and effective model for gene and stem therapies for LCA.

Poster: Molecular Basis of Mendelian Disorders
2882M

Genome-wide linkage analysis in conjunction with whole exome sequencing for identification of deafness-causing genes in multi-generational consanguineous families. R. Alkhalifa1, M. Al-Owain2, A. Khater3,4,5, Al-Masmoudi6, S. Blanton6, Q. Liu6,4,2, A. Raddawi6,4,2, XZ. Liu4,4,2, SM. Yang4,4,2, R. Mittal4,4,2, P. Dai4,4,2, K. Taibah1,2, J. Qing1,2, M. Al-Owain2,6, Y. Feng4,4,2, D. Yan4,4,2, S. Masmoudi6, S. Blanton6, QZ. Liu4,4,2, 1) Departments of Otologyngology-Head and Neck Surgery, University of Miami Miller School of Medicine, Miami, Florida 33136, USA; 2) Department of Ophthalmology, John T. Macdonald Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL 33136, USA; 3) Department of Otologyngology-Head and Neck Surgery, Institute of Otology, the Second Xiangya Hospital, Central South University, Changsha, Hunan, 410011, China; 4) Department of Otologyngology, Head and Neck Surgery, Chinese PLA, Beijing, 100853, China; 5) Department of Otologyngology, Xiangya Hospital, Central South University, Changsha, Hunan, 410008, China; 6) Microorganisms and Biomolecules Laboratory, Centre of Biotechnology of Sfax, Sfax University, Sfax, Tunisia.

Hearing loss (HL) is the most common sensory defect, affecting 1 in every 500 newborns around the world which half of it is due to genetic causes. Nonsyndromic HL (NSHL) accounts for more than 70% of all hereditary HL. Gene isolation for NSHL has been successful in isolated populations and consanguous families. However, even after the gene has been localized to a region on a chromosome, the process of positional cloning by Sanger sequencing to analyze each selected gene in the linked interval, is costly and time consuming and can be impractical, in the case of large candidate intervals. Now whole exome sequencing (WES) offers a valuable tool for the many challenges in the identification of mutations in heterogenic diseases such as deafness. In this study we report the genome-wide linkage in conjunction with WES analyses of a collection of 14 multi-generational families. We identified 3 genomic regions spanning 10 Mb each within consanguineous Saudi families. None of these regions harbor known deafness genes. In a collection of 4 consanguineous Tunisian families with autosomal recessive NSHL for finding deafness causative genes. The SureSelect human all exom 50Mb kit (Agilent Technologies) and the Hiq2000 instrument (Illumina) were used for WES. The Genome Analysis Application (GEMapp) was applied for data filtering. Computational functional prediction algorithms and conservation scores were also evaluated to possibly impact on protein function of the detected variations. By linkage analysis, we determined the chromosomal location of the disease to prioritize candidate genes for mutation screening. On average, 97% of RefSeq exons were captured and approximately, each exon in the 14 families had 97%, 85% and 70% of mappable bases of the Genome defined exome represented by coverage of at least 2, 10 and 20 reads, respectively. By our filtering strategy under the autosomal dominant or recessive model, we have rapidly identified novel genes with homozgyous or heterozygous missense variants common to affected siblings subjected to WES. Co-segregation of the variants with the disease phenotype is validated using Sanger sequencing in each family and in vitro studies to establish the pathogenetic nature of the variants are being performed. Our study shows that combination of genome wide linkage analysis with WES is a powerful strategy for identification of causative mutations in genetically heterogeneous condition such as deafness.

2882S

Homozygosity mapping of families with autosomal recessive intellectual disability and examination of WWOX, GFRA3, and PTBP1 genes. A. Alkhalifa1,2, S. Aburahma3, W. Habbab1, 1) Qatar Biomedical Research Institute, Doha, Qatar; 2) Jordan University of Science and Technology, Irbid, Jordan.

Intellectual disability is a relatively common disorder affecting all populations. In Arab countries, where consanguinity has a high prevalence, there is an increase in autosomal recessive intellectual disability. We tried to identify the genetic causes of intellectual disability in four consanguineous local Arab families. Families were recruited in the study after giving informed consent and obtaining institutional review board approval. All families had multiple affecteds, at least one of the affecteds was a female. Samples were genotyped by high-density HumanOmnExpress (700K, Illumina Inc). Genotypes were analyzed by HomozygosityMapper and multiple regions of homozygosity were found in the four families. No copy number variants were found. A single overlap between two families harbored WWOX (WW domain-containing oxidoreductase) gene. Wox was found previously to be mutated in mice with seizures and epilepsy and the two families had seizures and epilepsy as part of their phenotype. Sequencing WWOX gene detected three common SNPs with no indication of pathogenicity. In the third family, two other candidate genes (GFRA3 and PTBP1) localized within the homozygosity intervals were sequenced; no variants were found in the coding region. In conclusion, we recruited four local consanguineous families with multiple affecteds of intellectual disability. Those families showed different homozygosity intervals illustrating the heterogeneity of the phenotype. We examined the candidacy of WWOX, GFRA3, and PTBP1 genes within the homozygosity intervals but no pathologic variants were found. We are now in the process of doing whole exome sequencing for one patient of each family.

2885M

Santos syndrome is caused by homozygous mutation in WNT7A. L.U. Alves1, S. Santos2, R.S. Thiele-Aguiar3, F.A. Otto4, R.C. Mingroni-Netto5, 1) Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil; 2) Departamento de Biologia, Universidade Estadual da Paraíba, Campina Grande, PB, Brazil.

Santos et al. (Am J Med Genet 2008) described a new syndrome in six relatives living in a remote area in Northeastern Brazil. This syndrome is characterized by fibular agenesis/hypoplasia, hypoplastic femora and grossly malformed/deformed clubfeet with severe oligodactyly, upper limbs with severe ungual hypoplasia/anonychia sometimes associated with mild brachactyly and occasionally pre-axial polydactyly. This syndrome, named as “Santos syndrome” (OMIM 613005), was classified as a distinct syndrome from other previously described conditions exhibiting fibular agenesis/hypoplasia. An autosomal dominant inheritance model with incomplete penetrance was postulated to be responsible by Santos et al. (2008), but autosomal recessive inheritance was not discharged. Linkage analysis was performed based on results of genotyping by SNP-array (500 K - Illumina/Agilent) in samples from the six affected individuals. After multiple point LOD score calculation, two candidate regions were found to be statistically significant under the dominant (k=0.324) and recessive model hypothesis, in 3p26.1-p25.2 and in 3p13-q12.3. The maximum values of LOD score calculated under recessive model hypothesis were 2.856 and 3.235 for these two regions, respectively. Under the dominant model, positive LOD scores were observed in the same regions, but with smaller values than in the recessive model. The best candidate gene in the mapped regions was WNT7A, in which homozygous mutations had already been associated with two other limb defect syndromes (Fuhrmann syndrome, #228830 and AARRS, #276820). Sequencing of WNT7A revealed a novel homozygous c.934G>A (p.Gly312Ser) mutation in five of six affected individuals; the remaining affected individual in homozygous as to the mutation, and his genotype is consistent with such less severe than his relatives. The glycine 312 at WNT7A protein is highly conserved (UniProt) and the mutation was predicted to be probably damaging with score of 1.0 by PolyPhen-2. Thus, Santos syndrome can be explained by mutation in the WNT7A gene. We hypothesize that the affected heterozygous individual either has a different condition or his phenotype results from a mild clinical manifestation of Santos syndrome.
2886T
Mutation Screening in PRPF31 in an Autosomal Dominant Retinitis Pigmentosa (ADRP) Family with Incomplete Penetrance. S. Bhattacharya, S. Goyal, I.R. Singh, V. Vanita. 1) Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India; 2) Dr. Daljit Singh Eye Hospital, Amritsar.

Purpose: The objective of present study was to perform mutation screening in PRPF31 in a four generation ADRP family with incomplete penetrance.

Material and Methods: Present study involved a large four generation ADRP family with incomplete penetrance which was diagnosed and collected at the Dr. Daljit Singh Eye Hospital, Amritsar with six affected individuals and few asymptomatic carriers. Asymptomatic carrier passed the disease to next generations. Age of onset of the disease in this family was by birth. Ophthalmic examinations included visual acuity testing, intraocular pressure, fundus testing along with fundus photography and Optical Coherence Tomography (OCT) after pupil dilation and Electroretinography (ERG) testing.

We undertook mutation screening in the PRPF31 gene that is well reported to be linked with ADRP with incomplete penetrance. Bi-directional sequence analysis of amplified products of all the 14 PRPF31 exonic regions including splice sites, was performed in this family.

Results: Mutation screening revealed identification of a novel missense mutation in one of the exonic region of PRPF31 in this family. The identified missense mutation segregated completely with the disease in all the six affected members of this family as well as carriers of the disease. However, 19 tested unaffected members of this family and 100 control samples from the same population didn't carry the identified PRPF31 substitution, excluding its possibility as a polymorphism.

Conclusions: Present study identifies a previously unreported mutation in PRPF31. These findings further demonstrate the role of PRPF31 in relation to ADRP with incomplete penetrance.

2887S
Molecular Genetic Analysis in an Autosomal Recessive Retinitis Pigmentosa Family of Indian origin. S. Goyal, V. Vanita. Department of Human Genetics, Guru Nanak Dev University, Amritsar, Punjab, India.

Purpose: Present study aimed to identify underlying genetic defect in a nonsyndromal retina pigmentosa family (ARRP) of Indian origin by whole genome homozygosity mapping.

Material and Methods: It was a large 4 generation arrP family with 3 affected individuals and their parents being first cousins. Age of onset of the disease varied from 2 to 17 years. Ophthalmic examinations that included visual acuity, intra ocular pressure testing, electroretinography (ERG), ocular coherence tomography (OCT), fundus testing and fundus photography after pupil dilation, confirmed status of 3 individuals being affected by non-syndromic typical RP. Whole genome homozygosity mapping was conducted on 7 members (3 affected and 4 unaffected) using 8-sample HumanOmniExome Beadchip, consisting of >2,40,000 functional exonic markers. Genotyping data was subsequently used to identify the identical homozygous regions present in affected individuals using the online homozygosity mapping tool Homozygosity Mapper. Further, mutation screening in candidate genes at mapped homozygous region was done by sequencing of exonic regions of the amplified products of the affected family members using DNA sequencing. Sequence variations were verified and analyzed. Results: The homozygous region at 14q31.3 was identified that harbored the SPATA7 and Tetracopeptide domain 8 gene (TTC8). Sequencing of all the coding regions of SPATA7 indicated c.913-23T>G an already known SNP in all the 3 affected members as well as unaffected members of the family. However, mutation screening in all the 15 coding exonic regions including splice donor and acceptor sites of TTC8 revealed a novel missense substitution. This novel substitution segregated in all the three tested affected members in homozygous form whereas unaffected members of this family were either heterozygous or homozygous for the wild-type allele. Further, 50 tested control samples (100 chromosomes) from the same population did not carry the identified substitution, excluding its possibility as a polymorphism.

Conclusion: We identified a novel disease-linked missense mutation in TTC8 in an arrP family of Indian origin with nonsyndromic RP. TTC8 is previously reported to be linked with Bardet-Biedl syndrome (BBS) as well as with nonsyndromic RP. Our findings thus further expand the mutation spectrum of TTC8 and inherent genetic and phenotypic heterogeneity for arrP.

2888M

Achromatopsia is a rare autosomal recessive disorder that leads to color blindness, and can include photophobia, nystagmus, cataracts, reduced visual acuity and eccentric fixation. Achromatopsia can result from mutations in any of four genes: CNGB3, CNGA3, GNAT2, and PDE6C. The purpose of our study was to discover the genetic causes in consanguineous Pakistani families with achromatopsia. In four multiply affected families, where affected individuals had complete color blindness, reduced vision, photophobia and nystagmus, we first tested linkage to each of the four known genes, then fully sequenced any linked gene by Sanger sequencing. The first family evaluated includes 7 individuals with achromatopsia and 11 unaffected sibs from 3 generations. Parents of the affected individuals are members of the same extended kindred. CNGA3, which codes for the cyclic nucleotide-gated channel alpha-3 cone photoreceptor, is most likely responsible for achromatopsia in this family. Markers flanking CNGA3 were homozygous in affected family members. Sequencing of all 8 coding exons of CNGA3 revealed an in-frame 3 bp deletion at chr2:99,012,573-99,012,576 delATC, resulting in CNGA3 c.1357delATC (p.Ile312del). Deletion of isoleucine at residue 312 is predicted by bioinformatics tools to perturb the transmembrane structure of this photoreceptor. Genotyping this mutation in the four families revealed that all 17 affected family members were homozygous for the deletion and all 25 unaffected family members tested were either heterozygous or wildtype at this site. Multiple mutations in CNGA3, including single amino acid deletions, have been identified in persons with achromatopsia in the Middle East and elsewhere.

2889T
Mutations in ALDH1A3, FOXE3 and VSX2 cause ocular abnormalities in consanguineous Pakistani families. E. Ullah1,2, MA. Sajid1, N. Shaht2, S. Sajid1, R. Lao3, E. Wang1, P. Kwok3, M. Ansar1, MA. Saqib1,2, I.R. Singh2, Y. Ashhab2. 1) Department of Pediatrics, University of California San Francisco, San Francisco, CA; 2) Department of Biochemistry, Quaid-i-Azam University, Islamabad, Pakistan; 3) Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA.

Achromatopsia/microphthalmia (A/M) is a genetically heterogeneous birth defect for which the etiology remains poorly understood, with less than 50% patients receiving a molecular genetic diagnosis. We hypothesized that exome sequencing in A/M patients would be an efficient method to provide data to find novel candidate genes. We have collected 6 families with A/M but without any other phenotypic abnormalities from Pakistan. All families had multiple affected members and evidence of consanguinity, suggestive of autosomal recessive (AR) inheritance. We therefore performed Sanger sequencing to exclude mutations in known A/M genes with AR inheritance (ALDH1A3, RX, VSX2, FOXE3) and two genes with autosomal dominant (AD) inheritance (SOX2, OTX2). In the first family with non-syndromic bilateral anophthalmia, we found a nonsense mutation c.598C>T, p.Arg200* in VSX2, that has been published in a family with bilateral anophthalmia in the medical literature. In a second family with non-syndromic bilateral anophthalma, we found a nonsense mutation c.598C>T, p.Arg200* in VSX2, that has been published in a family with bilateral anophthalma in the medical literature. In a third family with bilateral microphthalmia, corneal opacity and anterior segment dysgenesis, we found a novel missense mutation c.289A>G, p.Ile97Val in FOXE3. This mutation is predicted as disease causing by SIFT, PolyPhen-2 and Mutation Taster. In a fourth family with bilateral microphthalmia, corneal opacity and anterior segment dysgenesis, we found a frameshift mutation c.21_24delGGAT, p.Me7Tfs*216 in FOXE3. This mutation has previously been reported in a Pakistani family with bilateral microphthalmia, sclerocornea, corneal opacity, buphthalmos and primary aphakia. The remaining two families were negative for the aforementioned disease genes and therefore were subjected to whole exome sequencing. We selected several autosomal recessive sequence variants as potential candidates that were predicted to be deleterious after exome data analysis by Inference of Variant Analysis. Although none of the candidate sequence variants segregated with the disease phenotype, we are currently analyzing whole exome data of both families to identify disease causing variant.
Two novel mutations in \textit{ABCG5} and \textit{ABCG8} genes in a Mexican family with sitosterolemia. A. Colima\textsuperscript{1,2}, E. Wong\textsuperscript{1,2}, M. Magnana\textsuperscript{1}, Laboratorio de Bioquímica I, Division de Genética, CIBO, IMSS, Guadalajara, JA, Mexico; 2) Doctorado en Genética, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara; 3) Hospital General de Tepic; 4) Universidad Autónoma de Nayarit, Mexico.

Sitosterolemia is an autosomal recessive inherited disorder characterized by enhanced intestinal absorption of vegetal sterols and cholesterol and a slow removal of them through the liver. Clinical features include xanthomas, arthrogryposis, anhidrotic, thrombocytopenia, accelerations of atherosclerosis and premature coronary heart disease. Sitosterolemia is caused by mutations in \textit{ABCG5} and \textit{ABCG8} genes. We studied a 12 years old girl from Nayarit, Mexico with biochemical and clinical characteristics of sitosterolemia. She showed slightly high total cholesterol values (295 mg/\textit{dl}), levels extremely high of sitosterol 681 µg/ml (normal range: 2.12±0.20 µg/ml) and campesterol 324 µg/ml (normal range: 3.09±1.65 µg/ml). She also had soft lumps on elbows and knees, thickening of the Achilles tendon and thrombocytopenia. Her treatment consisted of pravastatin (40 mg) and cholecalciferol. One of the two patient’s brothers had thrombocytopenia and high values of sitosterol (192 µg/ml) and campesterol (125 µg/ml). DNA was obtained with the CTAB-DTAB method for a molecular analysis. The primer pairs to PCR were design using the Oligo 6 software. Mutational screening of the \textit{ABCG5} and \textit{ABCG8} genes was performed by direct sequencing using an ABI PRISM310 genetic analyzer. In \textit{ABCG5} gene a detection was a heterozygous mutation c.1523 delC, rendering an amino acid change H510T and L511X, resulting in a truncated protein. This mutation was not present in the mother and in the sister. In addition, deletion of exon 2 of the \textit{ABCG8} gene in the index case and in her mildly affected brother was detected, in homozygous state. We found no information in the literature about the two mutations observed so that they could be considered as new mutations related to sitosterolemia.

Identification of hemizygous loss-of-function mutations in \textit{OFD1} in two unrelated male patients with a clinical phenotype of primary ciliary dyskinesia (PCD). W.B. Hannah\textsuperscript{1} and A.S. Rali\textsuperscript{2,3}. Laboratories of Molecular Genetics, New Haven, CT, USA.

Primary ciliary dyskinesia (PCD) is a rare, inherited and usually recessive disorder characterized by abnormalities of motility of the primary cilium. PCD presents with recurrent respiratory tract infections, macrocephaly, intellectual disability, and ciliary dyskinesia. Several mutations in \textit{OFD1} have been reported in male subjects presenting with recurrent respiratory tract infections, macrocephaly, intellectual disability, and ciliary dyskinesia. Whole exome sequencing revealed a hemizygous \textit{OFD1} deletion (c.3236+1delG) resulting in a frameshift followed by a premature stop codon.[p.Lys1079fs*28] This mutation has been associated with Hyperostosis corticalis OMIM 144750; Osteopetrosis, autosomal dominant 1 OMIM 607834; Osteopetrosis-pseudo-glioma syndrome OMIM 259770; Osteosclerosis OMIM144750; von Buchen disease, type 2 OMIM 607836; Bone mineral density variability 1 OMIM 601884; Osteoporosis OMIM 166710. Only missense mutations and splice site substitutions in \textit{LRRP5} have been associated with autosomal dominant and recessive PCD. This is the first report of an autosomal recessive \textit{LRRP5} splice-site deletion mutation causing a syndrome of FEVR.

A Mutation in \textit{SORBS2} Actin filament Adapter, Cell Adhesion, Migration & Intracellular Signaling Protein Causes Autosomal Recessive Hand & Foot Malformation Syndrome. H. El-Shanti\textsuperscript{1,3} and Y. Al-Sarraj\textsuperscript{1,3}. 1) Department of Pathology & Laboratory Medicine, University of Washington, Seattle, WA; 2) Department of Pediatrics, University of Iowa, Iowa City, IA, USA; 3) Yale University School of Medicine, Genetics Institute, New Haven, CT, USA.

A consanguineous Saudi Arabian family with two female siblings affected by an autosomal recessive condition resembling Familial Exudative Vitreoretinopathy [FEVR], but also with short stature, bone fragility with thin and wasted appearance was studied by homozygosity mapping and positional candidate gene screening to identify the offending gene and mutation. The gene was mapped to three possible homoyzogous genomic regions [2q, 4q, 11q], as the family structure did not allow identification of a single interval with a significant LOD score. Mutations in three genes (FZD4, TSPAN12, and NDF and LRP5) have been associated with FEVR. The LRP5 gene localizes within the 11q13.2 homozygosity interval in this family rendering it the positional candidate of choice. Screening by Sanger sequencing identified a novel homozygous one-base splice-site deletion mutation c.3236+1 delG in exon 14. LRRP5 is a low-density lipoprotein receptor (LDLR) a transmembrane protein that binds and internalizes lipids in the process of receptor-mediated endocytosis. The cDNA encodes a 1.615-amino acid protein containing conserved modules including a putative signal peptide, four epidermal growth factor (EGF) repeats with associated spacer domains, three LDL repeats, a single transmembrane-spanning domain, and a cytoplasmic domain. The extracellular domain contains 6 potential N-linked glycosylation sites. LRP5 has a unique organization of EGF and LDL repeats compared to other members of the LDLR family. Mutations in the 4q locus have been associated with Hyperostosis corticalis OMIM 144750; Osteopetrosis, autosomal dominant 1 OMIM 607834; Osteopetrosis-pseudo-glioma syndrome OMIM 259770; Osteosclerosis OMIM144750; von Buchen disease, type 2 OMIM 607836; Bone mineral density variability 1 OMIM 601884; Osteoporosis OMIM 166710. Only missense mutations and splice site substitutions in \textit{LRRP5} have been associated with autosomal dominant and recessive FEVR. This is the first report of an autosomal recessive \textit{LRRP5} splice-site deletion mutation causing a syndromic form of FEVR.
A Mutation in MYO1A Causes Autosomal Recessive Autism Spectrum Disease. M. Kambouris1*, V. Ilwyn1*, Y. Al-Sarraj1, H. Shaath1, F. Alshabani1, M. Tolefaro1, V. Chini2, H. El-Shanti2*. 1) Qatar Biomedical Research Institute, Medical Genetics Center, Doha, Qatar; 2) Yale University School of Medicine, Genetics, New Haven, CT, USA; 3) Carnegie Mellon University-Qatar, Doha, Qatar; 4) Shafallah Center for Children with Special Needs; 5) University of Iowa, Pediatrics, Iowa City, IA, USA.

A consanguineous family of Pakistani ethnicity with two female siblings (22 and 19 years of age), affected by a possibly novel autosomal recessive disorder, was studied by homozygosity mapping and whole Exome Next Generation Sequencing [NGS] of the two affected siblings and one parent to identify the responsible gene and mutation. The disorder is marked by intellectual disability, speech and motor delay, congenital malformations and possibly autism spectrum disorder (ASD). The malformations include microcephaly, microphthalmia, micrognathia and arachnodactyly with hyperextensibility and persistent fetal pads in fingers and toes. The offending gene was mapped to five possible homozygous genomic regions ([6q, 12q, 17p, 20p, 22q], as the family structure did not allow identification of a single interval with a significant LOD score. Comparative analyses of the NGS data for autosomal recessive inheritance and data mining for damaging variants within the homozygosity intervals identified a damaging homozygous c.C1675T / p.R559C mutation in the MYO1A gene, at 12q13.3. The mutation co-segregates with the disease phenotype within the family, is absent in known polymorphism databases and in 400 ethnically matched control chromosomes. Myosins are molecular motors that, upon interaction with actin filaments, utilize energy from ATP hydrolysis to generate mechanical force. The N-terminal motor domain contains both ATP-binding and actin-binding sequences. Following the motor domain is a light-chain-binding ‘neck’ region containing 1-6 copies of a repeat element, the IQ motif that serves in dimerization, membrane binding, protein binding, and/or enzymatic activities and targets each myosin to its particular subcellular location. Heterozygous mutations in MYO1A have been found in patients with sensorineural hearing loss, speculated to cause autosomal dominant sensorineural hearing loss. One of the family members who underwent exome sequencing and a novel gene which segregated with the RD phenotype was identified on chromosome 1q. The variant CDH3. One of the families for which the autosomal recessive retinal dystrophy was ascertained from different regions of Pakistan. DNA samples from these families were genotyped using 500,000 SNP markers which were analyzed using homozygosity mapping and linkage analysis. Two families are linked to regions containing known autosomal recessive retinal dystrophy genes; FAM161A and RDH12; one family mapped to a region with the autosomal dominant retinal dystrophy gene, PRCD; two families map to regions with syndromic genes e.g. CDH3 which include an eye phenotype and two families mapped to novel regions. For the FAM161A gene a novel variant c.782delA (p.Asp261ValfsX39) was found. No potentially pathogenic variants were identified in either PRCD or CDH3. We followed up interesting linkage peaks through haplotype analysis and studying potential candidate genes in the regions of interest. We also looked for indications of shared founder risk haplotypes between the different families.

PFAPA has previously been considered mainly a sporadic disease, but we have identified six families from the Oulu region of Finland with 2-4 confirmed cases and available DNA samples. We performed a genome-wide linkage analysis of PFAPA in these families, assuming a highly penetrant recessive model and a very rare mutation. We followed up interesting linkage peaks through haplotype analysis and studying potential candidate genes in the regions of interest. We also looked for indications of shared founder risk haplotypes between the different families.

We found a linkage peak with a parametric lodscore of 2.811 (using all families), contributed to mainly by three of the families. This region contains a plausible candidate gene for PFAPA, involved in innate immunity, and we are currently searching this gene for possible causative mutations. We find no indications that previously described possible PFAPA genes might contribute to the disease in the Finnish population.
Haploinsufficiency of a novel gene on 3p26.1, SMDD1, cause autosomal-dominant dentin dysplasia type I.

The hereditary dentin defects, dentinogenesis imperfecta (DGI) and dentin dysplasia (DDI) comprise a group of autosomal-dominant genetic conditions and the molecular basis of such dental disorders in all sub-groups except for DD Type I (DDI) is linked to mutations in dentin sialophospho-protein gene (DSPP). DDI shows an exclusive character in phenotypes featuring late-onset missing teeth with short dental roots and unknown genetic etiology, thus providing a useful model to unravel the mechanism for dentin formation involved in tooth development. Here we study the gene mapping and the molecular pathogenic mechanism of DDI. Using a cohort of a large Chinese family with 20 normal members and 14 DDI patients, we mapped the gene locus responsible for DDI to 3p26.2-3p24.3 by combining use of whole genome-wide SNP array and STR linkage analysis. We further identified a novel missense mutation, c.353 C>A (P118Q) in SMDD1 gene on 3p26.1 through targeted sequencing of 125 candidate genes by capture-based next-generation sequencing, and followed by Sanger sequencing and co-segregation analysis for confirmation. We showed that the mutant P118Q in dental pulp stem cells expressed at 50% of wild-type levels with exhibiting haploinsufficiency. In vivo zebrafish functional assay of a homology of human SMDD1, we determined that missing teeth similar to that of the index case phenotypes were replicated in the RNAi knock-down zebrafish and could be partially rescued by injection of normal human SMDD1 mRNA to the mutant zebrafish. We also observed that SMDD1 depletion in the zebrafish negatively regulates the expression of two major genes (bmp2 and pitx2) involved in odontogenesis. In addition, we generated the P118Q mutant knock-in transgenic (TG) mice and investigated the role of the P118Q mutation in heterozygous and homozygous TG mice. We found that dentin defects were radiographically evident in all teeth with homoygotes, to show incomplete obliteration of the nonmineralized pulp in TG mice, consistent with clinical characteristics in those patients with DDI. Moreover, high-resolution radiography, micro-computed tomography and scanning electron microscopy revealed a reduced zone of mineralized dentin with anomalies in the number and organization of dentinal tubules in TG mice. Our observations demonstrate that haploinsufficiency of SMDD1 disrupts dental formation and that this novel gene, together with other odontogenesis genes are involved in tooth development.

Mutations in CCNO identified in patients with a clinical phenotype consistent with primary ciliary dyskinesia (PCD) and defective mucociliary clearance reflecting reduced motile cilia generation. We identified two novel compound heterozygous mutations in CCNO in a large cohort of patients with PCD. These include a frameshift mutation ([c.875_897del23 [p.Asp292_Val]) and a previously known homozygous frameshift mutation (c.248_252dupTGCCC [p.Gly85Cysfs*11]). An affected sib-pair (UNC-136) harbored a previously known compound frameshift mutation (c.248_252dupTGCCC [p.Gly85Cysfs*11]) and an affected identical twin-pair (UNC-468) harbored previously known compound heterozygous mutations in CCNO. In addition, a novel homozygous frameshift mutation ([c.258_262dupGGCCC [p.Gln88Argfs*8]) mutations. To conclude, mutations in CCNO appear to be a common cause of clinical phenotype indistinguishable from PCD, reflecting RGMC. Large scale screening will be carried out in subjects suspected to suffer from PCD without a known genetic diagnosis to decipher the prevalence of CCNO in a large cohort. This abstract was funded by 5U54HL096458-06 (NIH/ORDR/NHLBI), 5R01HL071798 (NIH/NHLBI), UL1 TR000083 (NIH/NCATS), BESC-TICILIA (EU), and SYSCLILIA (EU).
identification of new genes and pathways for rare infantile forms of spinal muscular atrophy and neuromuscular disorders.

2899S

C.D. Balak 1, 2, J. Kieler 1, M.E. Ahearn 1, G. Lambert 1, D. Duggan 1, B. Winth 1, W. Temb 1, C. Legendre 1, W. Liang 1, L. Cuyugan 1, J. McDonald 2, J. Adkins 3, A. Kudoul 1, J. Corneveaux 1, M. Russell 1, M. Huettelman 2, D. Craig 1, J. Carpent 1, S.M. Bernes 1, J. Hall 1, L. Baumbach-Reardon 1, 2

1 Integrated Cancer Genomics, TGen, Phoenix, AZ; 2 Pharmaceutical Genomics, TGen, Phoenix, AZ; 3 Department of Pediatrics and Medical Genetics, University of British Columbia, Vancouver, Canada.

Spinal muscular atrophy (SMA) (MIM # 600354) is the leading genetic cause of infantile death. Mutations in the SMN1 gene account for a large percentage of SMA. However, numerous cases with SMA or related disorders do not have mutations in SMN1. Our goal is to use next-generation sequencing to identify the genetic cause of undiagnosed cases of motor neuron disease and muscular dystrophies in infants and children. Our journey began many years ago with the study of a rare form of X-linked (XL) form of SMA. In 2008 Dr. Lisa Baumbach-Reardon’s group discovered mutations in UBA1 as the cause of XL-SMA (MIM # 301850). Since that discovery, we have developed a custom sequencing panel for all coding regions of UBA1. Using Ion-torrent sequencing we have screened 10 affected cases and 18 relatives in our cohort. As expected, UBA1 sequencing revealed common variants in each individual, but surprisingly no new disease associated variants were identified. The molecular etiology of disease in these cases has remained undiagnosed for many years. At TGen, we have now been able to perform whole exome sequencing on these individuals as well as numerous other cases. We are excited to report that in many of these cases we have identified novel probable pathogenic mutations in genes known to cause disease as well as in genes not previously associated with disease. We highlight results from four cases. First, a family with two affected boys revealed novel compound heterozygous deletions in CHRND gene, a novel gene identified a novel frameshift mutation in exon 2 (c.361delC). Family 2 contains a known gene for LCA: CRB1 (MIM 604210). Screening of this gene identified a novel frameshift mutation in exon 2 (c.361delC). Family 3: a 4 generation family with 7 individuals affected by BP. The largest ROH was on chromosome 15q23-q25.1. The segregating variant rs28937873 (c.932G>A; p.R311Q) in NR2E3 (MIM 604485) has been previously reported in a Jewish family with RP. GeneSearch is a unique opportunity to diagnose rare genetic disorders within the Iranian community. The initial aim of this work was diagnostic and will benefit families and the community for genetic and carrier testing. This study provides the first example of this work. Genetic diagnosis is a crucial component of treatment and counselling for heterogeneous disorders such as these.
DUX4 induces FRG2 expression by directly activating its promoter in facioscapulohumeral muscular dystrophy. P.E. Thijssen1, J. Balog1, Z. Yao3, T.P. Kolb1, R. Tawil5, J.S. Tapscott2, S.M. Van der Maarel1, 1 Department of Human Genetics, Leiden University Medical Center, Leiden, Netherlands; 2 Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America; 3 Neuro-muscular Disease Unit, Department of Neurology, University of Rochester Medical Center, Rochester, New York, United States of America.

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common muscular dystrophy and the major form is associated with a genetic contraction of the polymorphic D4Z4 macrosatellite repeat on chromosome 4q35. Several candidate genes have been proposed as pathogenic entities in FSHD, including FRG2 located proximal to the D4Z4 macrosatellite repeat. For long it has been known that expression of FRG2 is induced specifically in differentiating myoblast cultures derived from FSHD patients, which was explained by a position effect emanating from the contracted D4Z4 repeat array. However, a contribution of FRG2 upregulation to FSHD pathology was never shown and the function of FRG2 remains unclear at this time. More recently, sporadic activation of the DUX4 retrogene, encoded within the D4Z4 unit, has emerged as the most prominent disease model. DUX4 is a double homeobox transcription factor and its expression has been shown to be toxic in myogenic cultures. We now show that FRG2 expression in FSHD derived differentiated myoblasts is a direct consequence of DUX4 protein activity. Overexpression of DUX4 in proliferating myoblasts and primary fibroblasts results in a steep upregulation of FRG2 transcription. Moreover, we identified DUX4 binding at the FRG2 promoter by chromatin immunoprecipitation followed by deep sequencing. Next, we confirmed the transcriptional efficacy of DUX4 on the DUX4 promoter by luciferase assays. Activation of luciferase was dependent on DUX4 expression and the presence of the DUX4 consensus binding sites identified in the FRG2 promoter. In conclusion we show that the FSHD specific upregulation of FRG2 can be explained by the protein activity of DUX4, confirming its central role in FSHD pathogenesis.

Expansion of the spectrum of nuclear envelopopathies: mutation in TOR1AIP1 associated with muscular dystrophy. P. Dincer1, G. Kayman-Kurekci1, B. Talim1, P. Korkusuz1, N. Sayar1, T. Sarioglu2, I. Oncel1, P. Sharafi1, H. Gundesli1, B. Balci-Hayta1, N. Purali1, P. Serdaroglu-Oflazer2, H. Topaloglu1, 1 Hacettepe University, Ankara, Turkey, 2 Bilkent University, Ankara, Turkey.

A consanguineous family with three individuals affected by a myopathic phenotype with joint contractures, proximal and distal weakness and atrophy with cardiomyopathy and respiratory involvement was analyzed by genome-wide homozygosity mapping using 250K NspI array. A single homozygous haplotype shared by the three affected individuals was detected. Homozygous c.186delG mutation in torsin A-interacting protein 1 (TOR1AIP1) gene encoding lamina-associated polypeptide 1B (LAP1B) was shown to cause a frameshift resulting in a premature stop codon (p.E62fsTer25). TOR1AIP1 mRNA level in the patient skeletal muscle was 5.88-fold lower than in the control sample. Expression of LAP1B protein was absent in the patient skeletal muscle fibres. Ultrastructural examination showed an intact sarcoplasmic reticulum. Inter-calibrarion showed an intact sarcoplasmic reticulum. Alterations of the nuclear envelope including nuclear fragmentation and degeneration, and altered chromatin condensation. LAP1B is an integral protein of the inner nuclear membrane that binds to A-type and B-type lamins, and is involved in the regulation of torsin A ATPase. Moreover, overexpression of luminal domain-like LAP1 (LULL1), the endoplasmic reticulum-localized regulatory partner of torsin A was overexpressed in the patient’s muscle. This suggests a compensatory effect between LAP1 and LULL1. The absence of LAP1B in muscle might influence the structural and mechanical stability of the nuclear envelope due to the impaired binding to the nuclear lamina. The muscle-restricted phenotype underlies a critical role for LAP1B in striated muscle and this study expands the spectrum of nuclear envelopopathy causing genes.
290SS

A missense mutation in hexokinase 1 (HK1) causes autosomal dominant retinitis pigmentosa (adRP). S.P. Daiger1,2,7, S.J. Bowne1, L.S. Sullivan1, D.C. Koboldt4, D.G. Birch3, D.K. Wheaton2, S.H. Blanton5, R.K. Koenekoop3, C.F. Chakarova6, R.S. Fulton7, R.K. Wilson2, G.M. Weinstock2, C.A. Garcia1, C.E. Avery1, E.D. Cadena1, R.A. Lewis6, 1) Human Genetics Center, Univ. of Texas Health Science Center, Houston, TX; 2) The Genome Institute, Washington Univ., St. Louis, MO; 3) The Retina Foundation of the Southwest, Dallas, TX; 4) Hussman Institute of Human Genomics, Univ. of Miami, Miami, FL; 5) McGill Ocular Genetics Laboratory, Depts of Paediatric Surgery, Human Genetics and Ophthalmology, McGill Univ. Health Center, Montreal, Quebec, Canada; 6) Institute of Ophthalmology, University College London, London, United Kingdom; 7) Dept. of Ophthalmology and Visual Sciences, Univ. of Texas Health Science Center, Houston, TX; 8) Depts of Ophthalmology and Molecular and Human Genetics, Baylor College of Medicine, Houston, TX.

Whole-genome linkage mapping in a seven-generation family with autosomal dominant retinitis pigmentosa (adRP) identified a linkage region on chromosome 10q21.3-q22.1 with a maximum LOD score of 3.6 at 0% recombination. All known adRP genes and X-linked RP genes were excluded in the family by a combination of linkage exclusion, Sanger sequencing, and next-generation sequencing (NGS). Whole-exome NGS revealed a missense mutation in hexokinase 1, HK1 c.2539G>A, p.Glu847Lys, tracking with disease in all affected individuals. One severely-affected family member who was homozygous for this region by linkage analysis has two copies of the mutation. No other potential mutations were detected in the linkage region nor were any likely candidates identified elsewhere in the genome. Subsequent testing revealed the identical mutation in four additional, unrelated families with adRP, for a total of five mutations in 459 probands tested (1.1%). Taken together, the families have a combined maximum LOD score of over 8.0. No further instance of the mutation was found in publically-available databases or in control samples tested in our laboratory. Of the five families, three are from the Acadian population in Louisiana, one is French Canadian and one is Sicilian. Analysis of the affected chromosome in each family and in the homozygous individual revealed multiple haplotypes with a maximum possible overlap of 600 kb, suggesting either independent origins of the mutations or an ancient founder mutation. HK1 is a widely-expressed gene, with multiple, abundant retinal transcripts, coding for hexokinase 1. Hexokinase catalyzes phosphorylation of glucose to glucose-6-phosphate, the first step in glycolysis. The Glu847Lys mutation is in a highly-conserved site, outside of the active site, and away from functional sites. Bioinformatic analysis of pathogenicity is inconclusive because of the large number of HK1 isoforms and homologs. Rare recessive null mutations in HK1 cause non-spherocytic hemolytic anemia, which was not observed in these families. However, a discrepancy between the phenotypes associated with dominant versus recessive mutations is common among inherited retinopathies. We conclude that a mutation, or mutations, in HK1 cause roughly 1% of adRP cases in these populations, revealing a novel retinal disease gene and possibly a common founder-effect mutation.

2906M

A novel disease-causing gene for Pelizaeus-Merzbacher disease. M. Nafisinia1, N. Sobreira2, W.A. Gold1,2, L. Riley1,3, R. Ouvrier3,4, C. Boehm2,3, J. Christodoulou1,2,6 1) Genetic Metabolic Disorders Research Unit, Western Sydney Genetics Program, the Children's Hospital at Westmead, Sydney, NSW, Australia; 2) Discipline of Paediatrics & Child Health, Sydney Medical School, University of Sydney, Sydney, NSW, Australia; 3) Institute of Genetic Medicine, Johns Hopkins Univ School Med, Baltimore, MD, USA; 4) Neurosurgery, Children's Hospital at Westmead, Sydney, NSW, Australia; 5) Children's Hospital at Westmead, Sydney, NSW, Australia; 6) Discipline of Genetic Medicine, Sydney Medical School, University of Sydney, Sydney, NSW, Australia.

Background: Pelizaeus-Merzbacher disease (PMD) is a rare Mendelian disorder characterized by central nervous system hypomyelination. The disease is caused by mutations in several genes, which are associated with delay in motor abilities and intellectual function. Aim: To identify the genetic cause of disease in two siblings with a clinical picture of PMD but negative for mutations in the known PMD genes PLP1 and GJC2. Patients & Methods: Two sibs with a clinical diagnosis of PMD were born to non-consanguineous parents of Maltese background. The affected brother had motor and expressive delay, bilateral horizontal nystagmus, an action tremor, head titulation, an ataxic gait and cognitive impairment. When last examined at 17 years of age, he had normal muscle bulk with slightly reduced strength, brisk deep tendon reflexes, ankle clonus and upgoing plantar reflexes bilaterally. His brain MRI scan with fluid attenuated inversion recovery sequence confirmed a dys- or demyelinating process in the central and non-myelination in the peripheral white matter. His sister was less severely affected than her brother with mild nystagmus, ataxic gait and anxiety attacks. When last examined at age 10, she had increased tone in her lower limbs, mild cerebellar signs, moderate cognitive impairment, truncal ataxia and pendular nystagmus. Whole exome sequencing (WES) was used to screen for likely pathogenic variants followed by western blotting, and temperature sensitivity and auxotropic studies of patient fibroblasts to confirm pathogenicity. Results: WES uncovered a homozygous mutation in the arginyl-tRNA synthetase (RARS) gene (c.5A>G, p.Asp2Gly), which has recently been associated with a hypomyelination syndrome. Protein levels of RARS, and those of a binding partner, KARS (Lysyl-tRNA synthetase), were found to be significantly reduced by 80%; and 65%; respectively using western blotting of patient fibroblast extracts. As RARS is involved in protein synthesis where it attaches arginine to its cognate tRNA, patient cells were tested to determine their ability to function without this essentiial amino acid. Patient fibroblasts cultured in limited arginine at 30°C, showed a significant reduction (p<0.001) in viability compared to control cells indicating the inefficiency of protein synthesis in the patient cells. Conclusion: Here we provide evidence for a novel PMD-causing gene. Screening of a larger PMD cohort is in progress to further delineate the RARS phenotype.
2907T

The NCLs or Batten Diseases are a group of progressive neurodegenerative disorders characterized by rapid death of cortical neurons and intracellular accumulation of autofluorescent lipopigment material (lipofuscin) in the body’s tissues. Fourteen genetically distinct human NCLs were identified. Clinically, NCLs are characterized by a combination of cerebellar ataxia, visual impairment, and variable association of seizures, behavioral disturbances, and cognitive deterioration. CLN5 mutations (13q21.1-q32) are reported in severe autosomal recessive forms of variant Late Infantile Neurolonal Cerebral Lipopiginoses (vLINCL). A recent clinical study included four families with CLN5 mutations and a patient with type B Seizure disorder with intellectual disability. Phenotypic assessment showed features of infantile neuroaxonal dystrophy. A previously unreported CLN5 mutation was identified. The patients presented with a homogenous phenotype characterized by onset between 5–6 years, atypical gait, dysarthria, dysphagia, seizures, developmental delay and intellectual disability. This is the first report of CLN5 mutation causing a late-onset atypical form of adult-onset NCL. The mutation in CLN5 was confirmed by whole-exome sequencing. CLN5 mutations have been previously reported in autosomal recessive NCL1, NCL2A, NCL2B and NCL5A. In particular, the three CLN5 mutations reported in this paper are novel. These findings suggest that CLN5 mutations may also cause late-onset ataxia with predominant cerebellar features, dominated by a motor sign and a non-progressive mental status. Additional genetic work is ongoing to identify additional mutations and to better characterize the spectrum of phenotypes associated with CLN5 mutations.

2908S

Recent reports show that loss of GJB2 expression in deaf probands can be linked to atypical phenotypes. Hereditary deafness is a genetically heterogeneous disorder, with wide range of symptoms through clinical variation, including difference in frequencies affected, hearing threshold changes and age of onset. The causes of hearing loss (HL) may be genetic, environmental, or multifactorial. Although tremendous progress has been made in our understanding of the molecular basis of hearing and HL, the identification of genes and gene defects that affect the process of hearing remains challenging. Up to date, there are more than 60 genes known to be involved in non syndromic HL (NSHL). Nevertheless, it is estimated that there are many more genes to be discovered. Moreover, for countless families too small for the conventional linkage analysis, the genetic cause of their HL is still an unsolved problem. Whole exome sequencing (WES) with focusing on the targeted region in the targeted sequencing of the protein-coding subset of the human genome, makes it possible to efficiently identify novel causative genes and mutations. In this study we perform WES on 4 families negative for common deafness genes mutations that are not amendable to conventional approaches. The SureSelect human all exon 50Mb kit (Agilent Technologies) was used for exons and flanking intronic sequences enrichment and the sequencing was performed on the Hisq2000 instrument (Illumina). The raw data were processed using Picard (https://picard.sourceforge.net) and GATK. Variant calling was carried out using the Genome Analysis Toolkit (GATK). Annotations were performed using Ensembl (http://asia.ensembl.org/index.html) and UCSC Genome Browser (http://genome.ucsc.edu). Filtering criteria were applied to identify the best candidates. The selected variants were verified by Sanger sequencing. The final list of variants was compared with the HGMD and OMIM databases. In total, 8 deaf probands were selected. Analysis detected all known variants in nuclear and mitochondrial genes. These results prove the accuracy and reliability of the custom capture exome. The search for the causative variations in candidate genes for the 8 unsolved NSHL hearing loss families are in progress.
Rubinstein-Taybi syndrome (RTS) is a rare congenital disorder. It affects ~1 in 100,000 individuals and is characterized by intellectual disability, growth delay, broad thumbs and big toes, organ malformations, behavioural problems, and specific facial features. Diagnosis of RTS is essentially based on clinical presentation and can be confirmed by genetic screening. Most cases are due to de novo mutations, and causative mutations can be identified in <50% of RTS patients, affecting a single copy of either CREBBP or EP300. The mutation distribution in RTS is uneven, with CREBBP mutations being significantly more frequent than EP300 mutations (~50% and 5% of cases respectively). The significant percentage of RTS cases without a genetic diagnosis suggests that RTS is a genetically heterogeneous disorder, and that there are additional genes involved in RTS. There are other syndromes characterised by clinical features overlapping with RTS, and some are caused by mutations in genes encoding CBP and/or P300-interacting proteins. Next generation sequencing (NGS) on the HiSeq 1500 (illumina) or the Ion Proton (Life Technologies) has been used to search for RTS-causing mutations by sequencing the entire protein-coding DNA sequence (exome) of six RTS patients. Data analysis showed one patient with a stop/gain SNV in CREBBP, which was confirmed by Sanger sequencing. No EP300 mutations were found in our patient cohort, and analysis is ongoing to generate a candidate gene list based on the Interactome and genes involved in diseases with similar characteristics to RTS.

2912M
The utility of clinical exome sequencing in identifying the genetic origins of eight unclassified developmental disorders in unique Canadian populations. S.M.K. Farhan1,2, J. Wang1, J.F. Robinson1, V.M. Sui3,4, C.A. Rupar3,4, R.A. Hegeler1-2 FORGE Canada Consortium, 1) Robarts Research Institute, Schulich School of Medicine and Dentistry, Western University; 2) Department of Biochemistry, Schulich School of Medicine and Dentistry, Western University; 3) Medical Genetics Program, Department of Pediatrics, London Health Sciences Centre; 4) Children's Health Research Institute, London Health Sciences Centre.

We are a part of two large national rare disease initiatives: FORGE and Care for Rare Canada, in which >300 rare diseases have been ascertained and diagnosed. Specifically, our group is investigating the genetic etiology of 8 developmental disorders, namely: (1) infantile mitochondrial complex II and III deficiency (IMC23D), (2) seizures, delay, and unusual facies (SDFU) syndrome, (3) epilepsy with ataxia, (4) osteopetrosis, (5) Angelman-like syndrome, (6) ichthyosis-microphthalmia, (7) ataxia, dystonia, and mental retardation, and (8) nocturnal seizures with developmental delay. Collectively, these affect 27 patients in 11 families from different Canadian communities. Their clinical presentations vary in severity, symptoms and tissues affected. Six disorders were found in communities with high rates of consanguinity; consequently, we applied homozygosity mapping to generate candidate loci. Next, whole exome sequencing (WES) was performed on DNA from 19 patients across the 8 disorders. Using a non-synonymous, rare variant analysis, WES with in silico analyses, was applied to identify mutations for each disorder. For some disorders, we performed population screening and functional studies to supplement the clinical description. Using our approach, we have identified the likely causative mutations of 6 disorders thus far. While each disorder is clinically distinct, the process used to identify the causative gene is the same. Importantly, we have identified 5 new genes heretofore never implicated in a human disease, which when mutated, can lead to developmental disorders. These include: (1) NFS1, an iron-sulfur cluster protein, which causes IMC23D; (2) EXT2, a heparan sulfate biosynthesis enzyme and a tumour suppressor, which leads to SDUF syndrome; (3) TMTC3, a protein whose function is still unknown, is depleted in nocturnal seizures with developmental delay; (4) FSD1, a centromere associated protein with specific brain tissue expression, is mutated in patients with Angelman-like syndrome; and (5) NUDCD2, which interacts with lissencephaly associated protein PAFAH1B1, is mutated in patients with ataxia, dystonia, and mental retardation. We have also identified a known disease causing gene, KCTD7 as the cause of progressive myoclonus epilepsy. These studies allow for an academic and a clinical focus by allowing us to better understand the dynamic function and mechanistic significance of these genes and importantly, how they can underlie human disease.
Whole genome sequencing of mummy DNA shows significant association with human disease phenotype. S. Bhattacharyya, J. Li, H. Lam, R. Lachman, N. Asadi, A. Butte, G. Nolan. 1) Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, 94305 USA; 2) Department of Bioinformatics, Bina Technologies, Redwood City, CA; 3) Stanford University Medical Center, CA; 4) Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305. 

In 2012 an analysis was initiated of a mummified specimen (Awa) found in the Atacama Desert of Chile, South America. The 6-inch specimen with human or primate features manifested only 10 pairs of ribs (humans have 12), a skull shape that did not appear human, and apparently prematurely ossified growth plates suggesting a greater age at time of death than the size of the specimen would indicate. To determine possible genetic drivers of the observed morphology, DNA from the specimen was subjected to whole genome sequencing using the Illumina HiSeq platform. 377,333,714 reads passed the set of filtering criteria. Using the BINA secondary analysis pipeline (ver 1.5.0-dev-217-ga8038cc), 97% reads successfully mapped; of these, 89.77% uniquely mapped to the human reference genome (GRCh37) with a set of decay sequences; 7.03% mapped to multiple locations, and 3.20% were unmapped. The sequence reads were also aligned to the Chimpanzee genome (88.01% uniquely mapped). In total, 3,356,569 single nucleotide variations (SNVs) were found as compared to the human reference genome, 518,365 insertions and deletions (INDELs) detected, of which 81.5% and 77.5% are in dbSNP, respectively. To identify putative causal variants associated with Awa’s phenotype, the SNVs were run through ANNOVAR pipeline and confirmed by BINA’s analytic platform. The SNV and INDEL mutations fell within the range of “human normal”, ruling out a non-human primate origin for the specimen. A preliminary phenotype enrichment analysis of a subset of variants in genes that carry one or more deleterious mutations, using the hypergeometric test, showed a significant association with several phenotypes, including: proportionate short stature (adjP=9.5e-02), and premature osteoarthritis (adjP=2.96e-02)—gene-disease association in several genes, and the identification of chromatin remodeling and regulation of transcription as key cellular pathways in the pathogenesis of these phenotypes. Finally, our data demonstrate the added value of studying sub-phenotypes of ID for identifying genes associated with this extremely heterogeneous and complex group of disorders.
Exome analysis of 116 patients supposed to be autosomal recessive hereditary spastic paraplegia established molecular diagnoses of 49 patients with broad genetic heterogeneities. H. Ishiura1, K. Koh2, H. Shimazaki2, J. Mitsui1, Y. Takahashi1, J. Goto1, K. Yoshimura2, K. Doi1, S. Morishita3, H. Sasaki3, Y. Takiyama1, S. Tsuj1,7, JASPAC (Japan Spastic Paraplegia Research Consortium). 1) Department of Neurology, The University of Tokyo, Tokyo, Japan Tokyo, Tokyo, Japan; 2) Department of Neurology, Yamanashi University, Yamanashi, Japan; 3) Department of Neurology, Jichi Medical University, Tochigi, Japan; 4) Department of Neurology, National Center of Neurology and Psychiatry, Tokyo, Japan; 5) Department of Computational Science, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan; 6) Department of Neurology, Hokkaido University Graduate School of Medicine, Hokkaido, Japan; 7) Medical Genome Center, The University of Tokyo Hospital, The University of Tokyo, Tokyo, Japan.

[Background] Hereditary spastic paraplegia (HSP) is a clinically and genetically heterogeneous neurodegenerative disorder. To date, SPG1-72 was identified as the disease loci, of which 56 causative genes have been identified. Autosomal recessive HSP (AR-HSP) is more clinically heterogeneous than autosomal dominant HSP (AD-HSP). A previous study revealed that causative mutations were found in only 17.2 % of patients who was suspected to have autosomal recessive (AR) inheritance (Ishiai et al. 2014). [Methods] We recruited 116 HSP patients who were supposed to have AR inheritance because 1. they had parental consanguinity with or without affected sibling(s), 2. they had affected sibling(s) without affecteds in the previous generations, or 3. they were sporadic but had complicated form HSP. The previous analyses with resequencing microarray for 13 genes, LGH for 16 genes, and Sanger sequencing for 3 genes were performed in 46 patients. Subsequently, one hundred and four patients underwent exome analysis. Variants with minor allele frequency of < 0.2 % in 56 causative genes for HSP and 203 causative/susceptibility genes for other neurodegenerative diseases were considered to be candidate mutations. Nonsense, insertion/deletion, splice site, and known missense mutations were considered to be pathogenic. In the case of novel missense mutations, we considered them as pathogenic if 1. more than one pathogenic missense mutation (PM) was reported in the vicinity of the region in genes such as SPAST or KIF5A where pathogenic mutations have been described to be clustered. [Results] We found 33 patients and 5 patients with mutations in AR-HSP and AD-HSP genes, respectively. SPG11 (12.1 %), SPG28 (4.3 %), SPG46 (3.4 %), and SPG15 (2.6 %) were frequent subtypes of AR-HSP. In addition, 8 patients with AR neurodegenerative disease such as spinocerebellar ataxias and leukencephalopathies and 3 patients with autosomal dominant neurodegenerative diseases were found. [Discussion] Exome analysis revealed pathogenic mutations in 42 % of patients suspected to have AR inheritance. Although exome sequencing is highly efficient for the molecular diagnosis of HSP, assessment of pathogenicity of missense mutations is still a challenging issue. Mutations were found in 23 genes, indicating clinical heterogeneity of these patients. The causative genes for the rest of the patients remain to be further elucidated.

New candidate genes associated with autosomal dominant partial epilepsy with auditory features identified by whole exome sequencing.

F.R. Torres1*1, P.A. Oliveira1,2, R. Secolin1,2, C.V. Soler1,2, M.G. Borges1,3, A.C. Coar3,3, M.E. Morita4,5, B.S. Carvalho1,2, F. Cendes1,6, I. Lopes-Cendes1,7, 1) Department of Medical Genetics, Unicamp, Campinas, SP, Brazil; 2) Department of Neurology, Unicamp, Campinas, SP, Brazil; 3) The Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, SP, Brazil.

Background: Epilepsy is a common chronic neurological disorder that affects approximately 1% of the population worldwide. Autosomal dominant parietal epilepsy with auditory features (ADPEAF) is a clinically well characterized syndrome. Deleterious mutations in LGII are present in about 50% of families segregating ADPEAF. Objective: To identify genes associated with ADPEAF in two large families with no mutations identified in LGII. Methods: Whole exome sequencing was performed in nine individuals from two ADPEAF families (F-1 and F-2), including seven patients and two unaffected parents. Exome was targeted with Nextera Rapid Capture Expanded Exome kit (Illumina®) and sequenced in a high-performance HiSeq Illuminia® 2500 sequencing machine (Illumina®) to obtain >50X average coverage per sample. Sequences were aligned using the BWA algorithm. Variant calling and in silico functional prediction of mutations was performed using VariantAnnotator and SniffTI tools. We prioritized non-synonymous, frameshift, splicing, and indel variants according to novelty, quality score, and putative pathogenicity. Results: We found a total of 1,007,575 and 1,006,376 variants in F-1 and F-2 families, respectively. After bioinformatics analysis, we identified 54 variants in F-1 and 83 in F-2, which are shared by all patients and are absent in unaffected individuals in unaffected the patients. Conclusions: Eight genes that are potentially associated with ADPEAF. Possible deleterious roles related to formation of axon connections, protein-protein interaction, neuronal differentiation and ionic channels may play the eight candidate genes identified relevant for ADPEAF.
2919T

Purpose Phenotypes of many hereditary tumour syndromes are known to overlap. In addition, different genes may underlie one syndrome. Therefore the number of genes to examine in a particular patient case can be relatively large. Next generation sequencing (NGS) makes it is possible to analyze large numbers of genes in parallel at relatively low cost. We studied the expediency of a targeted NGS gene panel approach in DNA diagnostics for a broad spectrum of either breast or colorectal cancer in whom a pathogenic mutation in the BRCA1/2 genes was already excluded by Sanger sequencing, or their tumours showed no signs of Lynch syndrome, respectively. Methods We developed and validated a sequencing panel based on Agilent Sure Select Target Enrichment® for mutation scanning in 71 genes known to be associated with tumour syndromes. The samples were sequenced using 151 base pair pair-end reads on an Illumina MiSeq® sequencer and analyzed using Softgenetics’ NextGENE® and Cartagenia’s Benchmark NGS® software. The 71 genes are divided in three virtual, non-overlapping gene subpanels, based on the levels of preventive options and strength of risk information, in which for genes in subpanel 1 the most information is available about preventive options and specific tumour risks. In-parallel genetic counselling the genes were tested were discussed as the 3 subpanels rather than individually. Patients can choose for any of the 3 subpanels to be tested and have the results returned to them. In addition, all 71 genes are tested in all patients for research purposes after de-identifying the patients. Results and Conclusion Approximately 80% of patients chose to have results returned. The enrichment rate was 30% for subpanels 1 and 2 and 5% for subpanel 1 only. In 5 cases a pathogenic mutation was detected, 4 times a heterozygous mutation in CHEK2 and one mutation in RAD51D. Heterozygosity for a pathogenic mutation in MUTYH was detected in 3 additional cases. In 13 cases one or two mutations classified as ‘likely pathogenic’ were detected in 13 different genes. Co-segregation and tumor analysis in the families may provide more information about pathogenicity of these mutations. In most cases the patient’s phenotype does not match with the currently known mutation spectrum associated with the mutated gene. Analysis of large gene panels may broaden our knowledge of tumour spectra associated with mutations in certain genes.

2920S
Mutation screening of retinal dystrophy patients by targeted capture from tagged pooled DNAs and next generation sequencing. M.E. Elasrag1, C.M. Watson2, D.A. Parry3, J.E. Morgan4, C.V. Logan5, I.M. Carr6, E. Sheridan2, R. Charter2, C.A. Johnson7, G. Taddio2, T. Toomes8, I.M. Carr1, A.H. van der Hout1, K. Grønskov1, A. Knoppe2, K. van Dijk-Bos1. 1) Section of Ophthalmology & Neurosciences, Leeds Institute of Biomedical & Clinical Sciences, University of Leeds, UK; 2) Section of Genetics, Leeds Institute of Biomedical & Clinical Sciences, University of Leeds, UK; 3) Yorkshire Regional Genetics Service, St. James’s University Hospital, Leeds, UK; 4) Department of Ophthalmology, St. James’s University Hospital, Leeds, UK.

Retinal dystrophies are genetically heterogeneous, resulting from mutations in over 200 genes. Prior to the development of massively parallel sequencing, comprehensive genetic screening was unobtainable for most patients. Identifying the causative genetic mutation facilitates genetic counselling, carrier testing and prenatal/pre-implantation diagnosis, and often leads to a clearer prognosis. In addition, in a proportion of cases, when the mutation is known treatment can be optimised and patients are eligible for enrolment into clinical trials for gene-specific therapies. The genomic DNA of twenty patients each from a different family was sheared, tagged and pooled in batches of four samples, prior to targeted capture and next generation sequencing. The enrichment reagent was designed against genes listed on the RetNet database. Sequence data were aligned to the human genome, and crossed asymmetry of the visual pathway in association with a variable hypopigmentation phenotype. The lack or reduction of pigment might affect the eyes, skin and hair (oculocutaneous albinism, OCA [MIM 203100]), or the eyes only (ocular albinism, OA [MIM 300500]). In a lightly pigmented population such as the Faroese, it can be difficult to distinguish between OCA and OA. Mutations in six genes (TYR [MIM 606933], OCA2 [MIM 611409], TYRP1 [MIM 115501], SLC45A2 [MIM 600202], SLC24A5 [MIM 609802] and C10orf11 [MIM 614537]) are known to cause autosomal recessive OCA and furthermore, one locus (OCA5 [MIM 615312]) has been identified. A founder mutation in C10orf11 was found in six out of nine Faroese families with OCA/OA. Mutations in one gene, GPR143 [MIM 300808], are known to cause X-linked OA. More genes are expected to be identified with OCA/OA. A consanguineous family from The Faroe Islands with two children affected with OCA/OA and two unaffected children were investigated. Mutation analysis of TYR, OCA2, SLC45A2 and C10orf11 in affected individuals showed no pathogenic changes. Homozygosity mapping showed 12 homozygous regions with approximately 100 genes. Exome sequencing will be performed.

2922T
Identifying novel genes that cause Rett syndrome by trio-based exome sequencing of MECP2-negative patients. S.A. Sajan1,2, S.N. Jianghianji2, D.M. Munzy3, R.A. Gibbs4, J.R. Lupski5, E. Glaze6, W.E. Kaufmann7, S.A. Skinner8, M. Frazz1, A.K. Percy1, J.L. Neul1,2,4. 1) Section of Child Neurology and Developmental Neuroscience, Department of Pediatrics, Baylor College of Medicine, Houston, TX; 2) Jan and Dan Duncan Neurological Research Institute, Texas Children’s Hospital, Houston, TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 4) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 5) Department of Pediatrics, Baylor College of Medicine and Texas Children’s Hospital, Houston, TX; 6) Department of Neurology, Boston Children’s Hospital, Boston, MA; 7) Greenwood Genetic Center, Greenwood, SC; 8) Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL.

Rett syndrome (RTT [MIM#312750]) is a neurodevelopmental disorder which is characterized by loss of speech and hand use, the development of gait problems and characteristic repetitive hand stereotypes. While some clinical sub-types of RTT, referred to as atypical RTT, have been shown to be caused by mutations in CASKL5 and FOXG1, classical/typical RTT still remains an unexplained monogenic disorder with de novo MECP2 mutations accounting for 95-97% of such cases. We sought to identify the molecular basis of some of the 3-5% of typical RTT cases that are MECP2-negative by trio-based exome sequencing. We used the Illumina platform to sequence the exome of 71 genes in 250 cases of RTT, which included RTT probands who had not harbored mutations in MECP2 as determined by Sanger sequencing of coding exons as well as by copy number variation analysis on a clinical array. Data analysis using the Genome Analysis Toolkit (GATK) revealed 8 heterozygous de novo protein truncation mutations in 6 different genes in 7 patients, with one patient having de novo mutations in two different genes. These two genes were MDMX and ZNF536 with missense and frameshift insertion mutations, respectively. Missense mutations were also identified and confirmed in GABRB2, IMPDH2, and SAPH2 and were predicted to be damaging by at least 2 of 6 mutation prediction algorithms. Two non-frameshift deletions were found in GRIN2B and IQSEC2, and a splice site mutation was found in TCF4. Other mutations in the two nervous system-expressed ion channel genes GRIN2B and GABRB2, as well as in TCF4 and IQSEC2, have previously been identified as causative. Overall, our results suggest that MECP2-negative RTT is heterogeneous and caused by mutations in genes that also result in other types of neurodevelopmental disorders. We recommend including these genes in a list of high priority genes when exome or targeted sequencing is performed on MECP2-negative RTT patients.
2923S
Baratela Scott Syndrome is a recessive skeletal dysplasia syndrome caused by disruption of the XYLT1 gene. K. Sol-Church1, M. Kircher2, D. Stabell1, K. Gripp1, M. Bober1. 1) Nemours Al duPont Hospital for Children, Wilmington, DE, USA; 2) Department of Genome Sciences, Seattle WA.

Baratela Scott Syndrome (OMIM 300881) is characterized by skeletal dysplasia, facial features, and developmental delay. Skeletal findings include patellar dislocation, short tubular bones, mild metaphyseal changes, brachymetacarpalia with stub thumbs, short femoral necks, shallow acetabular roofs, and platyspondyly. Facial features include: a flattened midface with broad nasal bridge, cleft palate or bifid uvula and synphysis. Though some of the cognitive delays are typically masked by a warm and engaging personality, all of the patients demonstrated preschool onset of a cognitive developmental delay with a shortened attention span. We previously described seven male patients from six different families with Baratela Scott Syndrome (BSS), and performed Whole Exome Sequencing on the patients and available unaffected family members. We identified in a single individual a homozygous c.1290 -1G>A splice variant in the Xylosyltransferase (XYLT1) gene located on chromosome 16p. Both parents were heterozygous for this splice variant. This mutation caused splicing out of exon 6 and premature termination. In 3 patients from 2 unrelated families we discovered a large 3 Mb deletion covering 16p13.11 to 16p12.3 which includes the XYLT1 gene. Because of poor capture performance for exon 1 of the XYLT1 gene, we used Sanger sequencing and identified other variants in exon 1 of XYLT1 in two additional patients: One patient carried a c.319 G>T. Gly107Ter mutation while the other carried a 26 bp deletion (c.281-306). We conclude that Baratela Scott Syndrome is a recessive disorder caused by disruption of the XYLT1 gene.

2924M
Whole exome sequencing a consanguineous family in search for a novel genetic cause of Charcot-Marie-Tooth (CMT) disease. S. Tej1, N. Shahrizaila2, K.J. Goh2, A.P. Drew3, M.L. Kennerson4, A. Ahmad Annuraa. 1) Department of Biomedical Science, Faculty of Medicine, University of Malaya, Malaysia; 2) Department of Medicine, Faculty of Medicine, University of Malaya, Malaysia; 3) Northcott Neuroscience Laboratory, ANZAC Research Institute, Sydney, Australia; 4) Molecular Medicine Laboratory, Concord Hospital, Sydney, Australia.

Charcot-Marie-Tooth (CMT) neuropathy is a clinically and genetically heterogeneous group of disorders affecting the motor and sensory neurons with an estimated prevalence of 1 in 2500 people [1]. Over eighty genes have been identified for CMT and related peripheral neuropathies however there are many cases yet to be resolved [2]. In this study, we have taken a whole exome sequencing (WES) approach to investigate the genetic aetiology of CMT in two affected brothers from a consanguineous marriage. The onset of symptoms was in their teens and clinical examination revealed bilateral pes cavus with weakness of dorsiflexion resulting in foot drop. The proband had associated brisk reflexes whereas in his brother, the reflexes were normal. We were not able to obtain WES from the unaffected parents. Nerve conduction studies indicated a demyelinating form of CMT. The segregation of the disease in the family suggests either an autosomal recessive or X-linked inheritance. WES was performed on two affected brothers, the parents and an unaffected father. Mutations in known genes associated with autosomal recessive and X-linked CMT were found to be negative after querying the WES data. This suggests that the family may present a novel genetic cause of CMT. By analysing the WES data of the family members, 52 candidate variants were identified in the affected siblings and 12 are not present in the unaffected family members. To narrow down the number of variants, we analysed our WES data using MERLIN [3] and PLINK [4] for linkage analysis and homozygosity mapping. We were able to reduce the number of candidate variants to only five. Currently, we are investigating the potential involvement of these candidate variants in disease pathology.


2925T

Shortness of stature is one of the most common pediatric concerns and has an incidence of 3 % in the general population. In 80 % of patients with growth deficit the etiology remains elusive in the absence of morphological details. Uncovering the genetic basis of short stature is therefore not only important for clinical diagnosis, prognosis and genetic counseling of affected individuals and their families, but is also a prerequisite for future development of therapeutic approaches. While Genome-wide association studies identified hundreds of common single nucleotide polymorphisms and copy number variants (CNVs) contributing to the height variation in the healthy population, we confirmed a frequent disease - rare variant hypothesis by the identification of pathogenic CNVs in 10% of patients with short stature. Our results implied a heterogeneity with more than 200 genes involved in short stature by power analysis. To address this hypothesis we thoroughly built a study group of more than 500 families with idiopathic short stature and performed whole exome sequencing (WES) in 60 trios. Thoroughly filtering and validation of the identified variants resulted in the identification of potential pathogenic variants in genes involved in epigenetic modification, cell cycle regulation, ubiquitination and protein synthesis. We found compound heterozygous inherited and de novo variants to be associated with short stature. In addition, including CNV analysis from WES further identified compound heterozygous variants. Thus, we were able to demonstrate candidate genes for short stature. Analysis of further individuals with short stature will lead to a more elaborate and detailed view on mechanisms involved in growth regulation.
2927M
Whole-Exome Sequencing and Linkage Analysis Reveal a Novel Genetic Locus for Autosomal Dominant Pattern Dystrophy of the Retinal Pigment Epithelium. A. Vincent1,2,3, N. Forster1, J.T. Maynes2,4, T.A. Paton5,6, G. Billingsley6, N. Roslin3,4, A. Ali3, J. Sutherland1, T. Wright1, C. Westall1, L. A.D. Paterson1,2,3, C.R. Marshall1,2,3, FORGE Canada Consortium, E. Heon1,2,3
1) Department of Ophthalmology, Hospital for Sick Children, Toronto, Canada; 2) Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Canada; 3) University of Toronto, Toronto, Canada; 4) Department of Anaesthesiology, Hospital for Sick Children, Toronto, Canada; 5) The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada; 6) FORGE Canada Consortium:

**Finding of Rare Disease Genes in Canada.**

**Purpose:** To identify the genetic cause of autosomal dominant pattern dystrophy (PD) of the retinal pigment epithelium (RPE) in two families.

**Methods and Results:** Eight members of a two generation family with PD (Family 1, five affected) were genotyped using Illumina’s Infinium Linkage-24 chip. Multipoint genome wide linkage analysis identified 7 regions of potential linkage; microsatellite markers were used to refine these regions in an extended version of Family 1 (n = 12, six affected) and a maximum LOD score of 2.09 was observed across regions of four chromosomes. Exome sequencing on two affected family members identified 15 shared non-synonymous coding sequence variants within the linked regions; candidate genes were prioritized and further analyzed. Sanger sequencing confirmed a novel heterozygous missense variant (E79K) in OTX2 that segregated with the disease phenotype. A second family (Family 2) with PD harbored the same missense variant in OTX2. A shared haplotype of 20.62 cM and 1,211 single nucleotide polymorphisms between the two families (HumanCoreExome chip) encompassing the OTX2 gene suggest a shared ancestry. All cases from both families had decrease in distance vision and myopia. All except one affected case across the two families (aged 13 years) had a “pattern” at the macula. In vivo structural retinal imaging showed discrete areas of RPE-photoreceptor degeneration at the macula in all affected individuals across both families; photoreceptor outer segments appeared disrupted in these regions. Electoretinogram testing showed generalized photoreceptor degeneration in three cases. Developmental anomalies observed include optic dysplasia (n = 1) and Rathke’s cleft cyst (n = 1).

**Conclusion:** This is the first report implicating OTX2 to underlie PD phenotype. This phenotype resembles conditional mice models that show slow photoreceptor degeneration secondary to loss of OtX2 function in the adult retina. This allelic OTX2 variant appears to spare pituitary hormonal abnormalities.

2929S
PNPLA6 mutations in Laurence-Moon Syndrome (LMS) illustrate its distinct genetic etiology from Bardet-Biedl syndrome (BBS) and suggest the classification as part of a newly described neurodegenerative spectrum. H. Dollfus1,2, M. Prasad3, E. Schaefer4, D. Bonneau5, C. Mutter2, C. Stoetzel1, 1) Laboratoire de Genétique Médicale, INSERM U1112, Strasbourg, France; 2) Centre de référence pour les Affections Rares en Génétique Ophthalmologique (CARGO), Hôpitaux Universitaires de Strasbourg, France; 3) Service de génétique, CHU d’Angers, Angers, France; 4) Centre d’Investigations Cliniques (CIC) Hôpitaux Universitaires de Strasbourg, France.

There has been a long-standing confusion concerning Bardet-Biedl Syndromes (BBS) and Laurence-Moon Syndrome (LMS) as, historically, their names were linked together because of overlapping features. We have been studying a large family, with 5 affected siblings, reported as having LMS (Chalvon-Demersay at al., 1993) because of early onset retinal dystrophy, obesity, hypogonadism and spastic paraplegia (the patients did not have polydactyly or kidney impairment). In-depth clinical exploration identified cerebellar vermis atrophy in two patients that were wheel chair bound with spastic paresis. Moreover, we show that the retinal phenotype is remarkably distinct compared to BBS retinal dystrophy. In order to identify the underlying molecular alteration we performed, in 2010, a next-generation sequencing exome scan, but failed to find any pathogenic mutations with the initial commercial analysis pipeline. In 2014, an in-house analysis pipeline revealed a novel [p.Arg1031Glnfs*38]; c. [3088_3091insAGCC] mutation in the PLNLA6 gene. Homozygous mutations in PNPLA6 were reported at the same time in Boucher-Neuhäuser and Gordon Holmes syndromes (Syntzlik et al, 2013). Due to poor coverage of the gene by exome sequencing we used exome capture using SureSelect to identify a second mutation, a novel missense mutation (p.Gly726Arg); c.[2176G>C] in a residue that is highly conserved through C. elegans. Both mutations segregated in the family. We suggest that Laurence-Moon syndrome enters the Boucher-Neuhäuser and Gordon Holmes syndromes spectrum.

2930M

**Purpose:** Axenfeld-Rieger Syndrome (ARS) is a genetically-heterogeneous autosomal dominant disorder that affects ocular development and puts patients at an elevated risk of acquiring glaucoma, a neurodegenerative blinding disorder. DNA variations that cause ARS have been discovered in two transcription factor-encoding genes, PITX2 and FOXC1, but only account for about half of the reported cases of ARS.

**Methods:** To investigate the missing heritability of the syndrome, we commissioned whole-exome sequencing (WES) for three members of a family with autosomal dominant inheritance of ARS who lack mutation of PITX2 and FOXC1. DNA variants detected by WES were analysed by PCR-based targeted sequencing of additional family members. Quantitative genomic PCR was used to examine gene dosage.

**Results:** No non-synonymous coding region mutations that fully segregated with the disease were discovered. However, the correct segregation of alleles of five DNA variants at the PITX2 locus (a missense single nucleotide polymorphism (SNP), three intronic SNPs and a nearby intergenic microsatellite marker), led us to examine PITX2 gene copy number in the family. As a result, we discovered half-dosage of a subset of conserved regulatory elements directly upstream of the PITX2 transcript, but normal dosage of PITX2 exons.

**Conclusions:** Our results suggest that these ARS patients harbour a deletion of no more than 360 kilobases, the smallest known deletion that is restricted to the PITX2 upstream region. Since deletion of the entire upstream region is sufficient to cause ARS, we conclude that this small upstream segment contains PITX2 regulatory elements essential for normal ocular development and function.
2931T
Mutations in DOCK7 in individuals with epileptic encephalopathy and cortical blindness. F.F. Hamdan, I. Perrault, M. Riol, J.M. Capo-Chichi, N. Botta-Delaili, J.C. Decaisne, B. mamma, P. Roch, M. Sylvestre, A. Lortie, P. Roux, E. Rossignol, X. Gerardin, G. Barioz, P. Berquin, A. Munnich, G. Rougeul, J. Kaplan, J. Michaud, J.M. Rozet. 1) CHU Sainte-Justine Research Center, MONTREAL, Canada; 2) INSERM UMR 1163, Laboratory of Genetics in Ophthalmology, 75015 Paris, France; 3) Paris Descartes - Sorbonne Paris Cité University, Imagine Institute, 75015 Paris France; 4) Department of Genetics, Hôpital Necker-Enfants Malades, 75015 Paris, France; 5) Department of Pediatric Radiology, Hôpital Necker-Enfants Malades, Geneva, Switzerland; 6) Department of Medical Imaging, Sainte-Justine Hospital, Montreal, Canada, H3T 1C5; 7) Division of Genetics, Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Canada, J1H SN4; 8) Department of Pediatric Neurology, Centre de Reference Epilepsies Rares, Hôpital Necker-Enfants Malades, APHP, Paris, Descartes University, 75015 Paris, France; 9) Division of Neurology, Centre Hospitalier Universitaire de Quebec, Quebec, Canada, G1V 4G2; 10) Institute for Research in Immunology and Cancer, University of Montreal, Montreal, Canada, H3C 3J7; 11) Department of Pediatric Neurology, CHU Amiens, 80054 Amiens Cedex, France; 12) Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada, H3A 2B4.

Epileptic encephalopathies are increasingly thought to be of genetic origin, although the exact etiology remains uncertain in many cases. We describe here three girls from two non-consanguineous families showing a clinical entity characterized by dysmorphic features, early-onset intractable epilepsy, intellectual disability, and cortical blindness. Brain imaging also showed marked pontobulbar sulcus as well as abnormal signals (T2 hyperintensities) and atrophy in the occipital lobe. Exome sequencing performed in the first family did not reveal any gene with rare homozygous variants shared by both affected siblings. In the second family, one gene variant, c.2510delA (p.Asp837Alafs*48), was found in both affected sisters. Exome sequencing performed in the proband of the second family also showed the presence of two rare hetozygous variants (c.2510delA (p.Asp837Alafs*48) and c.3709C>T (p.Arg1237*)) found in both affected siblings. Sanger sequencing confirmed that all three individuals are compound heterozygotes for these truncating mutations in DOCK7. These mutations have not been observed in public SNP databases and are predicted to abolish critical domains for DOCK7 function. DOCK7 codes for a Rac guanine nucleotide exchange factor that has been implicated in the genesis and polarization of new-born pyramidal neurons as well as in the morphological differentiation of GABAergic interneurons in the developing cortex. All together, these observations suggest that loss of DOCK7 function causes a previously unrecognized syndrome by affecting multiple neuronal processes.

2933M
Combined exome and targeted gene NGS panel identifies mutations in CCDC151 as a cause of Primary Ciliary Dyskinesia. A. Onouridad, R. Hjeij, A. Nadivi, T.R. Gaunt, G. Barcia, J. Kaplan, T. Menchen, S. Khenouj, N. Klungslova, R. Hjeij, B. Brotchi, R. Hjeij, K. Lemeke, Y. L, P. Pennekamp, T. Menchen, J. Martin, R. Werner, T. Burgoyne, C. Westermann, A. Rutman, I.M. Carr, E. Rossignol, C. O'Callaghan, E. Maya, V. Möller, EMK Chingu, UK10K, E. Shiel, H. Kamino, UK10K, M. Omar, H.M. Milchison. 1) Medical and Molecular Genetics, King's College London, London, United Kingdom; 2) Genetics and Genomic Medicine Programme, University College London (UCL) Institute of Child Health, London WC1N 1EH, UK; 3) Department of General Pediatrics, University Children's Hospital Muenster, 48149 Muenster, Germany; 4) Yorkshire Regional Genetics Service, James's University Hospital, Leeds, LS9 7TF, UK; 5) Leeds Institute of Biomedical and Clinical Sciences, University of Leeds, St. James’s University Hospital Leeds LS9 7TF, UK; 6) Department of Developmental Biology, University of Pittsburgh, Pittsburgh, PA 15201, USA; 7) Department of Molecular and Clinical Genetics, Institute of Human Genetics, Polish Academy of Sciences, Strzesszyzka 32, 60-479 Poznan, Poland; 8) International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland; 9) Danish PCD Centre and Pediatrics Pulmonary Service, Department of Pediatrics and Adolescent Medicine, Copenhagen University Hospital, Rigshospitalet, Denmark; 10) UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL, UK; 11) Gerhard-Domagk-Institut for Pathology, University Children’s Hospital Muenster, 48149 Muenster, Germany; 12) Centre for PCD Diagnosis and Research, Department of Infection, Immunity and Inflammation, RKCSB, University of Leicester, Leicester, LE2 7LX, UK; 13) Respiratory, Critical Care & Anaesthesia, Institute of Child Health, University College London, 1-32 Guilford Street, London, WC1N 1EH, UK; 14) Bradford Royal Infirmary, Bradford, West Yorkshire BD9 6R, UK; 15) General and Adolescent Paediatrics Section, Population, Policy and Practice Programme, University College London, 1-32 Guilford Street, London, WC1N 1EH, UK; 16) uk10k.org. 17) These authors contributed equally to the work.

Primary ciliary dyskinesia (PCD) is a heterogenous genetic disorder of ciliary/flagellar dysmotility characterized by chronic upper and lower respiratory infections and defects in laterality. The disease is recessively inherited and affects one per 15,000-30,000 births. So far, mutations in 27 genes have been identified to cause PCD including DNAH5, DNAH11, DNAI1, DNAI2 and DNAI1 which encode subunits of the axonemal outer dynein arm (ODA) components. ARMC4 which encodes an ODA docking complex component and ARMC4 which is essential for proper targeting and anchoring of ODA. Segregation of a panel of candidate ciliopathy genes was applied to an affected Bedouin-Arabic individual from a consanguineous marriag. Therefore, we focused on homozygous non-synonymous or splice-site substitutions or indels, that were novel or present in the 1000 Genomes Project with a frequency <0.01. This analysis revealed a homozygous protein truncating variant in CCDC151 (c.925G>T; p.Glu308*). In parallel, exome sequencing combined with amplicon sequencing was applied to an affected offspring from a UK-based consanguineous Pakistani-origin family. This strategy highlighted a large autosomal region in chromosome 19p13 harbouring an additional homozygous protein-truncating variant in CCDC151 (c.925G>T; p.Glu308*). These findings suggest that CCDC151 coding exons and their associated splice-sites in a cohort of 150 cases affected with PCD resulting from ODA defects identified an individual carrying the c.925G>T nonsense variant. Segregation analysis of the c.925G>T and c.1256C>T substitutions in all available members of the pedigree confirmed recessive inheritance of both variants. Transmission electron microscopy of respiratory cilia cross-sections from individuals carrying CCDC151 mutations showed loss of ODA. Immunofluorescence analysis showed that CCDC151 encodes an axonemal coiled coil protein, mutations in which abolish assembly of CCDC151 into respiratory cilia, and furthermore cause a failure in axonemal assembly of the ODA component DNAH5 and ODA-DC associated proteins CCDC114 and ARM C4. In summary, these data suggest that CCDC151 mutations cause PCD by disruption of the ODA docking complex formation.
2934T
Defective core protein IFT81 as a rare cause of a ciliopathy with neuro-
ological involvement. I. Perrault1, 2, J. Habtitter2, J. Porath2, X. Gérard2, 
H. Yung Gee2, D. Braun2, H. Fathy3, S. Thomas3, S. Saunier3, V. Cormier-
Daire2, T. Attié-Bitach2, N. Boddaert2, M. Taschner2, M. Schueler2, E. 
Lorentzen2, R. Lifton3, 4, E. Otto5, P. Bastin5, J. Kaplan5, 6, F. Hildebrandt5, 
6, J-M. Rozet5, 6, 1. Genetics in ophthalmology, IMAGINE, INSERM UMR 
1163, PARIS, France; 2 Paris Descartes - Sorbonne Paris Cité University, 
Imagine Institute, 75015 Paris France; 3 Division of Nephrology, Depart-
ment of Medicine, Boston Children’s Hospital, Harvard Medical School, Bos-
ton, MA 02115, USA; 4 Pediatric Nephrology Unit, University of Alexandria, 
Alexandria, Egypt; 5 INSERM UMR 1163, Molecular bases of hereditary 
kidney diseases: nephropathies and hypophysplasia, 75015 Paris, France; 
6 INSERM UMR 1163, Molecular and Physiopathological bases of osteo-
chondrodysplasia, 75015 Paris, France; 7 INSERM UMR 1163, Embryology 
and genetics of human malformation, 75015 Paris, France; 8 Department 
of Pediatric Radiology, Hôpital Necker-Enfants Malades, APHP, Paris, Descar-
tesUniversity, 75015 Paris, France; 9 Department of Structural Cell Biol-
omy, Max Planck Institute of Biochemistry, Martinsried, Germany; 10 Depart-
ment of Genetics and Howard Hughes Medical Institute, Yale University 
School of Medicine, New Haven, CT 06510, USA; 11 Departments of Pediat-
rics, University of Michigan, Ann Arbor, MI 48109, USA; 12 Trypanosome 
Cell Biology Unit, Institut Pasteur and CNRS, URA 2581, 75015, Paris, 
France; 13 Howard Hughes Medical Institute, Chevy Chase, MD, USA. 
Bidirectional intraflagellar transport (IFT) consists of two major protein 
complexes, IFT-A and IFT-B. In contrast to IFT-B, all components of IFT-
A have recently been linked to human ciliopathies when defective. By screen-
ing of 1,056 individuals with nephropathies-related ciliopathies for muta-
tions in all genes encoding IFT-B, we identified one family with an autosomal 
recessive splice site mutation in IFT81. Independently, by applying targeted 
cliome resequencing of 572 individuals with early-onset severe retinal dys-
function 1/3, we have identified 4 novel genes; only 16% of families now lack a 
pathognomonic finding of “molar tooth sign” on brain MRI. In about fifty 
patients in whom we found no mutation in known JSRD genes using standard 
methods, we performed whole exome sequencing; to facilitate the analysis, 
family members were included. To date, the analysis has revealed that 2- 
3 of these patients have mutations in known ciliopathy genes. In the remain-
ing 1/3, we have identified 4 novel genes; only 16% of families now lack a molecular basis for their JSRD. Identification of the underlying genetic 
causes of JSRD patients and further description of the full clinical spectrum 
of the related phenotypes will provide the groundwork for more focused 
clinical studies. We are now assessing cilia integrity and function at the cellular 
level as well as at the level of the whole organism, using animal models 
such as zebrafish and C. elegans. This study will help to define genotype-
phenotype correlations, improve diagnosis and prognosis, and stimulate 
development of improved supportive treatments as well as specific, novel 
and targeted therapies.

2935S
Genetic study of patients with Joubert Syndrome and Related Disor-
ders. T. Vloubu1, D. Yildirimli1, M. Vemulaalpi2, A.R. Cullinane1, M.C.V 
Malicdan1, Y.M. Chang1, 2, J. Bryant1, C. Sinclair1, B. Fischer1, B. Masken2, 
A.C. Young1, J. Mullikin2, M. Huizing1, W.A. Gahl1, 2, 3, 4, M. Suny-Aygun1, 2, 3, 
NISC Comparative Sequencing Program, NHGRI, NIH. 1, 2 MGB, NHGRI, 
NIH, Bethesda, MD; 2 NIH Intramural Sequencing Center, National Human 
Genome Research Institute, National Institutes of Health, Rockville MD 
20852, USA; 3 Office of the Clinical Director, NHGRI, National Institutes 
of Health, Bethesda, MD USA; 4 NIH Undiagnosed Diseases Program, 
Common Fund, Office of the Director, National Institutes of Health, Bethesda, 
Maryland, USA.

Joubert Syndrome and Related Disorders (JSRD) is a clinically and geneti-
cally heterogeneous group of ciliopathies defined based on a distinctive 
brain malformation (molar tooth sign) on brain imaging. Most JSRD patients 
display hypotonia, developmental delay, abnormal eye movements, and 
an abnormal respiratory pattern in infancy. Variable features include fibrocystic 
kidney disease, congenital hepatic fibrosis, retinal degeneration, retinal colo-
bomas, and polydactyly. To date, twenty-two genes have been identified; 
they account for approximately 50% of JSRD. Over the past 10 years, we have 
evaluated over 100 JSRD families at the NIH Clinical Center, under the 
NHGRI protocol “Clinical and Molecular Investigations into Ciliopathies” 
(www.clinicaltrials.gov, trial NCT0068224). The enrollment criterion was 
a clinical diagnosis of JSRD made based on brain MRI report describing the 
pathognomonic finding of “molar tooth sign” on brain MRI. In about fifty 
patients in whom we found no mutation in known JSRD genes using standard 
methods, we performed whole exome sequencing; to facilitate the analysis, 
family members were included. To date, the analysis has revealed that 2- 
3 of these patients have mutations in known JSRD genes. In the remain-
ing 1/3, we have identified 4 novel genes; only 16% of families now lack a molecular basis for their JSRD. Identification of the underlying genetic 
causes of JSRD patients and further description of the full clinical spectrum 
of the related phenotypes will provide the groundwork for more focused 
clinical studies. We are now assessing cilia integrity and function at the cellular 
level as well as at the level of the whole organism, using animal models 
such as zebrafish and C. elegans. This study will help to define genotype-
phenotype correlations, improve diagnosis and prognosis, and stimulate 
development of improved supportive treatments as well as specific, novel 
and targeted therapies.

2936M
The Role of a Rare Variants in Genetic Predisposition to Statin-Induced 
Myopathy. V. Stranecy1, K. Hodanova1, H. Hartmannova1, L. Pihova1, 
A. Pristoupilova1, M. Neroldova1, M. Vrablik2, M. Jirsa1, S. Kmoch1. 1) First 
Medical Faculty, Charles University of Prague, Prague, Czech Republic; 2) 
Laboratory of Experimental Hepatology, Institute for Clinical and Experimen-
tal Medicine, Prague, Czech Republic; 3) Third Medical Department, First 
Faculty of Medicine Charles University and General Faculty Hospital in 
Prague, Czech Republic.

Statin-induced myopathy is an important cause of statin intolerance and the 
most common cause of statin discontinuation. Statins exhibit a class-
wide side effect of muscle toxicity and weakness, which has led regulators 
to impose both dosage limitations and a recall. This study focuses on the 
genetic factors associated with increased statin muscle concentrations, 
especially the genes encoding an influx transporter (SLCO1B1), efflux trans-
porters (ABCB1 and ABCG2) and others, still unknown genes. We have 
analyzed qDNA of extreme cases of statin-induced myopathy. Patients had 
to fulfill the criteria: age over 18, clinically and biochemically determined 
myopathy, elevation of creatinin kinase and histological confirmation of mus-
cle disease, no intake of inhibitors of OATP1B1 membrane transporters. 
We have performed exon sequencing, whole genome genotyping, analyzed 
CNV’s and searched for large deletions. Here we present preliminary results 
of 11 patients. We found LOH (loss of heterozygosity) in SLCO1B1 loci in 4 
patients and a novel frameshift mutation c.209_210insATTGA (p.D70fs) 
in SLCO1B3 in one individual. Heterozygous variant c.521T>C (p.V174A) 
dbSNP rs4149056 in SLCO1B1 is associated with station myopathy (Link et al, 
NEJM 2008), was found in 5 patients. In two cases this variation was 
combined with LOH in the SLCO1B3 region. The incidence of this variant 
in our study is 5/22 alleles, which exceeds population frequency (MAF in 
1000Genomes 0.123). No potential pathogenic variants in genes SLCO1B1 
and SLCO1B3 was found in 3 of 11 patients. Furthermore in 823 genes, we 
found null mutations in 823 genes. The most relevant for the study are 
CYP3A43 (1 patient), CYP3A45 (1 patient) and AMPD1 (2 patients). Frequent occurrence of null mutations was found in CYP (CYP2F1 4x, CYP4B1 3x) gene family 
and transporters (SLC7A13 3x, ABCB10 2x). We found also null mutations in 
APOB1 (1 patient) and LPL (1 patient) which contribute to hyperlipoprotein-
emia. Theoretical outcome of our project lies in detailed knowledge of 
genetic predisposition to toxicity of statins. Our results will help to predict 
pharmacodynamics and kinetics of statins and other drugs. Individual 
approach based on the assessment of the patient’s genotype will improve 
the safety of the therapy.
2937T
Targeted resequencing identifies PTCH1 as a major contributor to ocular developmental anomalies. N. Chassaing1,2,3, E.E. Davis4,5,6, A. Causee7,8, V. David9, A. Desmaison9, A.R. Niedermayer2, S. Lamarre2,8, C. Vincent-Delorme10, L. Pasquier11, C. Coubes12, D. Lacombe13,14, M. Rossi15, J.L. Dulfer16, H. Dollfus17, J. Kaplan18, N. Katsanis3,4, H.C. Etcheverria19, S. Faguert-20, P. Calvas21. 1) Department of Medical Genetics, CHU Toulouse, France; 2) EA-4555, Toulouse III University, Toulouse, France; 3) Center for Human Disease Modeling, Duke University Medical Center, Durham, North Carolina, USA; 4) Department of Pediatrics and Department of Cell Biology, Duke University Medical Center, Durham, North Carolina, USA; 5) Service d’Ophthalmologie, Hôpital Purpan, CHU Toulouse, 31059 Toulouse, France; 6) Institut de Génétique et Développement, CNRS UMR6290, Université de Rennes 1, IFR140 GFAS, Faculté de Médecine, 35043 Rennes, France; 7) Laboratoire de Génétique Moléculaire, CHU Pontchaillou, 35043 Rennes Cedex, France; 8) INRA, UMR792, Ingénierie des Systèmes Biologiques et des Procédés, Toulouse, France; 9) Plateforme Biopuces de la Génopole de Toulouse Midi Pyrénées, INSA/DGBA 135, Toulouse, France; 10) Service de Génétiqunique, Hôpital Jeanne de Flandre, 59037 Lille, France; 11) Service de Génétique clinique, Hôpital Sud, 35200 Rennes, France; 12) Service de Génétique Médicale, Hôpital Arnaud de Villeneuve, 34295 Montpellier, France; 13) Service de Génétique Médicale, Hôpital Pellegrin, 33076 Bordeaux Cedex, France; 14) Université Bordeaux Segalen, Laboratoire MGRM, 33076 Bordeaux, France; 15) Service de Cytogénétique Constitutionnelle, Hospices Civils de Lyon, Groupement Hospitalier Est, 69667 Bron, France; 16) Service d’Ophthalmologie, Hôpital Necker enfants Malades, 75015 Paris; 17) Service de Généétique Médicale, Hôpitaux Universitaires de Strasbourg, 67091 Strasbourg, France; 18) INSERM U781 & Department of Genetics, Paris Descartes University, 75015 Paris, France; 19) INSERM, U910, Université de la Méditerranée Faculté de Médecine, 13385, Marseille, France; 20) INSERM unit 1048, I2MC, Team 12, 31432 Toulouse, France; 21) These authors contributed equally to this work.

Ocular developmental anomalies (ODA) such as Anophthalmia/Microphthalmia (AM) or anterior segment dysgenesis (ASD) have an estimated combined incidence of 3 in 10,000 births. Mutations in SOX2 are the most frequent contributors to ODA, but account for only 10% of ODA cases. We identified a frameshift mutation in patients affected with severe ODA. The c.1900T>G (p.Trp634Gly) mutation was identified in the EDN1 pathway or in craniofacial development. Our results suggested that mutations in EDN1 could be a potential contributor to ODA.

2938S
Deciphering the endothelin pathway in auriculocondylar syndrome and isolated question mark ears. C. Gordon1, F. Petit2, P. Kroisel3, L. Jakobsen4, R. Zocchi-Ceide5, M. Outadom1, C. Bole-Feyroud1, P. Nitschka4, A. Munlich2, S. Lyonnet1, M. Holder-Espinasse6, J. Amiel7. 1) Institut Imagine, INSERM U-1163, Paris, France; 2) Hôpital Jeanne de Flandre, Centre Hospitalier Regional Universitaire de Lille, Lille, France; 3) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 4) Copenhagen University Hospital Herlev, Denmark; 5) Hospital for Rehabilitation of Craniofacial Anomalies, University of Sao Paulo, Bauru, Brazil; 6) Department of Clinical Genetics, Guy’s Hospital, London, UK.

Auriculocondylar syndrome (ACS), a rare craniofacial disorder involving microglossia, condyl hypoplasia and question mark ear (QME), which involves a defect in the fusion of the lobe and helix, can occur as an isolated anomaly. Recently, mutations in PLCB4 (encoding a phospholipase C enzyme) and GNAI3 (encoding a G alpha protein) were identified in ACS. Both genes are predicted to function in the endothelin 1 (EDN1)-endothelin receptor type A (EDNRA) signalling pathway during the development of the pharyngeal arches. Following exclusion of PLCB4 and GNAI3 in a series of ACS and QME patients, we performed exome sequencing in four unsolved families. We identified a mutation in EDN1 in three cases, and a mutation in a gene known to regulate intracellular calcium release downstream of phospholipase C activity, in the fourth. Another EDN1 mutation was subsequently identified by direct sequencing. Two of the four EDN1 mutations occurred in patients affected with ACS, born to consanguineous, healthy parents; these patients harboured homozygous missense mutations in EDN1, each predicted to interfere with enzymatic cleavage of the EDN1 proprotein. The other two EDN1 mutations occurred in patients with dominantly inherited isolated OME; they harboured heterozygous EDN1 mutations — a premature stop in one case and a missense mutation affecting a highly conserved residue of the mature EDN1 peptide in the other. The nature of the mutations and the different modes of inheritance suggest that heterozygous EDN1 mutations are a novel cause of isolated OME.

2939M
Mutations in COG2 Encoding a Subunit of the Conserved Oligomeric Golgi Complex Cause a Congenital Disorder of Glycosylation. H. Kodera1, N. Ando2, I. Yuasa3, Y. Wada4, Y. Tsurusaki5, M. Nakashima1, N. Miyake1, S. Saitoh6, N. Matsumoto1, H. Saitou1. 1) Department of Human Genetics, Yokohama City University Graduate School of Medicine, Kanagawa, Japan; 2) Department of Pediatrics and Neonatology, Nagoya City University Graduate School of Medical Sciences, Aichi, Japan; 3) Division of Legal Medicine, Faculty of Medicine, Tottori University, Tottori, Japan; 4) Department of Molecular Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan.

The conserved oligomeric Golgi (COG) complex is involved in intra-Golgi retrograde trafficking, and mutations in six of its eight subunits have been reported in congenital disorders of glycosylation (CDG). Here we report a patient showing severe acquired microcephaly, psychomotor retardation, seizures, liver dysfunction, hypopigmentation, and hyperkeratoliposminemia. Analysis of his serum glycoproteins revealed defects in both sialylation and galactosylation of glycan termini. Trio-based whole-exome sequencing identified two heterozygous mutations in COG2: a de novo frameshift mutation [c.701dup (p.Tyr234*)] and an missense mutation [c.1900T>G (p.Trp634Gly)]. Sequencing of cloned reverse-transcription polymerase chain reaction products revealed that both mutations were located on separate alleles, as expected, and that the mutant transcript harboring the frameshift mutation underwent degradation. The c.1900T>G (p.Trp634Gly) mutation is located in a domain highly conserved among vertebrates and was absent from both the public database and our control exomes. Protein expression of COG2, along with COG3 and COG4, was decreased in fibroblasts from the patient. Our data strongly suggest that these compound heterozygous mutations in COG2 are causative of CDG.
2940T
De novo SOX11 mutations cause Coffin-Siris syndrome. N. Matsumoto1, Y. Tsurusaki1, E. Koshimizu1, H. Ohashi2, S. Phadke2, N. Miyake1. 1) Yokohama City Univ Grad Sch Med, Yokohama, Japan; 2) Division of Medical Genetics, Saitama Children’s Medical Center, Iwatsuki, Japan; 3) Department of Medical Genetics, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Uttar Pradesh, India.

Coffin-Siris syndrome (CSS) is a congenital disorder characterized by growth deficiency, intellectual disability, microcephaly, characteristic facial features and hypoplastic nails of the fifth fingers and/or toes. We previously identified mutations in five genes encoding subunits of the BAF complex, in 55% of CSS patients. To further address genetic causation, here we performed whole exome sequencing in additional CSS patients, identifying de novo SOX11 mutations in two CSS patients with mild intellectual disability, mild end of CSS phenotype, sox11a/b knockdown in zebrafish causes brain abnormalities, potentially explaining the brain phenotype of CSS. SOX11 is the downstream transcriptional factor of the PAX6-BAF complex, highlighting the importance of the BAF complex and SOX11 transcriptional network in brain development. Acknowledgment: Drs. Koushi I, Shina M, Suzuki T, Okamoto N, Imamura S, Yamashita M, Watanabe S, Yoshiura K-i, Kodera H, Miyatake S, Nakashima M, Saitou H, Ogata K are highly appreciated for their contribution to this work.

2941S
Copy Number Variations detection for Congenital Absence of bilateral ACL and PCL ligaments. Y. Liu1, R. Golhar2, ME. March1, Y. Guo1, Y. Li1, ME. March1, J. Li2, J. Zhang2, X. Xu2, MA. Deardoff1, B. Keating1, H. Hakonarson1. 1) Center for Applied Genomics, The children’s hospital of Philadelphia, Philadelphia, PA; 3) Department of Genetics, Saitama Children's Medical Center, Iwatsuki, Japan; 2) Division of Medical Genetics, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Uttar Pradesh, India.

Copy Number Variations (CNV) deletion of exons in the gene CEP57L identified in a mother of large familial case series of related conditions such as ACL agnesia and absence of menisci. rare congenital malformations that results in knee joint instability with a growth deficiency, intellectual disability, microcephaly, characteristic facial features and hypoplastic nails of the fifth fingers and/or toes. We previously identified mutations in five genes encoding subunits of the BAF complex, in 55% of CSS patients. To further address genetic causation, here we performed whole exome sequencing in additional CSS patients, identifying de novo SOX11 mutations in two CSS patients with mild intellectual disability, mild end of CSS phenotype, sox11a/b knockdown in zebrafish causes brain abnormalities, potentially explaining the brain phenotype of CSS. SOX11 is the downstream transcriptional factor of the PAX6-BAF complex, highlighting the importance of the BAF complex and SOX11 transcriptional network in brain development. Acknowledgment: Drs. Koushi I, Shina M, Suzuki T, Okamoto N, Imamura S, Yamashita M, Watanabe S, Yoshiura K-i, Kodera H, Miyatake S, Nakashima M, Saitou H, Ogata K are highly appreciated for their contribution to this work.

2942M

For almost 50 years, OMIM has been cataloging phenotype-gene relationships. Over time, variations have been ascertained by a variety of technologies and the certainty of these relationships is highly variable. The use of whole-genome/exome sequencing has enabled rapid growth in the publication of causative variants, but still does not confirm the certainty of the relationships. Here we present the criteria we adopted to establish a gene-phenotype relationship: (1) the existence of multiple, unrelated individuals with pathogenic variants in the same gene; (2) the variants segregate with the phenotype in multiplex families; and/or (3) the variants occur de novo in a statistically significant number of individuals. Functional data and/or animal models support the causality but are not required. A “qualified” gene-phenotype relationship is established based on the following: (1) only one multiplex family is reported to have variants in a single gene and the variants segregate with the phenotype in the family, and (2) there is supportive functional data such as in vitro enzyme activity, a comparable phenotype in a model organism experiment, or an animal model; in this case the gene-phenotype relationship is qualified by noting that the variant has been identified in only “1 family”. In rare instances, a similar gene-phenotype relationship may be established on the basis of a single patient if there is robust supporting phenotype and functional data. The morbid map listing of a qualified phenotype is preceded by a “?”. When other information is entered into OMIM that substantiates the phenotype-gene relationship, the qualifies is removed. No relationship is established if there is only 1 patient or family and provides no supporting functional data; however, an allelic variant may be created as a “variant of unknown significance” with a discussion of a possible relationship with the phenotype. The same criteria is being applied retrospectively to all existing OMIM gene-phenotype relationships that are based on a single allelic variant. To help users follow OMIM entries, we have implemented a new service, MIMmatch. MIMmatch members can designate entries or Phenotypic Series to follow and receive updates on these as they are updated. In addition, other researchers who share their interest in certain entries; and/or receive a daily update on new gene-phenotype relationships established in OMIM.
2946S Alleles of the reported deafness genes are major contributors to the etiology of moderate to severe hearing loss in Pakistani population. A. Iimura1,2, R. Bashir1, G. Mutbajia2, A. Masood2, I. Bukhari1, A. Rehman1, R.J. Morell1, T.B. Friedman3, S. Naz2. 1) NIDCD, National Institute of Health, Bethesda, MD; 2) School of Biological Sciences, University of the Punjab, Lahore, Pakistan.

Most of the genes that cause recessively inherited deafness are associated with profound, prelingual hearing loss, which occurs approximately 1/1,000 live births. The genetic causes of moderate, progressive, post-lingual hearing loss are not well understood, despite the fact that this type of hearing loss is much more prevalent in the general population. We studied the molecular genetic bases of recessively inherited, stable or progressive, moderate to severe hearing loss in 44 consanguineous Pakistani families. We used a stepwise experimental approach, starting with homoygosity mapping at reported DFNB loci, followed by targeted sequencing of the known deafness genes, and finally whole-exome sequencing (WES) to identify mutations in novel genes associated with this phenotype. 11 known and 17 novel mutations in the known deafness genes congregate with moderate to severe hearing loss in 68% of these families. The majority of these reported mutant alleles are thought to cause only profound deafness. Nevertheless, our finding of their association with less severe phenotype may suggest the presence of environmental or genetic modifiers (suppressors) that reduce phenotypic severity. SLC2A6, TMC1 and OTOF are three of the major contributors in this cohort. In 14 of the families we did not find causative variants in known deafness genes. Our findings indicate that the same genes and alleles known to cause profound deafness are also major contributors to the genetic basis of moderate to severe hearing loss.

2947S Mutations in a metabolic kinase gene lead to autosomal dominant retinitis pigmentosa. F. Wang1,2, Y. Wang1, B. Zhang1,2, H. Li1,3, L. Zhao1,2, K. Wang1,2, M. Xu1,2, Y. Li1,2, F. Wu1,2, C. Wen1,2, P. Bernstein1, H. Wang1,2, R. Sui1,2, C. Zhang1,2, R. Chen1,2,3,4, S. Rouleau1,2, D. Goudenege1, R. Coutant4, D. Bonneau1,2. 1) Department of Biochemistry and Genetics, CHU Angers, Angers; 2) UMR CNRS 6214 INSERM 1083, Angers, France; 3) Department of Pediatrics, CHU Angers, Angers, France; 4) Department of Genetics, CHU Nantes, Nantes, France.

Pseudo-Pendred syndrome (PPDS) is characterized by the association of sensorineural deafness, hypothyroidism due to iodide organification defect, absence of inner ear malformation and absence of mutation in TPO. In order to determine the cause of PPDS, we performed whole exome sequencing (WES) in a family with two children affected with deafness, developmental delay, positive perchlorate test and absence of inner ear malformation (75%). The initial mutation identified in the adRP family, suggesting that both alleles are indistinguishable from those of other typical pediatric CI recipients, genetic examination is indicated in all CI candidates prior to operation.

2945M Exome sequencing reveals TPO mutations in Pseudo-Pendred syndrome. A. Denommée-Pichon1, E. Coltin2, S. Marlin3, L. Jonard4, S. Rouleau5, S. Kury5, S. Dumont5, D. Goudenege1, R. Coutant4, D. Bonneau1,2. 1) Department of Biochemistry and Genetics, CHU Angers, Angers, France; 2) UMR CNRS 6214 INSERM 1083, Angers, France; 3) Department of Genetics, CHU Angers, Angers, France; 4) Department of Pediatrics, CHU Angers, Angers, France; 5) Department of Genetics, CHU Nantes, Nantes, France.

Pseudo-Pendred syndrome (PPDS) is defined by the association of sensorineural deafness, hypothyroidism due to iodide organification defect, absence of inner ear malformation and absence of mutation in TPO. In order to determine the cause of PPDS, we performed whole exome sequencing (WES) in a family with two children affected with deafness, developmental delay, positive perchlorate test and absence of inner ear malformation on CT-scan. Parents were healthy and non-consanguineous and direct sequencing of SLC2A6 was normal in both patients. In order to determine the cause of PPDS, we performed whole exome sequencing (WES) in a family with two children affected with deafness, developmental delay, positive perchlorate test and absence of inner ear malformation on CT-scan. Parents were healthy and non-consanguineous and direct sequencing of SLC2A6 was normal in both patients. In silico prediction tools determined a deleterious mutational impact of TPO in both patients. Screening of TPO genes were performed in 14 other cases of PPDS and a heterozygous missense mutations in TPO were detected in the 7 CI recipients with PDBH15 mutations, otocystic emissions were absent in both ears, and imaging findings were normal in all 7 implanted ears. DFNB3 or PCH15 mutations are major genetic determinants of poor CI performance, probably because of their role in the pathology of spiral ganglion neurons and/or brainstem auditory nuclei. Because children with DFNB3 or PCH15 mutations show clinical features indistinguishable from those of other typical pediatric CI recipients, genetic examination is indicated in all CI candidates prior to operation.
2948M
Expansion of the fibrosing poikiloderma phenotype caused by FAM111B to include cytopenia and pancreatic dysfunction. A. Seo1, T. Walsh2, M.K. Lee1, M-C. King1, A. Shimamura3. 1) Genome Sciences, University of Washington, Seattle, WA; 2) Division of Medical Genetics, University of Washington, Seattle, WA; 3) Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA.

Poikilodermia, a dermatological disorder including telangiectasia, uneven pigmentation, and epidermal atrophy, can be inherited recessively as part of Rothmund-Thomson syndrome (RTS [MIM 268400]), Kindler Syndrome (MIM 173700), and dyskeratosis congenita (DKC [MIM 305000]). Recently, heterozygous missense mutations in FAM111B (MIM 615584) were shown to be responsible for a newly defined autosomal dominant condition including poikilodermia, tendon contraction, myopathy, pulmonary fibrosis, heat intolerance, hypohidrosis, sparse hair, and elevated liver enzymes, apparently inherited in autosomal dominant fashion. The two youngest affected siblings also exhibited intermittent cytopenias, narrow hypocellularity, and exocrine pancreatic dysfunction, which are the characteristic features of Shwachman-Diamond Syndrome (SDS). A grandfather had a history of interstitial pulmonary disease, and several individuals had signs of tendon contracture. SDS, the causative gene for SDS, was wildtype in the affected siblings. Whole exome sequencing of these two siblings, the affected father, and the unaffected mother revealed a heterozygous in-frame deletion FAM111B c.1261_1263delAAG (p.Lys421del) in the affected individuals. Genotyping the family revealed the mutation to be present in about 10% to 50% of patients, only 1 recessive GJB2 mutation can be detected, constituting a diagnostic dilemma. In this patient, we hypothesize that there might be undetected mutations in: (1) the non-coding region of GJB2 (such as untranscribed exons, introns, promoter or enhancer) leading to compound heterozygous inheritance; (2) other gap junction genes leading to digenic inheritance; or (3) other modulating deafness genes. Accordingly, we designed a massively parallel sequencing (MPS) panel which target the whole GJB2 gene and the exons of 129 deafness genes (including all gap junction genes), and applied this panel to 12 patients with only 1 recessive GJB2 mutation detected by conventional Sanger sequencing of the GJB2 coding region. Causative mutations in other deafness genes were identified in 4 patients, including SLC26A4 mutations in 2 patients (genotypes: p.K369X/p.T410M and p.P8T/p.P8T, respectively), KCNO4 mutation in 1 patient (genotype: p.F182L/wt), and MYO15A mutations in 1 patient (genotype: p.W350G). A splicing site mutation (c.250C>A) in GJB2 was detected in 1 patient, contributing to compound heterozygosity of GJB2 mutations. Of note, we also found −67% mosaicism of GJB2 c.235delIC in another patient which was difficult to be identified by Sanger sequencing alone. In total, our MPS panel achieved genetic diagnosis in 6 (50%) of 12 patients with only 1 recessive GJB2 mutation detected by conventional Sanger sequencing, including 4 with SNHI attributable to mutations in other deafness genes and 2 with SNHL attributable to compound heterozygosity or mosaicism of GJB2 mutations. Our results demonstrate the utility of MPS in achieving genetic diagnosis in deaf patients with non-confirmative GJB2 genotypes on conventional genetic examinations.
Materials and Methods: Sixty patients with a clinical diagnosis of PAH (7 HPAH, 48 IPAH, 5 unknown) were evaluated for mutations in BMPR2 by Sanger sequencing and deletion/duplication analysis. BMPR2 mutation negative patients (n=47) were then tested for mutations in the other known PAH-associated genes (ACVR1, ENG, CAV1, and KCNK3) by Sanger sequencing. The ACVR1 and ENG genes were amplified in all patients in all exonic and intronic regions. Duplication analyses were not performed due to technical limitations. All remaining negative mutation patients (n=46) were then evaluated for mutations in EIF2AK4 by Sanger sequencing.

Results: Gene mutations were identified in a total of 14 patients. As expected, neither BMPR2 nor EIF2AK4 presented any mutation positive patients (n=13). Mutations were also identified in ENG (n=1) and variants of unknown significance were found in ACVR1 (n=1) and KCNK3 (n=1). No mutations were identified in CAV1 or EIF2AK4.

Discussion: EIF2AK4 is a potential candidate for a common cause of PAH in patients with a clinical diagnosis of PAH.
2955T PITUATORY HORMONE DEFICIENCY: HUNT FOR NOVEL CAUSATIVE GENES AND GENETIC CONTRIBUTIONS TO VARIABLE PENETRANCE Expression
Q. Eng1, L.J. Arndt2, A.F. Benedetto2, T. Brue2, L.R.S. Carvalho3, F. Castinetti2, F.A. Correa4, N. Foyouzi2, M.M. Franca2, J.Z. Liu1, Q. Ma1, B.A. Ozcel1, B.B. Mendonca4, M. Moreira2, A.P. Otto2, R. Reynaud2, A. Sadeghi-Neill2, S.A. Camper5. 1) Department of Human Genetics, University of Michigan, Ann Arbor, MI; 2) Department of Obstetrics and Gynecology, University of Michigan Medical School, Ann Arbor, MI; 3) University of São Paulo, Clinical Hospital, São Paulo, Brazil; 4) Aix-Marseille University, Center for Research in Neurobiology and Neurophysiology of Marseille (CRNB.M) – CNRS UMR 7326, Timone Hospital, Marseille, France; 5) Division of Pediatric Endocrinology, Tufts University School of Medicine, Boston, MA.

Pituitary hormone deficiency is genetically heterogeneous and fatal if hypothalamic and adrenal insufficiency are untreated. The known causal genes explain a minority of the cases. Predicting patient outcomes is difficult because some genes exhibit dominant inheritance with incomplete penetrance, and most mutant genes have variable clinical presentations. Mutations in the transcription factors OTX2 and HESX1 cause defects in craniofacial development involving the eyes, pituitary gland, or both. To identify novel causal genes and to search for genes that influence penetrance we sequenced the exomes of 22 individuals with hypopituitarism. We found two cases of mutations in transcription factor HESX1. A familial case of pituitary aplasia with neonatal crisis is a novel compound heterozygote for HESX1, R159W/R160H, which disrupts the DNA sequence recognition helix. We found a previously reported p.I26T lesion that impairs HESX1 repressor function in a consanguineous case of multiple pituitary hormone deficiency. We discovered a homozygous, recurrent p.L144H mutation in the growth hormone releasing hormone receptor (GHHR) gene in a consanguineous case of isolated growth hormone deficiency (IGHD). A proband with hypopituitarism, polydactyly, and normal eyes was ascertained in a large pedigree due to monogenic defects. Recent strategies to identify the genetic causes of PIDs include exome and whole genome sequencing. We performed RNA sequencing to identify the genetic cause of a proband with neonatal pituitary aplasia. Moreover, to confirm the pathogenicity of the detected mutation, punch biopsies from the patients and a control individual were used to culture fibroblasts and demonstrate the effects of the identified mutation through a series of functional studies. The average number of mapped reads was 67 million per sample with about 7 billion aligned bases. The average of variants found on these genes was 846 per patient. The majority of variants were placed in non-coding regions (introns, 5’ and 3’ UTRs). We prioritized the non-synonymous exonic variants and those characterized by loss or gain of stop codons, deletion or insertion of codons, frameshift, loss or gain of splicing sites, start site changes and stop gain variants. Using mRNA expression levels, we identified the gene as the only variant in a list of ~90 candidate genes. The presence of polydactyly in the father and the proband suggests a dominant negative in both heterologous (293T) and homologous GnRH neuronal cell lines (GT1-7) transfected with either a consensus OTX2 binding site or a truncated version of the gene of interest, the PDGFR receptor. The proband’s mother is a p.H230L carrier with short stature, and several unaffected individuals in the pedigree are heterozygous for this variant. The presence of polydactyly in the father and the proband suggests the gene may be a candidate for enhancing the effects of the OTX2 variant on hormone production and polydactyly. This exome sequencing pilot revealed that approximately 20% of probands have mutations in known genes, and nomimates novel genes and pathways that affect normal pituitary function.

2956S A Combined Exome sequencing and RNA-Seq Strategy Reveals a Novel Mutation in DOK8 That Results in Immunodeficiency and Radiosensitivity. S. Khan1, M. Kuruvillas1,2, B. Wakeland3, C. Liang1, K. Vishwannah1, R. Gatti1, T. Torgersen4, N. van Oers1, E. Wakeland1, M. Teresa de la Morena1,2. 1) UT Southwestern Medical Center, Dallas, TX; 2) Children’s Medical Center, Dallas, TX; 3) University of California, Los Angeles, CA; 4) Seattle Children’s Research Institute, Seattle, WA.

Primary immunodeficiency diseases (PIDs) are rare in the human population, with the causal genetic mutations difficult to identify. To date, more than 180 genes have been linked to diverse PIDs, with the vast majority due to monogenic defects. Recent strategies to identify the genetic causes of PIDs include exome and whole genome sequencing. We performed RNA sequencing and exome sequencing on RNA and DNA from fibroblasts with prominent manifestations of warts and cryptosporidial sclerosing cholangitis, profound T cell lymphopenia and radiosensitivity. Comparative analyses including an unaffected sibling and parents using exome sequencing revealed a novel homozygous 2 bp frameshift in the splice site region of DOK8 (p.H230L) in the affected siblings. RNA-SEQ analysis of their whole blood showed a significant reduction in mRNA levels of DOK8 gene as well as a cluster of aberrant transcripts within the splice junction. The consequence was a significant loss in DOK8 protein expression. Both affected siblings had an increased sensitivity to ionizing radiation. While mutations in the DOK8 gene have been reported as causal to immune dysfunction, this is the first report that indicates a role in radiosensitivity. Also this is the first report identifying a causal mutation to an intronic site by combining RNA sequencing with exome sequencing.
2959S
Mutation of CLPB, a human homologue of bacterial ClpB/yeast Hsp104 mitochondrial molecular chaperone, causes a novel form of autosomal recessive 3-methylglutacic aciduria. C. Saunders1,2, F. Wibrand1, P. Bross3, K. Ravn2, L. Smith1, I. Thiffault1, A. Atherton2, N. Miller1, E. Farrow1, S. Kingsmore1,2, E. Östergaard2. 1) Center for Pediatric Genomic Medicine, Children’s Mercy Hospital, Kansas City, MO; 2) Department of Pathology and Laboratory Medicine, Children’s Mercy Hospital, Kansas City, MO; 3) Department of Clinical Genetics, Copenhagen University Hospital Righospitalet, Copenhagen, Denmark; 4) Research Unit for Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark; 5) Department of Pediatrics, Section of Genetics, Children’s Mercy Hospital, Kansas City, MO. 
Methylglutacic aciduria comprises a heterogeneous group of five different syndromes of inborn errors of metabolism characterized by mitochondrial dysfunction (3-MGA-uria: MIM#250950). The metabolic landmark is urinary excretion of 3-methylglutacrylic acid and its derivative, 3-methylglutaric acid. Types I-III and V are caused by mutations in four different genes, however, the genetic cause of 3-MGA-uria type IV has not been established. Here we describe a novel, lethal 3-MGA-uria characterized by onset within the first nine months with severe hypotonia, myoclonus, bilateral cataracts, and lactic acidosis, associated with variants in the CLPB gene encoding the ClpB caseinolytic peptidase B homolog. Mutations were identified in five patients, including four Greenlandic individuals from three families not known to be related to Nuna-Arctic North. Homozygosity mapping in the Greenlandic patients identified a single 4.5 Mb region of interest on chromosome 11, encompassing 62 genes. Sequencing of putative mitochondrial genes in this region yielded the same homozygous CLPB variant, c.803C>T (p.Thr268Met), in all four patients. The variant changes a highly conserved threonine to a methionine, located in the ankyrin-repeat domain. It is predicted to be pathogenic by SIFT, PolyPhen2 and MutationTaster. Screening of Greenland controls showed a carrier frequency of 3.3%, similar to other severe recessive diseases in this population. Trio exome sequencing of the North American family revealed two nonsense variants in CLPB, p.Lys361* and p.Arg417*. Western blot analysis showed absence of CLPB protein in fibroblasts from the Greenland patients and liver from the North American patient. CLPB encodes a mitochondrial chaperone important for mitochondrial protein biogenesis and is required to maintain organelle structure in stress conditions. CLPB has mainly been studied in bacterial ClpB/yeast Hsp104, and belongs to the Clp/Hsp104 family, a group of sequence-related AAA+ proteins containing two consensus ATP-binding sites. Our data indicate that recessive variants in CLPB result in a CLPB homozygous phenotype including a new form of 3-MGA-uria. Further studies in these patients may provide new insights into mammalian mitochondrial dysfunction through protein aggregation in neurodegenerative disorders.

2960M
Leveraging Population Structure to Improve Cause Variant Identification in Exome Sequencing Studies of Mendelian Diseases. R. Brown1, B. H. Le2, A. Eskin2, G. Kichaev2, K. Lohmueller1,4, B. Reversade6, S. Nelson2,3, B. Pasaniuc1,2,1. 1) Bioinformatics Interdepartmental Program, UCLA, Los Angeles, CA; 2) Department of Pathology and Laboratory Medicine, Geffen School of Medicine, UCLA, Los Angeles, CA; 3) Department of Human Genetics, Geffen School of Medicine, UCLA, Los Angeles, CA; 4) Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia; 5) Institute of Medical Biology, A*STAR, Singapore.
Recent breakthroughs in whole exome sequencing technologies have made possible the identification of many causal variants of Mendelian (mono- genic) diseases. Although extremely powerful when causative variants and large family trios are simultaneously sequenced, exome sequencing of related individuals with Mendelian traits is often unsuccessful due to the large number of variants that need to be followed-up for validation. The standard approach for reducing the number of putative causal variants is to follow-up by 42% without filtering the true causal variants.

2961T
Truncating Mutation in CIB2 causes DFNB48 and not USH1J. K.T. Booth1, K. Kahrizi2, A.C. Simpson3, A.E. Shearer1, C.M. Sloan1, H. Najmabadi1, H. Azizie2, R.J. Smith1,1) Otolaryngology - Head & Neck Surgery, University of Iowa, Iowa City, IA; 2) Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.
Deafness is the most frequent sensory disorder worldwide affecting one out of every thousand persons. Currently there are over 160 genes causally related to hearing loss and of those 55 are associated with autosomal recessive non syndromic hearing loss (ARNSHL) making deafness extremely heterogeneous. Here we report a consanguineous family of Iranian origin in which we identified a homozygous truncating mutation in CIB2 that is taken into account (a 38% reduction). Most importantly, our proposed methods reduce the number of variants to be followed-up from an average of over 100000 to about 5000 variants in one affected individual. The deafness-causing gene in this family mapped to the DFNB48/USH1J locus and WES data revealed a novel homozygous large deletion encompassing exon 2 of CIB2 gene. Variants in CIB2 are known to cause ARNSHL. The deletion identified leads to a frameshift and premature stop codon at amino acid 24, c.52_68del, yielding a null allele. We used long-range PCR and Sanger sequencing to map the breakpoints to a 3.1 kb interval that spans CIB2 exon 2 and includes several regulatory elements (LINE, SINE...) on either side of the deletion. This type of sequence homology favors genomic rearrangement by non-allelic homologous recombination. CIB2 encodes the calcium- and integrin-binding protein 2 (CIB2), which is expressed in both the retina and inner ear, where it localizes to the tips of stereocilia. In the inner ear CIB2 interacts with several integral membrane proteins. To date, only four missense mutations have been described in CIB2. Two of them represent a major cause of ARNSHL in the Pakistani population. This study is the first to report of a copy number variation associated with CIB2-related deafness. This finding further uncovers the genetic spectrum of deafness in the Iranian population.

2962S
A Perturbed Transcription Underpins Cornelia de Lange Syndrome and Related Phenotypes. B. Yuan1, D. Peihlan1, E. Karaca2, N. Pater1, T. Garnbin1, C. Gonzaga-Jaureguy1, V.R. Sutton1, G. Yesil3, S.T. Bozdag4, T. Tos5, E. Kopo6, C.R. Beck7, S. Gu8, H. Aslan9, O.O. Yuregir10, K. Rubeaa11, D. Nakeeb4, M. Alshammari1, Y. Bayram1, M.M. Atik1, H. Aydin2, D.M. Kivrak11, S. Gu10, D. Boenewinkel11,12, B. Tuyus13, F.S. Alkuraya1,2, R.A. Gibbs1,1,1, R.J. Lupski11,14,15, B. Hopkins Center for Mendelian Genomics, 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; 3) Department of Medical Genetics, Bezmialem University, Istanbul, Turkey; 4) Department of Medical Genetics, Mersin University, Mersin, Turkey; 5) Department of Medical Genetics, Sami Ulus Children’s Hospital, Ankara, Turkey; 6) Department of Medical Genetics, Cerrahpasa Medical School of Istanbul University, Istanbul, Turkey; 7) Department of Medical Genetics, Adana Numune Hospital, Adana, Turkey; 8) University Diabetes Center, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 9) Department of Medical Genetics, Cerrahpasa Medical School of Istanbul University, Istanbul, Turkey; 10) Center of Genetics Diagnosis, Zeynep Kamil Maternity and Children’s Training and Research Hospital, Istanbul, Turkey; 11) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 12) University of Texas Health Science Center at Houston, Houston, Texas, United States of America; 13) Department of Pediatrics, Division of Medical Genetics, Cerrahpasa School of Medicine, Istanbul, Turkey; 14) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 15) Texas Children’s Hospital, Houston, TX, USA.
Cornelia de Lange syndrome (CdLS) is a genetically heterogeneous disorder manifesting extensive phenotypic variability. To date, mutations in NIPBL, SMC1A, SMC3, RAD23 and HDAC8, which encode subunits or regulators of the cohesin complex, are found in about 65% of patients. Furthermore, in families with a "cohesinopathy" diagnosis, whole exome sequencing (WES) revealed a hemizygous predicted-deleterious, missense mutation in SMC1A. Extensive clinical evaluation and WES of a Turkish cohort of 30 patients clinically diagnosed with CdLS revealed a de novo heterogeneous non-sense KMT2A mutation in one patient without characteristic WDS features. Moreover, a de novo heterozygous frameshift mutation in SMC3 was identified in a patient with combined CdLS and WDS features. In these families, the causative gene was ultimately identified to be KMT2A in two patients with CdLS, revealing a hemizygous missense mutation in TAF6F, which encodes a core component in a transcriptional regulation pathway. Our findings suggest CdLS and related phenotypes may result from a "transcriptomopathy" rather than a cohesinopathy: a conclusion supported by recent transcriptomic studies.

Posters: Molecular Basis of Mendelian Disorders

747

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Deep re-sequencing of CFTR bearing the common F508del mutation reveals a rare variant associating with variation in lung infection. B. Vecchio-Papain,3 M. Lee,4 M.R. Knowles3, S.M. Blackman,2 G.R. Cutting2
1 McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; 2 Division of Pediatric Endocrinology, Johns Hopkins University School of Medicine, Baltimore, MD; 3 Cystic Fibrosis/Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Common and rare intragenic variation has the potential to modify protein expression that may correlate with severity of disease phenotypes. Here we describe deep re-sequencing of a well-characterized locus, CFTR, in 482 patients homozygous for the most common disease mutation, F508del (rs113993960). Variants in CFTR cause the autosomal recessive disorder, cystic fibrosis (MIM #219700), in which the primary cause of mortality is lung disease complicated by infection. A custom designed capture of the 215kb region containing CFTR in 964 chromosomes harboring the F508del mutation revealed 272 common variants (MAF > 1%), and 307 rare variants. 301 rare SNPs and small INDELs present on 6 or more chromosomes were used to define linkage disequilibrium (LD) blocks and construct unique haplotypes in this population. Two major LD blocks were observed, and could be explained by a recombination event in intron 16. These results further refine previous reports of recombination occurring within intron 22. The most common haplotype in both of these blocks accounts for 63% of all F508del chromosomes, suggesting that this is the ancestral F508del background. The remaining 37% of F508del chromosomes are combinations of minor haplotypes from both blocks, indicating a wealth of genotypic diversity with the potential for modifying disease severity. Testing of 579 total variants (~32 independent tests due to LD) with cystic fibrosis disease-specific traits revealed association between the rs35453239 A allele and later age at first lung infection (p=9.16e-5, Bonferroni corrected multiplex families with ARNSD. Families in our sample set originated from Turkey, Iran, Saudi Arabia, Mexico, Ecuador, and Puerto Rico. Agilent SureSelect Human All Exon 50 Mb kits and an Illumina Hiseq2000 instrument were used. GATK (Genome Analysis Tool Kit) was used for detection of Single Nucleotide Variants (SNVs) and insertion/deletions (INDELs); CoNi-FER (Copy Number Inference From Exome Reads) and XHMM (eXome-Hidden Markov Model) were used for Copy Number Variation (CNV) detection. Autozygous regions as extended runs of homozygous genotypes in the studied samples were sought via WES data with AgileGenotyper and AgileVariantMapper. Candidate variants were confirmed with Sanger sequencing and evaluated for co-segregation with deafness in all available family members. We identified 15 rare variants in 0.6% of the families in our sample set. The 10 most common genes with either an SNV or an INDEL are MYO15A, SLC26A4, TMC1, TPMRSS3, ILDR1, LOXHD1, MYO7A, OTOF, PCDH15, and USH1C. Homozygous CNVs were detected in each of these 10 genes. In the remaining 84 families, 3453 SNVs and 203 CNVs were identified as pathogenic in novel deafness genes in known deafness genes is used to identify novel deafness genes. In conclusion, comprehensive analysis of SNVs, INDELs and CNVs via WES allows us to identify causative variants in both known and novel genes thus improving our ability to explain the underlying etiology in more families.

Resolving clinical diagnoses for syndromic cleft lip and palate phenotypes using whole-exome sequencing. A. Collins1, L. Arias, J. Martinez2, R. Upstill-Goddard4, R. Pengelly3, J. Gibson1, S. Ennis1, I. Briceno5, 1 Human Genetics, Univ Southampton, Southampton, United Kingdom; 2 Department of Biomedical Sciences, Medical School, Universidad de La Sabana, Bogota, Colombia.

Individuals from three families ascertained in Bogota, Colombia, showing unusual syndromic phenotypes which included cleft lip and/or palate were exome sequenced. In each case exome sequencing revealed the underlying causal variations confirming or establishing diagnoses. The findings include rare and novel variants which provide insights into genotype and phenotype relationships for the conditions. These include the molecular diagnosis of an individual with Nager syndrome and another family exhibiting an atypical Incontinentia Pigmenti phenotype for which we identified a previously reported missense mutation in the IKBKG gene. In this family this rare mutation is associated with a less severe phenotype in which affected male relatives, atypically, survive to full term. The third family exhibited unusual and variably penetrant phenotypic features suggesting the Pierre Robin Sequence (PRS). Affected individuals were found to share a novel mutation in the IRF6 gene which has not previously been associated with PRS. Exome sequencing is a powerful and increasingly cost-effective route to establishing molecular diagnoses for conditions showing phenotypic and/or genetic heterogeneity. It is particularly effective for conditions featuring cleft lip and/or palate for which a substantial proportion of underlying causal genes have not yet been identified.
2966M
Recurrent mutations cause Ablepharon-macrostomia syndrome and Barber-Say syndrome. T. Davis1, S. Marchegiani1, F. Tessadori2, C. Markello3, H. Huang4, D. Schanze5, G. van Haften6, V. Maduro7, F. Brancati8, B. de Vries9, C. Boerkoel10, S. Lin11, W. Gali12, M. Zenger13, T. Markello14, A. B. Barber15, C. Reilly15, R. M. Lachman15, and T. M. Olson16. 1) St. John's Institute of Dermatology, London, UK; 2) University of Cologne, Cologne, Germany; 3) Department of Internal Medicine, University of Washington, Seattle, WA, USA; 4) Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, USA; 5) Division of Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; 6) Department of Pediatrics, University of Florida, Gainesville, FL, USA; 7) Division of Pediatric and Adult Cardiology, Mayo Clinic, Rochester, MN; 8) Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX; 9) Department of Pediatrics, University of Tennessee, Knoxville, TN, USA; 10) Department of Pediatrics and Children's Research Institute, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 11) Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 12) Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 13) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA; 14) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA; 15) Department of Cardiovascular and Peripheral Nerve Research, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 16) Office of the Director, National Institutes of Health, Bethesda, MD, USA.

2968S
Homozgyous LRRRC10 Mutation in Sporadic Pediatric Diabetic Cardiomyopathy. P.A. Longi1, J.M. Evans2, B.T. Larsen2, A.C. Grimes3, Y. Lee4, R.G. Bush5, M. Osterholm5, T. Reis6, F. Tessadori7, E.V. Semina4, T. Markello1, and T. M. Olson6. 1) Department of Pediatrics and Children's Research Institute, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 2) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA; 3) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 4) Division of Cardiovascular and Peripheral Nerve Research, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 5) Office of the Director, National Institutes of Health, Bethesda, MD, USA; 6) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA; 7) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA.

2967T
Mutations in MAB21L2 result in ocular coloboma, microcornea, and cataracts. B. Dam1, A. Kaniminejad2, R.H.R. Boudjerrad3, S. Muhseni4, L. Reis5, E.V. Semina6,1. 1) Department of Pediatrics and Children's Research Institute, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 2) Department of Pediatrics, University of California-Los Angeles, Los Angeles, CA, USA; 3) Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; 4) Department of Genetics, National Human Genome Research Institute, Bethesda, MD, USA; 5) Department of Pediatrics, Children's Research Institute, Children's Hospital of Philadelphia, Philadelphia, PA, USA; 6) Department of Ophthalmology, University of Wisconsin-Madison, Madison, WI, USA.

2969M
Mutation in ANKFY1 as a Cause of Charcot-Marie-Tooth Neuropathy. M. Park1, B. Choi2, S. Choe3, H. Woo4, H. Jo5, K. Chung6, S. Koo1. 1) Division of Intractable Diseases, National Institute of Health, Chungcheongbuk-do, South Korea; 2) Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea; 3) Department of Microbiology, School of Medicine, Wonkwang University, Iksan, South Korea; 4) Department of Biological Science, Kongju National University, Gongju, South Korea.

LRRC10 displays no phenotypic evidence of DCM on screening echocardiography. Array comparative genomic hybridization detected no clinically reportable chromosomal aberrations. Whole exome sequencing of the family trio revealed approximately 40,000 variants per individual in the 50 megabase capture region. Ingenuity® Variant Analysis® was utilized to filter for rare, functionally significant variants and model homozygous recessive, compound heterozygous, de novo, and uniparental disomy modes of inheritance, resulting in only 3 candidate variants. Homozygosity for c.584T>C, p.Ile195Thr in the leucine rich repeat containing 10 gene (LRRC10 [MIM 610846]) was identified in the proband, the only variant predicted to be damaging by SIFT and PolyPhen2. Homozygosity for 1195T occurs at a frequency of 0.44% in the population. However, no instances of compound heterozygosity or homozygosity of missense, splice site, or frameshift variants in LRRC10 were identified in 7,956 exomes in Exome Variant Server and 1000 Genomes. While the molecular function of LRRC10 remains unknown, it was shown previously to have cardiac-specific expression in zebrafish, mice, and humans, with markedly increased expression at birth that is maintained through adulthood in mice. Knockout of Lmnc1 in mice results in an autosomal recessive, perinatal lethal phenotype.

LRRC10 exhibits an autosomal dominant mode of inheritance and remains clinically silent until adulthood. We sought to identify the genetic basis of severe, non-syndromic DCM diagnosed in a two-month-old girl who underwent cardiac transplantation two months later. Her non-consanguineous parents displayed no phenotypic evidence of DCM on screening echocardiography. Array comparative genomic hybridization detected no clinically reportable chromosomal aberrations. Whole exome sequencing of the family trio revealed approximately 40,000 variants per individual in the 50 megabase capture region. Ingenuity® Variant Analysis® was utilized to filter for rare, functionally significant variants and model homozygous recessive, compound heterozygous, de novo, and uniparental disomy modes of inheritance, resulting in only 3 candidate variants. Homozygosity for c.584T>C, p.Ile195Thr in the leucine rich repeat containing 10 gene (LRRC10 [MIM 610846]) was identified in the proband, the only variant predicted to be damaging by SIFT and PolyPhen2. Homozygosity for 1195T occurs at a frequency of 0.44% in the population. However, no instances of compound heterozygosity or homozygosity of missense, splice site, or frameshift variants in LRRC10 were identified in 7,956 exomes in Exome Variant Server and 1000 Genomes. While the molecular function of LRRC10 remains unknown, it was shown previously to have cardiac-specific expression in zebrafish, mice, and humans, with markedly increased expression at birth that is maintained through adulthood in mice. Knockout of Lmnc1 in mice results in an autosomal recessive, perinatal lethal phenotype.
Charaterization of mutation negative autosomal dominant polycystic kidney disease families using whole exome sequencing. B.M. Paul, K. Hoop, J.L. Sundsback, C.M. Hever, V.E. Torres, P.C. Harris. Division of Nephrology and Hypertension, Mayo Clinic, Rochester, MN.

Autosomal Dominant Polycystic Kidney Disease (ADPKD) is a common (frequency 1:1000), adult onset nephropathy accounting for 4-10% of patients requiring dialysis or renal transplant. An annual US cost of >$1 billion. Mutations to either PKD1 or PKD2 cause most ADPKD. However, the description of ADPKD families unlinked to either gene, and no mutations detected in ~8% of ADPKD families, suggests further genetic heterogeneity. We have identified ~130 (NMD) families out of ~1500 screened. Interestingly, they are characterized by milder kidney disease and more frequently have negative family history than is typical in ADPKD. Initially, we analyzed two families by whole exome sequencing (WES), P75 (3 affected, 2 unaffected) and M560 (3 affected). Library preparation and capture was done using the Agilent SureSelectXT Human All Exon V5+UTRs on the Agilent Bravo workstation and sequencing employed on Illumina HiSeq2000. Alignment was performed using Bowtie/BWA and GATK used for variant calling according to best standard practices. Data mining was performed using the Golden Helix SNP and Variation suite (SVS). Initial filtering was set by read depth (10X), genotype quality (≥20) and assignment of autosomal dominant inheritance according to the affected status of the individuals. The ESP6500 and 1000 Genomes databases were used to remove variants with a MAF of ≥0.1% and dbNSFP to characterize coding variants. Independent VAAST analysis was also employed. We found a PKD1 nonsense mutation in M660, previously missed by Sanger sequencing. In P75, 15 candidate variants were detected by both SBS and VAAST and six by VAAST alone. We observed six variants identified by at least one tool but not by the other two tools. We are prioritizing these ten candidates by their likely localization or because of poor consensus score in dbNSFP. Five variants were annotated as damaging by four prediction tools in dbNSFP and five others by VAAST alone. Eleven variants were removed due to misalignment or because of poorly mapping reads. These two tools. We are prioritizing these ten candidates by their likely localization or because of poor consensus score in dbNSFP. Five variants were annotated as damaging by four prediction tools in dbNSFP and five others by VAAST alone. Eleven variants were removed due to misalignment or because of poorly mapping reads. We are selecting large deletions/insertions, duplication/deletions, and large rearrangements detected by the Agilent HaloPlex in our affected families.

Whole Exome Sequencing to Uncover Causative Genes in Families with Inherited Autonomic Dysfunction. J.E. Posey,1 T. Gambin,1 S.N. Jhangiani,2 D.M. Muzny,4 R. Martinez,5 M.T. Muman,2 W. Wisniewski,1 R.A. Gibbs,1,2 J.L. Butler,4,5 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; 2) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA; 3) Division of Child and Adolescent Neurology, Department of Pediatrics, The University of Texas Health Science Center, Houston, TX 77030, USA; 4) Division of Cardiology, Department of Pediatrics, The University of Texas Health Science Center, Houston, TX 77030, USA; 5) Department of Pediatric Hematology/Oncology, Children’s Memorial Hospital, Chicago, IL 60660, USA; 6) Children’s Texas Children’s Hospital, Houston, TX 77030, USA.

Autonomic dysfunction can be a crippling medical condition, and is characterized by a variety of symptoms including dizziness, palpitations, syncope, temperature dysregulation, increased pain sensitivity, and gastrointestinal distress. These clinical manifestations may be associated with several well-defined genetic syndromes, including Hereditary Sensory and Autonomic Neuropathy, Ehlers-Danlos Syndrome or Familial Dysautonomia [OMIM 223900]. Familial Dysautonomia, also known as Riley-Day Syndrome, is an autosomal recessive type 3 Hereditary Sensory and Autonomic Neuropathy in which patients suffer from cardiovascular instability, gastrointestinal dysfunction, and altered sensitivity to pain and temperature. We identified 14 families in which autonomic dysfunction is inherited in an autosomal dominant manner. Extensive clinical characterization has demonstrated a very specific phenotype that includes postural orthostatic tachycardia syndrome (POTS), atrial septal defect (ASD), headaches, and gastrointestinal distress. These patients do not have decreased or absent lingual fungiform papillae, a hallmark of Familial Dysautonomia, and they do not respond to metanephrine or clonidine. In Jewish kindred, the only family with a known genetic diagnosis, the variant has been passed from generation to generation in a mendelian fashion. We performed whole exome sequencing and identified the mutation in SHK1 (p.Arg1163Trp) in a highly conserved protein domain of ZSWIM6. Sanger validation of the three trios confirmed its de novo status in the three probands. A single instance of unique and undiagnosed disorders. Genetic analysis of the literature, we are performing various follow-up studies on these and other genetic disorders to further study. To date, 550 families have participated, of which 57% represent families with similar phenotypes to develop a cohort to create a critical mass for further study. Enrollments is open to any family with a rare or unknown diagnosis that has a suspected, yet unidentified, genetic etiology despite extensive clinical testing. Following identification of a new family with a rare condition, we often ascertain and enroll additional families with similar phenotypes to develop a cohort to create a critical mass for further study. To date, 550 families have participated, of which 57% represent single instances of unique and undiagnosed disorders. Genetic analysis has been initiated for 263 families: 234 of which received whole exomic genotyping in a single institution in a CLIA-certified laboratory. For 20 families, variants were identified in novel genes without prior reports of association with human disease, and suspected to be pathogenic based on in silico analysis, gene expression data and pathway involvement. Following review of the literature, we are performing various follow-up studies on these cases, including clinical follow-up, and using a variety of tools, such as patients’ -derived cell or tissue samples, to evaluate causality. Whenever possible, functional assays are developed and tailored for each gene in question. We have successfully linked the phenotype to these novel genes in three of these 20 cases, and we are actively working on the remaining families.
2974S
Exome sequencing identifies the cause of a novel multiple pterygium syndrome and expands the spectrum of phenotypes caused by variants in MYH3, encoding embryonic myosin, cause Distal Arthro-gyposis type 1A (DA1), Sheldon-Hall syndrome (SHS or DA2A), and Freeman-Sheldon syndrome (FSS or DA2A). Each of these conditions is characterized by multiple congenital contractures (MCC) but have distinct clinical characteristics and natural histories that enable them to be distinguished from one another. Individuals with DA type 8 (DA8) likewise have MCC but also have short stature, severe scoliosis and multiple pterygia reminiscent of multiple pterygium syndrome (MPS). Sanger sequencing of candidate genes (TNNT2, TNN3, TPM2, CHRNG, and RAP55), which underlie DA1, DA2B, and MPS in a parent and two affected children with DA8 failed to identify a compelling causal variant. Subsequently exome sequencing revealed a novel 3-base pair duplication in exon 25 of MYH3, c.3214dup. While the protein of DA8 is similar to FSS and SHS, it was distinct enough for MYH3 to not have been considered a candidate gene for MCC. Exome sequencing enabled a more thorough description of variants in MCC that could explain the spectrum of phenotypes caused by variants in MYH3. Multiple rare homozygous and de novo variants in MYH3 were detected by our whole exome sequencing. The spectrum of phenotypes associated with MYH3 was expanded, and it appears that MYH3 is a disease gene that is more common than previously thought.

2974T
Clinical study of the effects of retinal dystrophy on vision.

2975M

Lysosomes are membrane-bound, acidic eukaryotic cellular organelles. As an enzyme container, they play important roles in the degradation of macromolecules. Monogenic mutations resulting in the loss of enzyme activities in the lysosome may lead to severe health problems, such as neurodegeneration, early death, etc. These conditions are categorized as lysosomal storage diseases (LSDs). The diagnosis of LSDs is typically straightforward, but in some cases, mutations that result in atypical clinical presentation or defects in previously undescribed lysosomal disease genes may complicate the identification of the underlying genetic defect. Here, we performed whole exome sequencing on 14 suspected LSD cases, with the goal of finding the causal mutations in each case. From the raw sequence data, we first identified DNA variants in each individual using three variant discovery pipelines: the Genome Analysis Toolkit, LifeScope and CLC Genomics Workbench. We then used the Variant Annotation Analysis Selection Tool (VAAST) to prioritize disease-causing mutations in 84 candidate LSD genes. As a probabilistic disease gene finder, VAAST integrates allele frequency, amino acid substitution severity and conservation information into a composite likelihood framework. Different from hard filtering methods, VAAST preserves all the candidates by listing them according to their disease-causing potential. So far, 14 number of candidate variants have been identified, and we are performing downstream mutation validation and proteome analyses to investigate the potential connection between our candidate variant and LSDs. Our project makes use of bioinformatics analyses to decode enormous exome sequencing data, narrowing down candidate lists and largely increasing the efficiency of downstream proteomic studies. Our results will shed light on the genetic basis of LSDs.
2977S
Pathogenic mutation of coagulation factor X deficiency may prevent atypical hemolytic uremic syndrome. F. Bu1,2, Y. Li1,*, L. Lin3, S. Lin3, J. Wang1, Q. Hu1, X. Wang1, S. Liu1, Y. Xu1,*, L. Zhang1, L. Wang4, M. Sun1, 2,5, Y. Shi1,2, R. Smith1,2,5, J. Shi1,2,5, 1) Interdepartmental PhD Program in Genetics, University of Iowa, Iowa City, IA; 2) Molecular Otolaryngology and Renal Research Laboratories, University of Iowa, Iowa City, IA; 3) Department of Biochemistry, University of Iowa, Iowa City, IA, USA; 4) Department of Otolaryngology, College of Public Health, University of Iowa, Iowa City, IA; 5) Rare Renal Disease Clinic, Departments of Pediatrics and Internal Medicine, Carver College of Medicine, University of Iowa, Iowa City, IA.

Introduction: Atypical hemolytic uremic syndrome (aHUS) is an ultra-rare renal disease caused by uncontrolled activation of the alternative pathway of complement at the cell surface. Approximately 22% of patients carry pathogenic mutations in CFH, the penetrance of which ranges from 12.5% to 100% in familial studies. We hypothesized that genetic modifiers exist that impact the penetrance of aHUS. Patients & Methods: Five families carrying a well-recognized aHUS mutation - c.3644G>A, p.Arg1215Gln - in CFH were studied. Using targeted genomic enrichment and massively parallel sequencing, we screened subjects for variants in coding exons of all genes in the complement and coagulation cascades based on the recognized cross-talk between these two pathways. Data were analyzed using locally implemented Galaxy software. Pathogenic variants were annotated based on computational predictions and the literature. The functional impact of select variants was confirmed using in vitro assays. Results: We identified a reported factor VII deficiency variant (‘Padua’ variant; F7 c.1091G>A, p.Arg364Gln) and a reported factor X deficiency variant (F10 c.424G>A, p.Glu142Lys) in two families that may impact the penetrance of aHUS. The factor X variant segregated in a three-generation family in which two children under 5 years of age developed aHUS. Neither child carried the F10.p.Glu142Lys mutation. All older family members, none of whom developed aHUS, segregated BOTH the F10.p.Glu142Lys variant and the FCP.h.Arg1215Gln variant.ab aHUS-affected individual. Protein structure simulations show the mutated positively charged lysine located at the second residue of a second messenger inositol-polyphosphate. Homozygous Impa1 knockout mice die in utero, but can be rescued by supplementation of myo-inositol. Deficiency of Impa1 might represent a new inborn error of metabolism, affecting inositol-phosphate biotransformation.

2979T
Prospecting genetic disorders in a highly inbred region of Brazil: two novel genes for AR intellectual deficiency. T. Figuiredo1,2,*, U. Souto1,2, M. Bussato1, R. Oliveira1,2, L. Azevedo1, H. de Azevedo1, E. Amaral1,2, R. Dantas1,2, 1) Núcleo de Estudos em Genética e Educação, Universidade Estadual da Paraíba, Campina Grande, Brazil; 2) Centro de Pesquisas sobre o Genoma Humano e Celulas-Tronco, São Paulo, Brazil; 3) Rede Nordeste de Biotecnologia, RENORBIO, Universidade Federal da Paraíba, João Pessoa, Brazil; 4) Faculdade de Medicina, Universidade de Fortaleza, Fortaleza, Brazil; 5) Faculdade de Medicina, Departamento de Neurologia, Universidade de São Paulo, São Paulo, Brazil.

Reduction of penetrance for a highly heterogeneous condition affecting 3% of the population worldwide. In a field study conducted in a highly inbred area of Northeastern Brazil, we investigated two large consanguineous families with ID. Genome-Wide Human SNP Array 6.0 (Affymetrix) microarrays have been used to determine the genetic defects in affected and normal individuals in each family. Whole exome sequencing (WES) was performed in one affected individual of each family using Nextera Rapid-Capture Exome and Illumina HiSeq2500. Potentially deleterious variants were studied in regions of homozygosity-by-descent and not present among 8,000 controls (including 600 Brazilians) were subject to further scrutiny and segregation analysis by Sanger sequencing. Family A has 9 affected descendants from four closely related first-cousin couples affected by severe non-syndromic ID associated to disrupted pastinac gene behavior. Homozygosity-by-descent analysis disclosed a 20.7 Mb region in 1q12.3-q21.2 (lod score: 3.11). WES identified a homozgyous deleterious variant in inositol monophosphatase1 gene (IMPA1), consisting of a 5 bp duplication (c.489_493dupGGGCT) leading to frameshift (p.Ser165Trpfs‘10). IMPA1 gene product, a C-terminal domain of the Mediator complex, involved in regulation of transcription of nearly all RNA polymerase II-dependent genes. Deleterious mutations in IMPA1 have been already associated with ID. These findings demonstrate that the combination of field investigation of large families in highly inbred regions with modern NGS technologies is an effective way to identify new genes and pathways which can be the target of future treatment.

2978M
Genetic testing with targeted exon enrichment and massively parallel sequencing for 272 Chinese cases with hearing loss. J. Cheng1, Y. Lu1,*, Z. Jiang1, C. Li2, X. Zhang1, H. Duan1, D. Han1, H. Yuan1, 1) Inst Otolaryngology, Chinese PLA Gen Hosp, beijing, china; 2) Center for Genetic & Genomic Analysis, Genesky Biotechnologies Inc., Shanghai, China.

Hereditary non-syndromic hearing loss (NSHL) is extremely heterogeneous. The large reservoir of known deafness genes precluded comprehensive genetic testing and population-scale sequencing. Until recently, the combination of targeted genomic capture and massively parallel sequencing (MPS) has become a promising tool for detecting novel and known mutations involved in hereditary hearing loss. In this study, we aimed to establish Chinese population-level frequencies of reported deafness-causing variants in known genes related hearing loss and to determine the genetic defects for 272 Chinese families and 170 congenital deafness cases. We performed a targeted enrichment that enables the capture of candidate genes in known regions of the human genome associated with hearing loss and some candidate genes in some important pathways related to hearing. With a single reaction for high throughput mutation identification by massively parallel sequencing, in the present study, we have sequenced 37 genes related to hearing loss by MPS. Mutations in 21 known DFNA genes were identified in 27 cases with progressive sensorineural hearing loss, and 22 known DFNB genes are identified in 44 cases with congenital deafness. For the potential candidates, the confirmation by Sanger sequencing and co-segregation analysis has been done. Besides the frequency of mutations in known gene was dissolved, novel genes and mutations are described. Discovery of these new genes will continue to help define compelling mechanisms for deafness.

2980S
Molecular Basis of Niemann-Pick A-B and Neimann-Pick C Diseases in the Aegean Region of Turkey: Identification of Three Novel Mutations. M. Zatz1,2,*, A. Brandao2,5, H. Yilmaz1, S. Kalkan2,5, E. Butun2,5, E. Onay2,5, A. Korkmaz1, E. Turan2,5, A. Korkmaz1, Y. Cakar2,5, H. Azaiez3,*, C. Thomas2,5, H. Yuan2,5, 1) Interdepartmental PhD Program in Genetics, University of Iowa, Iowa City, IA, USA; 2) Ege University Faculty of Medicine, Department of Medical Genetics, İzmir, Turkey; 3) Ege University Faculty of Medicine, Department of Pediatrics, İzmir, Turkey.

The Niemann-Pick disease group consists of two different entities: (1) acid sphingomyelinase-deficient forms which are caused by SMPD1 gene mutations and encompass type A and type B, as well as intermediate forms; (2) Niemann-Pick disease type C is a cellular lipid trafficking disorder caused by NPC1 or NPC2 gene mutations. Niemann-Pick disease type C patients present with enlarged livers and spleens, a failure to thrive, and neurological manifestations in the first 6 months of life resulting in death before 3 years of age. Niemann-Pick type B patients have little or no neurological involvement, and mutations affect a different class of lipids in the brain. Here, we report three novel mutations in Niemann-Pick disease type B and C in families originating from the Aegean Region of Turkey. Sequencing was performed using the direct sequencing method. There were 42 mutant SMPD1 alleles found with nine different types detected amongst these. The most prevalent mutation was p.L137P with an allelic frequency of 47.6 %. In total, 32 mutations, namely p.L137P and p.T189fsX65, accounted for 69% of the total mutated alleles. Three novel mutations (L161P, W176C, T397M) were detected in the group studied. Among these, L161P was found in 11.9% of the unrelated mutant alleles and was the third most common mutation of the patients studied. Of the patients studied, 31/32 with the novel mutation W176C and previously described I178N mutation. Another patient who had very mild symptoms and low enzyme levels was homozygous for the novel T397M mutation. Four different mutations were detected in NPC1 gene, with eight unaffected individuals and one affected individual, suspected of having NPC disease. One of the affected was homozygous for c.839delT(p.L280CfsX30) mutation, the other for p.N906Y mutation; and two were homozygous for G982R and A558T mutations. No NPC2 mutation was detected. The results were useful in genetic counseling and prenatal diagnostic services in our region.
2981M
Frataxin, a Fredrich's ataxia protein is defective in mitochondrial pro-
cessing peptidase-alpha (PMPCA) mutations. P.B. Agrawal1, M. Joshi1, I. Anselmi1, F. Gianini1, M. Towne2, K. Schmitz-Abel1, K. Markianos1, V.G. Sankaran1. 1) Department of Medicine, Boston Children's Hospital and Har-
vard Medical School, Boston, MA; 2) Department of Neurology, Boston Children's Hospital and Harvard Medical School, Boston, MA.

A 4-year old girl with an unknown mitochondrial disease of infant onset had multisys-
tem involvement including cardiomyopathy, hypotonia, respiratory insuffi-
ciency, myoclonic jerks and blindness. Her cousin (their fathers are brothers and mothers are first cousins), a boy died at 14 months age due to a similar presentation. A whole-exome sequencing was performed in the proband. A compound heterozygous PMPCA (MIM 613036) mutations (c.G1066A; p.G356S and c.G1128A; p.A377T) were identified in the proband. A comprehensive genomic analysis for mitochondrial respiratory chain (MRCD) disorders was performed in the remaining family members. We identified the genetic background in a total of 29 individuals of our precisely controlled strategy will be presented at the meeting. The diagnosis of mitochondrial disease currently involves multiple invasive procedures, which carry the risk of complications, and in many cases a diagnosis is made months or years after onset of symptoms. Identifying disease-causing genes is therefore of high clinical importance. Identification of disease-causing genes is therefore of high clinical importance.

2983S
A comprehensive genomic analysis for mitochondrial respiratory chain disor-
lar, Chiba, Japan; 6) Dept. of Pediatrics, Matsudo City Hospital, Chiba, Japan.

Mitochondria are small organelles and serve as the powerhouse of the living cells because they generate vital energy in their respiratory chain mechanisms of MRCD. A comprehensive genomic analysis for mitochondrial respiratory chain disorder (MRCD) is an intractable disease that develops in childhood. It is a highly frequent inborn errors of metabolism that occurs in at least one out of every 7,000 births. Prominent symptoms develop in such organs as the brain, heart, and muscles, where a great deal of energy is required. In most cases, effective treatment has not been established yet. It is known that various gene abnormalities cause defects of the protein complexes of the respiratory chain, which results in mitochondrial dysfunction. However, identification of the causative genes and their replacing of pathogenic mechanisms of MRCD remain largely unsolved.

2984M
Diagnosing mitochondrial disease: Is there an additive advantage of whole-exome sequencing? S.J. Mosch1, P.M. Gordon1, L. Dimmke1, S.T. Nakansid1,2, S. Hume1, D.S. Sinasar1,2, L.J. Parboosingh1,2,2, 1, A. Khan1,2,3, 1) Department of Medical Genetics, University of Calgary, Alberta, Canada; 2) Alberta Children's Hospital Research Institute, University of Calgary, Alberta, Canada; 3) Metropolitan Children's Diseases Clinic, Alberta Children's Hospital, Alberta, Canada; 4) Molecular Diagnostics Lab, Alberta Children's Hospital, Alberta, Canada; 5) Biochemical Genetics Laboratory, Alberta Children's Hospital, Alberta, Canada; 6) Department of Physiology and Pharmacology, University of Calgary, Alberta, Canada; 7) Department of Medical Genetics, University of Alberta, Alberta, Canada.

The diagnosis of mitochondrial disease currently involves multiple invasive procedures, which carry the risk of complications, and in many cases a diagnosis is made months or years after onset of symptoms. Identifying disease-causing genes is therefore of high clinical importance. The Calgary Metabolic Clinic reviewed 292 patients with a clinical suspicion of mitochondrial disease, of which 35% had an eventual diagnosis of mitochondrial disease (MDx), 24% had a diagnosis of non-mitochondrial disease (OtherDx), and 41% had no final diagnosis (NoDx). When muscle-extracted mitochondrial DNA (mtDNA) analysis was performed, a diagnosis was made in 66.7% of cases in the MDx group and none in either of the other groups (p<0.001). mtDNA analysis showed the highest specificity compared to muscle histology and enzyme biochemistry. Whole-exome sequencing (WES) was performed on 12 unrelated patients from the NoDx group, who had no mtDNA mutations, and 1 case that had an mtDNA deletion and a suspected causative nuclear mutation. A systematic workflow for the evaluation of variants identified through WES was developed using the Illumina TruSight Othogene list as the starting point. WES was found to be a first-line filter in an attempt to identify mutations in known disease-causing genes. Of the 13 patients, the variant prioritization workflow was able to identify the only known cause of mitochondrial disease, and putative disease causing genes. In 4 patients, a whole-genome approach was used for the molecular diagnosis of suspected mitochondrial disorders, with high potential to lead to specific diagnosis early in the diagnostic workup.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
2985T
Realignment of whole-genome and exome sequencing reads supports novel potassium channel (Kir2.x) isoforms that were formerly identified by Sanger sequencing as polymorphisms of a single channel gene in thyrotropic periodic paralysis locus. M.R. Dias da Silva,1 R.M. Paninka,1 I.S. Kuni,1 G.K. Funuza,1 M.M.L. Kiyos1, R.M.B. Maciel,2,1 M. Mitne-Neto,2 L.T. Cerdeira,1 S. Pinto,1 H. Rodrigues,1 S. Tutk1, D.R. Mazzotti,1 Laboratory of Molecular and Translational Endocrinology, Department of Medicine, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil; 2) Fleury Group, São Paulo, SP, Brazil; 3) Department of Psychobiology, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.

Thyrotropic periodic paralysis (TPP [MMIM 613239]) is the commonest cause of muscle weakness in adults suffering from hyperthyroidism, characterized by transient hypokalaemia and linked to mutations in the potassium channel KCNJ18 (Kir2.6 [MMIM 602323]) due to high homology among family members (95-99% at protein level). The exact positions of KCNJ17 (Kir2.5) and KCNJ18 (Kir2.6) remain unclear. Although a BAC containing KCNJ17 has not been found, both RP11-437N10 (KCNJ18) and RP11-728e14 (KCNJ12) have been mapped to 17p11.1-2. FISH originally localized KCNJ12 and KCNJ17 to 17p11.1. We aimed to identify possible misalignments of whole exome (WES) and whole genome sequencing (WGS) reads in the GRCh37/hg19 reference region of Kir2.2 by performing realignment of those reads to Kir2.2, Kir2.6 and Kir2.5 regions obtained by peripheral blood Sanger cDNA (direct and cloning) sequencing. WGS was performed with paired-end 100bp reads with the Illumina HiSeq2000 at 15X target coverage of 30X. Alignment to the reference genome and variant calling was carried out with a standard BWA/GATK pipeline. We identified a single Kir2.2 and Kir2.5 cDNA, and 3 major Kir2.6 sequences named as RRAI, QHEV and QHAV haplotype based on the amino acid changes at 35, 40, 56 and 249 residue. After inspection of reads aligned to the GRCh37/hg19 reference assembly within the pericentric region of 17p11.1-2 spanning KCNJ12 (chr17:21279699-21323179), a notable number of variants were detected, suggesting misalignment of reads due to low coverage in the GRCh37/hg19 assembly of that region. After realignment to the sequence identified by Sanger, we found 192 and 613 reads aligned to that in the WES and WGS. Among WES reads, 706 (35.4%), 750 (37.6%) and 536 (28.9%) aligned to Kir2.6, Kir2.5 and Kir2.2, respectively. Similarly, in WGS 195 (25.8%), 190 (31.0%) and 192 (43.0%) aligned to Kir2.5 and Kir2.2. Taking together mapping and functional data, our findings support the presence of novel isoforms of Kir channels not adequately mapped to the reference genome. Although pericentromeric duplications are thought to contain heterochromatic DNA and have fewer expressed genes, we have presented several lines of evidence suggesting that Kir2.x paralogues are distinct and may play a role in muscle plasticity as observed in TPP.

2986S
Novel SLC29A3 mutation causing H Syndrome in an Indian Adolescent. N. Kamath1, R.D. Shenoy2, M. Varma3, R. Khubchandani4, S. Babay2, A. Zlotogorski2,5, V. Molho Pesach2,3, 1) Department of Pediatrics, Kasturba Medical College, Manipal University, Mangalore, Karnataka, India; 2) Department of Pediatrics K.S. Hegde Medical Academy, Nitte University, Mangalore, India; 3) Private Rheumatology Clinic, Mumbai, India; 4) Pediatric Rheumatology Clinic, Jai Lok Hospital and Research Institute, Mumbai, India; 5) Center for Genetic Diseases of the Skin and Hair, Hadassah-Hebrew University Medical Center, Jerusalem, Israel; 6) Department of Dermatology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel.

Purpose: H syndrome (OMIM 6027828) is an autosomal recessive genodermatosis with systemic manifestations due to biallelic mutations in the SLC29A3 gene. The major clinical findings include hyperpigmentation, hypothyroidism, hearing loss, heart anomalies, hepatosplenomegaly, hypogonadism, low height (short stature), hallux valgus/flexion contractures and hyperglycemia/diabetes mellitus. It has recently been recognized as an inherited form of Histiocytosis. Genetic analysis of a case with clinical features consistent with H syndrome was performed in search for the causative mutation. Methods: A 16-year-old Indian adolescent, born out of non consanguineous parentage was referred due to short stature, sensorineural hearing loss, insulin dependent diabetes mellitus, delayed puberty and diffuse cutaneous hyperpigmentation. On physical examination she was noted to have height and weight below the third centile in the CDC 2000 growth charts, bilaterally symmetrically indurated sclerotadermatous hyperpigmented patches with accompanying hyperthercrosis overlying the trunk and lower extremities with sparing of the knees and buttocks, prepubertal sexual maturation, hepatosplenomegaly, dilated lateral scleral vessels, arcus senilis, infiltrated cheeks, lateral tibial torsion and hallux valgus. Laboratory evaluation revealed microcytic anemia, highly elevated ESR and hypogonadotropic hypogonadism. DNA was extracted from peripheral blood and analyzed for SLC29A3 mutations by Sanger sequencing. Results: We identified a novel missense homozygous mutation in the fifth exon of SLC29A3: c.677G>A (G209R). This variant was absent in dbsNP (http://www.ncbi.nlm.nih.gov/SNP), the 1000 Genomes project data (http://www.1000genomes.org/) and the Exome Variant Server (http://evs.gs.washington.edu/EVS/). This mutation was predicted to be disease-causing by MutationTaster with a score of 0.99 (http://www.mutationtaster.org/). Conclusions: We demonstrated a novel homozygous mutation in the SLC29A3 gene in an adolescent female of Indian ethnicity with clinical features consistent with H syndrome. Our patient emphasizes the pathognomonic cutaneous phenotype as well as the various extra cutaneous features of this autosomal recessive pleomorphic disorder.

2987M
An intergenic 9.4 kb microduplication at chromosome 5p13 as a cause of brachydactyly type A1. L. Racacho1, S.M. Nikkel2, J. MacKenzie3, C.M. Armour2, M.E. McCready4, Y. De Repentigny5, R. Kohlery5, L.A. Pennacchio5,6, D.E. Bulman1,2,1) Department of Biochemistry, University of Ottawa, Ottawa, ON, Canada; 2) Department of Pediatrics, University of Ottawa, Ottawa, Canada; 3) Department of Genetics, CHEO Research Institute, Children’s Hospital of Eastern Ontario, Ottawa, ON, Canada; 4) Department of Medical Genetics, Kingston General Hospital, Queen’s University, Kingston, ON, Canada; 5) Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada; 6) Regenerative Medicine Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada; 4) Genomics Division, Lawrence Berkeley National Laboratories, Berkeley, CA, USA; 5) US Department of Energy Joint Genomic Institute, Walnut Creek, CA, USA; 6) Newborn Screening Ontario, Children’s Hospital of Eastern Ontario, Ottawa, ON, Canada.

Brachydactyly type A1 (BDA1 [MMIM 112500]) is an autosomal dominant brachymesophalangeal trait, primarily characterized by a hypoplastic and/or aplastic pattern of the middle phalanges of digits 2-5 in the hands and feet. Missense mutations in either HHF or GDF5 have been associated with BDA1. We previously reported the linkage of a BDA1 locus to chromosome 5p13 (BDA1 [MM 607004]) in a single large family. The PCR-based sequencing of all the protein coding genes within the 3 Mb critical region did not reveal a mutation. We did not identify any large genomic rearrangements spanning the critical region with either a 3-color FISH or a 500K SNP array. In order to provide a higher sensitivity of mutation detection, we performed a targeted hybridization enrichment of the BDA1 region on two affected family members followed by high-throughput sequencing. The alignment of the sequence reads from both individuals to the reference genome revealed a novel and shared 9.4 kb intergenic tandem duplication. The microduplication falls within a ‘gene desert’ consisting of several vertebrate conserved sequence blocks. We were able to demonstrate the enhancer activity of these conserved blocks through the use of an in vivo transgenic reporter assay. Relative quantitative PCR on the patient’s fibroblast cDNA showed an up-regulation of distal genes when compared to unaffected controls. Our findings suggest that a cis-regulatory mutation is most likely involved in the pathology of BDA1 in these two individuals.
Novel molecular insights into severe congenital microcephaly through targeted next generation sequencing. G. Mirza1-2, V. Vasta1, S. Chris-tian1, O. Zhang3, S. Eun4, S. Collins5, S. Hahn1-2, W. Dobyns1-2 1) Human Genetics, Seattle Children’s Research Institute, Seattle, WA; 2) Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seat-tle, WA; 3) Fred Hutchinson Cancer Research Center; 4) Department of Pediatrics, Seoul National University College of Medicine, Seoul, Korea; 5) Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, OK USA. Microcephaly (MIC) is a developmental brain growth disorder that accounts for a significant fraction of childhood intellectual disability, autism and epilepsy. It occurs in isolation (primary MIC) and as part of a broad range of neurodevelopmental syndromes with or without other cortical malformations. MIC is genetically very heterogeneous and the rate of identification of novel MIC genes has increased dramatically with the advent of next generation sequencing (NGS). The majority of MIC genes are key regulators of critical processes including mitotic spindle assembly and structure, centrosome function, and DNA repair and damage response (DDR) pathways. We screened a cohort of 93 individuals with MIC using a targeted NGS panel that contains 377 known and candidate MIC genes. Most individuals had severe congenital or postnatal MIC (head circumference 3 standard deviations or more below the mean) with or without other structural brain malforma-tions, such as polymicrogyria, cerebellar and callosal abnormalities. We identified mutations in 26/93 (28%) of patients. Recessive mutations identi-fied in known MIC genes include ASPM (N=7; 7.5%), CENPJ (N=1), CEP135 (N=1), CASC5 (N=1), WDR62 (N=1), ASNS (N=2), RBBP8 (N=1), ADSL (N=1) and RAB3GAP2 (N=1). Heterozygous de novo mutations were identified in TUBA1A (N=2), DYNC1H1 (N=2), TUBB2B (N=1), DYRK1A (N=1), FOXG1 (N=1), MLZL2 (N=1), CASK (N=1), and HCCS (N=1). Our cohort expands a number of phenotypes including HCCS-associated features in males to include MIC, short stature and microopenis without severe eye defects. RBBP8-associated Jawad syndrome in a Pakistani family with MSG but without digital anomalies, adenosylsucinate lyase deficiency (ADSL) in a large sibship with MSG and increased extra-axial space, and Warburg Micro syndrome caused by mutations RAB3GAP to include postnatal MIC, hypoto-nia, and severe developmental encephalopathy without eye involvement. Our cohort also introduces new reports of rare phenotypes including asparagine-aspartyl transcarbamylase deficiency (NAMPT) and postnatal MIC and progressive cerebellar ataxia with mild white matter changes and growth retardation caused by mutations in the EUF1 gene.**

Denovo mutations in a novel disease causing gene cause Temple-Ba-ralitzer syndrome and non-syndromic epilepsy. C. Simons1, L. Rash1, J. Crawford1, K. Ru1, S.M. Grimmer2,2, D. Miller1,2, G. King1,1, J. McGau-thran1,1, M. Gabbett1,3, R.J. Taft1,2,6 1) Institute for Molecular Bioscience, University of Queensland, Jindalee, QLD, Australia; 2) Institute of Cancer Sciences, Translational Research Centre, University of Glasgow, Scotland; 3) Genetic Health Queensland, Royal Brisbane & Women’s Hospital, Austra-lia; 4) School of Medicine, The University of Queensland, Australia; 5) School of Medicine and Health Services, Departments of Integrated Systems Biol-ogy and of Pediatrics, George Washington University, USA; 6) Illumina, Inc., San Diego, CA USA. Temple-Ba-ralitzer syndrome (TBS) is a multi-system developmental disorder characterized by intellectual disability, epilepsy and hypo-or aplasia of the nails of the thumb and great toe. TBS is rare, with only five cases reported to date, although it is possible that that TBS is largely unrecognized and therefore under-reported. We recruited six unrelated TBS individuals and their parents for analysis by whole exome sequencing. Here we report damaging de novo mutations in a novel disease causing gene that is predomin-antly expressed in the central nervous system, in six TBS individuals. Functional characterization of two variants revealed that TBS-associated mutations lead to deleterious gain-of-function. Consistent with this result, we found that two TBS mothers with epilepsy, but who were otherwise healthy, were low level (12% and 30%, respectively) mosaic carriers of pathogenic mutations. This suggests that the etiology in some cases of non-syndromic epilepsy may be explained by low level mosaicism for mutations in this gene.

Molecular genetic characterization of an autosomal recessive Familial Essential Tremor. D. Monies1, E. Naim1, B. Al-Younes1, M. Al-Breaca1, M. Al-Saif1, S. Wakil2, K. Khabar2, B. Meyer1, S. Bohlega1 1) Genetics, RC, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia; 2) Department of Medical Genetics, College of Medicine, King Saud University, Riyadh, Saudi Arabia. Despite the presence of strong family histories the underlying genetics of familial essential tremor (FET) is not well defined. Although several loci have been associated with FET have been described, there exists only relatively poor linkage data and no causative mutations having been identified to date. This far the mode of inheritance for FET is best described as autosomal dominant with incomplete penetrance. We describe the molecular genetic characterization of an autosomal recessive familial disorder characterized by tremor of juvenile onset, dystonia and myoclonus with preserved cognitive, cerebellar and peripheral nervous system functions. Mild spasticity appeared with disease progression and some white matter changes were evident upon MRI. Linkage analysis of a consanguineous family with five affected individuals identified a locus on chromosome 17 with a LOD score >4.0. Homozygosity mapping confirmed a single homozygous region shared by all affected individuals only and consistent with the locus identified by linkage analysis. Whole exome sequencing was performed on 1 affected individual with variants being filtered based upon the linkage interval, absence in variant databases at a frequency >0.1%, presence in coding/flanking regions and homozygosity. A single candidate causative variant segregating with disease in the family was identified. We describe further the functional validation of this variant. This is the first description of FET inherited in an autosomal recessive manner and offers opportunities for the further investigation of tremor associated with other disorders including Parkinson’s disease.

High diagnostic success rate in a cohort of unresolved leukoencepha-lopathy patients investigated by whole exome sequencing. R.J. Taft1,2,3, G. Helman1, A. Pizzino1, A. Vanderver1, C. Simons3 1) Illumina Inc., San Diego, CA, 2) Institute for Molecular Bioscience, University of Queensland, St Lucia, Australia; 3) School of Medicine and Health Services, The George Washington University, Washington DC, USA; 4) Children’s National Medical Center, Washington DC, USA. Leukoencephalopathies are a heterogeneous group of heritable central-nervous system white matter disorders that are primarily characterized by abnormal signal on brain magnetic resonance imaging. Leukoencephalopa-thies remain a diagnostic challenge for medical practitioners, with nearly 50% of cases remaining unresolved despite recent advances in both radiological and biochemical testing. We performed whole exome sequencing (WES) on a cohort of 85 patients with unsolved leukoencephalopathies, who were prospectively collected by the Myelin Disorders Bioregistry Project, with non-specific radiologic findings or previously negative diagnostic testing. WES and downstream bioinformatics analysis was performed on trio or greater family groups in all cases, and Sanger sequencing was used to validate and confirm segregation of all candidate pathogenic mutations. Resolution of more than 40% of cases was achieved and included mutations in a wide variety of known and novel disease genes including those associ-ated with mitochondrial cytopathy, epileptic encephalopathy, hypomyelinating leukodystrophies and, unexpectedly, ataxia telangiectasia and dysker-atosis congenita. Overall, fewer than a quarter of the cases solved were due to mutations in genes classically associated with leukoencephalopa-thies. This study strongly suggests that clinical deployment of agnostic and untargeted high-throughput genomic screening approaches in the leukoencephalopathies will substantially reduce the number of unresolved cases and reduce the burden of the diagnostic odyssey.
Additionally, we detected 612 single nucleotide variants that occur in multiple degeneration in 13 of 26 consanguineous pedigrees of Pakistani origin. We reported causative mutations in genes known to be associated with retinal database. Conclusions: Here, we identify six novel and three previously variants present in five or more pedigrees were not reported in the dbSNP HapMap, 1000 Genome and NHLBI ESP6500 databases. Of these, 612 than 10, and mapping quality greater than 30 were further tabulated by playing no strand bias during sequencing, exhibiting a read depth greater ethnicity matched control chromosomes. All single nucleotide changes dis- variants were examined for segregation with the disease phenotype in their call format (VCF) files were generated using standard protocol and subse- Agilent or Nimblegen kits and sequenced on an Illumina HiSeq 2500. Variant Exomes of probands were captured using Nimblegen V3/V2 or Agilent or NHLBI ESP6500 databases. The identification of similar mutations in 2 unrelated kindreds sug- rial infections are caused by mutations in NEMO protein that affect NEMO factor- causes Haemophilus influenzae pericarditis at age 5 then diffuse cutaneous MAC at age 34. Full-length cDNA analysis of the major isoform showed a 110 base pair (bp) deletion in the 5' leader sequence due to a G>C transversion at the last base of the first exon, c.1-16G>C, weakening the consensus splice site and leading to utilization of a cryptic site 110 bases 5' to the end of the exon. Patient B.I.1 had the same change (c.1-16G>C) in 1IBKG with similar consequence. He had Mycobacterium marinum at age 6 and then chronic respiratory infections in childhood and then developed disseminated MAC at age 31. Intracellular staining of patient lymphocytes showed <50% of normal NEMO protein. Patients A.I.1 and B.I.1 were lymphopenic (total lymphocyte counts <0.60), with native CD4+ and CD8+ counts as well as CD20+/CD27+ memory B cells. Stimulation of PBMCs from patient A.I.1 showed decreased production of TNFα in response to LPS. Bioinformatic analysis revealed the deletion is wholly contained within a conserved 153 α show decreased production of TNFα, IL-1 and IL-6. The NEMOdel4_10 deletion, identified in 72% of IP patients, can be associated with a wide range of CNS symptoms also in the same IP family. The skewed X-chromosome inactivation could only partially explain this variability thus modifier loci may contribute to the severity of IP phenotype. We will present a trios-based exome sequencing approach to identify candidate genes for phenotypic variability of IP. We select three severe IP cases carrying the NEMOdel4_10 deletion, as well as 3-4 patients with extracentral-phenotypic variability: in the mother, only skin defects were present, in the children also a severe mental retardation with neuromuscular defects were reported; and two sporadic cases. We performed whole exome-sequencing of trios samples. From sequencing of an exome-enchriched library a list of single nucleotide and insertions/deletions variants were detected, filtered according to our filtering method we select SS/I variants and NS variants predicted to be damaging by polyphen software. We sorted all variants by two model of inheritance: de novo dominant, and recessive (homozygous, compound heterozygous). We will present a list of candidate genes and their associated pathways prioritizing variants fitting to recessive model of inheritance. The most representative category of altered genes is the metabolic pathway in which we identified two genes acting as modifiers in known diseases. Trio-based exome-sequencing has provided new candidate genes that may contribute to the phenotypic variability of IP. The identification of modifier genes of IP will be useful to anticipate the outcome of the IP disease in order to apply an early targeted therapy.
2996M
Inherited UNC13D or PRF1 Mutations in patients with PTLD and severe HHV viremia after HSCT. H. Liu1, Y. Zhang1, F. Wang1, W. Teng1, X. Chen1, P. Zhu2, C. Dong1, T. Wu1, Y. Zhao1, J. Zhang1, D. Lu1. 1) Medical Laboratory Division, Lushan Infection & Oncology Center, Beijing, China; 2) Department of Hematology, Peking University First Hospital, Beijing, China.

Background: Post-transplant lymphoproliferative disorder (PTLD) is a rare but life-threatening disease after hematopoietic stem cell transplantation (HSCT) and 90% are EBV positive. Severe human herpes virus (HHV) viremia is the most common form of HHV activation after HSCT. Familial hematopoietic lymphohistocytosis (FHL) is a mainly autosomal recessive inherited disorder and always triggered by HHV (esp. EBV), the most common mutated genes are UNC13D and PRF1. Methods and Cases: To identify the association of genetic defects and HHV disease after HSCT, 32 acute myeloid or lymphoblastic leukemia patients underwent HSCT were enrolled, including 7 PTLD cases and 25 suffered from severe HHV viremia. Genomic DNA was obtained from peripheral blood of the patients before and after HSCT and their donors. Qualitative and quantitative analysis of HHV in peripheral blood was performed by PCR. Molecular genetic studies were performed by direct sequencing of all coding exons and flanking sequences of UNC13D and PRF1. Results: 6 paired cases carried UNC13D or PRF1 mutations. Including 4 cases with monoallelic mutations in the recipients’ blood sample after HSCT and their donors, one with PRF1 p.P22RfsX29, one with UNC13D p.L202M and two with UNC13D c.2553+5C>G mutations. In the other 2 cases, mutations were inherited carried by the patients’ own somatic cells, one with UNC13D p.T1045C/m.c.2553+5C>G biallelic mutations, and the other with UNC13D c.2553+5C>G mutation. Conclusion: UNC13D or PRF1 mutations might be the inherited predisposing factors of severe HHV viremia or PTLD after HSCT.

2997T
Exome sequencing of a family with Wiskot-Aldrich syndrome reveals a mutation in the WIPF1 gene. A. Hawwari1, M. Dasouki2, L. Al Bak1, S. Al-Hiss3, O. Kheir1, H. Al-Mousa1, 2, 3, R. Arnaut1. 1) Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; 2) Department of Pediatrics, King Faisal Specialist Hospital & Research Center, Riyadh, Saudi Arabia; 3) Al-Faisal University, Riyadh, Saudi Arabia.

This family we are reporting in this abstract is the first family with WIP deficiency in Saudi Arabia and it would be the second WIP deficiency case performed by direct sequencing of all coding exons and flanking sequences of UNC13D and PRF1. Results: 6 paired cases carried UNC13D or PRF1 mutations. Including 4 cases with monoallelic mutations in the recipients’ blood sample after HSCT and their donors, one with PRF1 p.P22RfsX29, one with UNC13D p.L202M and two with UNC13D c.2553+5C>G mutations. In the other 2 cases, mutations were inherited carried by the patients’ own somatic cells, one with UNC13D p.T1045C/m.c.2553+5C>G biallelic mutations, and the other with UNC13D c.2553+5C>G mutation. Conclusion: UNC13D or PRF1 mutations might be the inherited predisposing factors of severe HHV viremia or PTLD after HSCT.

2998S
Erythroid Krüppel-like factor mutations are relatively more common in a thalassemia endemic region and ameliorate the clinical and hematological severity of β-thalassemia. X. Xu1, D. Liu1, Y. Zhang2, L. Yu1, R. Gai1, X. Ma1, C. Zheng3, Y. Zhou2, Q. Liu3, X. Wei1, L. Lin1, T. Yan2, J. Huang1, N. Mohandas1, X. An1. 1) Southern Medical University, Guangzhou, Guangdong, China; 2) Department of Hematology, 303rd Hospital of the People's Liberation Army, Nanning, China; 3) Department of Blood Health and Heredity, Luihzou Women and Children Care Hospital, Luihzou, China; 4) Prenatal Diagnostic Center, Guangxi Zunag Autonomous Region Women and Children Care Hospital, Nanning, China; 5) Department of Birth Health and Heredity, Zhumai Women and Children Care Hospital, Zhumai, China; 6) Key Lab for Experimental Teratology of the Ministry of Education and Department of Medical Genetics, School of Medicine, Shandong University, Jinan, China; 7) Red Cell Physiology Laboratory, New York Blood Center, New York City, NY, United States.

The erythroid transcriptional factor 1 (KLF1) has recently emerged as one of the key regulators of the γ- to β-globin gene switching. Mutations in human KLF1 have recently been reported to be responsible for increased fetal hemoglobin (HbF) and hemoglobin A2 (HbA2). As increased HbF and HbA2 levels are important features of β-thalassemia, we examined whether there is any relationship between KLF1 mutation and β-thalassemia in China. For this, we first studied the incidence of KLF1 mutations in two Chinese populations: 3839 individuals from a thalassemia endemic region in south China and 1190 individuals from a non-thalassemia endemic region in north China. Interestingly we found that the prevalence of KLF1 mutations is significantly higher in thalassemia endemic region than that in non-thalassemia endemic region (1.25% versus 0.08%). Furthermore, a total of 64 mutant alleles were documented in the present study, of which 41 are KLF1 heterozygotes alone, 11 are KLF1 mutations co-inherited with β-thalassemia homozygotes, and 12 are KLF1 mutations co-inherited with β-thalassemia homozygotes or compound heterozygotes. We identified seven functional variants including four previously reported (r.Gly176AlafsX179, p.Ala298-Pro, p.Thr334Arg and c.913+1G>A) and three novel ones (p.His299Asp, p.Cys341Tyr and p.Glu548Lys) in southern China. The two most common mutations, p.Gly176AlafsX179 and p.His299Asp, accounted for 90.6% of the mutant alleles. The study also provided diverse pathological features of individuals with KLF1 heterozygous mutations. In addition, we found it surprising that two KLF1 zinc-finger mutations were selectively represented in 12 β-thalassemia patients from southern China, who were identified by DNA sequencing as β-thalassemia patients with known β-globin genotypes and various modifier genes. Both multivariate and univariate analysis showed that the modifier of KLF1 mutations had an effect on amelioration of the severity in β-thalassemia. Our findings suggest that KLF1 mutations occur selectively in the presence of β-thalassemia to increase the production of HbF which in turn ameliorates the clinical severity of β-thalassemia. The knowledge gained from this study should help in clinical accurate diagnoses and genetic counseling, as well as enable designing of appropriate and personalized transfusion program for thalassemia patients with KLF1 mutations.
2999M

Homologous loss of DIAPH1 causes a rare, complex syndrome with epilepsy, blindness, immune deficiency and lymphoma. M. KausoG, R. Hinttala1,2,3, H. Almusa2, R. Renko2,3, H. Tuominen2, R. Herva2, J. Uusimaa2,3,4, J. Saarela4. 1 Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 2 Department of Pediatrics, University of Oulu, Oulu, Finland; 3 Department of Pediatrics, Oulu University Hospital, Oulu, Finland; 4 Medical Research Center, University of Oulu and Oulu University Hospital, Oulu, Finland. Using whole-exome sequencing (WES), we investigated a Finnish family with two affected siblings, who presented with epilepsy, complex syndrome including an early-onset intractable epilepsy, blindness, severe mental retardation, deficiency in cellular immunity and B-cell lymphoma. Muscle biopsy showed atrophy, mild increase of fat, a few SDH-COX negative muscle fibres, and an increased number of mitochondria supporting a mito-ochondrial disease. Brain MRI and spectroscopy revealed focal occipital leukoencephalopathy. The disease course was progressive leading to death in both siblings at the age of 5 and 18 years. The neuropathologic examination showed severe glosis, atrophy and vascular abnormality affecting the primary visual cortex. Through WES we identified a rare, homozygous splice donor variant (NM_005219: c.684+1G>A) in the gene DIAPH1, which encodes for the mammalian diaphanous-related formin mDia1. Capillary sequencing of reverse transcribed patient mRNA with amplicons spanning the mutated site showed that the mutation results in a frameshift in several different cryptic splice sites, all of which ultimately result in the introduction of a premature stop codon in the DrtGSD domain and likely cause lack of any functional protein. The absence of a protein product in patient cells or tissues was confirmed by western blotting and immunostaining for mDia1, which acts downstream of Rho GTPases promoting actin polymerization and microtubule stabilization, and it is known to play a role in various processes such as cell migration and tumour metastasis, neuronal development, mitochondrial trafficking and the function of several types of immune cells. Hence, the several aspects of mDia1 function seem to explain most if not all of the features of the phenotype observed in the affected individuals. The DIAPH1 splice donor variant is not present in databases such as the 1000 Genomes, the Exome Variant Server or dbSNP, but instead was found in the Sequencing Project Suomi (SISu) database with exome sequencing data from about 3300 Finnish individuals with a frequency of 0.075%. This leads us to hypothesize that this variant may be enriched in the Finnish population. Exome sequencing of reverse transcribed patient RNA with amplicons spanning the mutated site showed that the mutation results in the use of several donor splice sites.

3000T

Copy number variations in a cohort of Brazilian sickle cell anemia patients with and without cerebrovascular accident. P.R.S. Cruz1,2, G. Ananina1, F. Menaa3, A.S. Araujo4, G.P. Gil5, W.M. Avelar5, F. Cendes6, F.F. Costa6, K. Agematsu7, M. Renko2,3,4, H. Tuominen4. 1) Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; 2) Institute of Clinical Medicine, Department of Pediatrics, University of Oulu, Oulu, Finland; 3) Department of Pediatrics, Oulu University Hospital, Oulu, Finland; 4) Medical Research Center, University of Oulu and Oulu University Hospital, Oulu, Finland. Although sickle cell anemia (SCA) results from homozygosity for a single position at mutation 6 of gamma-hemoglobin locus, this disease presents high heterogeneity in phenotype, so that different patients may have significantly different clinical outcomes. Virtually all organs may be affected by sickle cell anemia, susceptible to stroke (CVA) could reduce the risk, possibly preventing the recurrence of infarcts and potentially reducing their incidence. Thus, we propose to investigate the presence of copy number variation (CNV) in this disease group. In this study, we analyzed the DNAs of 107 Thai adults revealed 11 reported mutations and 2 novel mutations so far. The 3D structure of the enzyme expression. Thus, not only a single variation but also haplotype of this putative novel disease variant in Finland by using genotyping data from about 3300 Finnish individuals with a frequency of 0.075%.

3001S

HLA confer the risk of familial Mediterranean fever in Japanese population. M. Yasunami1, H. Nakamura1, A. Agematsu1, M. Yaza2, K. Migtia2,3. 1) Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan; 2) Department of Infection and Host Defense, Shinshu University Graduate School of Medicine, Matsumoto, Japan; 3) Department of Medicine, Shinshu University School of Medicine, Matsumoto, Japan; 4) Clinical Research Center, National Hospital Organization Nagasaki Medical Center, Omura, Japan.

BACKGROUND Familial Mediterranean fever (FMF) has been considered to be an autosomal recessive trait which is characterized by self-limiting recurrent fever and serositis (OMIM #249100) and classified into a category of autoinflammatory diseases. MEFV, identified as the responsible gene for FMF, encodes cytosolic protein pyrin (also known as marionnet) which regulates the activity of NLRP3 inflammasome. Mutations in MEFV gene have been registered to “Infevers” database (http://infevers.rockefeller.edu/).

2969F

Poster numbers were significantly different in control samples with deletions, and 38% vs. 10% for duplications, respectively; p-value=9e-3.

3002M

A pathogenic haplotypes of the 66GDP gene correlating with enzyme activity. D. Nantakomol1, M. Chaowanakhikom1, P. Nuchnoi2,1) Department of Clinical Microscopy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand; 2) Department of Clinical Microscopy, Faculty of Medicine, Mahidol University.

The single nucleotide polymorphisms (SNP) in glucose-6-phosphate dehydrogenase (66GDP) gene caused enzyme deficiency leading to acute hemolytic anemia and neonatal jaundice. Over 400 SNPs in 66GDP have been detected in 66GDP database. Most of the 66GDP SNP showing clinical significa-
3003T

Defective Dimerization of STAT3 causes Autosomal Dominant Hyper-IgE Syndrome. M. Dasouki1, S. Keles2, T. Chatilla2. 1) Neurology, Univ Kansas Med Ctr. Kansas City, KS; 2) Pediatric Immunology, Harvard University School of Medicine.

Background: Autosomal dominant and autosomal recessive hyper-IgE syndromes are primary immune deficiency syndromes caused by heterozygous or homozygous mutations in STAT3, respectively. STAT3 is a transcription factor involved in immune response to cytokine stimulation, whereas pSTAT1 activation mediates cellular responses to interleukins, KITLG/SCF and other growth factors. It also binds to (IL-6)-responsive elements identified in the promoters of various acute-phase protein genes. Upon tyrosine phosphorylation and dimerization, in response to signaling by activated FGFR(1-4) it is translocated into the nucleus. Most previously reported STAT3 mutations involve its SH2 domain. Aim: To characterize the causal mutation and its functional effects in a 40 year old Dutch man with AD-HIES (severe chronic eczema, onychodystrophy, hyper-IgE syndrome, asthma, and bilateral cataract, lymphopenia, hypereosinophilia and hyper-IgE). Methods: STAT3 mutation analysis by DNA sequencing. Activation of STAT3 by IL-6 and IL-21 and STAT1 by IFN-γ was analyzed by intracellular staining with anti-phospho (p)STAT3 and -pSTAT1 antibodies. TH17 and TH1 cell differentiation was assessed by measuring the production of IL-17 and IFN-γ, respectively. Results: Patient had normal B&T cell flow cytometry profiles and normal TRECs count. A novel heterozygous STAT3 mutation (c.491G>A; p.R84Q) which is predicted to impair its dimerization was identified. Lymphocytes functional studies showed reduced phosphorylation of STAT1 in response to cytokine stimulation whereas pSTAT1 activation was unaffected. Also, impaired TH17 responses and (early steps) in TH17 differentiation was found. Conclusion: This novel amino terminal domain STAT3 mutation causes defective dimerization and impaired TH17 responses indicating a novel mechanism for AD-HIES.

3004S


Primary immunodeficiencies (PIDs) constitute a heterogeneous group of genetic diseases affecting the immune system. Depending on the genetic etiology, symptoms range from mild to severe and life threatening. Knowledge of the molecular genetic cause and disease mechanism is important and can direct targeted and curative therapy. However, subtype classification is difficult as patients often have overlapping phenotypes. In addition, more than 250 PID genes have been reported, and few are object of extensive immunological and genetic testing. We examined the utility of whole-exome sequencing (WES) to detect single nucleotide variants (SNVs) and copy number variations (CNVs) in the diagnosis of PIDs. As of June 2014, 275 patients with extensive immunological and genetic testing from 241 families have been recruited from Texas Children Hospital (Houston, USA) and Oslo University Hospital (Norway). Strategies for genetic analysis were tailored based on clinical data, immunophenotyping and family history, in 25 families more than one person WES tested, but for most families only the proband was subjected to WES. Initially, WES data were systematically screened for variants in reported and potential PID genes. In addition, a computational CNV prediction pipeline was applied to enable identification of potential disease-causing CNVs from the WES data. Analysis of the first 125 families identified PID relevant variants in 60 percent of the cases; half of these attaining a definitive molecular PID diagnosis. The other half had previously reported PID-causing variants, but with an unexpected or extended clinical phenotype, or heterozygote, potential deleterious variants in recessively inherited PID genes. In two families the patients had co-existing disease causing and modifying variants in two genes (ZAP70/RNF168 and SH2D1A/FANC6, respectively). In one family two different PID genes occurred (RAG1 SCID and XIC6 related Hyper IgE syndrome). Other interesting findings include PID-causing CNVs in 7 families, somatic revertant mosaicism (IL7R, FANCA, IL2RG), de novo AK2 lymphocyte mosaicism, and 13 novel disease genes (5 definitive NUDCD3, PG3M, COPA, MCM10, SIGIRR, and 8 potential).
3006T
IPEX and IPEX-like syndromes: FOXP3 and FOXP3-pathway related genes. M. Vignoli1,2, S. Cullini Manunta1,2, G. Colarusso3, F. Barzaghi3, R. Bacchetta4, A. Cecconi6, A. Tommasini6, A. Genney7, S.M. Hol7, A.J. Cant8, E. Gamberini1,2, 1) Dept of NEUROFARBA, section of Child’s Health, University of Florence; 2) Anna Meyer” Children’s Hospital, Florence, Italy; 3) Santo Stefano Hospital, Paediatric Unit, Prato, Italy; 4) San Raffaele Telethon Institute for Gene Therapy (HSS-TIGET), Milan, Italy; 5) Human Genetic Laboratories, Ospedale Galliera, Genoa, Italy; 6) IRCCS Burlo Garofalo, Paediatric Immunology Laboratory, Trieste, Italy; 7) Department of Paediatric Immunology, Newcastle upon Tyne Hospitals, Newcastle upon Tyne, UK; 8) Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA; 9) PRESTO, JST, Kawaguchi, Japan; 10) Human Genome and Developmental Biology Department, Tokyo Medical and Dental University, Tokyo, Japan; 11) Department of Microbiological, Immunological and Parasitological Sciences, University of Fukui, Fukui, Japan; 12) Pediatrics Department, Osaka University Medical School, Osaka, Japan; 13) Pediatrics Department, Kitano Hospital, Osaka, Japan; 14) Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA;

IPEX (MIM #304790) is characterized by severe early-onset enteropathy, endocrinopathy, dermatitis and other autoimmune phenomena. It is due to FOXP3 (Xp11.23) gene mutations. Interestingly an increasing number of patients show a phenotype consistent with IPEX but do not harbour FOXP3 mutations. IL2Ra (10p15.1) and STAT3b (17q11.2) are two genes involved in FOXP3 pathway whose mutations have been associated with IPEX-like phenotype, nevertheless many patients remain without a molecular diagnosis. Our cohort of patients is composed of 76 cases with a phenotype compatible with IPEX syndrome, that were referred to us during the past 10 years for molecular analysis. We decided to perform a retrospective study evaluating clinical and laboratory findings to better delineate the clinical spectrum of multiple autoimmune diseases and identify different subgroups of patients to be addressed to specific molecular investigations. We analyzed the clinical history, T regulatory cells (Tregs) and FOXP3 expression by flow cytometry of the gene sequencing results of FOXP3, IL2Ra, STAT3b, STAT1, IL10, IL10RA, IL10RB. In 16 cases we identified FOXP3 mutations confirming IPEX syndrome diagnosis. The 60 FOXP3-like patients were divided in three different groups. The first group includes 31 cases with a phenotype very similar to IPEX, they showed enteropathy with onset within six months of life, growth retardation, eczema and at least one autoimmune phenomena. The second group comprises 22 cases with a later onset of the clinical symptoms (mainly diarrhoea and at least two autoimmune signs). The third group includes 7 patients with severe early-onset diarrhoea, enteritis, anal abscesses or enteric fistulae, these cases are more similar to inflammatory Bowel Disease manifestation. Autoimmune enteropathy is the key clinical sign for IPEX-like patients of all groups, while in subjects without enteropathy, endocrinopathy and cytopenias are the main clinical symptoms (mainly diarrhoea and at least two autoimmune phenomena). High IgE levels is often reported. In the first group the molecular analysis resulted in 3 patients harboring IL2Ra mutations leading to the absence of CD25 expression, a patient with a STAT3b variation and one STAT1 deficiency case, confirming the association between mutations in FOXP3-related genes and clinical features very similar to IPEX. Therefore the clustering in distinct subgroups can help to address the molecular analysis of different genes and to select patients for whole exome sequencing analysis.

3007S
Aicardi-Goutières syndrome is caused by IFIH1 mutations. H. Oda1,2, K. Nakagawa1, J. Abe1,2, T. Awaya1, M. Funabiki4, A. Hikjata3, R. Nishikomori4, N. Morisada5, M. Funasaka6, Y. Okada7, T. Yamaoka8, T. Kato9, T. Shirai9, O. Ohara10, T. Fujita10, T. Heike10, 1) Department of Pediatrics, Kyoto University Graduate School of Medicine, Kyoto, Japan; 2) Integrative genomics, RIKEN-Institution for Integrative Medical Sciences, Yokohama, Japan; 3) Pediatric Department, Kitano Hospital, Osaka, Japan; 4) Laboratory of Molecular Genetics, Institute for Virus Research, Kyoto University, Kyoto, Japan; 5) Department of Biosciences, Nagahama Institute of Bio-Science and Technology, Nagahama, Japan; 6) Pediatric Department, Tokyo Children’s Medical University, Tokyo, Japan; 7) Faculty of Medical Sciences, University of Fukui, Fukui, Japan; 8) Pediatrics and Developmental Biology Department, Tokyo Medical and Dental University, Tokyo, Japan; 9) PRESTO, JST, Kawaguchi, Japan; 10) Human Genome Research Department, Kazusa DNA Research Institute, Kisarazu, Japan.

Aicardi-Goutières syndrome (AGS) is a rare, genetically determined earlyonset progressive encephalopathy. Individuals affected with AGS typically exhibit severe neurological symptoms associated with basal ganglia calcification and type I interferon elevation in the cerebrospinal fluid. To date, mutations in six genes (TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, and ADAR1) have been identified as etiologic for AGS. However, our previous Japanese nationwide AGS survey (Abe, 2013) as well as other reports showed a group of AGS individuals without mutations in the six genes above, suggesting the existence of additional AGS associated genes. We performed trio-based whole exome sequencing in three AGS individuals without molecular diagnosis. After the removal of common polymorphisms, de novo and parental mutations identified in the IFIH1 gene were identified. IFIH1 encodes MDA5, one of the cytosolic pattern recognition receptors that recognize double-stranded RNA (dsRNA). MDA5 activation by its oligomerization on dsRNA results in the enhancement of type I interferon transcription, resulting in viral activation. In the AGS mouse model (Funabiki, 2014), our results, currently under revision, showed that AGS individuals into Huh7 cells showed significantly stronger promoter activities of IFNβ1 as well as downstream IFIT1 transcription, suggesting that IFIH1 mutations induce activation of type I interferon pathway and are responsible for the phenotypes of AGS. Furthermore, we retrovirally reconstituted Ifih1-null mouse embryonic fibroblasts (MEFs) with the IFIH1 mutants, and subsequently infected these MEFs with encephalomyocarditis virus (EMCv), an MDA5-specific ligand. Interestingly, these MDA5 mutants lack ligand-specific responsiveness, which was similar to the lack of ligand responsiveness of p.Gly821Ser mouse MDA5 mutant, previously reported in the SLE mouse model (Funabiki, 2014). Our results, currently under revision, showed that AGS is caused by the mutations in IFIH1, which has also been recently reported by Rice et al (Rice, 2014). Further analysis remains to elucidate the precise mechanism how these IFIH1 mutations cause the type I interferon overproduction.

3008M
The comprehensive genetic analysis of congenital anomalies of kidney and urinary tract (CAKUT) in Japan. N. Morisada1, M. Taniguchi-Ikeda1, K. Nogu1, A. Shono2, K. Kame1, S. Ito1, R. Tanaka3, K. Illima1, 1) Pediatrics, Kobe University, Kobe, Hyogo Prefecture, Japan; 2) Nephrology and Rheumatology, National Center for Child Health and Development, Tokyo, Japan; 3) Nephrology, Hyogo Prefectural Children’s Hospital, Kobe, Japan.

Purpose: Congenital anomalies of kidney and urinary tract (CAKUT) are the most common cause of pediatric end-stage renal disease over the world. The detailed mechanisms of development of CAKUT were still unclear; however genetic mutations in various genes may play pivotal roles. We have conducted the comprehensive genetic analyses of CAKUT in Japan. Methods: This study was approved by the Institutional Review Board of Kobe University Graduate School of Medicine. Genomic DNA samples of patients were provided from peripheral blood mononuclear cells. We performed direct sequencing, multiplex ligation-dependent probe amplification (MLPA) analysis, array comparative genome hybridization (aCGH) analysis and/or next-generation sequencing (NGS). Results: We examined 206 patients from 187 families. Ninety-one patients out of 82 families were non-syndromic CAKUT (NSC) and 115 patients out of 105 families were syndromic CAKUT (SC). We identified the responsible genes in 62 patients from 44 families. The detection rates in NSC and SC families were 14.6% and 30.5%, respectively. The most common responsible gene found in this cohort was PAX2 (11 families). Other responsible genes were EYA1 (9), HNF1B (7), UMOD (1), OFD1 (1), SALL1 (1) and CHD7 (1). MLPA analysis enabled us to identify TSC2-PKD1 contiguous gene deletions in a patient in our cohort. aCGH is useful to reveal the responsible chromosomal lesion in CAKUT with intellectual disability (ID), and we have identified copy number variations in patients with CAKUT and ID such as 22q11 microdeletion, 16q microdeletion and 1q21.1 microdeletion. We identified several responsible genes in the patients by NGS. Conclusions: This is the first nation-wide study for the genetic approach of CAKUT in Japan.
3009T


The Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome describes women with congenital aplasia of the uterus, cervix and the upper two-thirds of the vagina due to absent paramesonephric ( Müllerian) ducts. Women with MRKH have a normal female karyotype (46, XX), normal female external genitalia, but aplasia or hypoplasia of the upper two-thirds of the vagina. Some women have correctly timed pubarche and thelarche. The incidence of the MRKH syndrome is one in 4,500 female newborns. MRKH patients represent the second most frequent cause of primary amenorrhea after Turner syndrome and commonly have associated malformations, such as skeletal and renal abnormalities. The MRKH syndrome occurs isolated as a genital malformation (type 1), but also with associated malformations (especially malformations of the kidneys and urinary tract (type 2) and MURCS association (Müllerian renal and cervicothoracic somite abnormalities; partial or complete absence of the spinal column, ribs, or arms; asymmetric or improperly developed ribs or arms). To date the genetic or molecular etiology of the MRKH syndrome is completely unknown. We performed copy number variation analysis with high density SNP microarrays, whole exome sequencing and sequenced the enriched methylome from two pairs of discordant monozygotic twins and their family members. Since the results of this integrative analysis of these three high-throughput analysis data sets did not result in clear causative variants even with a combined genetic-epigenetic disease model for the differential diagnosis, we here made the deep sequencing analysis. We now used two in depth analysis methods: sequencing whole genomes and bisulfite converted whole genomes in parallel. The first method enables in addition to comprehensive SNV and small InDel analysis also a thoroughly analysis for structural variants like large InDels, inversions, translocations, CNVs etc. The latter method is the gold standard for DNA methylation analysis and provides single nucleotide resolution map of 5-mC of the genome. So far the first tier of data analysis is very promising and a complete comparative analysis is ongoing. We believe that the methodology presented in the present study will help to unravel the genetic or molecular causes of MRKH and be of clinical relevance in order to help clinicians for a accurate molecular diagnos- is and council patients and their families in order to understand the syn- drome.

3010S


Introduction: Congenital anomalies of the kidney and urinary tract (CAKUT) are the most common cause of pediatric end stage renal disease (ESRD). The discovery of underlying genetic etiologies would not only improve diagnostic precision for non-syndromic forms, optimize management of affected patients, and improve outcomes of neonates and infants with in utero urinary tract abnormalities, but also aid prenatal diagnosis in forms that result in early-onset ESRD. Here, we show the utility of whole exome sequencing (WES) in a diagnostic setting and in the genomic research of CAKUT. Methods: We completed whole exome sequencing in 65 exomes of affected individuals and their family members. Our data consists of 20 genes that are known to cause CAKUT and another 20 candidate genes related to the kidney and urinary tract. Results: Mutation analysis was performed and the following deleterious or possibly deleterious mutations were identified: PAZ2 (G24fs, S171fs), HNF1B (G378fs), RET (S401C), EYA1 (c.867–8G>A), SIX2 (P241L), and DSTKY (R52Q, D120N). No mutation was identified in GDNF, SIX1, SOX17, or GATA3. Benign variants or variants of unknown significance were identified in BMP7, CDCC5L, CHD1L, SALL1, SIX5, ROBO2, UPK3A, BMP4, KAL1, and TXNBP. The G24fs mutation occurs in a family with multiply renal dysplasia and membranous nephropathy that has an autosomal dominant inheritance. We identified a gene with another defective variant in HNF1B. Phenotypes of some of these families expand our current knowledge about the corresponding genes. Conclusion: This is the first report utilizing WES for diagnosis in patients with CAKUT. At least 12% of individual patients with CAKUT have variants in known genes that can be identified by WES. Therefore, WES should be considered as a diagnostic tool in such patients, especially in the families with complex medical histories. Identification of novel mutations and expansion of the phenotype may aid clinicians to better understand WES as a potential test. Acknowledgement: Supported by: a-K12 DK0083014, the Multidisciplinary K12 Urologic Research (KURE) Career Development Program to DPL (MRB is a KURE Scholar). b- NHGRI Mendelian grant to JRL c- RAL is a Senior Scientist Investigator of the Boston Children’s Research Institute, New York, whose unrestricted funds supported part of these investigations.

3011M

Focal segmental glomerulosclerosis exomes reveal candidate variants highly enriched in cell movement and cell adhesion related genes. J. Suh, G. Genovese, V. Charommatana, A. Knob, M. Pollak, Israel Deaconess Medical Center, Boston, MA; 2) Broad Institute of MIT and Harvard, Cambridge, MA.

The findings of focal segmental glomerulosclerosis (FSGS)-causing genes elucidate the role of specific loci in the development of FSGS during recent two decades. However, in majority of families even with the evidence of strong inheritance, underlying genetic cause is unknown. Here, we analyzed exome of families and sporadics which are unexplained by known FSGS genes in an extended effort to find the gene responsible for the development of FSGS. We performed exome sequencing of 403 probands from 311 familial and sporadic cases of FSGS. We also sequenced 96 unaffected family members. After variants calling, variants are filtered against common variants from the 1,000 Genomes Project, dbSNP, and Exome Sequencing Project except known FSGS-causing variants, and then novel non-synonymous or splice variants are kept. Exome analysis revealed that 30.3 percent of all cases contain known FSGS-causing variants with 47.9 percent of familial cases and 16.0 percent of sporadic cases, respec- tively. After excluding explained cases, we filtered genes showing more than 4 novel non-synonymous and splice variants, and grouped genes according to 1) their cellular component, molecular function, and biological process using gene ontology, 2) AMPG2, and 2) their protein-protein interaction using DAPPLE. GO terms of genes from unexplained cases were significant enriched in cellular component movement and cell adhesion both with p-value less than 0.0001. DAPPLE results consistently showed two big protein-protein interactions each involving cell movement and cell adhesion, respectively. Here we show sporadic FSGS cases are less explained by known FSGS-causing genes compared to familial cases. These data show that novel non-synonymous or splice variants are highly enriched in cell movement- and cell adhesion-involved genes in the exome of FSGS patient samples, suggesting the importance of podocyte movement and interaction with neighboring cells and extracellular matrix.

3012T

The role of MAZ in the regulation of genitourinary development via modulation of WNT signaling. M. H. Derfer, D. J. Lamb. Baylor College of Medicine, Department of Molecular and Cellular Biology, Center for Reproductive Medicine, Houston, TX.

Genitourinary (GU) birth defects comprise some of the most common yet least studied congenital malformations and range in severity from conditions such as descended testes (cryptorchidism) and ventrally misplaced urethral meatus (hypospadias) to highly complex malformations such as bladder exstrophy epispadias complex (BEEC) and ambiguous genitalia. Congential establishment of the kidney and ureteral tract are also common and include pheno- types such as duplicated tract components, kidney agenesis, congenital hydronephrosis, horseshoe kidney, and cystic kidneys. Genomic aberrations such as copy number gains and losses can result in congenital malformations of the GU tract among other organ systems. Genomics define wide array comparator- generic genomic hybridization (aCGH) together with an extensive literature review allowed identification of genomic hotspots of GU development and delineated the smallest CNV regions of maximum overlap. CNV mapping identified over 30 patients with GU defects harboring duplications or deletions in the syndromic genomic region, 16p11.2-1.2 the most common known patho- genic gene dosage region in humans. The only gene covered collectively by all the mapped CNVs in patients with GU defects was MYC-associated zinc finger (MAZ). MAZ encodes a transcription factor with a similar consen- sus sequence to that of WT1, and is implicated in WNT signaling. In situ hybridization on mouse embryos and isolated mouse GU tracts defined the expression profile of Maz during development. In situ experiments showed normal expression of the GU tract including the kidneys, ureters, testes, bladder and genital tubercle. Immunohistochemistry confirmed this expression pat- tern at the protein level. Analysis of the lab’s cohort of genomic DNA from GU-abnormal patients and GU-normal fertile controls via CNV qPCR deter- mined the frequency of MAZ CNVs in these groups. Validated by two inde- pendent probe sets, 6% of our GU affected cohort displayed CNVs in MAZ (n=258) compared to 0% in controls (n=57). As predicted based on its consensus sequence, knockdown or overexpression of MAZ in human embryonic kidney (HEK293) cells results in differential expression of several WNT target genes such as DACT1 and DACT2, as well as FRZB, SFRP4/5, WNT4, WNT11 and GSK3A. MAZ, once thought to be a housekeeping gene, encodes a dosage sensitive transcription factor that may play a key role in urgenital development and contribute to the congenital malformations of the 16p11.2 phenotype.
3013S
Functional Characterization of Renin Variants Identified in African Americans in Exome Sequencing Project. S. Knioh1, M. Zvina1, K. Hodanova1, P. Vyletal1, V. Stranecky1, B.I. Freedman2, A.J. Bleyer1. 1) First Medical Faculty, Charles University of Prague, Prague, Czech Republic; 2) Section of Nephrology, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.

Inheritance (Oligogenic and homozygosity in two Cilia genes) are likely always associated with cystic kidneys and Encephaloceles indicating Molecular karyotype to exclude aneuploidy was negative for all cases. we recruited forty fetuses presenting with PKD diagnosed by ultrasound. mutations of genes expressed in the primary cilium of renal epithelial cells. Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, M. Al-Nemer

Exact population frequencies and clinical effects of these two variants should analysis suggest that P8A and R33W variants significantly affect renin bio-
gonadism. Y. Bayram1, T. Gurian2, G. Yesil3, Z. Atay4, E. Karaca5, T. Gambin6, S. Turan7, D. Pehlivan8, S.N. Jiang9, D. Muzz3, B. Haliloglu10, A. Bereke11, R.A. Gibbs12, J.R. Lupski12. 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Department of Pediatric Endocrinology and Diabetes, Marmara University Hospital, Istanbul, Turkey; 3) Department of Medical Genetics, Bezmialem University Hospital, Istanbul, Turkey; 4) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 5) Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; 6) Texas Children's Hospital, Houston, TX, USA.

Patients with hypergonadotropic hypogonadism (HH) present with primary or secondary amenorrhea as well as puberty secondary to primary gonadal dysfunction and characterized by elevated gonadotropins and low sex steroids. HH is always a pathological state; therefore, assessment should include history of surgery, irradiation and chemotherapy, examination for features of the Turner syndrome and consideration of a karyotype. Several genetic defects have been shown to cause HH including genes having a role in gonadal development or maintenance and defects in sex steroid synthesis, or end organ resistance to gonadotropins. However, elucidation of novel genetic defects causing HH will provide major insights into the biology of this condition - especially knowledge of the regulation of human reproductive function. Here we report two sisters who were born to a first degree cousin marriage presenting with HH. Both of them were referred to a pediatric endocrinology clinic because of gonadotropin elevation and late appearance of secondary sex characteristics at 16 and 14 years of age, respectively. Further hormonal investigations revealed elevated FSH and LH and decreased estradiol concentrations. On the pelvic ultrasound imaging both patients revealed polycystic ovaries and small uterus without fibroids. Karyotype analyses were normal and no other significant clinical or dysmorphic features were identified; the patients were diagnosed with non-syndromic HH. We performed whole exome sequencing (WES) for both affected patients which identified a homozygous frame-shift deletion (c.705delT, p.Lys236Arg3) in the SOHLH1 gene. Sequencing within the family revealed that both parents were heterozygous carriers while the unaffected sister was wild type for the same variant consistent with Mendelian expectations. SOHLH1 encodes a basic helix-loop-helix transcription factor with homologues in humans and other placental mammals. Previous mouse studies revealed that SOHLH1 is preferentially expressed in mouse oocytes during early folliculogenesis and required for oogenesis. In a Sohlh1 knockout mouse model, Sohlh1−/− female mice have been observed to be infertile with atrophic ovaries and poorly formed oocytes. We show that mice carrying Sohlh1+/- heterozygous mutations as a single factor can also cause ovarian dysfunction. In conclusion; we describe the first recurrent SOHLH1 mutation in humans and clinical features in our patients together with findings observed in a mouse model strongly suggest that homozygous mutations in SOHLH1 cause ovarian dysfunction.

3015T

BACKGROUND: Zeb2 is a zinc finger E-box-binding homeobox transcription factor. Mutations in Zeb2 cause the Mowat-Wilson syndrome (MIM #235730), an autosomal dominant disorder characterized by multiple congenital anomalies including kidney anomalies. However, the role of Zeb2 during kidney development is unknown. METHODS: Zeb2 expression was analyzed by immunostaining of the embryonic kidneys during mouse development. Kidney specific Zeb2 conditional knockout mice were generated by crossing Zeb2flx/flox mice with Pck1cre and Six2cre alleles. The conditional knockout mice were analyzed from embryonic stage E14.5 to postnatal 8 weeks old. The mouse kidneys were analyzed by H&E staining and immunostaining for apoptosis. Gene expression analy-
sis was performed by TaqMan assays. Kidney function was assessed by measuring urine protein levels, serum creatinine and blood urea nitrogen.

RESULTS: Zeb2 is highly expressed during mouse kidney development. Both Zeb2flx/flox;Pax2cre and Zeb2flx/flox;Six2cre conditional knockout mice developed kidney glomerular cysts starting at E15.5 days. Zeb2flx/flox;Pax2cre mice died at birth and Zeb2flx/flox;Six2cre mice developed abnormal kidney function at 5 weeks old. The cysts originate from glomeruli with dilated Bowman’s capsules and collapsed glomerular tufts. Reduced apoptosis was detected in the S-shaped and C-shaped branches of Zeb2 knockout mice compared to the wild type controls. Gene expression analysis revealed increased levels of Pdk1 mRNA in the developing kidney of both Zeb2 conditional knockout mice compared to the wild-type controls. Interest-

3016S
Whole exome sequencing identifies a homozygous mutation in SOHLH1 in two sisters with non-syndromic hypergonadotropic hypogonadism. Y. Bayram1, T. Gurian2, G. Yesil3, Z. Atay4, E. Karaca5, T. Gambin6, S. Turan7, D. Pehlivan8, S.N. Jiang9, D. Muzz3, B. Haliloglu10, A. Bereke11, R.A. Gibbs12, J.R. Lupski12. 1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 2) Department of Pediatric Endocrinology and Diabetes, Marmara University Hospital, Istanbul, Turkey; 3) Department of Medical Genetics, Bezmialem University Hospital, Istanbul, Turkey; 4) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 5) Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; 6) Texas Children's Hospital, Houston, TX, USA.

Patients with hypergonadotropic hypogonadism (HH) present with primary or secondary amenorrhea as well as puberty secondary to primary gonadal dysfunction and characterized by elevated gonadotropins and low sex steroids. HH is always a pathological state; therefore, assessment should include history of surgery, irradiation and chemotherapy, examination for features of the Turner syndrome and consideration of a karyotype. Several genetic defects have been shown to cause HH including genes having a role in gonadal development or maintenance and defects in sex steroid synthesis, or end organ resistance to gonadotropins. However, elucidation of novel genetic defects causing HH will provide major insights into the biology of this condition - especially knowledge of the regulation of human reproductive function. Here we report two sisters who were born to a first degree cousin marriage presenting with HH. Both of them were referred to a pediatric endocrinology clinic because of gonadotropin elevation and late appearance of secondary sex characteristics at 16 and 14 years of age, respectively. Further hormonal investigations revealed elevated FSH and LH and decreased estradiol concentrations. On the pelvic ultrasound imaging both patients revealed polycystic ovaries and small uterus without fibroids. Karyotype analyses were normal and no other significant clinical or dysmorphic features were identified; the patients were diagnosed with non-syndromic HH. We performed whole exome sequencing (WES) for both affected patients which identified a homozygous frame-shift deletion (c.705delT, p.Lys236Arg3) in the SOHLH1 gene. Sequencing within the family revealed that both parents were heterozygous carriers while the unaffected sister was wild type for the same variant consistent with Mendelian expectations. SOHLH1 encodes a basic helix-loop-helix transcription factor with homologues in humans and other placental mammals. Previous mouse studies revealed that SOHLH1 is preferentially expressed in mouse oocytes during early folliculogenesis and required for oogenesis. In a Sohlh1 knockout mouse model, Sohlh1−/− female mice have been observed to be infertile with atrophic ovaries and poorly formed oocytes. We show that mice carrying Sohlh1+/- heterozygous mutations as a single factor can also cause ovarian dysfunction. In conclusion; we describe the first recurrent SOHLH1 mutation in humans and clinical features in our patients together with findings observed in a mouse model strongly suggest that homozygous mutations in SOHLH1 cause ovarian dysfunction.
3017M

IMAGe syndrome mutations in the PCNA-binding site of CDKN1C cause in increased protein stability and inhibition of the cell cycle. V.A. Arboleda1,2, K.S. Borges3, E. Vilain1,4,5, 1) Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA; 2) Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA; 3) Department of Genetics, Ribeirão Preto Medical School, University of São, Ribeirão Preto, Brazil; 4) Department of Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, CA; 5) Department of Urology, David Geffen School of Medicine at UCLA, Los Angeles, CA.

Mutations in CDKN1C (a.k.a. P57KIP2) can cause opposing genetic syndromes: the undergrowth syndromes IMAGe Syndrome (Intrauterine growth retardation, Metaphyseal dysplasia. Adrenal hypoplasia and Genital anomalies) and Russell Silver Syndrome (RSS) and the overgrowth syndrome Beckwith-Wiedemann syndrome (BWS). Gain-of-function mutations located in the PCNA-binding domain of CDKN1C result in a spectrum of undergrowth disorders ranging from IMAGe Syndrome and RSS, while loss-of-function mutations in CDKN1C cause BWS. CDKN1C is a cyclin-dependent kinase inhibitor that functions as a negative regulator of cell proliferation. Here, we investigate the effects of gain-of-function CDKN1C mutations on protein stability during cell cycle progression and on cell proliferation to understand the molecular mechanisms underlying IMAGe syndrome.

293T cells transfected with wild type and IMAGe-mutant (p.K278E or p.F276V) and treated with de novo protein synthesis inhibitor, cyclohexamide (CHX), were subject to immunoblotting and demonstrated an increased stability of IMAGe-mutant CDKN1C compared to wild-type CDKN1C. To demonstrate that IMAGe-mutant CDKN1C resulted in impaired progression through the cell-cycle, we transfected them with wild-type, IMAGe-mutant (p.K278E), or BWS-mutant (p.L42P), synchronized the cells, and then assayed the levels of CDKN1C protein over a 24-hour period. Both the wild-type and BWS-mutant cells showed peak levels of CDKN1C at 3 and 15-hour post thymidine release while the IMAGe-mutant showed stable CDKN1C protein expression over 24-hours, indicating a lack of progression through the cell cycle. Cell-cycle phase distribution at the time T0 showed that in IMAGe-mutant transfected cells, a larger proportion of cells were in the G1 phase (36.94%) compared to wild-type CDKN1C, empty plasmid, and BWS-mutant (32.91%). Further, gene expression analysis at 24 hours of CHX onstrated that cells transfected with IMAGe-mutant CDKN1C had smaller and fewer colonies than cells transfected with wild type CDKN1C plasmids. In conclusion, gain-of-function mutations in the PCNA-binding site of CDKN1C increase protein stability by inhibiting degradation by the proteasome during cell cycle progression. Therefore, cells are stalled in G1-phase and have decreased clonogenic potential. Our findings shed light on the molecular mechanism behind IMAGe mutations on cell cycle, proliferation, and undergrowth disorders in humans.

3018T


Mutations in PROP1 and several other transcription factors, including POU1F1 (PIT1), HESX1, and OTX2, cause Multiple Pituitary Hormone Deficiency (MPHD) in humans and mice. PROP1 is co-expressed with stem cell markers including the growth factor receptor, GFRα2 and transcription factor SOX2, SOX9, and OCT4 (1, 2). In mouse fetal development Prop1 expression coincides spatially and temporally with the delamination, migration, and differentiation of progenitor cells in Rathke’s Pouch (RP) into hormone producing cells of the adenohypophysis. Prop1 mutant progenitors fail to undergo these processes, leading to organ dysmorphology, poor vascularization, and postnatal pituitary hypoplasia (3). Prop1 is necessary for regulating expression of two other MPHD genes; repressing Hes1 and activating Pou1f1. Mutations in Prop1 are the most common known cause of MPHD, but most of the patients have no molecular diagnosis. The aim of this work is to identify additional downstream target genes of Prop1 to reveal new candidate genes for the cases of hypopituitarism that remain unexplained and to better understand the pathophysiology of MPHD. We chose an in vitro approach of identifying DNA sequences bound by Prop1 in an immortalized pituitary cell line by ChIP-Seq technology. Because the efficiency and specificity of ChIP depends on the quality of the antibody, we developed a biotin-tagged PROP1 system that permits precipitation of PROP1 with an avidin based detection assay. We identified enrichment of PROP1 binding at the Pou1f1 gene and in genes that encode components of the pathways of cell junction signaling and regulation of epithelial to mesenchymal transition (EMT). We examined expression of several candidate genes in Prop1 mutant mice and confirmed that their expression was altered. These data suggest that Prop1 promotes the transition of progenitors to differentiation by suppressing expression of progenitor markers like SOX genes and repressing claudin gene expression, in order to release tight junctions and permit EMT. In addition, Pou1f1 activates expression of cyclin E, a marker of the differentiation state, and Pou1f1, a marker of the transition state. This study establishes new mechanisms underlying PROP1’s role in hypopituitarism.


3019S

Cell Specific Biochemical Changes in Snyder Robinson Syndrome. J. Albert1,2, L. Wolfe1,2, W. Bone1, T. Markello1,2, D. Adams1,2, C. Schwartz3, R. Gahl4, M. Collins4, L. Tosi4, N. Bhattacharyya4, W. Gahl1,2, C. Boerkoel1, 1) Undiagnosed Diseases Program, Common Fund, Office of the Director, National Institutes of Health, Bethesda, MD, USA; 2) Medical Genetics Branch, National Human Genome Research Institute, Bethesda, MD, USA; 3) J.C. Self Research Institute, Greenwood Genetics Center, Greenwood, SC 29646, USA; 4) Skeletal Clinical Studies Unit, Craniofacial and Skeletal Disease Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD 20892, USA; 5) George Washington University School of Medicine, Washington, DC, USA; 6) Children’s National Medical Center, Washington, DC, USA.

Spermidine and spermine are simple positively charged ubiquitous molecules synthesized by spermidine synthase and spermine synthase (SMS), respectively. Spermidine and spermine homeostasis is crucial for ion channel regulation, transcription and translation. Deficiency of the SMS enzyme disturbs this homeostasis in Snyder-Robinson syndrome (SRS), a form of X-linked intellectual disability. Individuals with SRS have developmental delay, dysmorphic facies, osteoporosis, seizures, and hypotonia. A major cause of morbidity in these individuals is bone disease. To elucidate the mechanism of osteoporosis, we obtained osteoblasts from an individual with SRS and our results demonstrate that there is an osteoblast specific effect of deficient SMS enzyme activity and/or aberrant polyamine homeostasis. SMS fibroblasts and osteoblasts have similar decreases in SMS protein compared with normal controls. However, normal control osteoblasts have less SMS protein compared with normal control fibroblasts suggesting that osteoblasts may be more sensitive to more subtle changes in SMS protein. In both cell lines, we observed a slight decrease in both cytoplasmic and nuclear SMS protein as detected by cellular fractionation and western blot.

Furthermore, the differences in spermine and spermidine levels were significantly lower and higher, respectively in the osteoblasts. These osteoblast specific changes in polyamine metabolism may exemplify the mechanism underlying the cellular specificity of disease in SRS.
3020M Genetic and enzymatic characterization of the anti-oxidant enzyme GPx1 in sickle cell disease patients. M. Beaudoin1, G. Galarmeau1, A. Binding2, M. Reid3, I. Hambleton4, C. McKenzie4, J.N. Hirschhorn4,5, M. Warner2, G. Lettre1. 1) Montreal Heart Institute and Université de Montréal, Montreal, Quebec, Canada; 2) McGill University, Montreal, Quebec, Canada; 3) Tropical Medicine Research Institute, The University of the West Indies, Mona, Jamaica; 4) The University of the West Indies, Cave Hill, Barbados; 5) Children’s Hospital Boston, Boston, MA, USA; 6) Broad Institute, Cambridge, MA, USA.

Introduction: Sickle cell disease (SCD) is a genetic blood disorder characterized by anemia, pain crises, strokes and early death. Affected individuals inherit two copies of a mutation in the gene encoding β-globin, one of the two subunits of hemoglobin. Despite of this simple mode of inheritance, SCD is characterized by patient-to-patient clinical heterogeneity. Alpha-thalassemia and fetal hemoglobin levels are two known modifiers of SCD severity. The goal of our study was to identify new genetic modifiers of SCD using whole-exome sequencing followed by functional characterization.

Methods: We sequenced the whole-exome of 19 patients selected from the Jamaican SCD birth cohort. Over 18 years of follow-up observations, these patients presented with severe SCD-related complications when compared to the rest of the cohort (0.07 vs. 0.57 severe events per year, p=0.001). Analysis of the whole-exome sequence data identified a common missense variant in GPx1 (Pro200Leu, rs1050450) that is enriched in Jamaican patients with mild SCD. We obtained DNA and red blood cell (RBC) membranes from 40 SCD patients recruited at the Royal Victoria Hospital in Montreal, Canada. We genotyped GPx1 Pro200Leu using a high resolution melting protocol. We measured GPx1 protein level and enzymatic activity with ELISA and absorbance assays, respectively. Results: The GPx1 Pro200Leu allele is enriched in 19 “mild” Jamaican SCD patients when compared to populations of African ancestry from the 1000 Genomes Project (47% vs. 25-34%). In 95 patients from the Jamaican SCD birth cohort, GPx1 Pro200Leu is associated with a composite score that summarizes severe SCD-related events collected from birth until adulthood (p=0.02). In RBC membranes from SCD patients, GPx1 Pro200Leu is not associated with GPx1 protein amount (p=0.95), but is correlated with GPx1 enzymatic activity (p=0.049). Compared to GPx1 enzymatic activity per RBC, Pro200Leu is an enzyme that protects RBC from free radical damages. Previous studies have implicated free radical production in SCD severity, and in particular in the etiology of vaso-occlusive crises. Our results indicate that a common missense variant in SCD patients regulates the activity of an important enzyme implicated in the RBC defense against oxidation. Larger studies are required to determine how this enzymatic variation may impact clinical heterogeneity in SCD.

3021T Structural, molecular and cellular impact of the Ogden syndrome Ser37Pro mutant N-terminal acetyltransferase Naa10. G.J. Lyon1, L. Myklebust2, P. Van Damme3, S. Stove2, N. Reuter1, M. Doerfel1, Y. Wu1, G. Liszczak1, A. Abboud2, T. Kalvik2, C. Grauflisch2, V. Jonckheere2, H. Kaasa2, R. Marmorstein2, K. Gevaert3, T. Arnesen2. 1) Cold Spring Harbor Laboratory, New York City, NY; 2) Department of Molecular Biology, University of Bergen, N-5020 Bergen, Norway; 3) Department of Medical Protein Research, VIB, B-9000 Ghent, Belgium; 4) Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA.

The infantile X-linked lethal Ogden syndrome was the first reported human genetic disorder associated with a mutation in an N-terminal acetyltransferase (NAT) gene. The eight affected and now deceased males in two unrelated families had a Ser37Pro mutation in the gene encoding Naa10, the catalytic subunit of NatA, the major human NAT. NATs transfer an acetyl moiety from acetyl coenzyme A (Ac-CoA) to the primary α-amino group of nascent polypeptides as they emerge from ribosomes, and proteomic analyses have revealed that 50-90% of yeast, plant, fruit fly and human proteins are Nt-acetylated. In order to understand the detrimental impact of Naa10 Ser37Pro in humans, we performed structural, molecular and cellular investigations. Structural modeling and molecular dynamics simulations of the human NatA and its Ser37Pro mutant suggest differences in regions involved in catalysis and at the interface between Naa10 and the auxiliary subunit Naa15. In agreement, biochemical data demonstrate a reduced catalytic capacity and a reduced complex formation of Naa10 Ser37Pro with Naa15 and Naa50, indicating a disruption of the NatA complex. Patient derived Naa10 mutant cells show reduced Nat-acylation for a subset of NatA-type substrates compared to wild type Naa10 cells. More specifically, N-terminal acetylation analysis of B-cells and fibroblasts reveal an in vivo perturbation of Naa10 (NatA) mediated Nt-acylation in Ogden syndrome males as some NatA-type substrates are less Nt-acetylated, with some of these substrates previously found to be affected in their Nt-acylation status by siNatA-mediated knockdown. Ogden syndrome fibroblasts displayed reduced cell proliferation and migration capacity as well as abnormal growth patterns. Introduction of human NatA into yeast that lack endogenous NatA could rescue a specific heat stress phenotype, whereas NatA Ser37P failed to. Additional proteomic analyses of these strains revealed a reduced expression of ribosomal proteins in the Ogden strains. The impaired NatA Nt-acetylation and catalytic capacity, together with reduced in vivo Nt-acylation and abnormal cellular phenotypes provide new insights into the underlying molecular and cellular mechanisms of Ogden syndrome. More broadly, these studies broaden our understanding of the role of N-terminal acetylation of proteins, which is a vastly understudied but very prevalent co- and post-translational modification.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3022S

Analyzing KMT2D/MLL2 missense mutations as causative in Kabuki syndrome and testing U1snRNAs-based approaches to revert KMT2D/MLL2 splicing defects. L. Micali, B. Mandrian, P. De Nittis, B. Augello, C. Fusco, A. Romanaro, B. Piccinii, M.T. Pellicci, C. Rinaldi, A. Di Lauro, T. Verri, G. Merla. 1) Medical Genetics Unit, IRCCS Casa Sollievo Della Sofferenza Hospital, San Giovanni Rotondo, Italy; 2) Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy; 3) Department of Biological and Environmental Sciences and Technologies, University of Salento, Lecce, Italy.

Histone methylation is an epigenetic mechanism by which spatial and temporal expression of distinct genes and pathways are regulated at precise developmental stages. The discovery of histone methyltransferase KMT2D and demethylase KDM6A genetic alterations in Kabuki syndrome patients expanded and highlighted the role of histone modifiers in causing congenital anomalies and intellectual disability syndromes. Kabuki syndrome is a multiple congenital malformation syndrome characterized by facial features, skeletal anomalies, dermatoglyphic abnormalities, mental retardation, and postnatal growth deficiency. KMT2D and KDM6A play important roles in the epigenetic control of active chromatin states modulating the expression of genes essential for embryogenesis and development. We performed a mutational screening on about 400 Kabuki patients by sequencing, MLPA, and qPCR. Among the KMT2D mutations we identified 49 missense variants across the entire length of the KMT2D gene and 9 splicing site variants that result in a frameshift with the generation of premature stop codons. In this study we ascertain the pathogenicity of KMT2D missense mutations through an integrative analysis of bioinformatics tools and biochemical and cellular assays. We used an innovative in silico approach that combines comparative genomics, evolutionary search tools with cell lines, IMFs, and adapted U1snRNAs complementar to the mutated splice site, allowed us to revert the splicing defects for all the analysed mutations. The experimental approaches proposed offer a valuable support to estimate the real deleterious effect of KMT2D missense variants, a main issue in the field.

3023M

Infantile myofibromatosis: Gene discovery leads to novel treatment paradigm based on PDGFRB gain-of-function mutations. K. Oishi, R. Evans, T. Silvers, J.A. Martinetti. 1) Dept Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY; 2) Dept Pediatrics, Icahn School of Medicine at Mount Sinai, New York, NY.

Background: Infantile myofibromatosis (IMF; MIM 228350) is an autosomal dominant disorder characterized by solitary or multiple tumors of skin, muscles, viscera and bones. Tumor growth may begin during fetal development and can be present throughout adult life. Occasionally, these tumors may spontaneously regress without treatment. Recently, our group and others identified germline mutations in the phosphotyrosine-dependent growth factor receptor beta gene (PDGFRB) as causing the majority of IMF cases. The effect of IMF mutations on PDGFRB signaling and the molecular pathway(s) which are dysregulated and result in IMF tumors are unknown but will need to be understood if a targeted therapy is to be devised. Objective: To characterize the molecular functions of the IMF-causative mutant PDGFRB. Design/methods: Wild-type (WT) human PDGFRB cDNA was cloned into pcDNA3.1 expression vector. Two IMF-causing mutations, R561C and P660T, were introduced using site-directed mutagenesis. Mutant proteins were transiently expressed in COS-7 cells, which lack PDGFRB, and the residual cell lines were stimulated with PDGF-BB for 30 min after serum starvation. Phospho-PDGFRB and phospho-AKT antibodies were used for western blot. Results: Cells expressing mutant PDGFRB proteins had an increase in basal phosphorylation of PDGFRB at Y751 and Y587, compared to cells expressing WT protein. Following PDGF-BB stimulation, all cells demonstrated an increased phosphorylation of PDGFRB, but the IMF mutant cells still maintained markedly higher phosphorylation than control cells. Mutant PDGFRB-expressing cells had increased activation of AKT following PDGF-BB stimulation. Levels of phospho-PDGFRB and phospho-AKT were both consistently higher in P660T than R561C. Imatinib, a tyrosine-kinase inhibitor, suppressed the PDGFRB activation. Conclusions: IMF-causing PDGFRB mutations result in constitutive gain-of-function (GOF). It was suggested that the mutations induce excessive kinase activity and binding capacities with other signaling molecules through increased phosphorylation of PDGFRB. Hence, we propose a new therapeutic approach for IMF.

3024T

Biochemistry of UBA1 Mutations that Cause Infantile X-Linked Spinal Muscular Atrophy. C.D. Balak, J.M. Hunter, G. D’Urso, D. Wiley, L. Baumbach-Reardon. 1) Translational Genomics Research Institute, Phoenix, AZ; 2) University of Miami, Miller School of Medicine, Miami, Florida.

Spinal muscular atrophy (SMA) is the leading genetic cause of infantile death. Mutations in the SMN1 gene account for a large percentage of SMA. However, in 2008, Dr. Lisa Baumbach-Reardon’s group discovered that both a rare and lethal form of X-linked SMA can be caused by mutations in the UBA1 gene. UBA1 is an essential non-redundant gene found highly conserved in all organisms from yeast to man; coding for the Ubiquitin Activating Enzyme 1 (UBA1). UBA1 is the pinnacle enzyme in the Ubiquitin Proteasome System (UPS) that is responsible for the turnover and degradation of most proteins in every living cell. The role of Uba1 in the UPS is to activate and transfer a Ubiquitin (Ub) molecule to an E2 Ubiquitin-conjugating enzyme. These Ubiquitin-E2 complexes, in conjunction with E3 ubiquitin ligases, transfer the activated Ub to targeted proteins “tagging” them for degradation by the proteasome. This process of tagging and degrading proteins is essential for the survival of every cell. Our goal is to understand how mutations in Uba1 after its biochemical function and the disease mechanism in XL-SMA and developing treatments for XL-SMA. Uba1 activates Ub in a 2 step process. First, Uba1 adenylates Ub in a reaction that transfers the energy of ATP to Ub. The activated Ub is then covalently bound to a cysteine residue of Uba1 to form a thiolester bond in the second step. The Ub is then transferred to a cysteine residue of an E2 enzyme. To date, all of the mutations that have been found in Uba1 cluster in one domain of Uba1 thought to be important for ATP and Mg2+ binding and adenylation, or for protein-protein interactions. Therefore we have developed a continuous kinetic assay to measure the adenylation activity of wild type (WT) and mutant (MT) forms of Uba1. Our initial experiments demonstrate Uba1 mutants retain adenylation activity, thus suggesting alterations in other catalytic activities or protein-protein interactions. We are investigating the rate of thiolester bond formation between Ub and Uba1, and triacontiation of activated Ub to E2 enzymes. Finally we are measuring interactions between mutant forms of Uba1 and key SMA proteins, Smn and gigaxonin. The results of our studies will identify the causal biochemical defect in XL-SMA and pave the way forward for identifying pharmaceutical targets and treatments for XL-SMA.
NIPBL plays a cardinal role in transcriptional regulation, whose disruption suggests the differential functional roles of NIPBL isoforms, and full-length maintenance of sister chromatids cohesion during cell division. Our results suggest that full length NIPBL primarily possesses gene expression regulation, while the presence of hitherto un-described shorter NIPBL isoform(s). These data seen in CdLS patient samples, this NIPBL null clone recapitulates some full length NIPBL is not required. RNA-seq using this clone discovered 3732 proteins expressed in this clone, antibody demonstrated the complete absence of full length NIPBL. Interestingly, sister chromatid separation defects were not obvious in this clone, frameshift mutations using Cas9/CRISPR system. Among the clones common cell line originating from human embryonic kidney, with NIPBL mutation in exon 3 (NIPBL null clone). Western blot with N-terminus NIPBL antibody demonstrated the complete absence of full length NIPBL. Interrestingly, sister chromatid separation defects were not obvious in this clone, suggesting that for the prevention of premature sister chromatid separation, full-length NIPBL is not required. RNA-seq analysis of this clone disclosed genes (out of 35636) whose expression was significantly altered from wild type 293FT line. The expression of many HOX genes was altered in the NIPBL null clone. Since abnormal expression of HOX genes were often seen in CdLS patient samples, this NIPBL null clone recapitulates the some features of transcriptional alteration seen in CdLS. Surprisingly, RNA-seq demonstrated the normal expression level of NIPBL transcript, suggesting the presence of hitherto un-described shorter NIPBL isoform(s). These data suggest that full length NIPBL primarily possesses gene expression regulation, while the absence of full length NIPBL appears not to be required for the maintenance of sister chromatids cohesion during cell division. Our results suggest the differential functional roles of NIPBL isoforms, and full-length NIPBL plays a cardinal role in transcriptional regulation, whose disruption causes CdLS.
Disruption of FMR1, RA1, MBDS, and TCF4 results in abnormal expression of circadian rhythm genes and sleep disturbances in fragile X, Smith-Magenis, 2q23.1 deletion, and Pitt-Hopkins syndromes. S.V. Mullerlegana, L. Pugliesi, B. Burns, R. Tahir, Z. Shah, G. Gu, D.L. Nelson, W.-H. Tan, S.H. Elsea. 1) Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Human and Molecular Genetics, Virginia Commonwealth University School of Medicine, Richmond, VA; 3) St. Joseph's Regional Medical Center, Paterson, NJ; 4) Division of Genetics, Boston Children's Hospital; Harvard Medical School, Boston, MA USA.

Sleep disturbance is prevalent in neurodevelopmental disorders such as fragile X syndrome (FXS, FMR1), Smith-Magenis syndrome (SMS, RA1), 2q23.1 deletion syndrome (del 2q23.1, MBDS), and Pitt-Hopkins syndrome (PTHS, TCF4). The frequent comorbidity of sleep disturbance with neurodevelopmental disorders implies overlapping pathways that exist to lead to shared sleep phenotypes. We hypothesize that sleep phenotypes in FXS, SMS, del 2q23.1, and PTHS have common etiologies and that FMRP, RA1, MBDS, and TCF4 function in overlapping circadian pathways that lead to these sleep disturbances. While sleep disturbance is well described in SMS and FXS, we sought to delineate the sleep anomalies in del 2q23.1 and PTHS. Through caregiver surveys and review of the literature, we found that difficulty maintaining sleep was a feature of both disorders. Furthermore, del 2q23.1 patients have difficulty falling asleep and experience night or early awakenings, while PTHS patients have frequent night awakenings. Thus, common sleep issues in these 4 disorders include reduced total sleep time and frequent nighttime and early morning awakenings. We then evaluated expression of the causative genes for these 4 neurodevelopmental disorders in patient lymphoblastoid cell lines (LCLs), which revealed altered levels of MBDS, TCF4, RA1, and FMR1 across each of the 4 syndrome cell lines, suggesting a possible common etiology in all 4 syndromes. To support involvement of these genes in circadian gene expression, we identified putative E-boxes in MBDS, RA1, and FMR1, while TCF4 regulates gene expression at E-box containing promoters. Expression levels of circadian genes (NR1D2 and CRY2) were reduced in patient LCLs, suggesting key roles for MBDS, TCF4, RA1, and FMRP in the regulation of circadian gene expression. Pathway analysis of MBDS, RA1, and TCF4 knockdown microarray data show dysregulation of pathways associated with sleep disturbance in these disorders. We hypothesize that sleep phenotypes in FXS, SMS, del 2q23.1, and PTHS are associated with ciliary-centrosomal protein TOPORS. B. Czub, A.Z. Shah, P. Kuczeg, G. Alfano, C. Chakarova, S.S. Bhattacharya. Institute of Ophthalmology, University College London, London, United Kingdom.

Purpose: TOPORS (MIM 609507) is a ubiquitously expressed gene implicated in autosomal dominant retinitis pigmentosa (RP [MIM 268000]); mutations are known to cause only RP with no systemic symptoms. We performed yeast two-hybrid (Y2H) screens of human retinal cDNA libraries to identify proteins interacting with TOPORS, which could explain the restricted phenotype. This led to the isolation of a soluble fragment of integral membrane protein 2B (ITM2B [MIM 603004]), mutations in which are associated with inherited forms of dementia and, most recently, a dominant form of retinal dystrophy. Methods: Matchmaker™ Gold Y2H System (Clontech) was used for library screening and testing direct protein-protein interactions (PPIs). Results were validated in HeLa and hTERT-RPE1 (RPE1) cell extracts by co-immunoprecipitation. Immunofluorescence methods were used for co-localisation studies in RPE1 cells and mouse retina cryo-sections. ProteoExtract® Subcellular Proteome Extraction Kit (Merck) was used to prepare cellular fractions (HeLa). Results: The TOPORS-ITM2B PPI, identified by screening, was selected for further study due to the role of ITM2B in retinopathy. It was detected in complexes with TOPORS precipitated from HeLa and RPE1 extracts, and both proteins co-localised at the centrosome in RPE1 cells. In mouse retina sections the strongest ITM2B signal was observed in rod photoreceptor inner segment, while this localization was lost for p.Arg106Pro [p.Gln67del] and c.810+1G>T) in an isolated case with cone-rod dystrophy. Mutations (c.199_201del POC1B) associated could explain their roles and the phenotypes resulting from their mutations. Furthermore, evaluation of their relationship could lead to a better understanding of neurodegeneration mechanisms overall. Subsequent work will aim to delineate the roles of the various ITM2B peptides, generated by physiological proteolytic cleavage of the membrane-bound precursor protein, which could clarify the centrosomal co-localisation with TOPORS.
3032M

**Purpose:** TOPORS (MIM 609507) is a ubiquitously expressed gene, encoding a protein localising to the nucleus and centrosome of dividing cells, and to the basal bodies of ciliated cells (including connecting cilium of photoreceptor cells in the retina). Interestingly, mutations in TOPORS are only implicated in dominant retinitis pigmentosa (MIM 268000). This study was undertaken to better understand the role of TOPORS in the cell. Methods: Subcellular localisation of TOPORS was examined in HTERT-RPE1 (RPE1) cell line by confocal immunofluorescence in cultures of increasing confluence. Stress experiments were performed to test observations. Co-immunoprecipitation was performed on RPE1 total cell extract; cellular fractions were prepared using ProteoExtract® Subcellular Proteome Extraction Kit (Merck); both analyses were performed using RT-PCR on cDNA from human cell lines and retina was performed to identify alternatively spliced variants. TOPORS antibody was validated using RNAi knockdown. All experiments were performed on endogenous proteins. Results: Confluenge studies showed TOPORS localises primarily to the nucleus but is found in the cytoplasm, centrosome and close to the cell membrane with increasing cell density. Furthermore it localises along actin filaments. Stress conditions (such as changes in osmolarity) do not appear to effect TOPORS localisation. Results suggest TOPORS could interact with actin and other cytoskeletal components. There is one alternatively spliced isoform reported for TOPORS. The commercially available TOPORS antibody is able to detect both long and short isoforms, as confirmed by RNAi knockdown experiments. The short isoform for TOPORS was detected by Western blotting in normal primary human lymphocytes, and could account for the cytoplasmic localisation observed. However, to date it has not been detected in human retina. Preliminary evidence suggests differential expression of both TOPORS isoforms in different cellular compartments. Conclusions: In addition to being a nuclear and centrosomal protein, TOPORS is a novel component of the actin cytoskeleton. This could be due to the short TOPORS isoform. Our findings suggest a potential role of TOPORS in mediating actin-related cellular changes through its E3 ubiquitin ligase activities and could indicate a role in cell signalling. This work highlights the increasing complexity of protein function and the importance of understanding their role(s) in health and disease.

3033T
Loss of the ribosome-associated factor Sbds in murine models of Shwachman-Diamond syndrome leads to aberrant polysome profiles. H. Liu1, 2, M.E. Tourlakis3, 4, R. Gandhi1, P. Hu1, J.M. Rommens1, 2, 3, 1) Program in Genetics & Genome Biology, SickKids Research Institute, Toronto, Ontario, Canada; 2) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 3) Department of Biochemistry & Medical Genetics, University of Manitoba, Winnipeg, Manitoba, Canada.

Shwachman-Diamond syndrome (SDS) is an autosomal recessive disease characterized by growth retardation, exocrine pancreatic dysfunction, skeletal dysplasia, cognitive impairment and hematological abnormalities with increased susceptibility for leukemia. SDS is caused by mutations in Sbds. A current hypothesis proposes that SBDS/Sbds functions together with EFTUD1, EFTud1 and is involved in the release of EIF6/EIF6 from the pre-60S complex, enabling translation initiation. Mouse embryos deficient for Sbds (SbdsR1267>T) show notable growth retardation and cell culture models reveal reduced protein synthesis. To assess translation deficiency, polysome profiles of fetal organs from this SDS murine model (E18.5) were examined and found to show anomalies compared to matched controls (Sbds+/+ and SbdsR1267>T), with reduction in 80S peaks as well as preservation of polysomes. The low 80S peak suggested impairment of translation initiation; however the prominent polysomes were unexpected and appeared inconsistent with deficiency in protein synthesis. To investigate how the translateome is affected in SDS, total and polysomal mRNAs of mutant and control fetal liver samples were studied using cDNA microarray analyses. By comparing individual polysomal transcript to respective total transcript level, we could establish which genes were being actively translated and/or retained by the polysomes with loss of Sbds function. With normalisation, 173 genes showed reduced association with polysomes, while 666 showed increased association. Transcript long transcripts with lower GC content were favored for polysome association. However, the changes of polysome association did not correspond to respective steady state protein levels, as shown by western immunoblotting and label-free mass spectrometry. Further, the proteome revealed that relative levels of ribosomal proteins and translation factors are unaltered in SDS fetal livers. The observations of prominent polysomes and the large number of transcripts with high polysome association in absence of corresponding increase in the protein levels suggest that loss of Sbds function leads to the formation of ribosomes which may be deficient in elongation or ribosome recycling, possibly due to impaired removal of EIF6 prior to subunit joining.

3034S
Examining the Molecular Mechanisms Underlying SRCAP Mutations in Floating-Harbor Syndrome. R.L. Hood1, 2, K.M Boycott1, 2, 3, 4, W.L. Stanhope1, 2, D.E Bulman1, 2, 1) Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada; 2) Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa; 3) Department of Genetics, Children’s Hospital of Eastern Ontario; 4) Department of Pediatrics, University of Ottawa; 5) The Sprott Centre for Stem Cell Research, Regenerative Medicine Program, Ottawa Hospital Research Institute; 6) Department of Cellular and Molecular Medicine, University of Ottawa.

Floating-Harbor Syndrome (FHS; OMIM# 136140) is a rare genetic disorder which is characterized by short stature, delayed osseous maturation, language deficits, and unique dysmorphic facial features. FHS typically occurs sporadically; however a few instances of autosomal dominant transmission have been reported. Since our first report of mutations in SRCAP as the cause of FHS, we have identified a total of 50 mutation positive patients. The mutations identified were shown to be de novo in all cases where parental DNA was available. SRCAP encodes a SNF2-related chromatin-remodeling factor which is known to be a coactivator of CREB-binding protein (CREBBP or CBP) and CBP-mediated transcription. Additionally SRCAP is part of a large complex which catalyzes ATP-dependent substitution of the variant histone H2A.Z into nucleosomes. Based on the existence of patients with deletions encompassing all of SRCAP who do not have FHS, we postulate that the mechanism of disease is not due to hapsininsufficiency. Furthermore we expect that FHS mutations result in widespread gene dysregulation; however the mechanisms and biological pathways underlying the FHS phenotype are currently unknown. We are directing our efforts at elucidating the molecular mechanisms underlying the bone pathogenesis of FHS. Taken together these studies will increase our understanding of how truncating mutations of SRCAP alter its function and gain insight into the pathways which contribute to the development of Floating-Harbor syndrome.
3035M  
Functional characterization and pharmaceutical correction of eleven novel mutations identified in Indian CF population. R. Prasad1, H. Sharma2, M. Jollivet Souchez2, F. Becq2, P.M. Kruczek1  
1) Department of Biochemistry, PGIMER, Chandigarh, Chandigarh, India; 2) Laboratoire Signalisation et Transports Ioniques Membranaires, Université de Poitiers, CNRS, Poitiers, France.

Cystic Fibrosis is an autosomal recessive disorder is usually considered as rare disease for Indian population hence much less is known about this disease in context with Indian sub continent. Recently we have established a spectrum of mutations in CFTR gene from classical Cystic Fibrosis as well as for CF patients in India. For this purpose we have screened 11 novel mutations, among them S549N, L69H in classical CF and F508del in male patients and L69H in female patients. Whereas S549N mutation can be categorized into a class II/III mutation observed in F508del mutation. This defect is rescued by the corrector VX809. Mutations G126S, F87I, S118P, H139Q, F157C, F494L, E543A, Y852F, D1270E in CBADV males were among the rare missense mutations (Sharma et al., 2009). In this study we have attempted to conduct in vitro expression analysis to well establish genotype and phenotype correlation and to characterize these four rare missense mutations according to the mechanism that disrupt CFTR protein function. All four mutations from Indian population were characterized by expressing pEGFP-CFTR constructs in BHK-21 cells via 3 step technique viz: CFTR cellular localization was determined by confocal microscopy, where as Western blot analysis and automated iodic efflux assays was used to determine CFTR maturation processes and its channel activity respectively. In Western blot analysis only b-band is obtained for L69H substitution as for F508del whereas in the case of other mutants both b and c bands were found, indicating L69H mutation impair CFTR maturation process, the finding was again confirmed by confocal imaging. Iodic Efflux assay revealed significant decrease in channel activity for L69H and S549N mutants whereas c and b bands were observed in cases of substitutions in channel activity when cells were incubated at 27 °C. When the effect of CFTR correctors was checked on different mutants, it was found that VX809 significantly ameliorate the defect caused by L69H mutation. Mutations G126S, F87I, S118P, H139Q, F157C, F494L, E543A, Y852F, D1270E have no impact on CFTR maturation and function. Thus we can conclude that L69H mutation is a class II CF mutation causing impaired maturation leading to protein degradation and Cl- ions impermeability as observed for F508del mutation. This defect is rescued by the corrector VX809. Whereas S549N mutation can be categorized into a class II/III mutation causing impaired maturation and reduced channel activity.

3036T  
Understanding the role of EYS by identification and characterisation of its retinal interacting partners. P.M. Kruczek1, G. Alfano1, C. Czub1, A.Z. Shah1, A.C. Zelhof2, S.S. Bhattacharya1  
1) Department of Genetics, UCL Institute of Ophthalmology, London, United Kingdom; 2) Department of Biology, Indiana University, 1001 East Third St, Bloomington, IN 47405, USA.

Purpose: The eyes shut homolog (Drosophila) gene (EYS [MIM 612424]) is one of the largest genes known to be expressed in the human retina. Mutations in EYS are the most common cause of autosomal recessive retinitis pigmentosa (arRP) (RP25 [MIM 602772]; RP [MIM 268000]). EYS encodes a protein named EYS whose biological role in humans is presently unclear, however, a Drosophila ortholog of EYS has been shown to be essential for the organization of photoreceptor rhabdomere. Prominin, a transmembrane protein associated with microvilli of the Drosophila photoreceptor rhabdomere. Prominin is highly conserved throughout evolution and interestingly, mutations in human Prominin (PROM1 [MIM 604365]) cause retinopathies. This study was undertaken to investigate the relationship of EYS and Prominin-1 in humans, and to identify novel interacting partners for EYS in the retina.

Methods: Putative interaction of EYS and Prominin-1 was examined by immunohistochemistry in cultured Y70 and ARPE19 cells using novel potential binding partners of EYS were identified using Yeast 2-hybrid technology (Y2H) - baits, consisting of full-length EYS and three fragments covering N-terminal, middle and C-terminal parts of the protein, were screened against human retinal cDNA libraries. Results: Preliminary data suggest that EYS and Prominin-1 may co-localise in Y70 cell line. Furthermore, exogenous expression of EYS and Prominin-1 in ARPE19 cell line suggests that Prominin-1 might be required for proper localisation of EYS. Y2H screens revealed six proteins that potentially interact with EYS: AIPL1, SERF2, FBNI, PEA3, AIPL1A and COX7C. Conclusions: Investigation of interaction between EYS and Prominin-1 suggests that these proteins may possess similar functions to their Drosophila orthologs and could play a role in maintaining the integrity of human photoreceptors. Y2H screens have identified six potential interaction partners of EYS which may help us elucidate the role of the EYS gene in human health and disease, but it will also shed light on physiology and maintenance of human photoreceptors.

3037S  
A pathologic genomic rearrangement in Incontinentia pigmenti locus reveals a transcriptional regulation of NEMO gene by p63 family proteins. M. Conte1, M. Paciolla1, E. Candidi1, A. Zoppino1, G. Melino2, M.V. Usini1, F. Fusco1  
1) Institute of Genetics and Biophysics, Naples, Italy; 2) Universita’ di Tor Vergata, Rome, Italy.

Incontinentia pigmenti (IP, MIM308300, 1/10.000) is an X-linked dominant neurocutaneous disease, lethal in males and caused in female by a mutation in the IKBKG/NEMO gene that encodes for NEMO/IκBγamma regulatory protein required for the activation of the NF-kappaB pathway. The typical IP skin defects, always present in IP patients, are due to the mosaic expression of NEMO/IKBKG mutation related to X-chromosome inactivation in heterozygous females. NEMO/IKBKG is transcribed by the bidirectional promoter B, shared by the two overlapping IKBKG/NEMO and G6PD genes and promoter A, unidirectional for NEMO expression. We recently described an IP de novo case due to a genomic deletion, named IKBKgdel, that eliminates the NEMO principal promoter B and preserves its expression by the secondary promoter A. The patient with IKBKgdel, showed only skin abnormalities suggesting that an early X-inactivation of mutated allele occurred in the other extracutaneous tissues, or the promoterA activity of NEMO prevented the IP alterations. We analysed promoterA transcription in vivo on RNA from different layers of the skin from healthy donors and in vitro on RNA from differentiated HaCat keratinocytes to obtain a better insight into its role in the skin. Bioinformatic sequence comparison using the multiple alignment software m-Vista of the 5’ of the NEMO promoter A in different species predicted several p53-like binding sites in the highly conserved regions. Here, we present our findings by in vivo and in vitro assays on the different role of p53-related proteins, p63 and p73. Data suggest that p63 regulates promoter activity and expression of p53 family, Dnpp63alpha and TaP63alpha, which are key transcription factors in epithelial development and differentiation, are able to regulate the expression of promoter A of NEMO during the skin differentiation. Moreover, we found that the transcription of NEMO promoter A is not affected by p53-related proteins, p63 and p73 on NEMO and G6PD will be illustrated.

3038M  
1) Guangzhou Women and Children’s Medical Center, Guangzhou, China; 2) Department of Psychiatry, The University of Hong Kong, Hong Kong SAR, China; 3) Department of Surgery, The University of Hong Kong, Hong Kong SAR, China; 4) Department of Pathology, The University of Hong Kong, Hong Kong SAR, China; 5) Center for Genomic Sciences, The University of Hong Kong, Hong Kong SAR, China.

Diffuse oesophageal leiomyomatosis (DOL) is a rare disorder characterised by proliferation of smooth muscle cells in the oesophageal wall. DOL is present in 5% of Alport syndrome (AS) patients. AS is a hereditary disease that involves varying degrees of hearing impairment, ocular changes and progressive glomerulonephritis leading to renal failure. In DOL-AS patients, the genetic defect consists of a deletion involving the COL4A5 and COL4A6 genes on the X chromosome. We present a two-generation family with two members (mother and son) affected with DOL and none of the features of AS, Exome sequencing and Copy Number Analyses revealed a new 40kb deletion encompassing from intron 2 of COL4A5 to intron 1 of COL4A6 at Xq22.3. The breakpoints were also identified. Possible confounding pathogenic exonic variants in genes known to be involved in other extracutaneous matrices disorders were also shared by the two affected individuals. Mosaic analysis of cell lines and ancillary lines of analysis were performed to investigate the presence of mosaic deletions involving heterozygous and homozygous deletions of the 40kb deletion. This is the first case where a “classic” DOL-AS COL4A6 deletion is described in patients with no AS and the first report of gonosomal mosaicism associated to this condition. COL4A5/COL4A6 genes should also be investigated in cases of isolated DOL.
3039T

A large number of conditions have been described where severe intellectual disability and ataxia are found in patients with cerebellar hypotrophy. The etiology of most of these syndromes is unknown, and molecular genetic testing is not available.

Here we describe three similar but unrelated consanguineous families with autosomal recessive cerebellar ataxia and intellectual disability syndrome, associated with early onset cerebellar atrophy and relative macrocephaly. P. Stanley1, A.C. Thomas1, H. Williams1, N. Seidl-Salvia1, C. Bacchelli1, M. O’Sullivan1, L. Ocak4, K. Mengrelia1, M. Ishida2, G. Anderson2, D. Morrough1, M. Ryten2, J.M. Saravia5, F. Ramos4, B. Farren2, D. Saunders2, P. Gissen1, A. Straatmaarm-Iwanowska5, F. Baas6, N. Wood4, R. Robinson1, J. Hershey1, H. Houlden1, R. Hennekam2, J. Hurst3, R. Scott1, M. Bilner-Glindzicz1, G. Moore4, S.B. Sousa1.

1) University College London, Institute of Child Health, London, United Kingdom; 2) Great Ormond Street Hospital, London, United Kingdom; 3) University College London, Institute of Neurology, London, United Kingdom; 4) Hospital Pediátrico de Coimbra, Portugal; 5) University of Amsterdam, Holland.

3042T
Identification of a novel gene causing a recognizable and distinct autosomal recessive cerebellar ataxia and intellectual disability syndrome, associated with early onset cerebellar atrophy and relative macrocephaly. P. Stanley1, A.C. Thomas1, H. Williams1, N. Seidl-Salvia1, C. Bacchelli1, M. O’Sullivan1, L. Ocak4, K. Mengrelia1, M. Ishida2, G. Anderson2, D. Morrough1, M. Ryten2, J.M. Saravia5, F. Ramos4, B. Farren2, D. Saunders2, P. Gissen1, A. Straatmaarm-Iwanowska5, F. Baas6, N. Wood4, R. Robinson1, J. Hershey1, H. Houlden1, R. Hennekam2, J. Hurst3, R. Scott1, M. Bilner-Glindzicz1, G. Moore4, S.B. Sousa1. 1) University College London, Institute of Child Health, London, United Kingdom; 2) Great Ormond Street Hospital, London, United Kingdom; 3) University College London, Institute of Neurology, London, United Kingdom; 4) Hospital Pediátrico de Coimbra, Portugal; 5) University of Amsterdam, Holland.

A large number of conditions have been described where severe intellectual disability and ataxia are found in patients with cerebellar hypotrophy. The etiology of most of these syndromes is unknown, and molecular genetic testing is not available.

Here we describe three similar but unrelated consanguineous families with autosomal recessive cerebellar ataxia and intellectual disability syndrome, associated with early onset cerebellar atrophy and relative macrocephaly. P. Stanley1, A.C. Thomas1, H. Williams1, N. Seidl-Salvia1, C. Bacchelli1, M. O’Sullivan1, L. Ocak4, K. Mengrelia1, M. Ishida2, G. Anderson2, D. Morrough1, M. Ryten2, J.M. Saravia5, F. Ramos4, B. Farren2, D. Saunders2, P. Gissen1, A. Straatmaarm-Iwanowska5, F. Baas6, N. Wood4, R. Robinson1, J. Hershey1, H. Houlden1, R. Hennekam2, J. Hurst3, R. Scott1, M. Bilner-Glindzicz1, G. Moore4, S.B. Sousa1.
3043S

The role of SRSF10 in SMN1/2 splicing. S. Brøner1, Y. Hua2, T.K. Doktor1, G.H. Bruun1, M.R. Larsen1, A.R. Krainer1, B.S. Andresen1. 1) Department of Biochemistry and Molecular Biology, University of Southern Denmark, 5230 Odense M, Denmark; 2) Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA.

SR-proteins are important regulators of pre-mRNA splicing through sequence-specific binding to motifs in splicing silencers and enhancers, but their in vivo binding motifs are still insufficiently characterized. The intrinsic splicing silencer, ISS-N1, located in intron 7 of the SMN genes is the target of splicing correction by Antisense Oligonucleotide blocking, which is currently in clinical trials for treatment of the neurodegenerative disease Spinal Muscular Atrophy (SMA). Until now, only members of the heterogeneous nuclear ribonucleoprotein family, like hnRNPA1, are known to bind to the ISS-N1 silencer and inhibit splicing. We have performed RNA affinity binding assays combined with iTRAQ labelling and LC-MS/MS analyses in order to elucidate the binding of different splicing factors to this ISS-N1 motif. One new candidate is the splicing factor SRSF10, which belongs to the well-characterized SR-protein family. SRSF10 was found to bind significantly stronger to the wild type ISS-N1 motif compared to a mutant motif in which two A-G substitutions destroy the two hnRNPA1 motifs. Our preliminary data using over expression of SRSF10 and SMN minigens have shown that the long isoform of SRSF10 binds to the ISS-N1 motif and can affect the splicing of exon 7 in the SMN2 gene similar to hnRNPA1. Interestingly, the presumed SRSF10 binding site overlaps the hnRNPA1 binding sites located in the ISS-N1 silencer. Whereas the hnRNPA1 A1 is more constitutively expressed in all tissues SRSF10 shows a more neuronal tissue specific expression pattern. Thus, SRSF10 could be an important factor affecting the splicing of the SMN genes in a neuronal context. We are currently trying to elucidate the role of SRSF10 in determining SMN splicing efficiency and we employ a more genome-wide approach in order to characterize the role and the binding motifs of SRSF10 in other disease genes.

3044M

A functional role for BDNF in Familial Dysautonomia. N. Nilbratt1, M. Salani1, E. Morini1, F. Urbina1, G. Lee2, S. Haggaard3, S. Slaugenhaupt4, 1) Neurology, Massachusetts General Hospital, Boston, MA; 2) Institute for Cell Engineering, John Hopkins University, Baltimore, MD; 3) Familial Dysautonomia (FD), or Riley-Day syndrome, is an autosomal recessive disorder with extensive sensory and autonomic nervous system involvement present at birth. FD is characterized by recurrent episodes of hypertension with tachycardia, impaired thermoregulation, and decreased sensitivity to pain and temperature. Neuropathological studies show a marked reduction of neurons in the sympathetic and sensory ganglia. All FD patients have an intrinsic splice site mutation in the IKBKAP gene, the scaffolding member of the Elongator protein complex involved in transcriptional elongation. This mutation results in tissue-specific skipping of exon 20 in the mRNA with aberrant splicing most pronounced in neuronal tissues. The alternative splicing defect leads to reduced production of IKAP protein. Complete embryonic loss of IKK alpha in mice causes early embryonic lethality, proving that this gene is required for effective transcription of genes involved in early neural development. The molecular mechanisms leading to FD are poorly understood. Neurotrophic factors are implicated in the survival and differentiation of neuronal populations. However, there is evidence for a role of neurotrophins in various human neuropathies, their relevance to the disease process is not fully understood. Previous in vitro studies have indicated reduced neurotrophic activity in human fibroblasts from FD patients and defective expression of genes encoding neurotrophic factors in mice leads to similar neuronal phenotypes to those observed in FD patients. Here we show that reduced IKAP expression impacts brain-derived neurotrophic factor (BDNF) transcription in FD patient fibroblast cells, and we demonstrate a functional consequence of BDNF misregulation when using inducible neuronal-like (iPS) cell model system of human neural development. We found insufficient neurotrophic support from BDNF cells on the development of human iPS cell-derived neurons. Compromised development of human neurons was also observed by inhibition of neurotrophic activity from normal fibroblasts by neurotrophin antagonists. Interestingly, the reduced biological activity from FD patient cells was rescued pharmacologically using the plant cytokinin kinetin, which corrected aberrant mRNA splicing of IKBKAP and restored BDNF protein expression. Our data suggest that aberrant regulation of BDNF by IKAP reduction may contribute to the developmental defects in neural development observed in FD patients.

3045T

Mutations in LAMA1 cause cerebellar dysplasia and cysts with and without retinal dystrophy in Poretti-Bolthausen syndrome. K.A. Aldinger1-3, B.J. Mosca4, M. T Dupont5, J.C. Dempsey6, Q.E. Ishak7, T. Hartley5, I. Phelps5, R.E. Lamont7, D. Base6, K.W. Gripp5, J. St. Jacek8, M.J. Stephen9, F.P. Bernier10, K.M. Boycott5, J. Majewski5, J.S. Parboosingh11, A.M. Innes12, D. Doherly11, 1) Seattle Children’s Research Institute, Seattle, WA; 2) Department of Pediatrics, University of Washington, Seattle, WA; 3) Department of Medical Genetics, University of Calgary, Alberta; 4) Department of Human Genetics, McGill University, Montréal, Quebec, Montréal, Quebec; 5) Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa, Ontario, Care4Rare, Canada; 6) Division of Genetics, Medical College of Wisconsin, Milwaukee, WI; 7) Division of Medical Genetics, A.I. duPont Hospital for Children, Wilmington, DE; 8) Department of Pediatrics, Madigan Army Medical Center, Tacoma, WA; 9) Department of Radiology, Seattle Children’s Hospital, University of Washington, Seattle, WA.

Cerebellar dysplasia with cysts (CDC) is an imaging finding typically seen in combination with cobblestone cortex and congenital muscular dystrophy in patients with dystroglycanopathies. More recently, CDC was reported in a series of seven children without neuromuscular involvement (Poretti-Bolthausen syndrome). Using a combination of homozygosity mapping and whole exome sequencing, we identified biallelic LAMA1 mutations as the cause of CDC in seven additional affected individuals from five families. Most of these patients also have high myopia, and some have retinal dystrophy and patchy increased T2/FLAIR signal in the cortical white matter. Two siblings with truncating LAMA1 mutations have retinal dystrophy and mild cerebellar dysplasia without cysts, further expanding the phenotype associated with CDC. These results add to the spectrum of laminopathy disorders and highlights the tissue-specific roles played by different laminin genes.

3046S

Whole genome sequencing of a balanced translocation reveals new gene candidates for epilepsy, learning difficulties and risk of acute myeloid leukaemia. S.-K. Chung1, A.R. Krainer2,3, S. Haggarty4, B.S. Andresen5, 1) Neurology and Molecular Neuroscience, College of Medicine, Swansea University, Swansea, UK; 2) Wales Epilepsy Research Network, College of Medicine, Swansea University, UK; 3) Institute of Medical Genetics, University Hospital Wales, Cardiff, UK; 4) Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, UK.

Whole-genome sequencing (WGS) was commissioned in a patient with generalised epilepsy, learning difficulties (LD), a familial history of treatment-resistant acute myeloid leukaemia (AML) and a constitutional balanced translocation t(1;14) (q32;q31). WGS revealed the exact translocation co-ordinates and confirmed breakpoints on 1q32 and 14q31 that were validated by Sanger sequencing. On chromosome 1 the breakpoint interrupts a gene of undetermined function (TAT protein family) which is poorly-documented in the literature. Pathway analysis indicates a ‘DNase domain-containing protein’ function with important endonuclease role in chromosome segregation and eye development in zebrafish. On chromosome 14, a more familiar gene was interrupted (FOX gene family) whose function includes neurodevelopmental regulation of the cell cycle and protein synthesis. Other FOX family genes have been linked to seizure disorders, cognition and haematological malignancies. Further mutation screening of these genes in unrelated epilepsy and LD cohorts (n=60) has revealed 4 non-synonymous variants that are not in controls or control dataset websites. Structure / function assays are underway to confirm pathogenic status of these variants as well as gene-fusion studies in the original index-case and screening these genes in cytogenetic-negative AML. These findings have predictive value for the at-risk members of the family and have further neuro-genetic impact in LD and genetic generalised epilepsy.

1) Institute of Genetics and Biophysics "Adriano Buzzati Traverso" CNR, Naples, Italy; 2) Department of Paediatrics, University of Adelaide, South Australia 5006, Australia; 3) Centre for Epigenetics, University of Copenhagen, Copenhagen DK 2000, Denmark; 4) J.C. Self Research Institute, Greenwood Genetic Center, Greenwood, SC, USA.

Lysolecithin-specific demethylase SC (KDM5C) is a chromatin remodelling regulator with histone demethylase activity for di- and trimethylated histone 3 lysine 4 (H3K4me2 and H3K4me3), acting as transcriptional repressor during brain development and neuronal maturation. Mutations in KDM5C are emerging as frequent causes of a spectrum of X-linked intellectual disorders (XLID), with cognitive handicap alone or in association with Epilepsy (XLID/Epilepsy). With Regulatory Element-1 Silencing Transcription factor (REST), a critical regulator of the spatio-temporal transition of neural progenitors to neurons, KDM5C co-occupies the promoters of a subset of REST target genes. We recently identified an epimutation path, linking functionally KDM5C to another XLID/Epilepsy gene, encoding the homeotic transcription factor ARX, whose mutations severely impair KDM5C transcript regulation. Expanding our study, we analysed two additional XLID proteins that also bind the KDM5C promoter. They are PHD Finger Protein 8 (PHFB), a H3K9 demethylase belonging to JmJC protein family, and Zinc Finger Protein 711 (ZNF711), a transcriptional factor, whose role is unclear. We observed that PHFB and ZNF711, which co-occupy the target promoter, cooperatively induce KDM5C transcription. We therefore tested the impact of ZNF711 mutants on KDM5C trans-activation in order to establish their functional effects. This activity seems to be ARX-independent and we propose that the transcriptional induction by ARX does not synergize with the action of the PHFB/ZNF711 complex and consequently they represent distinct levels in KDM5C regulation. Starting from these data, we are now able to break up a unique epigenetic road involved in ID and epilepsy. However, the framework of this complex regulation may become more clear once we will decode the role of other key elements, such as ncRNAs, chromatin modifiers and cofactors. Ongoing efforts to define this transcriptional path may help to identify useful tools towards research and drug discovery for XLID/Epilepsy phenotypes and many other companion disorders with malignant seizure.

Genetics of Joubert syndrome in the French Canadian population. M. Strour1, F.P. Hamdan1, J. Schwartz-Zentner2, L. Patry3, C. Nassif1, L.H. Ospina1, E. Lemyre1, C. Massicotte1, D. Armont1, E. Andermann1,2, R. Laframboise3, B. Maranda1, D. Labuda1, J.C. Decarie3, F. Rypens3, C. Fallet-Bianco1, J.F. Soucy1, K. Boycott1, B. Brais1, R-M. Boucher1, G.A. Rouleau1, J. Majewski4, J.L. Michaud5, D. Labuda6, K. Helin3.

1) Centre Hospitalier Universitaire de Québec, Québec, PQ, Canada; 2) CHU Sainte-Justine, Montréal, PQ, Canada; 3) Department of Neurology and Neurosurgery, McGill University, Montréal, PQ, Canada; 4) FORGE Canada Consortium.

Joubert syndrome (JBTS) is a primarily autosomal recessive disorder characterized by a distinctive mid-hindbrain/cerebellum malformation, ocular-motor apraxia, irregular breathing, developmental delay, hypotonia and ataxia. JBTS is considered a childhood disease; most of the known 21 causal genes are implicated in cerebellar structure or function. We sought to characterize the genetic landscape associated with JBTS in the French Canadian (FC) population. We performed mutation analysis in 40 FC JBTS patients from 32 families using a stepwise approach, combining screening of recurrent FC mutations by Sanger sequencing and exome sequencing. We identified causal mutations in 28 families. Twelve families had mutations in CSORDF42, 9 in CC2D2A, 2 in TMEM231, 2 in NPHF1, one in TCTN1, one in TMEM67. Interestingly, we documented a complex founder effect with multiple recurrent mutations in 3 genes (CSORDF42, CC2D2A, TMEM231). In the remaining families, we identified potentially pathogenic mutations in known JBTS genes (CEP290, OFD1) and in cilia genes, which are therefore strong candidates for JBTS. We conclude that JBTS has substantial locus and allelic heterogeneity in FC populations.


1) NIH Undiagnosed Diseases Program, National Institutes of Health, Bethesda, MD; 2) Office of the Clinical Director, NHGRI, National Institutes of Health, Bethesda, MD; 3) Children’s National Medical Center, Washington, DC.

Hereditary diffuse leukoencephalopathy with spheroids (HDLS [MIM 164770]) is an uncommon, dominantly inherited, adult-onset neurological disorder with characteristic clinical and MRI features. Most often, patients present in their third to fifth decade of life with progressive personality changes, dementia, spasticity, gait difficulties and depression. For decades, the diagnosis was only possible by brain biopsy or autopsy. Recently, however, mutations in the protein tyrosine kinase (PTK) domain (exons 13 thru 19) of the colony stimulating factor receptor 1 (CSF1R) gene have been associated with this disorder. As part of the NIH Undiagnosed Diseases Program, we evaluated 8 patients in 5 families with typical clinical and imaging features of HDLS. All patients underwent clinical phenotyping, brain and spinal cord MRI and CSF examination. One patient had undergone a brain biopsy and subsequent autopsy brain examination, which confirmed the clinical diagnosis of HDLS. All patients underwent whole exome sequencing or direct Sanger sequencing of all CSF1R coding exons. Previously unreported, predicted damaging mutations in exons 13 and 19 of CSF1R were identified in 2 families. One family with documented vertical transmission and autopsy-proven HDLS did not harbor mutations in CSF1R, indicating locus heterogeneity for HDLS. Another patient harboring a novel predicted damaging mutation in exon 19 (NM_005211.3:c.2512G>T:p.V838F), who is also founder, was age 75 when he was clinically unaffected, indicating incomplete penetrance. Our study enhances understanding of the genetic basis of this devastating disorder and highlights the need to continue searching for additional causative genes; this could be aided by and improved insight into the cell biological mechanisms responsible for the emergence of signs and symptoms.


1) Section of Medical Genetics, Department of Laboratory Medicine, Telemark Hospital, Skien, Norway; 2) Head and Neck Research Group, Research Centre, Akerhussen University Hospital, Lørenskog, Norway; 3) Institute of Clinical Medicine, Campus Akershus University Hospital, University of Oslo, Nordbyhagen, Oslo, Norway.

Charcot-Marie-Tooth disease (CMT) is a heterogeneous inherited neuropathy. The number of known CMT genes is rapidly increasing mainly due to next generation sequencing technology, at present more than 60 CMT-associated genes are known. We investigated whether variants in the DCTN2 gene could cause CMT. Material and methods: Fifty-nine Norwegian CMT families from the general population with unknown genotype were tested by targeted next-generation sequencing (NGS) for variants in DCTN2 along with 33 CMT genes and 18 other genes causing other inherited neuropathies or neuropathies, due to phenotypic overlap. Results: Targeted NGS identified in one family a variant of DCTN2, c.337C>T, segregating with the phenotype in five affected while it was not present in the three unaffected members. The DCTN2 variant c.337C>T; p.(His113 Tyr) was not found in in-house controls nor in SNP databases. The result is further strengthened by the fact that 33 known CMT genes and 18 other peripheral neuropathy genes did not harbor pathogenic mutations. Conclusions: This is the first time a DCTN2 variant has been linked to an inherited neuropathy in man.
A missense mutation in the PISA domain of HSAS-6 causes autosomal recessive primary microcephaly in a large consanguineous Pakistani family. V. M. Rupp,2, M. Khan,1 M. Ordonez,1 J. Altmüller,4,5 M. O. Steinmetz,2 C. Enzinger,2 H. Thiele6, W. Höhne6,5, G. Nürnberg4, S. M. Baig2, M. Ansar10, P. Nürnberg5,11,12, J. B. Vincentt12, M. R. Speicher12, P. Göröczi,1 Windpassinger1.1 Institute of Human Genetics, Medical University of Graz, Graz, Styria, Austria; 2) Gomai Centre of Biochemistry and Biotechnology, Gomai University D.I.Khan, Khyber-Pakhtoonkhwa, Pakistan; 3) Swiss Institute for Experimental Cancer Research (ISREC), Swiss Federal Institute of Technology (EPFL), School of Life Sciences, University of Lausanne, Lausanne, Switzerland; 4) University of Cologne, Cologne, Germany; 5) Institute of Biochemistry I, Medical Faculty, University of Cologne, Cologne, Germany; 6) Institute of Human Genetics, University of Cologne, Cologne, Germany; 7) Laboratory of Biomolecular Research, Department of Biology and Chemistry, Paul Scherrer Institut, Villigen PSI, Switzerland; 8) Department of Neurology, Medical University of Graz, Graz, Austria; 9) Human Molecular Genetics Laboratory, Health Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan; 10) Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan; 11) Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany; 12) Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany; 13) Molecular Neuroepidemiology and Development (MIND) Lab, The Campbell Family Brain Research Institute, The Centre for Addiction & Mental Health (CAMH), Toronto, Ontario, Canada.

Asymmetric cell division is essential for normal human brain development. Mutations in several genes encoding centrosomal proteins that participate in accurate cell division have been reported to cause autosomal recessive primary microcephaly (MCPH). By homologously mapping including three affected families from a consanguineous MCPH family from Pakistan, we delineated a critical region of 18.53 Mb on chromosome 1p21.3-1p13.1. This region contains the gene encoding HSAS-6, a centrosomal protein primordial for seeding the formation of new centrioles during the cell cycle. Both neurodevelopment and Sanger sequencing revealed a heterozygous c.1857>T missense mutation in the HSAS-6 gene, resulting in a p.Ile62Thr substitution within a highly conserved region of the PISA domain of HSAS-6. This variant is neither present in any single nucleotide polymorphism or exome sequencing databases nor in the control of 1000 Genomes Project. The variant was observed in tissue culture cells revealed that the p.Ile62Thr mutant of HSAS-6 is substantially less efficient than the wild-type protein in sustaining centriole formation. Together, our findings demonstrate a dramatic impact of the mutation p.Ile62Thr on HSAS-6 function and add this component to the list of genes mutated in primary microcephaly.

Consilvus seizures and SUDEP in a mouse model of SCN8A epileptic encephalopathy. M. L. Wiegand,1 H. W. Korn1,2, R. Parent1,2, B. Hammer4, G. G. Murphy10,3, J. M. Parent5,6, M. H. Meisler10,2. 1) Human Genetics, University of Michigan, Ann Arbor, MI; 2) Neurology, University of Michigan, Ann Arbor, MI; 3) Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI; 4) Arizona Research Laboratories, VA Puget Sound Health Care System, Seattle, WA; 5) Division of Biotechnology, University of Arizona, Tuscon AZ; 6) Molecular and Behavioral Neuroscience Institute, University of Michigan, Ann Arbor MI; 6) VA Ann Arbor Healthcare System, Ann Arbor, MI.

Mutations of the voltage-gated sodium channel gene SCN8A have recently been recognized as a cause of epileptic encephalopathy, which is characterized by refractory seizures with developmental delay and cognitive disability. We previously described the heterozygous SCN8A missense mutation p.Asn717Asp in a child with epileptic encephalopathy that included febrile seizures, ataxia, autism, and sudden unexpected death in epilepsy (SUDEP). This mutation results in increased persistent sodium current and hyperactivity of transferred neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons. We introduced the gain-of-function mutation into the mouse genome to investigate the pathology of the altered channel in vivo and in transfected neurons.
3055S

De novo mutations in NALCN cause a new syndrome with congenital contractures, hypotonia, and early death. K.M. Shively1, J.X. Chong1, M.J. McMillin1, A.E. Beck2, C.T. Marin1, K.J. Buckingham1, J.R Armen- teros1, A. Montell1, J. Shendure3, D.A. Nickerson4, M.J. Bamshad3,4, University of Washington Center for Mendelian Genomics. 1) Department of Pediatrics, University of Washington, Seattle, WA, USA; 2) Seattle Children’s Hospital, Seattle, WA, USA; 3) Department of Physiology, Institute for Func- tional Genomics, Montpellier, France; 4) Department of Genome Sciences, University of Washington, Seattle, WA, USA.

Mutations in multiple genes, including TPM2, TNNI2, TNNT3, and MYH3, cause distal arthrogryposis (DA) but ~60% of cases remain unexplained. Analysis of exome sequencing data from six kindreds with an atypical form of DA with neurological abnormalities including hypotonia and developmental delay identified six different missense mutations in the gene sodium leak channel, non-selective, or NALCN. All six mutations were confirmed by Sanger sequencing to be de novo. Recently, it has been reported that mutations in NALCN cause autosomal recessive infantile hypotonia with psychomotor retardation and characteristic facies (IHPRF) and infantile neu- roaxonal dystrophy (INAD) with facial dysmorphism in consanguineous Turk- ish and Saudi Arabian families. Two of these mutations were predicted to result in a truncated protein (or no protein if subject to nonsense-mediated decay), while the third mutation was a missense variant located in the 3rd helix of domain IV. In contrast, all six of the mutations we discovered are in the helices of the pore-forming regions of NALCN. These findings demon- strate that mutations in NALCN cause both recessive and dominant condi- tions, perhaps depending on the domain(s) affected by the mutations. More- over, while there is some phenotypic overlap between the individuals with atypical DA with facial dysmorphism or IHPRF and the individuals with typical DA, there are unique clinical features that distinguish the individuals with de novo mutations including congenital contractures of the hands, mouth, and feet. Last, the widely varied severity of developmental delay and con- tractions in the six kindreds we studied suggests that one of her two sons as well as in the maternal grandmother we previously have been reported in this family. Array painting performed by using laser translo- cation in the six kindreds we studied suggests that 5' exons and/or the syntaxin family. Array painting performed by using laser translo- cation in the six kindreds we studied suggests that 5' exons and/or the

3056M

Application of array painting and next generation mate-pair sequencing (MPS) for improved mapping of chromosomal breakpoints in a familial translocation segregating with a particular phenotype. P.M. Kroisel1, M. Auer1, K.M. Rooter2, R. Birnbacher3, K. Wagner1, M.R. Speicher1, M.M. Hessel ICONY1, M. Bak1, J.B. Geigl1, N. Tonnemans1, 1) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 2) Hanusch-Kranken- haus, Vienna, Austria; 3) Department of Pediatrics, Regional Hospital Vil- lach, Villach, Austria; 4) Wilhelm Johannsen Centre for Functional Genome Research, Department of Cellular and Molecular Medicine, Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark.

In a familial translocation t(1;6)(p36.33;p21.2) identified in a mother and one of her two sons as well as in the maternal grandmother we previously confirmed by array CGH that no genomic unbalance is present. The translo- cation cosegregates with distinct phenotypic features including a mild to moderate intellectual disability and microcephaly. No other genetic disorders have been reported in this family. Array painting performed by using laser micro dissected and amplified chromosomal fragments carrying the translo- cation breakpoints followed by hybridization to standard 60k CGH arrays allowed the characterization of both chromosomal breakpoints at an achieved resolution that made it possible to identify genes potentially affected by the translocation. A number of less than ten chromosomal frag- ments was sufficient to obtain an adequate signal to noise ratio for this analysis. The CAMTA1-gene (Calmodulin binding transcription activator 1) at 1p36.31 and the BTBD9 or KIAA1880 gene at 6p21.2 were found to be located very close to or at the chromosomal breakpoints. Next generation mate- pair sequencing (MPS) will further improve the resolution of the break- point mapping by pin-pointing the location of the breakpoints at base pair level, allowing a detailed characterization of the genotype-phenotype relation- ship and of the DNA repair mechanism involved in the rearrangement. This approach might allow a more immediate and better interpretation of the molecular and functional effects of particular structural genomic aber- rations.

3057T

Deletion of the 5' exons of the TCF4 gene in patients without classical Pitt-Hopkins syndrome. S. YU1,2, D. GAP1, N. Jillian1. 1) GENETIC Medi- cine, SA Pathology, North Adelaide, South Australia, Australia; 2) Royal Melbourne Hospital, Parkville, Victoria, Australia; 3) School of Paediatrics and Reproductive Health, University of Adelaide, South Australia, Australia. Pitt-Hopkins syndrome (PHS) is characterised by severe intellectual disa- bility (ID) and typical facial features. Other features include episodic hyper- ventilation, seizures, eye abnormalities and constipation. The TCF4 gene was identified as disease-causing. Hall-pen-sufficiency of TCF4 has been reported to cause classical PHS. We describe two patients with deletions of the first two 5' exons of TCF4. One patient is a 16 year old boy with high functioning autism spectrum disorder and constipation. Array CGH identified a 128 kb deletion [chr18: 53,278,787-53,526,749 (hg19)]. He also has a 227kb duplication of 11p15.5 (chr15:1,448,960-1,222,378) with unknown clinical significance. His mother does not carry the TCF4 deletion but the father is not available for testing. The second patient is a 14 year old girl who had mild ID with perceptual reasoning problems. She had a similar deletion [chr18: 51,429,785-51,677,747]. Parents are not available for testing. No dysmorphism was reported in either patient. DECIPHER database recorded at least four cases with a deletion only involving the 5' exons of the TCF4 gene without classical PHS. Two cases (269234 and 269230) had mild intellectual disability. The third (250488) had microcephaly, constipation, and seizures. The fourth (1024) had ID, microcephaly, delayed speech and other features. In addition, disruption of the TCF4 gene proximal to exon 4 by a translocation was reported in a girl with ID but without classic PHS (Kalscheuer et al). TCF4 has 20 exons with multiple isoforms using alternative 5' exons. Mutations and deletions of the TCF4 gene cause a loss of function. These results suggest that coding exons sparing the 5' exons (exons 1-3) (Whalen et al). TCF4 mRNA expres- sion study showed that the isoforms containing the 5'exons were restricted to tests, prostate and plaenca, and not expressed in brain, while other isoforms expressed in multiple brain regions (Sepp et al). It is possible that deletions of the 5' exons of the TCF4 gene only affect isoforms not expressed in brain and therefore lead to a less severe phenotype. This is the first report to suggest that deletion of the 5' exons of the TCF4 gene may not cause classical PHS. This suggests that particular isoforms may contrib- ute to the milder phenotype. 1.Kalscheuer VM et al. (2008) Am J Med Genet Part A 146A:2053-2059. 2. Sepp M et al. (2011) PLoS ONE 6:e22138. 3.Whalen S et al. (2011) Human Mutation 33:64-72.

3058S

A genetic dosage study of DYT1 Dystonia using an inducible knock- in ΔE-Tor1a mouse model. C. Weisheit1, W. Dauer1,2,3, 1) Graduate Pro- gram in Cellular and Molecular Biology, University of Michigan, Ann Arbor, MI; 2) Department of Cell and Reproductive Biology, University of Michigan, Ann Arbor, MI; 3) Department of Neurology, University of Michigan Medical School, Ann Arbor, MI.

DYT1 dystonia is a common inherited form of primary dystonia caused by a dominant mutation in the TorsinA gene that removes a single glutamic acid (“ΔE”) residue from torsinA, the encoded protein. We generated a novel line of inducible knock-in (I-KI) ΔE-Tor1a knock-in mice to evaluate whether the ΔE mutation acts via a gain- or loss-of-function mechanism. The ΔE mutation is located in the last exon (exon 5) of Tor1a. I-KI mice contain a floxed exon 5, and an additional downstream exon 5 containing the ΔE mutation; this allele is converted from a wild type to DYT1 allele following Cre deletion, allowing spatial and temporal induction of the pathogenic protein from the endogenous locus. We intercrossed these mice with mice carrying a floxed Tor1a gene to generate a gene dosage series, hypothesizing that if the ΔE mutation causes a gain of toxic function, animals harboring two mutant alleles would fare worse than littermates with one mutant and one null allele. This hypothesis has not been met. This suggests that during normal development, the ΔE mutation causes a new syndrome with congenital contractures, hypotonia, and feet. Last, the widely varied severity of developmental delay and con- tractions in the six kindreds we studied suggests that one of her two sons as well as in the maternal grandmother we previously have been reported in this family. Array painting performed by using laser translo- cation in the six kindreds we studied suggests that 5' exons and/or the syntaxin family. Array painting performed by using laser translo- cation in the six kindreds we studied suggests that 5' exons and/or the

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
Whole exome sequencing identified the first STRADA point mutation in a patient with polyhydramnios, megalencephaly, and symptomatic epilepsy syndrome (PMSE). To date we have completed mutation detection in 703 patients in 21 genes. While de novo mutations occurred in 26 known genes, we identified a novel mutation in STRADA. This mutation is predicted to lead to a frameshift change and is located within one of the highly conserved exons of STRADA. This finding suggests that STRADA is a potential candidate gene for PMSE.

Glutathione-S-transferase gene polymorphisms (GSTM1, GSTT1, GSTM3 & GSTP1) and its correlation with GST enzyme activity in DM1 patients.

Glutathione-S-transferase gene polymorphisms (GSTM1, GSTT1, GSTM3 & GSTP1) and its correlation with GST enzyme activity in DM1 patients.

Glutathione-S-transferase gene polymorphisms (GSTM1, GSTT1, GSTM3 & GSTP1) and its correlation with GST enzyme activity in DM1 patients.
3063T

The whole genome sequences from a Rottweiler and Black Russian Terrier with overlapping neurological syndromes contain the same RAB3GAP1 frame shift mutation. T. Mhlanga-Mutangadura 1, G.S. Johnson 1, G.C. Johnson 2, L. Hansen 1, G.V. Tamassia 1, J.F. Taylor 1, R.D. Schnabel 2, D.P. O’Brien 1. 1) College of Veterinary Medicine, University of Missouri, Columbia, MO; 2) College of Agriculture, Food and Natural Resources, University of Missouri, Columbia, MO.

We generated whole-genome sequences (WGSs) with DNA from a 5-month-old male Rottweiler diagnosed with neuronal vacuolation and laryngeal paralysis and a 3-month-old female Black Russian Terrier diagnosed with juvenile laryngeal paralysis, polyneuropathy, microphthalma, persistent pupillary membranes and cataracts. Initially we thought these two dogs had overlapping but distinct disease syndromes; however, both of their WGSs contained the same homozygous single base pair deletion: RAB3GAP1:c.7-43delC. The frame shift produced by this deletion predicts a premature stop codon and a truncated gene product RAB3GAP1:p.P248Lfs*3 missing 730 C-terminal amino acids. In addition to the Rottweiler and Black Russian Terrier with the WGSs, the DNA samples from all 3 other affected Rottweilers and all 4 other Black Russian Terriers in our collection were homozygous for the c.743delC allele; whereas, 85 clinically normal Rottweilers and 41 clinically normal Black Russian Terriers tested either heterozygous (n = 18) or homozygous for the ancestral allele (n = 108). Truncating mutations in human RAB3GAP1 have been found in patients with Warburg micro syndrome. This severely debilitating developmental disorder is characterized by microcephaly, microphthalma, microcornea, congenital cataracts, optic atrophy, cortical dysplasia, mental retardation, and hypogonadism. In contrast, RAB3gap1 knockout mice only exhibit subtle abnormalities. The disease phenotype produced by the frame shift mutation in canine RAB3GAP1 appears to be less severe than that in the children with Warburg micro syndrome but much more severe than that of Rab3gap1 knockout mice. We are currently re-evaluating the disease phenotypes from affected Rottweilers and we are, in fact, identical and to determine if they have other clinical signs in common with the Warburg micro syndrome patients.

3064S

The missing factors influencing spinal and bulbar muscular atrophy: evaluation of genetic polymorphisms. C. Bertolín 1, G. Querín 2, M. Penuto 3, E. Pegoraro 1, C. Geller 1, D. Pareyson 3, C. Mariotti 4, G. Soraru 5. 1) Biology, University of Padova, Padova, Padova, Italy; 2) 2Dulbecco Telethon Institute Laboratory of Neurodegenerative Diseases, Centre for Integrative Biology (CIBIO), University of Trento, Italy; 3) Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico “Carlo Besta”, Milan, Italy.

Spinal and bulbar muscular atrophy (SBMA) is caused by a pathological expansion of over 38 of a CAG repeat in the first exon of the androgen receptor (AR) gene on chromosome X, coding for a polyQ tract (La Spada et al, 1991). SBMA is an androgen-dependent disorder, with males with full disease penetrance. We have recently identified a polymorphism within the SBMA locus. In an Italian cohort. We genotyped AR polymorphisms in 132 molecularly defined SBMA phenotype, we genotyped AR polymorphisms in 132 molecularly defined patients belonging to different polyG allele classes showed genotyping of polyG stretch is polymorphic in our cohort. Interestingly, the genotyping of polyG stretch is polymorphic in our cohort. Interestingly, the genotyping of polyG stretch in a different level of correlation between polyQ length and disease’s severity. SBMA phenotype, we genotyped AR polymorphisms in 132 molecularly defined patients belonging to different polyG allele classes showed that patients belonging to different polyG allele classes showed different levels of correlation between polyQ length and disease’s severity. SBMA phenotype, we genotyped AR polymorphisms in 132 molecularly defined patients belonging to different polyG allele classes showed different levels of correlation between polyQ length and disease’s severity.

3065M

Abnormalities in neuronal architecture and synaptic activity impairment in mice heterozygous for different deletions of the Williams-Beuren syndrome locus. C. Borrallos 1, 2, M. Bosch-Moral 2, B. Guiuver 1, M. Bravo-Garmendia 2, L.A. Pérez-Jurado 1, 2, 3, F.J. Muñoz 1, V. Camargo 1, 2, 3. 1) Hospital del Mar Research Institute (IMIM), Barcelona, Barcelona, Spain; 2) Universitat Pompeu Fabra, Barcelona, Spain; 3) Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain.

Williams-Beuren syndrome (WBS, [MIM 194050]) is a neurodevelopmental disorder caused by a spontaneous 1.55-1.83 Mb deletion at chromosome 7q11.23. This deletion contains 29 genes encoding 2-31-41 family members. WBS is characterized by distinct dystrophic features, mild growth retardation, cardiovascular abnormalities, and a specific cognitive profile with mild intellectual disability, visuospatial deficits, hypersensitivity to sounds and hypersociability. Neurologic examination and brain imaging in patients have shown changes in brain function and brain anatomy. Haploinsufficiency at GTF2I has been shown to play a major role in the neurodevelopmental abnormalities of WBS. To clarify the involvement of GTF2I in neurocognitive features we characterized four WBS mouse models deleting Gtf2i with and without additional genes: PD (proximal deletion, from Gtf2i to Limk1), CD (complete deletion, from Gtf2i to Kbp6), ΔGtf2i/-, and ΔGtf2i/-. We examined the morphology of a transgenically labeled subset of cortical and hippocampal pyramidal neurons in YFP-/- mice crossed with the different mutant lines. The number of YFP+ cells was decreased by up to 87% in WBS mouse models in various cortical regions. Lower number of YFP+ cells was present as early as PD14-16 developmental stages. Spine density of apical dendrites in the CA1 area of the hippocampus was reduced in Gtf2i/-, PD and CD mice by 15%, 12% and 17%, respectively. Moreover, more CA1 pyramidal neurons in the transgenic lines showed alterations in glutamatergic synaptic activity. As GTF2I has a role in PI3K/AKT pathway (involved in synaptic plasticity) and is a regulator of Dlx genes (involved in the differentiation of GABAergic projection neurons), we analyzed the expression of these genes in our mutant mice. Mutant mice showed deregulation of these genes, suggesting impairment in synaptic plasticity and reduced inhibitory synaptic activity. We suggest that GTF2I plays an important role in neurodevelopmental abnormalities in WBS. Abnormalities in neuronal architecture and perturbations in synaptic activity could be associated with the cognitive impairment seen in WBS patients.

3066T

A Homozygous PIGN Missense Mutation in Soft Coated Wheaten Terriers with Paroxysmal Dyskinesia. A. Kolicheskii 1, G.S. Johnson 1, D. Gilliam 1, T. Mhlanga-Mutangadura 1, J.F. Taylor 1, R.D. Schnabel 2, D.P. O’Brien 2, 3. 1) College of Veterinary Medicine, University of Missouri-Columbia, Columbia, MO; 2) College of Agriculture, Food and Natural Resources, University of Missouri, Columbia, MO.

The paroxysmal dyskinesia (PD) we studied is a previously undescribed movement disorder in the Soft Coated Wheaten Terrier (SCWT) movement disorder characterized by multiple episodes of dyskinesia that sometimes last for only a few minutes and other times for over an hour. Episodes begin in young adult SCWTs and include hyperextension or hyperflexion primarily of the pelvic limbs with difficulty moving forward and arrhythmic hopping. Between episodes the dogs move normally. We generated a whole genome sequence (WGS) with 21-fold average coverage and 6.9 million sequence variants (differences from the reference canine genome sequence) with DNA from a PD-affected SCWT and another WGS with 18-fold average coverage and 5.5 million sequence variants with DNA from a second PD-affected SCWT that had no known relationship to the first. We filtered the sequence variants to identify those that were homozygous in both case WGSs, predicted to alter the 3065M protein, and absent from 100 other canine WGSs. The only variant that met these criteria was a PIGN:c.398C>T transition that predicts a PIGN:p.T133I substitution. All 13 PD-affected SCWT (including the 2 with WGSs) were homozygous for the PIGN:c.398T allele. The genotyping of archived SCWT DNA samples of unknown phenotype identified 469 PIGN:c.398C homozygotes, 7 heterozygotes and no PIGN:c.398T homozygotes. Samples of 505 dogs of varied breeds were also genotyped and all were homozygous wild type. PIGN encodes an enzyme in the biosynthetic pathway for glycosylphosphatidylinositol anchors that attach many different proteins to cell surfaces. Given that a human PIGN, cause multiple congenital anomalies-hypotonia-seizures syndrome 1 (MCASH1 [MIM 6110480]), with characteristic neonatal hypotonia, facial dysmorphism, and psychomotor underdevelopment. A minority of MCASH1 patients have exhibited extreme behavioral changes that suggest PIGN plays an important role in neurodevelopmental abnormalities in WBS. Abnormalities in neuronal architecture and perturbations in synaptic activity could be associated with the cognitive impairment seen in WBS patients.
Congenital hypotonia: Two rare diseases in one family. T. Falik-Zaccai1,2, M. Baydany1,2, L. Kallon1, N. Nasser Samra1,2, Y. Shoval1, D. Savidze3, S. Irry2, Z. Zion2, H. Mandel1. 1) Institute of Human Genetics, The Galilee Medical Center, Nahariya, Israel; 2) The Galilee Faculty of Medicine, Bar Ilan University, Safed, Israel; 3) Department of Pediatric Neurology, The Galilee Medical Center, Nahariya, Israel; 4) Department of Anesthesiology, The Galilee Medical Center, Nahariya, Israel; 5) Pediatric Intensive Care Unit, The Galilee Medical Center, Nahariya, Israel; 6) Meyer Children’s Hospital, Metabolic Unit, Rambam Medical Center, Haifa, and Rappaport Faculty of Medicine, Technion, Haifa, Israel.

Background: Congenital hypotonia is a non-specific physical finding, common to a heterogeneous group of diseases. We present an extended highly consanguineous family of Druze origin, with multiple affected children born with severe congenital hypotonia. We hypothesize that congenital hypotonia in this family is a monogenic autosomal recessive disorder. We aim to determine its causative gene and mutation, and to characterize its molecular and biochemical basis. Methods: Clinical investigations, homozygosity mapping of patients and their parents, linkage analyses and whole exome sequencing, were performed to find the novel sequence variation segregating with the disease in two nuclear families within the extended pedigree. mRNA and protein levels were further determined. Results: We have identified 4 novel sequence variations within two genes in the two nuclear families that present both with multiple affected children with congenital hypotonia. In the first family we have identified a novel nonsense mutation in the gene NGLY1. A recently reported gene with only 8 patients described worldwide. RNA levels were significantly reduced among the patients compared to controls in both fibroblasts and leukocytes. In the second nuclear family which is closely related to the first, we have identified compound heterozygosity for three novel RYR1 mutations in two affected sisters. Two mutations inherited from the father in cis and a third mutation inherited from the mother. We screened 100 healthy ethnically matched individuals from the same population for the disease causing mutations. Four of the mutations were absent in the control population. Conclusions: We have identified a genetic isolate for this high risk population. The family presented here demonstrates the critical role of comprehensive clinical evaluation along with whole exome sequencing for accurate determination of the genetic basis of congenital hypotonia and other rare genetic disorders.

Identification of a novel autophagy-related gene mutation in a canine storage disease. K. Kyöstilä1,2,4 V. Jagannathan5, P. Syrjä1, T. Jokinen5, G. Chandrasekar1, J. Kere1,2,4,7,8 E.H. Steppäla1,2,3,4, D. Becker5, M. Drögemüller2, E. Dietrich2, C. Drögemüller2, L. Lang9, F. Steffen10, C. Rohdin11, P. Wohlsin12, D. Henke13, A. Oevermann13, H. Lohi1,2,3,4, T. Leeb1. 1) Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland; 2) Department of Clinical Veterinary Medicine, Division of Clinical Radiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland; 3) Center for Biosciences, Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden; 4) Institute of Genetics, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland; 5) Institute of Genetics, Faculty of Medicine, Technion, Haifa, Israel; 6) Department of Equine and Small Animal Medicine, University of Helsinki, Helsinki, Finland; 7) Department of Molecular Genetics, Folkhälsoinstitutet, Institute of Genetics, Helsinki, Finland; 8) Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland; 9) Department of Pathology, University of Veterinary Medicine Hannover, Hannover, Germany; 10) Neurology Service, Department of Small Animals, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland; 11) University Animal Hospital, Swedish University of Agricultural Sciences, Uppsala, Sweden; 12) Department of Pathology, University of Veterinary Medicine Hannover, Hannover, Germany; 13) Division of Neurological Sciences, Vetsuisse Faculty, University of Bern, Bern, Switzerland.

Purebred dogs suffer from different types of inherited neurodegenerative diseases similar to humans. As a large animal, the dog is physiologically closer to human than traditional model organisms and therefore a better replicate for human phenotypes. We have studied clinical, pathological and genetic aspects of a storage disease in the Lagotto Romagnolo dog that progressively affects the nervous system. The affected dogs show early-onset nystagmus and behavioral changes. The onset of clinical signs varies from 4 months to 4 years of age. The main histological findings in the central and peripheral nervous system are profound cytoplasmatic vacuolization in neurons along with granular spheroid formation. Cytoplasmatic vacuolization of several secretory epithelia is also present. To identify the genetic cause of the disease, we genotyped three pathologically confirmed cases and 306 dogs from 40 other breeds. The gene has not been previously associated with the disease (p=3.8 x 10^-6). Furthermore, the variant was absent in 642 dogs from 40 other breeds. The gene has not been previously associated with inherited diseases in any species. The encoded protein is poorly characterized but proposed to function in a macroautophagy pathway, which has been implicated in neurodegenerative diseases. We are currently performing experiments to further delineate the function of the gene in the zebrafish model. In conclusion, our study identifies a novel mutation associated with a previously uncharacterized canine storage disorder. This animal model may be useful to further characterize the interplay between dysfunctional autophagy and neural degeneration, and to explore possible therapeutic options. Meanwhile, dog breeders will benefit from genetic testing.
3069T

Charcot-Marie-Tooth disease (CMT) comprises a group of heterogeneous motor and sensory neuropathies caused by mutations in several genes. Abnormal accumulation of neurofilaments and disruption of neurofilament network is a common cause of motor neuron degeneration in CMT. In particular, mutations in the neurofilament light chain (NEFL) and heat shock protein, HSPB1, leads to neurofilaments pathological aggregates. Neurofilament aggregation is also the hallmark of neuronal dysfunctional associated with other neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), Parkinson and Alzheimer. Applying exome sequencing, we have identified 2 novel dominantly inherited CMT mutations in the neurofilament heavy (NEFH) subunit. Both mutations caused frameshifts in the conserved c-terminal domain, resulting in the loss of the stop codon and consequently producing an extra 40 amino acid long extension at the c-terminal tail. NEFH mutations identified in the patients were also evaluated at the functional level, where we observed prominent abnormal aggregates in neuro2a cells transfected with the mutant protein. Therefore, we concluded that abnormal aggregation of the mutant NEFH suggests a possible toxic gain of function effect that leads to the progression of CMT in patients. Moreover, zebrafish eggs injected with mRNA encoding the mutant NEFH resulted in embryos with significantly decreased motor neuron lengths. Our studies support the hypothesis of neurofilament aggregation and dysfunction as a disease mechanism for axonopathies.

3070S
MUTATIONS IN THE TRICARBOXYLIC ACID CYCLE ENZYME, ACONITASE2, CAUSE EITHER ISOLATED OR SYNDROMIC OPTIC NEUROPATHY WITH ENCEPHALOPATHY AND CEREBELLAR ATROPHY. JM. Rozet1, M. Metodiev1, S. Gerber1, L. Hubert1, D. Chereten1, X. Gerard1, P. Amati-Bonneau2, MC. Jacomoto3, N. Boddart4, A. Kamińska5, t. Desguerre6, M. Río7, J. Kaplan1, A. Mumlich8, A. Rotig1, C. Besmond1. 1) Inserm UMR1163, UPD - Sorbonne París Cité, Imagine-Institute of genetic diseases, Paris, France; 2) Department of Biochemistry and Genetics, CHU d’Angers, Angers, France; 3) Department of Ophthalmology, Polyclinique du Maine, Laval, France; 4) Departments of Pediatrics, Radiology and Genetics, Hôpital Necker-Enfants Malades, Paris, France.

Inherited optic neuropathy has been ascribed to mutations in mitochondrial fusion/fission dynamic genes, nuclear and mitochondrial DNA-encoded respiratory enzyme genes or nuclear genes of poorly known mitochondrial function. On the other hand, enzymopathies of the tricarboxylic acid cycle (TCA) have been reported to cause severe encephalopathies or isolated retinitis pigmentosa, but no TCA-cycle enzyme deficiency has been hitherto reported in isolated optic neuropathy. Studying a series of five patients with optic atrophy, we found homonymous or compound heterozygous missense and frameshift mutations in the gene encoding mitochondrial aconitate (ACO2), a TCA-cycle enzyme, catalyzing interconversion of citrate into isocitrate. Retrospective studies using patient-derived cultured skin fibroblasts revealed various degrees of deficiency in ACO2 activity but also in ACO1 cytosolic activity. Our study shows that autosomal recessive ACO2 mutations can cause either isolated or syndromic optic neuropathy. This observation identifies ACO2 as the second gene responsible for non syndromic autosomal recessive optic neuropathies and provides evidence for a genetic overlap between isolated and syndromic forms, giving further support to the view that optic atrophy is a hallmark of defective mitochondrial energy supply.
Whole genome sequences from two individual dogs with neuronal ceroid lipofuscinosis contain novel truncating mutations: one in CLN8 and the other in UNC13D. The 2 dogs had a clinical history suggestive of neuronal ceroid lipofuscinosis and brain dysmorphology. Whole genome sequencing identified truncating mutations in CLN8 and UNC13D, respectively. These findings suggest a role of CLN8 and UNC13D in the pathogenesis of neuronal ceroid lipofuscinosis and highlight the importance of whole genome sequencing in the identification of genetic causes of rare Mendelian disorders.
3076S

Protein synthesis regulation via the mammalian target of rapamycin (mTOR) signaling pathway plays key roles in neural development and function, and its dysregulation is involved in neurodegeneration. Here we show that mTORC1 regulates translation initiation in vitro and in cell-based assays using neuronal precursor cells (NPCs) and glioblastoma cells. Our data indicate that mTORC1 mediates the translation of proteins involved in neuronal function, including proteins that are dysregulated in neurodegenerative diseases. These findings suggest that targeting mTORC1 may be a promising therapeutic strategy for neurodegenerative diseases.

3077M
De novo mutations in the motor neuron protein KIF1A cause a severe static or progressive encephalopathy with cerebral and cerebellar atrophy. S. Esmaeili Nieh1, M. Madon1, B. Freague1, A.J. Barkovitch2, J.G. Pappas3, C.G. Spaeth3, M. Madou1. 1) Department of Pediatrics, University of California, San Francisco, CA; 2) Department of Radiology, University of California, San Francisco CA; 3) Department of Pediatrics, Children’s Hospital of Philadelphia, Philadelphia, PA;

The causes of developmental encephalopathies are complex and often remain poorly understood. Arriving at a diagnosis can have important implications for clinical management and prognosis. Clinical indications for apparent de novo mutations in KIF1A are suggestive of severe static or progressive encephalopathy with cerebral and cerebellar atrophy. We speculate that KIF1A mutations may be associated with an underappreciated phenotype, and that a comprehensive approach to the study of de novo mutations in KIF1A may provide novel insights into the biological mechanisms underlying developmental encephalopathies.

3078T
Clinical and molecular characterization of progressive encephalopathies in children. J.R. Helle1, D. Misceo1, T. Barety1, C. Ramane1, A. Magriko1, T. Hughes1, M.D. Viegland1, S. Thirukkakaran1, S.S. Amundsen1, T.H. Gamage1, A. Torgersbraten1, N. Skau1, I. Akkhov1, E.O. Carlsen1, M. Fannemel1, P. Stramme2, E. Frengen1. 1) Department of Medical Genetics, Oslo University Hospital and University of Oslo, Oslo, Norway; 2) Women and Children’s Division, Department of Clinical Neurosciences for Children, Oslo University Hospital Ullevål, Oslo, Norway.

BACKGROUND AND OBJECTIVES: Progressive encephalopathies (PE) in children consist of clinically and genetically heterogeneous disorders. They may present with developmental delay, seizures, speech delay, and a variety of neurological features. The identification of the causative gene for PE and lead to developmental arrest before subsequent regression of acquired skills. Neurodegenerative disorders often result in atrophy of already existing neural cells. Clinical and radiological imaging provide a means to assess progression, identify clinical features, and guide decisions for therapy. Early diagnosis and treatment can improve outcomes.

3079S
Low levels of CHIP in fibroblasts derived from patients with autosomal recessive cerebellar ataxia can be caused by mutations in STUB1. S. Johansson1, K. Hjelmaas1, M. Sanchez-Guixé1, I. Aukrust2, J. Bollerslev2, A.K. Eriksen2, E. Gude2, J.A. Koht3, S. Erdaht3, T. Fiskerstrand3, B.I. Haukanes4, L.G. Bjorkhaug5, C.M.E. Talakkas6, P.M. Knappskog7, M.D. Vigeland7. 1) Department of Medical Genetics, Oslo University hospital, Oslo, Norway; 2) Department of Medical Genetics, University of Oslo Hospital, Oslo, Norway; 3) Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway; 4) Section of Specialized Endocrinology, Department of Pediatrics, Haukeland University Hospital, Bergen, Norway; 5) Department of Clinical Genetics, Oslo University Hospital and University of Oslo, Oslo, Norway; 6) Department of Cardiology, Oslo University Hospital Rikshospitalet, Oslo, Norway; 7) Department of Neurology, Vestre Viken Hospital, Drammen; 8) Department of Neurology, Oslo University Hospital, Oslo, Norway.

A subset of hereditary cerebellar ataxias is inherited as autosomal recessive traits (ARCs). Recent studies have reported that mutations in the endoplasmic reticulum stress gene STUB1 are responsible for ARCs. In the present study, we have characterized the protein CHIP in fibroblasts derived from patients with ARCs and healthy controls. CHIP is a protein that regulates ubiquitination and protein degradation. We found that CHIP levels were decreased in ARCs, and that this was associated with decreased levels of the protein heat shock protein 70 (hsp70). These findings suggest that CHIP is a key player in the regulation of hsp70 levels and that this may be a pathogenic mechanism in ARCs.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3080M
Truncating mutations in the negative feed-back regulator of interferon 1 signalling, USP18 gene causes pseudo-TORCH syndrome. G.M.S. Mancini1, M.E.C.M. Meuwissen1, R. Schol1, S. Tinschert2, L. van Unen1, D. Heijmans3, W.F.J. van Lekken3, J.M. Kros4, R. Willemsen4, F.W. Verheijen1. 1) Clinical Genetics, Erasmus University Medical Center, Rotterdam, Netherlands; 2) Bioinformatics, Erasmus University Medical Center, Rotterdam, Netherlands; 3) Pathology, Erasmus University Medical Center, Rotterdam, Netherlands; 4) Medical genetics, University of Innsbruck, Austria.

Inappropriate stimulation of the type 1 interferon (IFN1) response pathway has been proposed as the disease mechanism in autoimmune disorders such as systemic lupus erythematosus. Also disorders resembling congenital viral infections such as pseudo-TORCH and Aicardi-Goutières syndrome have been ascribed to mutations in receptors and enzymes that lead to cytosolic accumulation of nucleic acids as triggers, indirectly linking these disorders to the IFN1 pathway. The general term “type I interferonopathies” is now used to indicate these disorders. We report on two unrelated families with five affected individuals with a pseudo-TORCH-like phenotype, presenting in the perinatal period with severe brain hemorrhage, destruction of ependymal layer and early demise. In the first consanguineous family including 3 affected sibs we used linkage analysis, whole exome sequencing and Sanger sequencing of the non-covered exons in the linkage area and identified a homozygous truncating mutation in USP18 causing total absence of USP18 transcript in cultured fibroblasts. In the second family including two sibs the same truncating mutation in heterozygous form was found. However, no transcript was amplified from cells of the two patients, suggesting that a second pathogenic mutation is present in areas not covered by Sanger sequencing. Usp18 knockout mice have cerebral hemorrhage and hydrocephalus with ependymal necrosis. Microscopy of patient brain tissues shows severe ependymal abnormalities, similar to the mouse. USP18 encodes USP43, a protein with a dual independent function in cleavage of ubiquitin-like protein 15 (ISG15) conjugates, through its carboxypeptidase domain, and, through its SH2 domain, a function as negative feed back regulator of the intracellular interferon-dependent JAK-STAT signalling pathway. USP18 mutations therefore result in a sustained, deleterious and deregulated interferon response. USP18 mutations are a novel genetic cause of pseudo-TORCH syndrome and severe cerebral hemorrhage. While the other interferonopathies are linked to inappropriate stimulation of the IFN1 response pathway, this is the first example of defective negative feed-back regulation of the IFN1 pathway, which might have consequences for the design of future therapies.

3081T
The Expanding Phenotype of TRPV4 Related Neuropathies With Notable Intrafamilial Variability. L. Medne1, C. Bonnemann2, S. Scherer2-4, R. Finkel2, X. Ortiz-Gonzalez2, A. Glianzman3, T. Estlow3, A. Moll4, R. Leshner4, Y. Wang2, T. Winder2, S. Yumi2. 1) Division of Neurology, The Children’s Hospital of Philadelphia, Philadelphia, PA; 2) NINDS, NIH, Bethesda, MD; 3) Division of Neurology, The Hospital of the University of Pennsylvania; 4) School of Medicine, The University of Pennsylvania; 5) Nemours Children’s Hospital, Orlando, FL; 6) Department of Physical Therapy, The Children’s Hospital of Philadelphia, Philadelphia, PA; 7) Department of Occupational Therapy, The Children’s Hospital of Philadelphia, Philadelphia, PA; 8) Division of Neurology, Rady Children’s Hospital, San Diego, CA; 9) Prevention Genetics Laboratory, Marshfield, WI.

TRPV4 gene mutations are known to cause 5 skeletal dysplasias, 1 arthropy and 3 forms of neuropathies: congenital distal spinal muscular atrophy (CDSMA), scapuloperoneal spinal muscular atrophy (SPSMA) and Charcot-Marie-Tooth disease type 2C (CMT2C). We report 5 patients with TRPV4 related neuropathies, three of whom have positive family histories. Three patients share the common p.Arg269His mutation but have varied phenotypes. Patients 1 and 2 presented with arthrogryposis multiplex congenital (AMC), hip dysplasia, scoliosis, distal-proximal weakness in lower extremities (LE) with limited walking, and weakness of proximal upper extremities (UE). Patient 3 with the p.Arg269His has CMT2 with pure axonal neuropathy, which she inherited from her subclinically affected mother with evidence of mild chronic denervation on EMG. Her maternal uncle and cousin, also with p.Arg269His, had AMC, motor axonal neuropathy and severe course: uncle lost ambulation in his 40s; her cousin never walked. Patient 4 has AMC with limited walking since 2.5 years, distal-proximal weakness in LE, mild proximal UE weakness and prominent scapulae. He has 2 mutations: p.Val620Ile - seen in CMT2C and brachyolmia type 3, but not in CDSMA; and p.Arg151Trp seen with low frequency in dbSNP. Patient 5 is a 3 y/o with gait abnormalities and axonal neuropathy with a similarly affected father and grandmother. His p.Arg186Gln is a known CMT2C mutation. All five probands reported here had clear motor axonal neuropathies on nerve conduction and EMG studies. All had muscle fasciculations noted during their muscle ultrasound examinations that showed an overall neurogenic pattern of denervation and reinnervation. There is no definitive genotype : phenotype correlation, particularly as TRPV4 related neuropathies represents both a diagnostic and clinical management dilemma. There is no definitive genotype : phenotype correlation, particularly as demonstrated by the presence of both CMT2C and CDSMA phenotypes in the same family (patient 3) reported here. All three affected congenital arthrogryposis presentation also had proximal upper extremity weakness and patients 2 and 4 had scapular winging, showing an overlap between CDSMA and SPSMA phenotypes rather than completely distinct clinical entities. TRPV4 gene mutations should be considered in all cases of AMC presentation with evidence of axonal neuropathy.
3082S WWOX and severe early onset epileptic encephalopathies: description of two additional patients and new clinical insights. C. Philippel-1,2, C. Mignon1, P. Hasquier1, J. Lambert1, L. Allouch-Cegla1,2, P. Jouveaux1,2,1) Laboratoire of Medical Genetics, Centre Hospitalier Universitaire, Vandoeuvre les Nancy, France; 2) Laboratoire INSERM US054-NGERE-Nutrition Génétique et Exposition aux Risques Environnementaux, Faculté de Médecine, Université de Lorraine, Vandoeuvre les Nancy, France; 3) Département de Génétique, Centre de Référence "Déficiences Intellectuelles de Causes Rares"; APHP; Groupe Hospitalier Pitié Salpêtrière, Paris, France; 4) Service de Généétique Clinique, Centre de Référence CLAD Ouest, CHU Rennes, Rennes, France; 5) Department of Neurology, University Hospital, University of British Columbia, Vancouver, Canada.

Wwox on chromosome 16q was initially identified as a putative tumor suppressor gene considering that it is frequently altered by deletions or translocations in many neoplastic lesions. The 46-kDa WWOX protein contains two N-terminal WW domains and a central short-chain dehydrogenase/reductase domain. It has been implicated in many biological processes including tumor suppression, metabolic disorders, immune defects, bone tumors, neurodegenerative diseases, and early onset severe epileptic encephalopathies. The description of individuals carrying germline WWOX mutations or deletions is recent. The first germline rearrangement of WWOX in a heterozygous state was reported in a child with ambiguous genitalia (White et al. Eur. J Hum Genet 2012). Recently, WWOX was implicated in autosomal recessive spino-cerebellar ataxia 12 (SCAR12) in two consanguineous families with homozygous missense mutations (Mallaret et al., Brain 2014). In another consanguineous family, a homozygous early stop codon was identified in an Egyptian girl affected by a severe lethal neurodevelopmental phenotype characterized by intrauterine growth retardation at age 8 months and late onset severe neurological impairment. Rare WWOX mutations have been reported by us and other groups, the in vitro functional characterization as detected by immunocytochemistry and functional domain specific silencing as detected by Crispr-Cas9, has shown a correlation with neuronal cell-cycle delays in the G2/M phase as indicated by flow cytometry. This varied collection of mechanisms may underlie the clinical and structural heterogeneity observed in individuals with tubulopathies, and provides context to detail before the design of in vivo models for interventional research.

3084T A novel pathogenic mechanism in Hereditary Spastic Paraplegia. R. Schuler1,2,3, A. Rebole4, M. Bonin4, E. Battaglia5, G. Woehlke1, L. Schols1,2,3, S. Zuchner1, A. Caballero Oteyza2,3,1) Hussman Institute for Human Genetics, Miami, FL; 2) Hiert Institute for Clinical Brain Research, Tubingen, Germany; 3) German Center for Neurodegenerative Disease, Tubingen, Germany; 4) Institute of Medical Genetics and Applied Genomics, Tubingen, Germany; 5) Bogazici University, Department of Molecular Biology and Genetics, Istanbul, Turkey; 6) Department of Physics E22 (Biophysics), Technical University Munich, Garching, Germany.

Hereditary Spastic Paraplegias (HSP) are characterized by a length dependent lower motor neuron disease which may be genetically caused by mutations in LRRK2 and DNM3. Findings suggest the regulation of protein sorting and trafficking in neurons is central to the etiology of PD.

3085S DN3: a genetic modifier of LRRK2 parkinsonism. J. Trinh1, E. Gustavsson1, I. Guella1,2,3, W.B. Doherty1, D. Doherty1, D. Pütz1,2,1) Neurology & Molecular Neuroscience Research, College of Medicine, Swansea University, UK; 2) Wales Epilepsy Research Collaboration, Department of Clinical Genetics and Expertise for Neurodevelopmental Disorders, Erasmus University Medical Center, Rotterdam, The Netherlands; 3) Departments of Pediatrics and Biomedical Genetics, University of Rochester Medical Center, Rochester, NY, USA; 5) Center for Integrative Brain Research, Seattle Children’s Hospital and Department of Pediatrics, University of Washington, Seattle, USA; 6) Institute of Medical Genetics, University Hospital Wales, Cardiff University, UK.

A novel pathogenic mechanism in Hereditary Spastic Paraplegia, R. Schuler1,2,3, A. Rebole4, M. Bonin4, E. Battaglia5, G. Woehlke1, L. Schols1,2,3, S. Zuchner1, A. Caballero Oteyza2,3,1) Hussman Institute for Human Genetics, Miami, FL; 2) Hiert Institute for Clinical Brain Research, Tubingen, Germany; 3) German Center for Neurodegenerative Disease, Tubingen, Germany; 4) Institute of Medical Genetics and Applied Genomics, Tubingen, Germany; 5) Bogazici University, Department of Molecular Biology and Genetics, Istanbul, Turkey; 6) Department of Physics E22 (Biophysics), Technical University Munich, Garching, Germany.

Hereditary Spastic Paraplegias (HSP) are characterized by a length dependent lower motor neuron disease which may be genetically caused by mutations in LRRK2 and DNM3. Findings suggest the regulation of protein sorting and trafficking in neurons is central to the etiology of PD.
Novel de novo sequence variation in **HNRNPU** gene is associated with generalized epilepsy responsive to ketogenic diet. R. Veith\(^1\), L. Carey\(^1\), D. Helbling\(^3\), M. Tscharnren\(^2\), J. Wendt-Andrae\(^2\), M. Waknitz\(^2\), G. Scherer\(^2\,\(^3\).

1) Pediatrics-Genetics, Children’s Hospital of Wisconsin, Milwaukee, WI, United States; 2) HMGC, DNL-Seq Laboratory, Dept. of Pediatrics, Medical College of Wisconsin, Milwaukee, WI, United States.

Infantile and childhood onset seizure disorders and epileptic encephalopathies (EE) are a heterogeneous group of disorders for which there is established evidence of genetic contribution. Novel or de-novo sequence changes play an important role in seizure disorders and have been described in several known and new genes through the use of whole exome sequencing (WES). Our patient is a 27 month old Caucasian female with onset of generalized seizures between 10-12 months of age, which initially appeared to be single febrile events. Mild developmental regression was noted following the second and third seizure. At 12 months the patient experienced her first cluster (afebrile) seizures, was admitted with report of “normal” MRI and EEG (with diffuse slowing), and at that time was started on Phenobarbital. Several antiepileptic drugs were administered, but did not control seizures adequately. During the next 12 months she developed a severe EE and at peak experienced 20-25 seizures per day. The patient empirically began a ketogenic diet at 22 months of age and experienced significant relief of seizures within a month of initiation. This has dramatically improved her clinical course and she now has only 3-4 seizures per day; and has gone for up to 10 days without seizures. Extensive genetic evaluation (including comprehensive epilepsy gene panel), and metabolic work-up did not reveal a diagnosis. WES (SureSelect® v4) was pursued and identified a novel deleterious sequence variant in **HNRNPU** (p.ARG770TER). This de-novo variant was not detected in both parents or the healthy sister. The variant was also not identified in 10,000 reference exomes or control chromosomes. **HNRNPU** has been recently published in association with EE via point mutations and microdeletions. Individuals with deletions of chromosome 1q44 expected that an EE gene resides in this region. **HNRNPU** would be a likely candidate. The known predisposition genes for the EE network and **HNRNPU** directly interacts with other well documented EE genes including **PNKP**. Interestingly, none of the previously documented EE candidate. The known predisposition genes for the EE work in a complex network and **HNRNPU** directly interacts with other well documented EE genes including **PNKP**. Interestingly, none of the previously documented EE gene. Preliminary data should be confirmed by an outlier sampling strategy, in the future.
3088S
Defining the presence of GTF2IRD1 in epigenetic complexes as a means to understand features of Williams-Beuren syndrome. P. Carmona-Mora1, F. Tomaselgi1, C.P. Canales1, A. Alishaw2, M. Dottori2, E.C. Hardeman1, S.J. Palmer2. 1) Cellular and Genetic Medicine Unit, School of Medical Sciences, UNSW Australia, Sydney, New South Wales 2052, Australia; 2) Centre for Neural Engineering, The University of Melbourne, Melbourne VIC 3010, Australia.

GTF2IRD1 is a member of the GTF2I gene family, which lies on chromosome 7 within a 1.8 Mb region that is prone to duplications and deletions leading to a series of physical and neurobehavioural abnormalities. Hereditary deletions cause Williams-Beuren syndrome (WBS) and duplications cause WBS duplication syndrome. Genotype-phenotype correlations of patients with atypical deletions and knockout mouse models show GTF2IRD1 as responsible for the craniofacial abnormalities, mental retardation, visuospatial construction deficits and hypersociability of WBS. In order to understand the basis of the phenotypes related to GTF2IRD1, we have studied its molecular function in vitro. We evaluated its expression in human ES cell-derived neurons, to correlate it with specific neuronal types and stages of differentiation. Furthermore, a detailed analysis in cell lines of endogenous GTF2IRD1 shows a specific nuclear speckled pattern of expression, colocalizing with markers of gene silencing complexes. To define GTF2IRD1 functional relationships, we used a yeast two-hybrid approach to isolate interacting partners and validate many interactions in mammalian cells. Most of the proteins are involved in chromatin modification and transcriptional regulation in agreement with the nuclear localization data. In addition, several of these partners, such as ZMYM2, ZMYM3, ZC4H2 and DCAF6, were shown to occupy the same nuclear compartment as GTF2IRD1 in chromatin modifier complexes and identified direct associations in histone deacetylase (HDAC) complexes related to neuron-specific gene regulation and others are directly linked to mental retardation through mutations in their encoding gene. We have demonstrated the presence of GTF2IRD1 in chromatin modifier complexes and identified direct associations with HDACs, allowing us to expand its interactional network and propose functional links with specific transcription factors and epigenetic silencing complexes. Within the same context, we also explored GTF2IRD1 capabilites to modify nuclear acetylation and tested the dependence of these proteins are involved in histone deacetylase (HDAC) complexes related to neuron-specific gene regulation and others are directly linked to mental retardation through mutations in their encoding gene. We have demonstrated the presence of GTF2IRD1 in chromatin modifier complexes and identified direct associations with HDACs, allowing us to expand its interactional network and propose functional links with specific transcription factors and epigenetic silencing complexes. Within the same context, we also explored GTF2IRD1 capabilites to modify nuclear acetylation and tested the dependence of these proteins on histone acetylation status. Further, we analyzed the influence of GTF2IRD1 on transcriptional activity, revealing a complex balance of transcriptional activation and repression. Overall, these findings expand our understanding of the role of GTF2IRD1 in neurodevelopmental processes and suggest potential therapeutic targets for neurodevelopmental disorders.

3088M
Generation of a comprehensive panel of patient-derived pluripotent stem cells to dissect oligodendrocyte dysfunction in the pediatric myelin disorder Pelizaeus-Merzbacher Disease. Z. Nevin1, R. Karl1, G. Hobson2, P. Tesar3,4. 1) Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH; 2) Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE; 3) New York Stem Cell Foundation, New York, NY.

Patient-derived pluripotent stem cells (iPSCs) have become integral tools for parsing the molecular biology of complex human genetic disorders. Pelizaeus-Merzbacher Disease (PMD) (MIM 312080) is a pediatric leukodystrophy affecting central nervous system myelination that results in severe motor impairment, intellectual delay, and premature death and has no available treatment. Though PMD is a single-gene disorder, it exhibits a spectrum of phenotypes and clinical severities attributed to over 200 different mutations in the essential myelin gene PLP1 (MIM 300041), which encodes for two protein isoforms, proteolipid protein 1 (PLP1) and DM20. Protein misfolding and endoplasmic reticulum stress are linked in pathogenesis of PMD in transgenic animal models overexpressing PLP1, but this may not hold true for the many variant point mutations and gene deletions identified in patients. Detailed investigations of PLP1’s normal function and the etiology of PMD are further complicated by a lack of access to primary human oligodendrocytes, the only cell type in which PLP1 protein is known to be expressed. To address this need, we have generated a large panel of patient-derived iPSCs from children with unique duplication, triplication, deletion, and point mutations in PLP1 that encompass the genotype and phenotypic variation seen in patients. Using a method that has been validated in our lab, we have derived pure populations of oligodendrocytes and oligodendrocyte progenitors from these iPSCs in order to dissect individual patient phenotypes at a cellular level. Access to this unique resource is now available to understand features of Williams-Beuren syndrome. 1,3, P. Tesar.

3090T
The NINDS Repository collection of patient-derived fibroblasts and induced pluripotent stem cells for neurodegenerative disease research. C. O. Pearsall1, S. Heilig1,4, C. A. Pérez1,4, C.A. Garcia-Ballesteros1, K. Hodges2,4, K. Hughes3, M. Sutherland3, R.A. Corriveau3, C. Tam1,4. 1) NINDS Repository; 2) Stem Cell Biobank; 3) Cell Culture Laboratory; 4) Coriell Institute for Medical Research, Camden, NJ; 5) National Institute of Neurological Disorders and Stroke - NIH, Bethesda, MD.

Induced pluripotent stem cells (iPSCs) reprogrammed from patient-derived primary fibroblasts have become an increasingly utilized and needed resource for the study of human disease and have proven especially valuable for the identification of novel stem cell-based therapeutic targets and disease-modifying strategies. Despite rapid progress in the field, there exists difficulty to establish the National Institute of Neurological Disorders and Stroke (NINDS) Repository as a public resource established in 2002 aiming to provide a centralized and open collection of biological samples (DNA, lymphocytes, cell lines, fibroblasts, biologics such as plasma, serum, cerebrospinal fluids, and urine) and associated de-identified clinical data from a diverse population of patients and normal controls. Since 2011, the NINDS Repository has added to its web-based catalog (http://ccr.coriell.org/NINDS) close to 50 iPSC and 165 fibroblast lines. Most iPSC lines are contributed by Parkinson’s disease (PD), amyotrophic lateral sclerosis (ALS) or Huntington’s disease (HD) investigators from NINDS-sponsored Stem Cell Consortia. By making available to the research community neurodegenerative disease iPSCs and fibroblasts, the NINDS Repository continues to fulfill the NINDS mission of reducing the burden of neurological disease - a burden borne by every age group, by every segment of society, by people all over the world. To ensure the quality of these valuable resources, all iPSCs and fibroblasts submitted to the NINDS Repository by iPSC Consortiums or Stem Cell Biobanks are tested for quality, including karyotype, pluripotency, potential, karyotyping, differentiation status, gene expression analysis, sterility prior to distribution by the NINDS repository. The results are summarized in a Certificate of Analysis and/or displayed on the web-based catalog along with recommended culturing protocol. The NINDS Repository fibroblast and iPSC collections include mostly cell lines bearing specific genetic mutations associated with PD, ALS, HD, frontotemporal degeneration, or Alzheimer’s disease, as well as samples derived from neurologically normal controls. Elucidating the effects of affected and parental fibroblasts, corresponding DNA, and whole blood DNA are available. The NINDS Repository serves as a unique and effective centralized resource where these iPSCs, fibroblasts and their de-identified phenotypic data, are available to basic and applied research investigators worldwide.

3091S
Human iPSC model of the Ras/MAPK pathway role in neurodevelopmental disorders. E. Yeh1,2, Y. Wu1, C. Tom1, B. Adviento1, K.A. Rauen1, L.B. Weiss1. 1) Institute of Medical and Molecular Genetics, UCSF, San Francisco, CA; 2) Department of Pediatrics, UC Davis, Sacramento, CA.

Germline mutations in the primordial Ras/MAPK pathway cause a class of human genetic syndromes termed RASopathies characterized by over-activation of this pathway. Cardiofaciocutaneous syndrome (CFC) is a rare RASopathy characterized by heart defects, a distinctive facial appearance, and ectodermal abnormalities. Neurological abnormalities include low IQ, learning disability, and speech delay. Autistic traits are present in 54-64% of affected individuals. To better understand the role of the Ras/MAPK signaling pathway in neurodevelopmental disorders, we established induced pluripotent stem cell (iPSC) lines from skin fibroblasts from 3 CFC patients carrying the most common mutation Q257R in BRAF. Interestingly, we found that the iPSC lines with the most common mutation in patients have the same Q257R mutation in BRAF. The iPSC lines generated from the CFC patients show increased proliferation compared to wild-type (WT) NPC, as evidenced by higher percentage of KI67 positive cells (74.9±6.6% vs 50.7±7.5%, p=0.028) and shorter doubling time (41.9±7.8 vs 85.6±20.8 h, p=0.038). Seven days after induction, BRAFQ257R NPC showed higher proliferation rates compared to wild-type (WT) NPC, as evidenced by higher percentage of Ki67 positive cells (74.9±6.6% vs 50.7±7.5%, p=0.028) and shorter doubling time (41.9±7.8 vs 85.6±20.8 h, p=0.038). Seven days after induction, BRAFQ257R and WT cultures expressed similar levels of the neuron-specific markers Beta III tubulin, MAP2 and Tau protein. However, image analysis revealed that BRAFQ257R neurons had more neurites (3.5±0.1 vs 2.9±0.1 neurites/cell, p=0.0047) and longer neurites (156.2±17.6 vs 79.8±9.5 µm, p=0.0096). The early maturation in BRAFQ257R neurons was further evidenced by a higher number of cells positive for TBR1, a layer 5 marker (44.07±5.6% vs 8.90±1.8%, p=0.0001). The higher NPC proliferation rate we observe is particularly interesting as >50% of CFC patients have macrocephaly. Previous literature shows that BRAF null mice have shorter neurites and kinase-activated BRAF mice have longer neurites, while BRAF deficient mice have normal neurites. Therefore, the number of neurites and longer neurites are indicators of neuronal differentiation and maturation, which was supported by increased expression of a cortical layer 5/6 marker. Thus, we observed the presence in first time in human cells that a CFC mutation in BRAF is associated with altered neurite morphology. This study will help us obtain insight not only into the pathophysiology of CFC, but also into the role of Ras/MAPK pathway in common disorders, like Autism Spectrum Disorders.
3092M
New insights into the biological role of VAX2 in human, in health and
generation. G. Altara1, A.Z. Shah1, N.H. Waseem1, A.R. Webster1,2, S.S. Bhattacharya1, 1) Institute of Ophthalmology, UCL, London EC1V 9EL, UK; 2) Moodyfield Eye Hospital, London EC1V 2PD, UK.
Purpose: VAX2 (MIM 604295) is a transcription factor expressed in the
ventral region of the prospective neural retina in vertebrates and is required for
telecephalon specification. Its genetic inactivation in mice leads to an
incompletely penetrant ocular coloboma. Based on the similarities between
the mutant mouse phenotype and the clinical features of isolated coloboma
in human, VAX2 represents a good candidate gene for this human condition.
Thus far, mutations affecting VAX2 in coloboma patients have not been
reported. Interestingly, recently it has been shown that VAX2 in mouse
plays a key role in ensuring appropriate gene expression in cone photorecep-
tor cells. Despite a wide range of studies in vertebrates its function in human
is still poorly understood. This study was undertaken to better elucidate the
role of VAX2 in human in health and disease. Methods: VAX2 transcripts
were obtained by RT-PCR on cDNA from human tissues and cultured cells.
These transcripts were sub-cloned into mammalian expression vectors for
in vitro analysis. Variant proteins were generated by site-directed mutagene-
sis. A cohort of cone-rod dystrophy patients was screened for variant identifi-
cation using Sanger sequencing. Results: It was observed that VAX2 is
highly expressed in neuronal tissues and cultured cell lines. Two spliced
variants were detected: isoform-1 (NM_012476), and a novel transcript
(isoform-2), mainly expressed in retina and predicted to encode a shorter
protein of 150 amino acids. In vitro exogenous expression of tagged protein
isoforms showed differential localization potentially underlying distinct bio-
logical roles. Isoform-1 localizes to the nucleus while isoform-2 is widespread
throughout the cell. The involvement of the isoform-2 in specifically being
expressed suggested screening of patients with cone-rod dystrophy. Sequence
analysis revealed a heterozygous change c.416T>G (p.Leu139Arg) in one
patient. Characterization of this variant showed damaging effects on protein
localisation and is therefore suggestive of a functional defect, with potential
involvement in the disease phenotype. Conclusions: Our data suggest that
in addition to being a transcription factor VAX2 could be involved in other
molecular pathways due to the short isoform in the human. Moreover, VAX2
may be involved in cone photoreceptor cell biology. In summary, our findings
raise new questions about the VAX2 gene and its biological role in health
and disease.

3093T
Digital gene expression differences in the OV737 ovine model of Hun-
tington’s Disease. S.J. Reid1, R. Brauning1, P. Maclean1, S. Patassini1, R.R. Handley1, P. Tsai2, H.J. Waldvogel2, J.F. Gusella3, R.L.M. Faull1, R.G. Snell1, and the H.D.S.C.R.G.1, 1) Centre for Brain
Research, The University of Auckland, Auckland, Auckland, New Zealand; 2) AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand; 3) AgResearch, Ruakura Research Centre, Hamilton, New Zealand; 4) School of Biological Sciences, The University of Auckland, Auckland; 5) Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston; 6) Huntington’s Disease Sheep Collaborative Research Group (10.3233/jhd-130067).
In order to better understand the early molecular mechanisms of Hun-
tington’s Disease, we examine the OV737 transgenic sheep for progressive
histopathological and molecular changes. These animals express at modest
levels the full length human huntingtin protein with 73 glutamine repeats
[1], demonstrate reduced striatum levels of GABA_A receptors, cortical
huntingtin inclusions [2] and an early circadian abnormality [3]. The sheep
are healthy and visibly indistinguishable from control South Australian
Merino. They present an opportunity to better understand and utilise the
pre-symptomatic phase of Huntington’s Disease.

To identify genome-wide gene expression differences, Illumina RNA-seq
data (2.3 x 10^6 50bp PE reads) was collected from ribosomal depleted total
RNA isolated from anterior striatum tissue of twelve 5-year old OV737 sheep
(6 transgenic, 6 control). For assembly of a de-novo transcriptome, further
reads (1.3 x 10^7 100bp PE) were collected from brain and liver total RNA
from fetal and neonate sheep. Differentially expressed transcripts in OV737
were identified using both genome based approaches and the assembled
transcriptome. A selection of these annotated transcript findings will be
presented, including OXTR and AVPR1A which may relate to the circadian
abnormality seen in the OV737 animals.

2. HDSRCG, et al., Further Molecular Characterisation of the OV737 Transgenic Sheep Model of Huntington’s Disease Identifies Cortical Aggre-
3. Morton, A.J., et al., Early and progressive circadian abnormalities in
Huntington’s disease sheep are unmasked by social environment. Hum Mol Genet. 2014.

3094S
De Novo mutations in TEAD1 and OCEL1 in non-X linked Aicardi Syn-
drome. I. Schrauwen1,2, S. Szelinger1,2, A.L. Siniard1,2, J.J. Come-
evaux1,2, A. Kurgopol1,2, R. Richholt1,2, M. De Both1,2, I. Malenica1,2, S. Swaminathan1,2, S. Rangasamy1,2, N. Kulkarni1,2, S. Bernes4,5, J. Buch-
halter6, K. Ramsey7,2, D.W. Craig1,2, V. Narayanan1,2, M.J. Huentelman1,2.
1) Dorrance Center for Rare Childhood Disorders, Translational Genomics
Research Institute, Phoenix, AZ; 2) Neurogenomics Division, Translational
Genomics Research Institute, Phoenix, AZ; 3) Department of Medical
Genetics, University of Antwerp, Antwerp, Belgium; 4) Phoenix Children’s
Hospital, Phoenix, AZ; 5) MiHs integrated program, Phoenix, AZ; 6) Alberta
Children’s Hospital, Calgary, Alberta, Canada.
Aicardi syndrome (AIC) is a congenital neurodevelopmental disorder char-
derminated by infantile spasms, agenesis of the corpus callous and chorioret-
inal lacunae. There is a wide range in severity of these classic symptoms
however, with chorioretinal lacunae as the most constant feature of the
disease. Although cytogenetic studies have linked AIC to the X-chromo-
some, no gene has been identified to date, and little is known about the
pathogenic mechanisms underlying the disease. Recently, XY males with
the disease have been described as well, suggesting the presence of genetic
heterogeneity. In this study, we investigated 10 female children diagnosed
with AIC, and their parents, by exome/genome sequencing. In two cases,
we identified de novo mutations in the autosomal genes TEAD1 and OCEL1,
both expressed and important in the retina and brain. The mutation in TEAD1
is a nonsense mutation (p.Trp206Ter; CADD score = 40), and the mutation
in OCEL1 is a non-synonymous mutation (p.Ala167Thr; CADD score = 15.3),
which is predicted damaging according to MutationTaster, LTR, Plo-
pyhen2, and affects a highly conserved residue (PhyloP = 2.1). A missense
mutation in OCEL1 (p. Gly129Arg) is predicted damaging according to
Sift, Polyphen2, and affects a highly conserved residue (PhyloP = 3.1). Both
mutations were rare, absent from 6,439 VUS cases and 125,653 controls.
It is noteworthy that these mutations are located in the C2 domain of
the VAX2 protein, which is responsible for repression of target genes at a
pre-mRNA level. These results raise the important possibility that mutations in
the genes detailed here may be present in male patients with similar phenotypic presentations. Investigation of the pathological mechanisms that lead to disease (retinal and neuronal) through
OCEL1 and TEAD1 dysfunction might reveal novel therapeutic avenues for
treatment and symptomatic management of the disease.
Absence of ER cation channel TMEM38B/TRIC-B causes recessive osteogenesis imperfecta by dysregulation of collagen post-translational modification. W.A. Cabral1, E.N. Makareva2, M. Ishikawa3, A.M. Barnès4, M.A. Weis4, F. Lacobawan6, D.R. Eyre6, Y. Yamada5, S. Leikin7, J.C. Marini8, 1) Bone & Extracellular Matrix Branch, NICHD, NIH, Bethesda, MD, USA; 2) Section on Physical Biochemistry, NICHD, NIH, Bethesda, MD, USA; 3) Molecular Biology Section, NIDCR, NIH, Bethesda, MD, USA; 4) Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA, USA; 5) Department of Medical Genetics, Children’s National Medical Center, Washington DC, USA.

A novel form of recessive osteogenesis imperfecta (OI) has been reported in Israeli and Saudi Bedouins, caused by homozygous deletion of exon 4 and surrounding intronic sequence of TMEM38B. TMEM38B encodes TRIC-B, an integral ER membrane protein proposed to facilitate Ca\(^{++}\) influx from intracellular stores. However, the molecular mechanisms through which this founder mutation causes OI are unknown. We identified a 26-month-old girl with moderately severe OI who is the offspring of consanguineous parents from Saudi Arabia, and is homozygous for the TMEM38B founder mutation. The proband presented with growth deficiency, osteopenia, recurring long bone fractures and blue sclerae. Proband TMEM38B transcripts are 25% of normal control level and include six spliced forms. Although one minor transcript is in-frame, absence of TRIC-B protein was confirmed by immunoblot. Consequently, proband cells have decreased intracellular and ER luminal Ca\(^{++}\) concentrations, which were more rapidly depleted upon ATP-induced Ca\(^{++}\) efflux from the ER. The lower Ca\(^{++}\) level does not involve changes in expression or stability of SERCA2a or IP3R, the ion channels involved in ER Ca\(^{++}\) loading and release, respectively. Furthermore, depletion of Ca\(^{++}\) is associated with ER stress, activation of ATF4 translation, and dysregulation of type I collagen synthesis. Although proband cells have normal PDI levels, the 20-minute delay in procollagen chain assembly into heterotrimeric molecules is consistent with sequestration of PDI by calreticulin in low Ca\(^{++}\) conditions. Abnormal modification of procollagen is evident by increased electrophoretic migration of alpha chains on SDS-Urea PAGE, consistent with a 30% reduction in collagen lysine hydroxylation, evident by increased electrophoretic migration of alpha chains on SDS-Urea PAGE, consistent with a 30% reduction in collagen lysine hydroxylation, indicating abnormal procollagen conformation. However, extracellular matrix deposited by proband fibroblasts contained only collagen with normal thermal stability, and was associated with 30% reduction of cross-links. These data suggest a role for TRIC-B in maintaining intracellular Ca\(^{++}\), and demonstrate that absence of TRIC-B causes OI by dysregulation of multiple Ca\(^{++}\)-regulated collagen-specific chaperones and modifying enzymes in the ER. Therefore, the mechanism of TRIC-B absence falls within the collagen-related paradigm of OI.

Exome sequencing identifies locus heterogeneity in multiple epiphysseal dysplasia. K. Balasubramanian1, M.J. Bamshad2, D.A. Nickerson3, R.S. Lachman3, D.H. Cohn1,4, 5) University of Washington Center for Mendelian Genomics. 1) Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, Los Angeles, CA; 2) Department of Pediatrics, University of Washington, Seattle, WA; 3) Department of Genome Sciences, University of Washington, Seattle, WA; 4) Department of Orthopaedic Surgery and Orthopaedic Hospital Research Center, University of California, Los Angeles, Los Angeles, CA; 5) International Skeletal Dysplasia Registry, University of California, Los Angeles, Los Angeles, CA. Multiple Epiphysseal Dysplasia (MED) is a relatively common skeletal dysplasia associated with mild short stature and early-onset osteoarthritis. Mutations in five genes have been associated with dominant forms of the disease: cartilage oligomeric matrix protein (COMP), matrilin-3 (MATN3), and the three type IX procollagen genes (COL9A1, COL9A2, and COL9A3). All of these genes are selectively expressed in cartilage and encode structural molecules of the cartilage extracellular matrix. A clinically distinct recessive form of MED results from mutations in a widely expressed gene, Solute Carrier Family 26 Member 2 (SLC26A2), which encodes a protein required for post-translational sulfation. In about 15% of MED cases the molecular basis of disease remains unknown, implying further genetic heterogeneity. To identify additional disease loci, exome sequencing of eight MED families was conducted. This cohort included four sporadic cases hypothesized to represent dominant MED, two recurrent cases thought to represent recessive MED, and two large families with clear dominant inheritance. In three cases, mutations in the known genes were identified. A novel de novo c.949C>T mutation in COMP that results in the protein change p.Asp317Asn in the calmodulin-like repeats was identified in one sporadic case. A recurrent c.518G>T mutation resulting in the p.Ala173Asp amino acid change and a novel c.326A>T mutation resulting in the p.Ile105Lys protein change in MATN3, both of which target the Von Willebrand Factor-A domain in which all previous MATN3 MED mutations have been found, were identified in the two large families with dominant MED. The remaining five cases were negative for mutations in the known genes, but heterozygosity for two novel de novo variants was identified in one case: a c.571G>T (p.Ala191Val) variant reported in a PTH2R protein change and a c.734delIT frameshift variant in PRICKLE1, a protein involved in the WNT5A noncanonical planar cell polarity pathway. Screening of 32 additional MED cases did not detect mutations in either PTH2R or PRICKLE1, suggesting that mutations in one of these genes may be a rare cause of MED. Causative variants were not found in the remaining 4 cases, demonstrating further locus heterogeneity in MED.

1) Department of Medical Genetics, Kasturba Medical College, Manipal University, Manipal, India; 2) Pediatric Orthopedics service, Department of Orthopedics, Kasturba Medical College, Manipal University, Manipal, India; 3) Clinical Genetics Unit, Christian Medical College, Vellore, India; 4) Department of Medicine, Kasturba Medical College, Manipal University, Manipal, India; 5) Department of Medical Genetics, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, India; 6) Department of Pediatric Genetics, Amrita Institute of Medical Sciences and Research Center, ALIMS Ponekkara, Cochin, India; 7) Centre for Human Genetics, Bangalore, Bangalore, India; 8) Division of Genetics, Department of Pediatrics, AIIMS, New Delhi, India; 9) Centre of Medical Genetics, Sir Gangas Ram Hospital, New Delhi, India; 10) Department of Pediatrics, J JM Medical College, Davanagere, India; 11) Department of Pediatrics, PGIMER, Chandigarh, India; 12) Vivekananda Institute of Medical Sciences, Kolkata, India; 13) Department of Medical Genetics, Nizam’s Institute of Medical Sciences, Hyderabad, India; 14) Department of Rheumatology, Madras Medical College, Chennai, India; 15) Department of Pediatrics, SAT Hospital, Government Medical College, Trivandrum, India; 16) Diagnostics Division, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India.

A small subset of arthropathies is monogenic. We aimed to analyze the mutations underlying all the conditions with arthropathy listed under the different groups in the Nosology and Classification of Genetic Skeletal Disorders (2010). The list includes progressive pseudohypertrophic muscular dystrophy; chronic infantile neurologic cutaneous articular syndrome; steroid multifocal osteochondromalacia, periositis, and pustulosis; familial hyperphosphatemic tumoral calcinosis; infantile systemic hyalinosis; familial digital arthropathy with brachydactyly; multicentric carpal-tarsal osteosclerosis with and without nephropathy; sticker spondylopathy type 1, 2, 3 and type 4. The clinical and molecular spectrum of these syndromes is being studied. We recruited 43 families with progressive pseudohypertrophic muscular dystrophy (DMD) carriers, 5 families with multicentric carpal-tarsal osteosclerosis, 2 families with familial hyperphosphatemic tumoral calcinosis; one family each with chronic infantile neurologic cutaneous articular syndrome and infantile systemic hyalinosis. We identified 6 mutations: 5 are novel (p.S304Pfs*115 and p.N430Tfs*68, p.R230Lfs*4, p.C247Lfs*31, p.C337Y) in LOX. In addition, we identified 1 known (p.R101H). These are the preliminary results of our cohort of patients with inherited arthropathies. We wish to extend this study by collaborating with national and international investigators and hope to explore the genetic bases of Mendelian arthropathies.
3100S  
Identification of mutations in patients with osteogenesis imperfecta from Russia. D.D. Nadyrshina1, R.I. Khusainova2, E.K. Khusratdinova1,2.  
1) Biological Department, Bashkir State University, Ufa, RE., Russian Federation; 2) Laboratory of human molecular genetics, Institute of Biochemistry and Genetics, Ufa, Russian Federation.  
Osteogenesis imperfecta (OI) is a clinically and genetically heterogeneous brittle bone disorder. Whereas dominant OI is mostly due to heterozygous mutations in either COL1A1 or COL1A2 encoding type I procollagen, recessive OI is caused by mutations in genes encoding proteins involved in type I procollagen processing or chaperoning. The aim of our study was to identify mutations in COL1A1, COL1A2, CRTAP, LEPRE1, PPIB and SERPINF1 genes in Russian OI patients. We examined 78 patients with OI and 100 healthy controls corresponding by age, gender, ethnicity and place of residence. We sequenced the coding and exon-flanking regions of COL1A1, COL1A2, CRTAP, LEPRE1, PPIB and SERPINF1 genes. We identified five distinct mutations, undescribed before. For the first time previously unreported nonsense mutation c.967G>T (p.Gly323X) in COL1A1 gene was identified in heterozygous state in two Russian patients, the novel frameshift mutation c.3541insC (p.Gly1181ArgfsX38) in COL1A1 gene was observed in Yakut patient, the splicing mutation c.1724+4G>A in LEPRE1 gene was identified in two patients from Tatar population. And novel compound heterozygous mutations (c.913C>G (p.Leu305Val) of SERPINF1 gene and c.641T>C (p.Val214Ala) of CRTAP gene) in OI type III was observed in three patients from Bashkir population. We also detected three previously described nonsense mutations in five Russian patients: c.1081C>T (p.Arg361X), c.1243C>T (p.Arg415X) and c.2896C>T (p.Gln957X), two frameshift mutations in two Tatar patients: c.579delT (p.Gly194ValfsX71) and c.2444delG (p.Gly815AlafsX293) and one splicing mutation in one Yakut patient: c.4005+1G>T. One frameshift mutation (c.579delT (p.Gly194ValfsX71), three nonsense mutations (c.967G>T (p.Gly323X), c.1081C>T (p.Arg361X), c.2896C>T (p.Gln957X)) and one splice mutation (c.4005+1G>T) were identified in patients with OI type I. Two mutations (c.2444delG (p.Gly815AlafsX293) and c.3540_3541insC (p.Gly1181AlafsX38)) occurred in sporadic cases of OI type I, whereas c.1243C>T (p.Arg415X) mutation - of OI type III. The mutation c.1724+4G>A were identified in patients with OI type VIII. In conclusion, the present study identified nine novel mutations in OI genes, five of them was not observed before and no mutations in COL1A2 and PPIB genes in Russian patients with OI. Future research will focus on other genes responsible for OI development in Russian patients.

3101M  
Two Distinct Mutations in IFITM5 Causing Different Forms of Osteogenesis Imperfecta Using Reciprocal Mechanisms. A. Reich1, AM. Barnes1, CR. Farber2,3, P. Becerra4, F. Rauch2, WA. Cabrall1, A. Bae1, FH. Glorieux5, TL. Clemens2, JC. Marin1.  
1) NICHD, NIH, Bethesda, MD; 2) Center for Public Health Genomics, Univ Virginia, Charlottesville, VA; 3) Depts Public Health Sciences, Biochemistry & Mol Genet, UVA, Charlottesville, VA; 4) Section on Protein Structure and Function, LROMB, NEI, NIH, Bethesda, MD; 5) Shriners Hosp for Children and McGill Univ, Montreal, Canada; 6) Dept Orthopaedic Surgery, Johns Hopkins Sch Med, Baltimore, MD.  
Osteogenesis imperfecta (OI) type V is caused by a recurrent dominant mutation (c.-14C>T) in IFITM5, which encodes BRIL, a transmembrane Ifitm-like protein most strongly expressed in osteoblasts, whose expression coordinates with mineralization. Patients with type V OI have distinctive clinical manifestations with overactive bone mineralization and include mesh-like lamellation on bone histology. Recessive null mutations in SERPINF1 cause OI type VI, which impairs mineralization. Type VI OI patients have absence of serum PEDF, elevated alkaline phosphatase (ALPL) as children, and bone histology with broad unmineralized osteoid and feathered pattern. We identified a 25-year-old woman with severe OI, who had multiple features of type VI OI. Her dermal fibroblasts (FB) and cultured osteoblasts (OB) expressed very little PEDF, and she had normal bone mineralization. Exome sequencing yielded a de novo mutation in IFITM5 in one allele of the proband, causing a p.S40L substitution in the BRIL intracellular domain, despite her not having type V OI characteristics. In OB with the IFITM5 p.S40L mutation, SERPINF1 expression was decreased in p.S40L OB, and PEDF was barely detectable in conditioned media of S40L cells. OB with the S40L mutation also had decreased ALPL, observed by microscopy. Notably, SERPINF1 expression was increased in p.S40L OB, and PEDF was barely detectable in conditioned media of S40L cells. OB with the S40L mutation also had decreased expression of ALPL and osteoectacin (OCN), as seen in primary OB cultures. Conversely, OB from type V OI, with 5 residues added to the N-terminus of BRIL, have increased SERPINF1 expression and PEDF secretion during differentiation, and increased ALPL and OCN expression. The IFITM5 mutations also have opposite effects on OB mineralization in culture - OB with the 5'-terminal mutation have increased mineralization during differentiation, while mineralization was decreased in S40L OB. OB from both mutations share a collagen related defect, with decreased expression and secretion of COL1A1. Together, these data show that 2 distinct mutations in IFITM5 generated different forms of OI with distinctive phenotypes using reciprocal mechanisms; the type V OI and p.S40L BRIL are gain- and loss-of-function mutations, respectively. Furthermore, BRIL and PEDF have a relationship that connects the genes for types V and VI OI and their roles in bone mineralization.
Clinical features of rare disorders are very often poorly understood due to their low prevalence. Quite a few times, these rare disorders remain uncharacterized, or patients are misdiagnosed and get poor medical attention. A mysterious skeletal disorder that remained unidentified for decades and rendered many people physically challenged and disabled for life has been reported in an isolated remote village 'Arai' of Poonch district of Jammu and Kashmir state, India. The typical phenotypic characteristics shared among the affected included: pain, kyphoscoliosis, fatigability, muscular weakness, progressive restriction of joint movement, stiffness and swelling at several joints, including the proximal and distal interphalangeal joints. This village is located deep in mountains, and the population residing in the region is highly consanguineous. Combining multiple lines of evidence (familial histories and genetic data, clinical features, radiological and biochemical data and phenotypic features), we identified the disorder as a recessive hereditary skeletal disease "Progressive Pseudorheumatoid Arthropathy of Childhood" (PPAC) also known as "Spondyloepiphyseal Dysplasia (SEDPA)" with a novel splice site mutation in the gene "*" are known to be responsible for causing of PPAC. Therefore, individuals, we observed about 70 affected people. Mutations in one reported disorder [for example, with an estimated frequency one per million in UK], Arthropathy of Childhood" (PPAC) also known as "Spondyloepiphyseal Dysplasia (SEDPA)" have been extensively studied. The typical clinical features of PPAC include: skeletal abnormalities resembling fibroma, irregular woven bone, and marked deformity, and fractured major limb bones. She received calcitonin and then rehydrated and got improved. A skeletal survey revealed thickened cortices of tubular bones, but a normal skull. Biopsy of a tibia and ileum showed unremarkable lamellar bone, slightly increased vascularity, numerous osteoblasts, and normal marrow. Serum alkaline and acid phosphatase were elevated. An atypical form of JPD was diagnosed. At age 14 years, she had normal stature, slight knock-knee deformity, and fractured major limb bones. She received calcitriol and then a bisphosphonate. At age 16 years, biopsy of the maxilla and jaw showed areas resembling cementifying fibroma, irregular woven bone, and marked osteoblastic and osteoclastic activity in a highly cellular fibrovascular stroma resembling osteolytic Paget's disease. Teeth were eroded and replaced by osteocementified and fibrous tissue. She was edentulous by age 27 years. At age 30 years, she had a prominent forehead and nasal bridge, deep set eyes, and anterior tibial bowing. Then, ultrasound of her first pregnancy at 18 weeks showed abnormal fetal bones with angulated femurs. Sequencing of coding exons of the RANK and OPG genes of the mother and abortus was negative. However, in both, microarray-based copy number analysis showed amplification of exons 4-9 of the RANK gene. Confirmatory qPCR showed 3 copies of exons 4-9 in the mother, and 3 in the fetus. An in-frame tandem duplication of exons 4-9 creates a RANK fusion protein of one extracellular RANKL-binding domain combined with double intracellular activation domains. Thus, we have identified a new genetic basis within the family of disorders featuring constitutive RANK activation.
3105T

The molecular mechanisms underlying the progressive fibrosis and failure of regeneration in Duchenne muscular dystrophy have remained elusive. Analysis of 166 patient muscle biopsy mRNA profiles (test and validation sets) identified a 56 member network centered on TGFβI associated with severe pathology (fibrosis and failed regeneration). Superimposing this pathology-related network on a 27 time point murine normal muscle regeneration series showed stage-specific regulation of each network member during normal staged regeneration, but at distinct time points (temporal parsing into subnetworks). From this, we developed an asynchronous remodeling model for fibrosis and failed regeneration. This model predicted that the normal 2 week regeneration cycle of muscle is disrupted through inappropriate crosstalk between neighboring cells in different time points of the 2 week-cycle. To test this hypothesis, we developed an experimental model of focal asynchronous bouts of muscle regeneration, with laser capture microscopy of marked tissue regions (first bout, second bout, and in between [crosstalk] areas). mRNA profiling and immunohistochemical validations showed that the crosstalk areas in between staged bouts of regeneration became inappropriately fixed in the developmental time point by which the initial bouts were separated. This led to a chronic inflammatory state and mitochondrial insufficiency in bouts separated by 4 days, and a chronic fibrotic state in bouts separated by 10 days. Molecular networks associated with these localized areas of pro-inflammatory states were suppressed by treatment with glucocorticoids and VBP15. In summary, synchronous regeneration of muscle is successful, whereas neighboring asynchronous bouts create inappropriate crosstalk between cells in different stages of the regenerative process that results in failed regeneration and the pro-fibrotic and pro-inflammatory state. Our data provide a unifying model for failed regeneration and pathological fibrosis in muscular dystrophies that is likely generalizable to chronic inflammatory states in other regenerative tissues. This model also provides a novel mechanism of action for glucocorticoids in many of these disorders; they serve to re-synchronize remodeling, much as diastolic cortisol fluctuations do in animals.

3106S
Novel mutations in GNAI3 associated with Auriculocondylar Syndrome strengthen a common dominant negative effect. V.L.R. Tavareś, C.T. Gordon, R.M. Zechi-Çeide, N.M. Kokitsu-Nakata, N. Vosini, T.Y. Tan, A.A. Heggie, S. Vendramini-Pittolo, E.J. Propst, B.C. Papsin, T.T. Torres, H. Buermans, L.P. Capelo, J.T. den Dunnen, M.L. Guion-Almeida, S. Lyovern, J. Amy, M.R. Passos-Bueno. 1) Centro de Pesquisas do Genoma Humano e Célula, Universidade de São Paulo, São Paulo, SP, Brazil; 2) INSERM U1163, Université Paris Descartes-Sorbonne Paris Cité, Institut Imagine, Paris, France; 3) Department of Clinical Genetics, Hospital de Referência e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil; 2) INSERM U1163, Université Paris Descartes-Sorbonne Paris Cité, Institut Imagine, Paris, France; 4) Victorian Clinical Genetics Service, Murdoch Children’s Research Institute, Royal Children’s Hospital, and Department of Paediatrics, University of Melbourne, Melbourne, Australia; 5) Department of Plastic and Maxillofacial Surgery, Royal Children’s Hospital, Melbourne, Australia; 6) Department of Otolaryngology - Head & Neck Surgery, The Hospital for Sick Children, Toronto, Canada; 7) Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brazil; 8) Leiden University Medical Center, Leiden, The Netherlands; 9) Instituto de Ciencias e Tecnologia, Universidade Federal de São Paulo, São José dos Campos, Brazil; 10) Département de Génétique, Hôpital Necker-Enfants Malades AP-HP, Paris, France.

Auriculocondylar syndrome is a rare craniofacial disorder comprising core features of micrognathia, condyly dysplasia and question mark ear. Causative variants have been identified in PLCB4, GNAI3 and EDN1, which are predicted to function within the EDN1-EDNRB pathway during early pharyngeal arch patterning. To date, two GNAI3 variants in three families have been reported. Here we report three novel GNAI3 variants, homozygous with affected members in a family previously linked to 1p21.1-q23.3 and two de novo variants in simplex cases. Two variants occur in known functional motifs, the G1 and G4 boxes, and the third variant is one amino acid out of frame. We have not been able to identify any affected members. The variants are biallelic and cluster in a region involved in GTP/GDP binding. We hypothesize that all GNAI3 variants lead to dominant negative effects.

3107M
Mouse Model with Mutant Type I Collagen C-propeptide Cleavage Site has Brittle Bones and Increased Osteoblast Mineralization. A.M. Barnes, J.E. Perosky, M.H. Rajag, K.M. Koziofi, J.C. Marin. 1) BMBF, NICHD/NID, Bethesda, MD; 2) Orthopedic Surgery and Biomedical Engineering, University of Michigan, Ann Arbor, MI.

Classical dominant osteogenesis imperfecta (OI) is caused by mutations in type I collagen. Mutations in the C-propeptide cleavage site of both COL1A1 and COL1A2 were shown to cause high bone mass OI, characterized by bone hypermineralization and normal to increased DXA T-scores. We generated a mouse with a heterozygous C-propeptide cleavage site (high bone mass, HBM) in which the Ala-Asp residues of the pro(1)I C-propeptide cleavage site were mutated to Thr-Asn to prevent cleavage by BMP1. We utilized this mouse to investigate the role of pro(1)I C-propeptide processing in bone formation. Collagen from HBM calvarial osteoblasts (OB) had normal biochemistry and the collagen was incorporated normally into trimers. Conditioned media from HBM fibroblasts showed decreased processing of the C-propeptide. HBM OB deposited only about 50% of WT matrix. An in vitro mineralization assay showed that HBM calvarial OB had a 15% increase in mineralization (p=0.006), b > BMP2. Dermal fibril diameters were smaller and more homogeneous in HBM than WT, with a loss of large fibrils. HBM mice are growth deficient, remaining smaller than WT from 4-16 weeks. At 2 months, male HBM mice are significantly smaller in weight (77%) and length (92%) and have shorter femurs (22%). Femoral area BMD in HBM mice is decreased 25% (p=0.001); vertebral BMD is normal. All 2 month HBM mice have pelvic deformities, and 40% have kyphosis. On µCT, HBM femora have a thinner cortex with decreased cortical area; the distal femoral metaphysis is notably thinner. Cortical TMD is modestly reduced, while trabecular TMD, trabecular bone volume and trabecular number and thickness are decreased. Four point bending revealed an extremely brittle phenotype; post-yield displacement is only ~10% of WT (p<0.001). HBM femoral stiffness, yield load, and ultimate load are also significantly decreased. The bone phenotype of HBM mice is similar to Bmp1<sup>1-2</sup>/TibI<sup>1-2</sup> mice (Muir, et al. HMG, 2014, 23:3085-3101), in which cleavage of multiple procollagens as well as bone-related factors is defective. Bmp1<sup>1-2</sup>/TibI<sup>1-2</sup> mice have small size, thin cortices, reduced failure energy and maximum load and a dramatic defect in post-yield displacement. Our data suggest that the essential elements of the broader BMP1 enzyme deficiency are reproduced by a substrate defect in type I C-propeptide cleavage. These data show the importance of the pro(1)I C-propeptide to both collagenous and mineral properties of bone.

3108T

Recently a group of craniofacial and skeletal disorders caused by dominant mutations in spliceosomal genes has emerged. The first of these to be identified is mandibulofacial dysostosis type Guion-Almeida (MFDGA), caused by haploinsufficiency of the US snRNP component EFTUD2. We recently identified haploinsufficiency of SF3B4 as the cause of Nager syndrome, an acrofacial dysostosis (AFD). Mutation of SF3B4 was subsequently reported in AFD Rodriguez type (AFDR). SF3B4 is part of the U2 snRNP of the major spliceosome. Herein we also present heterozygous mutations in a highly conserved regulatory exon of SNRPB as the cause of cerebro-costo-mandibular syndrome (CCMS). Our experimental data show that these mutations disturb highly conserved exonic splicing silencer sequences crucial to the regulation of the gene’s expression. SNRPB encodes a component of the Sm complex of the spliceosome, which is present in all splicing reactions. Pierre Robin sequence and ear anomalies are present in all four syndromes. However, malar hypoplasia is present only in Nager, MFDGA and AFDR. Microcephaly is typically present only in MFDGA. Other disease-specific features include posterior rib gaps in CCMS, tracheoesophageal fistula in MFDGA, pre-axial anomalies in Nager and severe limb defects in AFDR. Together, these findings invite questions on the sensitivity of craniofacial and in particular pharyngeal arch development to spliceosomal defects, and on the role of the spliceosome in the regulation of development. We propose a model in which spliceosomal defects and alterations in splicing might cause group of disorders with a particular emphasis on the role of SF3B4, EFTUD2 and SNRPB in disorders of craniofacial and skeletal development.
Novel COL1A2 Gene Mutation in Czech Osteogenesis Imperfecta Patient: Case report. S. Mazurova1, L. Hruškova1, I. Marko2, P. Martasek1, I. Mazura1. 1) Department of Pediatrics and Adolescent Medicine, First Faculty of Medicine, Charles University in Prague, Ke Karlovu 2, 12100 Prague 2, Czech Republic; 2) Ambulant Centre for Defects of Locomotor Apparatus, Oštřanská 7, 13000 Prague 3.

Osteogenesis imperfecta (OI) is an inherited disorder of connective tissue. About 90% of affected patients with OI type I-IV show causative mutation in one of the two genes, COL1A1 and COL1A2, coding 2 alpha 1 chains and 1 alpha 2 chain of collagen type I protein. Clinical effect of these mutations results in mild to lethal phenotype regard on type of osteogenesis imperfecta and position of the nucleotide change. We continue with analyses of the Czech patients with OI type I-IV and COL1A1 and COL1A2 genes of which have already been described earlier. This paper is focused on one mutation of COL1A2 gene and correlation of genotype-phenotype parameters in Czech osteogenesis imperfecta patient. Mutation observed in the COL1A2 gene is localized in exon 40 and we describe clinical symptoms of patient correlating with Gly814Trp substitution. We used molecular genetic analysis of genomic DNA extracted from peripheral blood leucocyte. DNA quality and quantity were confirmed by spectrophotometry and horizontal electrophoresis and subsequently target amplification and two-directional sequencing was performed. We demonstrate a case of 25 year old woman affected by OI type III which is clinically followed from 3 years and 4 months of life. Last clinical examination was performed in 17 years of life. This woman was born from second gravidity of unaffected parents and perinatal fractures of both low extremities were described. This work was supported by following grants: grant no.: SVV-2014-260034 supported by Charles University Prague, grant no.: UNCE 204011 supported by Ministry of Education, Youth and Sport, Czech Republic, grant no.: PRVOUK P24/1LF/3 supported by Ministry of Education, Youth and Sport, Czech Republic.
Transcriptional Dysregulation Associated with Somatic Neurofibromin Deficiency in Tibial Pseudoarthrosis with Neurofibromatosis Type 1. J.J. Rios, N. Panja, T.J. Cho, H. Cho, N. Kamiya, K. Kayembe, R. Mao, R.L. Margraf, G. Obermosser, I. Oxendine, D.W. Sant, M.H. Song, D.A. Stevenson, D.H. Viskochil, C.A. Wise, HKW. Kim. 1) Research, Texas Scottish Rite Hospital for Children, Dallas, TX; 2) Division of Pediatric Orthopaedics, Seoul National University Children's Hospital, Seoul, Republic of Korea; 3) Baylor Institute for Immunology Research, Dallas, Texas; 4) ARUP Laboratories, Salt Lake City, Utah; 5) Department of Pediatrics, University of Utah, Salt Lake City, Utah; 6) Department of Orthopaedic Surgery, Jeju National University Hospital, Jeju, Republic of Korea.

NF1 haploinsufficiency causes neurofibromatosis type 1 (NF1) and results in constitutive Ras activation, predisposing individuals to neurofibroma and malignant peripheral nerve sheath tumor (MPNST) formation. Children with NF1 are also predisposed to tibial dysplasia with high risk of fracture, pseudoarthrosis and chronic nonunion. Previous studies described somatic loss-of-heterozygosity (LOH) at the NF1 locus in tibial pseudoarthrosis tissue. This study sought to confirm and expand this observation by characterizing the somatic mutation spectrum and transcriptional dysregulation in pseudoarthroses from 16 individuals with NF1. Genomic microarray analysis and whole-exome sequencing identified somatic variants in the NF1 gene in 12 samples, and unlike neurofibromas and MPNSTs, no other recurring somatic variants were identified. Somatic NF1 variants included a single large deletion, six individuals with sequence variants and five individuals with LOH encompassing the entire long-arm of chromosome 17. We did not identify somatic variants for four individuals. Whole-transcriptome sequencing (RNA-seq) was performed using cells cultured from pseudoarthrosis tissue from five individuals, and somatic variants were confirmed in the mosaic cell cultures. RNAseq analysis identified significant upregulation of a tumor-promoting transcriptional pattern in pseudoarthrosis samples, as compared to iliac crest cells from two affected individuals and control samples from three individuals without NF1. Upregulated genes included EGFR (2.5-fold, p=4.8e-4), KITLG/SCF (3.4-fold, p=1.9e-4) and EREG (55.7-fold, p=2.49e-11). Because NF1 haploinsufficiency results in activation of the ERK signaling pathway, we investigated whether pseudoarthrosis cells with neurofibromin deficiency had higher levels of ERK activation compared to haploinsufficient iliac crest cells. Flow cytometry identified no difference in the levels of ERK activation, suggesting upregulation is associated with neurofibromin deficiency rather than quantitative differences in pathway activation. Despite the transcriptional dysregulation, NF1 pseudoarthroses lacked other obvious oncogenic variants, and tumor formation from pseudoarthrosis tissue was not observed nor previously reported in the literature. Results from this study suggest receptor tyrosine kinase inhibitors should be investigated as a non-surgical treatment for tibial pseudoarthrosis to promote bone union in individuals with NF1.
3113M

Rare TBX6 null mutations in congenital scoliosis. N. Wu1, Z. Wu1, X. Ming1, J. Xiao2, X. Chen3, J. Liu4, S. Liu5, Y. Ming2, L. Jin1, X. Zhang1, G. Ooi1, F. Zhang3 & M. Department of Orthopedic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China; 2) State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China; 3) Tufts Children’s Hospital, Boston, MA, USA; 4) Shanghai Children’s Medical Center, Shanghai Children’s Medical Center, Tongji University School of Medicine, Shanghai, China; 5) Public Health Research Institute, New York, USA.

Introduction: Congenital scoliosis (CS) is a common disorder with an estimated incidence of around 0.5-1 in 1000 live births. Genetic mutations have been implicated in CS pathogenesis, but the genetic basis of sporadic CS remains unclear. Human 16p11.2 deletions are rare but recurrent mutations. The CS conditions have been recently observed in a small portion of 16p11.2 deletion carriers, suggesting a potential involvement of 16p11.2 in CS. Methods: Subject Information. We enrolled 161 unrelated sporadic CS patients of Han Chinese from PUMCH between October 2010 and November 2012. Their CS conditions were confirmed by radiological imaging. The patients with known syndromes were excluded. A total of 166 unrelated healthy participants were used as population controls. High-density Oligonucleotide Comparative Genomic Hybridization Microarray. A high-density genome-wide copy number analysis (CGH) array analysis was performed in 20 trios. The genomic DNA was extracted from peripheral blood leukocytes. Quantitative Polymerase Chain Reaction (qPCR) Analysis. qPCR analysis was conducted to screen for 16p11.2 TBX6 deletion in additional 141 sporadic CS patients. Any 16p11.2 deletion candidate suggested by qPCR was further confirmed by the CGH microarrays. DNA Sequencing of the TBX6 gene. The entire TBX6 gene and its approximately 1-kb upstream region were amplified using long-range PCR in both CS patients and healthy controls. The PCR products were sequenced by Sanger sequencing. The Fisher’s exact test (two-sided) was used to investigate the potential significant difference in TBX6 mutation/variant frequencies between the CS patients and controls. The differences with P<0.05 were considered as statistically significant. Results: Thirty-six de novo recurrent deletions in proximal 16p11.2 in two patients (2/20). Furthermore, both deletions were de novo. Notably, ten more sporadic CS patients were found to have this recurrent 16p11.2 deletion. However, we did not observe any 16p11.2 deletion in 166 Han Chinese controls (12/161 vs. 0/166; P=0.0002). The DNA sequencing of the TBX6 gene revealed additional one nonsense and four frameshift mutations of TBX6 in these group of CS patients. Conclusions: Up to our knowledge, we reported the largest genetic study of sporadic CS, suggesting that the TBX6 null mutation is responsible for CS (approximately 10% of sporadic CS). Our findings will facilitate molecular diagnostics of CS.

3115S


Diacylglycerol O-acyltransferase 1 (DGAT1) catalyzes the transfer of fatty acid-CoA to diacylglycerol in the terminal step in triglyceride synthesis. DGAT1 mice exhibit increased longevity, decreased body fat, resistance to diet-induced obesity and a failure to lactate. The absence of detectable negative effects outside the mammary gland in Dgat1-deficient mice, and the enzyme’s role in human storage lipid synthesis, have made pharmacological inhibition of DGAT1 an attractive approach to the management of obesity. Interestingly, a family has been reported with a contrary phenotype of growth retardation and diarrhea due to homozygous for a splice site mutation in DGAT1, abolishing the production of active enzyme [Haas et al., 2012, PMID: 23114594]. In an endeavour to select for altered milk production in dairy cows we have discovered 2.5 million cows and identified an extremely unusual animal producing milk with a 40% reduction of saturated fat (4 SD below breed mean). To identify the genetic basis of this extreme phenotype, we bred a three-generation pedigree from the unique founder. Linkage mapping and genome sequencing revealed a de novo heterozygous A>C transversion in exon 16 of the DGAT1 gene (g.8078>C; p.M435L). The mutation caused skipping of exon 16 in 90-95% of mutant mRNA molecules, and the mutant enzyme was unable to transfer fatty acids from CoA to diacylglycerol. Growth retardation, weight, and life span of heterozygous cows were normal. However, calves homozygous for the mutation developed severe, non-bloody diarrhea within 2-3 days after being apparently born healthy and without visible growth defects. Mutant calves of similar weights. Our discovery of the mutation in animals supports the identification of the causal mutation in the family above and a potential role for DGAT1 in malabsorption and diarrhoeal disorders. Finally, our observations underscore concerns for DGAT1 inhibition as a strategy to combat obesity in humans, and suggest the use of a large animal model to understand the role of DGAT1 in the intestine.

3116M

Hennekam syndrome can be caused by FAT4 mutations and be allelic to Van Melderg syndrome. M. Alders1, L. Al-Gazali1, I. Correio3, B. Dallapiccola2, L. Garavello2, B. Tuyisuz4, F. Salehi1, M. Haagmans1, O. Mook1, C. Majoc1, M. Mannens1, R. Hennekam1, 1) Academic Medical Center, Amsterdam, The Netherlands; 2) United Arab Emirates University, Al-Ain, United Arab Emirates; 3) Hospital Santa Maria, Lisboa, Portugal; 4) Arcipielaga S. Maria Nuova, Reggio Emilia, Italy; 5) Cerrahpasa Medical School, Istanbul University, Istanbul, Turkey.

The Hennekam lymphangiectasia-lymphedema syndrome is a genetically heterogeneous disorder. It can be caused by mutations in CCBE1 which are found in approximately 25% of cases. We used homozygosity mapping to identify mutations in a family with multiple affected individuals in whom no CCBE1 mutation had been detected, and identified a homozygous mutation in the FAT4 gene. Subsequent targeted mutation analysis of FAT4 in a cohort of 24 CCBE1 mutation negative Hennekam syndrome patients identified homozygous or compound heterozygous mutations in 14 additional families. Mutations in FAT4 have been previously associated with Van Melderg syndrome. Detailed clinical comparison between van Melderg and Hennekam syndrome patients shows that there is a substantial overlap in phenotype, especially in facial appearance. We conclude that Hennekam syndrome can be caused by mutations in FAT4 and be allelic to Van Melderg syndrome.
3117T
Novel variant of TNNI2 causes an atypical Distal Arthrogryposis syndrome. C.T. Marvin1, J.X. Chong1, K.J. Buckingham1, K.M. Shively1, A.E. Beck2, H.I.S. Gildersleeve1, M.J. McMillin1, D.A. Nickerson2, J. Shendure2, R.E. Stevenson1, M.J. Bamshad1,2,3, University of Washington Center for Mendelian Genomics. 1) Dept. of Pediatrics, University of Washington, Seattle, WA, USA; 2) Seattle Children's Hospital, Seattle, WA, USA; 3) Dept. of Genome Sciences, University of Washington, Seattle, WA, USA; 4) J.C. Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, SC, USA.

Distal arthrogryposis (DA) type 2B, or Sheldon-Hall syndrome, can be caused by pathogenic variants in any one of several genes that encode proteins of the contractile apparatus of skeletal muscles. Causal variants in four genes (MYH3, TNNI2, TPM2 and TNNT3) each explain approximately 10% of families, so the genetic basis remains unknown in most circumstances. We were referred a single affected male who presented as a newborn with a history of reduced fetal movement accompanied by polyhydramnios, severe contractures of the upper and lower limbs, multiple dysomorphic facial characteristics, hyperelastic skin, multiple rib and long-bone fractures, and hypotonia born to apparently unaffected parents. Despite an exhaustive clinical evaluation, no specific diagnosis could be confirmed. However, evaluation by a group of expert dysmorphologists considered atypical DAB2 to be one possible diagnosis. We screened exons of each gene known to underlie DAB2 including exon 8 of TNNI2, exon 10 of TNNT3, and all coding exons of MYH3 and TNNI2 without finding a compelling candidate variant. Subsequently, we performed exome sequencing of the proband and his parents and identified a de novo missense variant (c.337C>T, predicted p.Arg113Trp) in exon 7 of TNNI2. To date, variants causing DAB2 have been reported only in exon 8 of TNNI2. Additionally, we identified a rare (MAF 0.024%) in-frame deletion (c.272_274del, predicted p.Glu92del) in exon 6 of TNNI2. This variant was also found in the mother of the proband who in retrospect was noted to also have somewhat hyperelastic skin. No candidate functional variants were found in genes potentially associated with hyperelastic skin. These results suggest that the atypical DA phenotype and in particular, the hyperelastic skin, may be caused by the occurrence of more than one disruptive variant in TNNI2, either on the same allele or affecting both alleles. Exome sequencing of additional atypical DAB2 cases is underway to determine whether causal variants exist in regions other than those encoded by exon 8 and whether the phenotype in some individuals is caused by multiple disruptive variants in TNNI2.

3118S

The mutations in metabolic pathway has been observed to be associated with higher birth frequency of Down syndrome. The most common mutations include MTHFR C677T (1p36.3) and CBS 844ins68 (21q22.3). MTHFR helps in converting homocystein to methionine while cystathionine beta synthase (CBS) helps in converting homocystein and serine to cystathionine thus regulating the homocystein concentration. Mutations in MTHFR and CBS gene impair the activity of enzymes, leading to increased homocystein concentration, which has been reported to be the possible risk factor for Down syndrome. The present study investigated 126 mothers of DS children and 58 mothers having normal children for CBS 844ins68 polymorphism whereas 110 case mothers and 111 control mothers for MTHFR C677T polymorphism. Peripheral blood from mothers was collected in EDTA-coated vacutainer for CBS 844ins68, while DNA was extracted from peripheral blood of DS children and 58 mothers having normal children for CBS 844ins68. Peripheral blood from mothers was collected in EDTA-coated vacutainer for MTHFR C677T. The DNA was extracted from peripheral blood of DS children and 58 normal children. The genotyping was done using polymerase chain reaction. Results were electrophoresed in 2% agarose gel using ethidium bromide. In CBS 844ins68, presence of 68bp insertion produced 242bp band while wild type produced 174bp band. Using polymerase chain reaction. Products were electrophoresed in 2% agarose gel using ethidium bromide. In CBS 844ins68, presence of 68bp insertion produced 242bp band while wild type produced 174bp band. Among cases, 88.9% (112) were having -/- genotype and 11.1% (14) were having +/- genotype. However, +/- genotype was not observed among any case or control mothers. The chi square value (χ², 0.065, p<0.79) suggested non-significant association between birth of DS child and presence of CBS 844ins68 polymorphism. In case of MTHFR C677T polymorphism, homozygous wild genotype (CC) was observed in 78.2%, heterozygous mutant (CT) in 20% and homozygous mutant (TT) in 1.8% mothers of DS children. Among cases and controls, the chi square value (χ², 2.064) revealed non-significant association between cases and controls. We are working on larger sample size to understand the complex mechanism behind the presence of maternal polymorphisms and maternal risk of having DS child.

3119M
Investigation of a missense in NOTCH4 in autosomal dominant scleroderma. C.J. Cardinale1, D. Liu1, P.M. Chiavacci2, L. Tian1, S.J. Bumham3, H. Hakonarson4. 1) Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, PA; 2) Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA.

We encountered an 8-year-old girl presenting with systemic sclerosis (SSc), i.e., scleroderma, and a positive family history. This patient had an affected maternal aunt and a maternal grandfather who died of SSc. The pedigree suggests an autosomal dominant inheritance model with incomplete penetrance. We performed whole-exome sequencing on the proband and aunt as well as the unaffected maternal grandmother. We examined missense, nonsense, splice-altering, and coding indels matching the inheritance model. Results were filtered to exclude synonymous variants, variants with minor allele frequency (MAF) greater than 0.5%, and variants previously identified in controls by our in-house exome variant database. Relevant candidates were taken forward for manual curation. A p.Met1415Ile substitution in NOTCH4, previously implicated in the pathways by which TGF-β induces pulmonary fibrosis—one of the most severe clinical manifestations of SSc—was identified as the most likely candidate. We have cloned an expressed AcGFP1-tagged clone of full-length NOTCH4 in HEK 293T cells and examined the impact of the mutation through immunoprecipitation and luciferase reporter assays.

3120T
Molecular Spectrum of Mutations in CFTR gene: First Report from the Aegean region of Turkey and definition of three novel mutations. A. Aykut1, H. Onay1, I.M. Tekin1, F. Gulen2, F. Ozkinay1. 1) Ege University Faculty of Medicine, Department Medical Genetics, Izmir, Turkey; 2) Ege University Faculty of Medicine, Department of Pediatric Allergy and Respiratory Diseases, Izmir, Turkey.

Cystic fibrosis (CF) is an autosomal recessive disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Up to date, more than 1900 different mutations in the CFTR gene have been reported. The aim of this study was to evaluate the spectrum of CFTR gene mutations in the Aegean region of Turkey. In this study, a total of 251 mutated CF alleles were identified in 194 individuals referred to our center for CFTR molecular analysis between January 2005 and May 2014. Among 251 mutant CF alleles, 55 different mutations, 3 of them being novel, were detected. The most common five mutations were F508del (32.8%), I148T (8%), 2183AA>G (7%), N1303K (4.7%) and G542X (4.7%). The novel mutations identified were p. S101del, c.339delG and p.R128G. The frequencies of CFTR mutations found in our region were slightly different from the frequencies found in general Turkish population.

Posters: Molecular Basis of Mendelian Disorders

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3121S
Molecular Analysis of Dystrophic Epidermolysis Bullosa in Iran. H. Vahidnezhad1-3, L. Youssefi2, M. Barzgar1, S. Lotoureh1, O. Li4, N. Mozaffari1, A. Ismail1, M. Daneshyazhnoh2, M. Tabrizi5, S. Zeinali1, J. Uitto6. 1) Biotechnology Research Center, Department of Molecular Medicine, Pasteur Institute of Iran, Tehran, Iran; 2) Tehran University of Medical Sciences, Tehran, Iran; 3) Department of Dermatology and Cutaneous Biology, Thomas Jefferson University, Philadelphia, Pennsylvania, USA; 4) Shahid Beheshti University of Medical Sciences, Tehran, Iran; 4) Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Epidermolysis bullosa (EB) is a clinically and genetically heterogeneous group of genodermatoses characterized by trauma induced mucocutaneous blistering, fragile skin, and devastating consequences for the patient and family. Many EB loci have previously been reported in European and US families. In our study, we have observed to be modifiers of a different CF phenotype, SLC6A14 and SLC9A3 solute carrier genes. The international consensus classification defines four types of EB based on the ultrastructural level of tissue separation: EB simplex (EBS), junctional EB (JEB), dystrophic EB (DEB), and the Kindler syndrome. Numerous studies, mainly performed in European and US families, with EB, have revealed a number of characteristic epidemiological and genetic features, which form the basis for current diagnostic and counseling strategies. However, there is currently no comprehensive study about the molecular epidemiology of EB in Iranian populations. In this study, we registered 250 EB families and here we report the evaluation of 150 EB families with ~200 affected individuals in Iran. Clinical examination and immunofluorescence mapping showed that EBS, JEB, and DEB represented 13, 15, and 66% of the cohort, respectively. Due to high rate of consanguineous marriages this ratio is different from EB patients in Europe and the USA in which EBS, JEB, and DEB were reported to represent 67, 27, and 35%, respectively. We have assessed a cohort of 67 DEB families for pathogenic sequence alterations in COL7A1 gene. Our results from 13 families showed recurrent novel mutations, including p.Gly2040Val, p.Arg2063Trp, p.Gly2031Ser, c.6265delC, c.4958-4959delAA and IVS84+2T>G (c.6714+2T>G). Our data indicate the need for population-specific diagnostic and management approaches. Identification of specific mutations in the candidate genes and elucidation of the consequences of such mutations can provide a basis for the development of novel therapeutic approaches, including gene therapy, protein replacement, or cell based therapies, some of which are in early clinical trials for EB. Finally, DNA-based prenatal testing and preimplantation genetic diagnosis can be applied to families at risk for recurrence of this, currently intractable disease.

3122M
Genome-wide association meta-analysis of 6,365 subjects replicates EHF-APIP and identifies new modifier loci of lung disease severity in Cystic Fibrosis. H. Corvol1, G.R. Cutting2, M.L. Drumm3, M.R. Knowles4, J.M. Rommens5 on behalf of the International CF Gene Modifier Consortium. 1) AP-HP, Hôpital Trousseau, Pediatric Pulmonary Department; Inserm U938, Paris, France; 2) Institute of Genetic Medicine, Johns Hopkins University, Baltimore MD; 3) Dept of Pediatrics, Case Western Reserve University, Cleveland, OH; 4) Dept of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC; 5) Program in Genetics & Genome Biology, SickKids Research Institute and Univ of Toronto, Toronto, ON, Canada.

Lung disease is responsible for 80% of mortality in cystic fibrosis (CF). Genetic modifiers play a prominent role in variation in lung function measurements (H2~0.54). The North American Cystic Fibrosis (CF) Modifier Consortium had previously reported genome-wide significant association with lung function measures for SNPs positioned between APIP and EHF on chromosome 11p13 (GWAS1; n = 3,444; Wright et al, Nat Genetics, 2011). We now report replication of the EHF-APIP locus and identification of new modifier loci of lung disease severity in CF. Genome-wide SNP typing was performed on a second cohort of CF subjects from North America and the French CF Gene Modifier Consortium (GWAS2; n=2,921). Imputation using an admixed 1000 Genomes reference population generated a common set of 8,520,458 SNPs that were screened for association using linear mixed effects modeling with meta-analysis across patient groups and genotype platforms. Support for the chromosome 11p13 region was evident in a meta-analysis of GWAS1 and GWAS 2 (P = 4.8 x 10^-9; GWAS1+2) and in the subgroup homogenous for the common CF mutation (P = 1.9 x 10^-10; GWAS1+2 p.Phe508del p.Phe508del). Regions on chromosomes 5, 6 and X that previously indicated suggestive evidence for associations (GWAS1) met the genome-wide significance criterion (P = 6.8 x 10^-12, 1.7 x 10^-9 and 1.8 x 10^-9, respectively; GWAS1+2) in the meta-analysis. Two new loci showed significant evidence for association, chromosome 3 (P = 3.3 x 10^-11; GWAS1+2) and chromosome 17 (P = 1.8 x 10^-7, GWAS2; and P = 2.4 x 10^-9, GWAS2 p.Phe508del/p.Phe508del). The intervals of the six identified loci that contribute to lung function in CF patients suggest a series of biologically candidate genes that include EHF and APIP, the solute carrier SLC9A3, the muciones MUC4 and MUC20, class II MHC genes, the IncRNA CASC17 and an angiotensin receptor ATR2 that is proximal to another solute carrier SLC6A14. Pleiotropic effects are evident as SLC9A3 and SLC6A14 have been observed to be modifiers of a different CF phenotype, neonatal intestinal obstruction. Together, these loci provide new molecular targets for therapeutic intervention of the major cause of mortality in CF.

3123T
A Homozygous NIPAL4 Mutation In A Case With Ichthyosis And Deafness. E. Arslan Ates1, H. Onay2, I. Ertam2, I. Uitto3, N. E. Arslan Ates1. 1) Dermatology, Ege University Faculty of Medicine, Izmir, Turkey; 2) Medical Genetics, Ege University Faculty of Medicine, Izmir, Turkey; 3) Dermatology, Ege University Faculty of Medicine, Izmir, Turkey.

Ichthyosis-Deafness (KID; OMIM ID#148210) syndrome is one of the syndromic ichthyosis disorders. Its cardinal features are ichthyosis, sensorineural deafness and ophthalmological findings. Heterozygous Connnexin 26 (GJB2) gene mutations have been described as the underlying defects of KID. Autosomal Recessive Congenital Ichthyosis (ARI) is a nonsyndromic ichthyosis form, without any systemic involvement. TMG1, NIPAL4, ALOXE3, CYP4F22, ABCA12, PNPLA1 and LIPN genes have been described associated with ARCI. Here, we present a 6 year-old girl with ichthyosis and deafness. She was born to consanguineous parents. Scaling and redness of skin were recognised at birth. On physical examination she had typical skin findings of ichthyosis. Histopathological evaluation of skin biopsy was compatible with vulgar ichthyosis. She had sensorineural deafness. Ophthalmologic examination was normal. Because she was considered to have KID Syndrome, CX26 gene was sequenced and no mutation was found. Subsequently, GJB1 gene which have been known to be the most common cause of ARCI has been sequenced but no mutation was found in this gene as well. The second most common cause of ARCI are the defects found in NIPAL4 gene. Therefore NIPAL 4 gene was sequenced as the third gene and a homozygous c.527C>A (p.A176D) mutation was identified. To date there have been no reported ARCI case having deafness and NIPAL4 mutations in the literature. Although there may be a coincidence in this patient, NIPAL4 mutation analysis should be considered in patients with KID syndrome or in patients having both ichthyosis and deafness.
3124S
The ALK1IVS3-35A>G polymorphism is associated with arteriovenous malformations in hereditary hemorrhagic telangiectasia patients with ENG mutations, but not in patients with ALK1 mutations. L. Pawlikowski1,2, J. Nelson1, D.E. Guo1, C.E. McCulloch3, M.T. Lawton1, W.L. Young1,2, H. Kim1,2,3, M.E. Faughnan4,5,6, Brain Vascular Malformation Consortium HHT Investigator Group, 1) Center for Cerebrovascular Research, Dept of Anesthesia and Perioperative Care, University of California, San Francisco, San Francisco, CA; 2) Institute for Human Genetics, University of California, San Francisco, San Francisco, CA; 3) Dept of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA; 4) Dept of Neurological Surgery, University of California, San Francisco, San Francisco, CA; 5) Division of Respiratory, Department of Medicine, University of Toronto, Toronto, Canada; 6) Division of Respiratory, Dept of Medicine and Li Ka Shing Knowledge Institute, St. Michael’s Hospital, Toronto, Canada.

Background. Hereditary hemorrhagic telangiectasia (HHT) is caused by mutations in several TGFβ signaling pathway genes. Most HHT patients have a mutation in ENG or ALK1 (ACVR1) and have diverse phenotypes including chronic bleeding, hemorrhage, stroke, heart failure and other complications related to skin and mucosal telangiectases and organ arteriovenous malformations (AVM). The phenotypic heterogeneity of HHT suggests a possible role for genetic modifier effects. The common polymorphisms ALK1 IVS3-35A>G and ENG 207G>A have been reported to be associated with sporadic brain AVM (BAVM); the ALK1 IVS3-35A>G association has been replicated in 2 independent cohorts. We hypothesized that these polymorphisms are also associated with AVM in HHT patients. Methods. We genotyped ALK1 IVS3-35A>G and ENG 207G>A in 716 HHT patients (95.9% Caucasian) recruited by the Brain Vascular Malformation Consortium. We evaluated association of genotype with AVM, defined as presence of any BAVM, liver AVM (LAVM) or pulmonary AVM (PAVM), and separately with each AVM type. Logistic regression analyses were adjusted for age, gender and family structure and further stratified by HHT mutation status (ALK1 or ENG). Results. Among HHT patients with complete clinical data, 71% had at least one AVM, 50% had PAVM, 20% had BAVM and 20% had LAVM. Among 436 (61%) patients with mutation status available, 48% had ENG mutations, 18% had ALK1 mutations, 30% had no ALK1 or ENG mutations. Among all 716 patients, neither polymorphism was significantly associated with AVM, although ALK1 IVS3-35A>G showed a trend toward association with PAVM (OR=1.48, 95%CI=0.90-2.22, p=0.062). When stratifying by HHT mutation, ALK1 IVS3-35A>G was significantly associated with AVM among ENG mutation carriers (OR=2.66, 95%CI=1.15-6.13, p=0.022), but not among ALK1 mutation carriers (OR=0.79, CI=0.38-1.63, p=0.52). ALK1 IVS3-35A>G was also significantly associated with PAVM (OR=2.45, p=0.016) and LAVM (p<0.05, all 42 LAVM-positive patients carried the A risk allele) among ENG mutation carriers. There were no significant associations between ENG 207G>A and AVM, but a trend toward association with BAVM (OR=3.09, p=0.16) was observed among ALK1 mutation carriers. Conclusions. ALK1 IVS3-35A>G was associated with AVM in HHT patients bearing ENG but not ALK1 mutations. These results suggest that common polymorphisms in TGFβ pathway genes other than the mutated HHT gene act as genetic modifiers and contribute to HHT phenotypic heterogeneity.

3125M
Neutral Lipid Storage Disease with Myopathy: disease modeling using patients’ hiPSC, D.A. Coviello1, S. Missaglia1, M. Castagnetta1, E.M. Pennisi2, M. Mogni3, D. Tavian2,4, 1) Laboratory of Human Genetics, E.O. Ospedali Galliera, Genoa, GE, Italy; 2) Laboratory of Cellular Biochemistry and Molecular Biology, CRIBENS, Catholic University of the Sacred Heart, Milan, MI, Italy; 3) UOC Neurologia, A.C.O. San Filippo Neri, Rome, RM, Italy; 4) Psychology Department, Catholic University of the Sacred Heart, Milan, MI, Italy.

Mutations in the PNPLA2 gene cause the onset of Neutral Lipid Storage Disease with Myopathy (NLSD-M), a rare autosomal recessive disorder characterized by an abnormal accumulation of triacylglycerol into cytoplasmic lipid droplets (LDs). In most tissues the LDs are cellular organelles for the triacylglycerol storage. LDs metabolic functions are mediated by proteins bound to their surface. In particular, the lipase that catalyzes the removal of the first acyl chain from triacylglycerol is the adipose triglyceride lipase (ATGL), also known as patatin-like phospholipase domain-containing protein 2 (PNPLA2). To our best knowledge, twenty six different PNPLA2 mutations have been described in thirty two NLSD-M patients. NLS-D-M patients are mainly affected by progressive myopathy, cardiomypathy and hepatomegaly. However, their clinical severity appears to be highly variable. Other clinical symptoms may include diabetes, chronic pancreatitis and short stature. NLSD-M has, at present, no specific therapy. We have previously reported clinical and genetic findings of some NLSD-M patients obtaining dermal biopsies from them. Here we report the development of hiPSC (human induced pluripotent stem cell) from patients’ fibroblasts harboring different PNPLA2 mutations. The first patient was found to be homozygous for a deletion at nucleotide 542 (c.542delAG). This deletion caused a premature stop codon at position 212. The molecular analysis of patient 2 showed a missense mutation in PNPLA2 (c.662G>C (p.R221P)). Initial hiPSC colony selection was based on morphologic evaluation and on detection of pluripotency surface markers (SSEA-4 and TRA-1-81). hiPSC also expressed undifferentiated ES cell markers (NANOG, SOX2 and OCT4). Moreover, embryoid bodies (EBs) have been generated from NLSD-M-iPSCs to assess the pluripotent properties of these cells. Karyotypic analysis of hiPSC lines indicated a normal complement of chromosomes. Immunohistochemical evaluations of LDs have revealed that they recapitulate pathological hallmark of the disease. We propose use of inherently patients- and disease specific hiPSC to study the pathogenetic mechanisms leading to NLSD-M and as a potential model for therapeutic evaluation. Communications to: daniela.tavian@unicatt.it.
3126T

British Ectopia Lentis (EL) patients with novel ADAMTSL4 mutations: 2 homozygotes, 1 compound heterozygote & 1 compound heterozygote involving a splice site - J.A. Aragon-Martín1, V.E.A. Wan1, L. Collins3, P. Mazadza1, G. Armo2, A. Chandra2, D.G. Charteris1, A. Saggar1, A.H. Child1. 1) Cardiovascular Sciences Research Centre, St George’s University of London, London, England, United Kingdom; 2) Institute of Ophthalmology, University College London, London, United Kingdom; 3) Vitreoretinal Unit, Royal Victorian Eye and Ear Hospital, Melbourne, Australia; 4) Vitreoretinal Outcomes of EL, A. Bennassar1,2 and first for appar- 3127S

Xerodema Pigmentosum diagnosed in adulthood: atypical clinical presen- 3128M

Mutation spectrum of the ABCA4 Gene in Greek patients with Stargardt disease: Identification of two novel mutations and evidence of three prevalent mutated alleles - S. Kamakari1,2, G. Koutsodontis1, P. Stamatiou1, T. Panagiotoglou1, I. Datseris1, M. Tsilimbaris3. 1) OPHTHALMIC GENETICS UNIT, OMMA OPHTHALMOLOGICAL INSTITUTE OF ATHENS, ATHENS, GREECE; 2) SCHOOL OF MEDICINE, UNI- Versity of Athens, Greece; 3) DEPARTMENT OF OPHTHALMOLOGY, SCHOOL OF MEDICINE, IRAKLION, GREECE.

Purpose: Mutations in the ABCA4, photoreceptor-specific ATP-binding cassette transporter 4 gene (MIM 601691) have been associated with auto- somal recessive Stargardt disease (STGD; MIM 248200) characterized by central vision impairment. This is the first systematic study investigating the frequency and pattern of disease-associated mutations among Greek STGD patients, never previously genetically characterized for sequence variations in the ABCA4 gene. Methods: 34 unrelated and 2 related STGD patients were analyzed using the combined methodology of ABCR000 microarray analysis and direct sequencing of 4 selected or all 50 exons and flanking intronic regions of the ABCA4 gene. Results: Disease-associated mutations, including two novel splice defects, c.4352+1G>A and c.5714+1G>C in introns 29 and 40, respectively, were detected in 28/34 unrelated patients analyzed (82.35%). Both mutant alleles were found in 16/28 cases (57.14%), whereas in 12/28 (42.85%) only one allele was identified. The major patho- genic allele c.5714+5G>A accounted for 28.57% of the mutant alleles. Other frequently mutant alleles were c.1622T>C (Leu541Pro) alone or as a complex c.1622T>C, c.3113C>T / (Leu541Pro), p.(Ala1038Val) and c.5882G>A/ (p.(Gly1961Glu), each accounting for 23.8% of the mutated alleles. Notably, a young patient with unaffected parents whose paternal mutation c.1622T>C / (Leu541Pro) was found homozygous for the p.(Leu541Pro) (Ala1038Val) and compound heterozygous for the p.(Leu541Pro)/ (Ala1038Val) and (Gly1961Glu) alleles, respectively. Conclusions: This first report of the ABCA4 mutation spectrum underlying STGD disease in Greece further elucidates the distribution of ABCA4 mutations in European populations. The detection rate by the combined methodology was 82.35%. We identified two novel, potentially pathogenic ABCA4 splice mutations and two novel disease-causing alleles. Further evaluation of the genotype-phenotype correlations will advance our knowledge of STGD’s etiology in Greek patients.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3129S  
NLZ1 is required for cilia formation in zebrafish embryogenesis. S. Dutta1, S. Sriskanda1, E. Boobalan2, R. Alur1, M. Gunay-Aydin2, A. Ekhadlioum3, M. Asai-Coakwell1, O. Lehmann4, F. Valente5, A. Micalizzi5, J. Landef1, M. Romani5, S. Ware5, N. Katsanis6, B. Brooks1. 1) Unit on Pediatric, Developmental & Genetic Ophthalmology, Ophthalmic Genetics and Visual Function Branch, National Eye Institute; 2) Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA; 3) Microarray Core, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA; 4) Department of Ophthalmology, University of Alberta, Edmonton, AB, Canada; 5) IRCCS Casa Sollievo della Sofferenza, Mendel Laboratory San Giovanni Rotondo, Italy; 6) Cincinnati Children’s Hospital, Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA; 7) University School of Medicine, Wells Center for Pediatric Research, Indianapolis, IN; 8) Department of Cell Biology, Center for Human Development Modeling, Duke University, Durham, NC.

The formation of cilia is a fundamental developmental process affecting diverse functions such as intracellular signaling, mechanotransduction, tissue morphogenesis and body patterning. Defects in ciliogenesis result in a diverse spectrum of human diseases, known as ciliopathies. However, the precise mechanisms of ciliogenesis during vertebrate development are not fully understood. In this report we describe a novel role of the NLZ1 protein in ciliogenesis. We demonstrate morpho-function-related knockdown of NLZ1 in zebrafish caused abnormal specification of the cells of Kupffer’s vesicle (KV, similar to the mammalian node); a severe reduction of the number of cilia in KV, the pronephros, and the neural floorplate; as well as a spectrum of later phenotypes reminiscent of human ciliopathies. NLZ1 is expressed in KV and in the nucleus of dividing cells. In vivo and in vitro data indicate that NLZ1 acts downstream of the ciliary “master transcription factor”, Fox1a and, NLZ1, is upregulated by canonical Wnt signaling. Mutation screening of NLZ1 in humans with a variety of ciliopathies and left-right patterning defects are underway. Together, our data suggest a novel role of NLZ1 in ciliogenesis and the morphogenesis of multiple tissues.

3130S 
A primary ciliopathy protein plays an extra-ciliary role in neurodevelopmental disease. N. Nuangchamnong1, C.S. Carter2,3, Q. Zhang2,3, T. Vogel4, V.C. Sheffield2,3. 1) Obstetrics and Gynecology, University of Iowa, Iowa City, IA; 2) Pediatrics, University of Iowa, Iowa City, IA; 3) Howard Hughes Medical Institute, University of Iowa, Iowa City, IA; 4) Neurosurgery, University of Cincinnati, Cincinnati, OH.

Hydrocephalus affects 1 to 3 of every 1000 live births and is a significant cause of perinatal and pediatric morbidity and mortality despite current surgical interventions. The majority of congenital cases are idiopathic, but hydrocephalus is increasingly considered to be a neurodevelopmental disorder. Bardet-Biedl Syndrome (BBS) is an autosomal recessive human ciliopathy in which neurological defects are common findings, and in which ventriculomegaly occurs at a higher incidence than in the general population. Primary cilia are now recognized for playing a critical role in development of the central nervous system. We have previously shown that a mouse model with a mutated component of the BBSome complex (BBS1) has aberrant cell signaling at primary cilia that causes decreased survival and proliferation of a specific class of neural progenitor cells and leads to mild communicating neonatal hydrocephalus that can be partially rescued with lithium. Further exploration of the molecular mechanisms underlying hydrocephalus has the potential to make noninvasive pharmacologic treatments possible. In this study, we use a hydrocephalic mouse model that lacks BBS3, a GTPase that is necessary for localization of the BBSome to the ciliary gate, but that, unlike BBS1, is not a component of the BBSome complex. We compare the development of hydrocephalus in the BBS3 knockout mouse to wild type littermate controls and to mice with a mutated component of the BBSome complex (such as BBS1). We find late fetal onset of hydrocephalus in the BBS3 knockout mouse model with ventriculomegaly as early as embryonic day 17 (E17). We also find decreased apoptosis along the cerebellar ventricular zone and increased apoptosis at the midbrain ventricle overlying the Sylvian aqueduct of BBS3 knockout mice at E17. Similar to the BBS1 mouse model, there is increased apoptosis and decreased proliferation of cells in the subventricular zone of the lateral ventricles in the neonatal period. Unlike the BBS1 model, however, there is fetal rather than neonatal onset of ventricular dilation and postnatal progression to severe disease that is consistent with a non-communicating (obstructive) type of hydrocephalus. The knockout BBS3 mouse is a useful model of obstructive congenital hydrocephalus, as BBS3 may have an extra-ciliary function that prevents progression of hydrocephalus from mild to severe.

3131S 
Congenital heart disease associated to PCD. J. Wallmeier1, H. Obrich2, P. Pennekamp3, C. Edelbusch1, N.T. Loges1, C. Werner1, D. Kececioglu4, B. Stiller1, H. Omran1. 1) Department of Pediatrics, University Hospital Muenster, Muenster, Germany; 2) Clinic fuer Angeborene Herzfehler; Herz- und Diabeteszentrum NRW; Georgstrasse 11; 32545 Bad Oeynhausen,Germany; 3) Department of Pediatric Cardiology, University Hospital Freiburg, Freiburg, Germany.

Primary ciliary dyskinesia (PCD) is a rare genetic disorder, that is characterized by recurrent infections of the upper and lower respiratory airways, sinusitis and otitis media, and is often associated with reduced fertility of males. In 50% of PCD-patients laterality defects occur, mostly situs inversus totalis (Kartagener’s syndrome). About 6% of all patients show heterotaxy. There are certain organ abnormalities that are associated with heterotaxy such as asplenia and polysplenia. An important consequence associated with heterotaxy are complex cardiac defects. Dysfunction of motile cilia and flagella explains the complex PCD phenotype involving various organ systems. Respiratory epithelia covered with cilia move extracellular mucus towards the throat with a coordinated beating (mucociliary clearance). Loss of this mucociliary clearance causes recurrent respiratory infections. Dysfunction of nodal cilia during early embryogenesis causes randomization of left/right body asymmetry. So far 28 genes have been identified to cause PCD either coding for axonemal proteins or ciliary assembly factors. One of the first identified and most often mutated genes DNAH5 causes an outer dynein arm defect leading to cilia with just residual ciliary beating. We sequenced all coding exons of DNAH5, by sanger sequencing including the exon-/intron boundaries in a cohort of 101 patients. All patients were followed up in cardiologic centres, suffered from heterotaxy and/or congenital heart defects. We found mutations in 6 of the 101 patients. In the second cohort, we further evaluated for PCD. We refer to heterotaxy as the instance of situs inversus totalis, which means a mirror image of the common situs (situs solitus) as well as the instance of situs ambiguous, an abnormal arrangement of the thoracic or abdominal organs. Interestingly, two mice-of-function mutations in DNAH5 (deletion/-stop-mutation) were identified. The results of these genetic analyses established the diagnosis of PCD. Even though PCD is a rare disorder 1% of our cohort with heterotaxy showed a mutation in one of the most common PCD causing genes. Because this accounts for 28% of our patients for 28% of PCD cases the results suggest that other genes might play a role. Future studies on the analysed heterotaxia cohort will suffer from PCD. Therefore we like to emphasize the importance of the differential diagnosis of PCD in patients with CHD and respiratory disease.


DPM with at least one patient presenting with CD, and identified 2 compound heterozygous mutations of PKHD1, including in cis

Our results show that early activation of SRY transgene could be activated differentially with the respective sex of the Ddx4-Cre mice. Our results show that early activation of SRY transgene during embryonic development results in postnatal growth retardation and lethality in transgenic mice. T. Kido, Z. Sun, Y. Li, Y. Lau. Dept Med, VA Med Ctr 111C5, Univ California, San Francisco, San Francisco, CA.

Sexual dimorphisms are prevalent between the sexes, particularly in neurodevelopment. Currently, the role of the Y chromosome in such phenomena has not been clearly defined. Among the Y chromosome genes, the sex-determining gene, SRY, could be a significant candidate capable of exerting male-specific effects on sexual dimorphisms. SRY is the founder of the SRY-box (SOX) genes, which are key regulators for various developmental processes. SRY and SOX proteins harbor a conserved HMG box DNA-binding domain. An ectopically expressed SRY could compete with the SOX factors in regulating the respective SOX targets, thereby disrupting the corresponding developmental processes. To test this hypothesis, we have established a transgene activation mouse system, in which a SRY transgene could be activated by a tissue-specific Cre. The Ddx4-Cre transgenic line expresses the Cre recombinase exclusively in the germ lines of both sexes. When a female Y responder is crossed with a male SRY transgene, the Cre recombinase is transferred to the single-cell embryo from the fertilized oocyte, thereby activating the SRY transgene in all tissues.

Although there is strong evidence for the long-term effects of in utero exposure to maternal diabetes and obesity, and potential fetal overnutrition, the exact mechanisms by which these effects are transmitted from human mother to offspring remain poorly understood. Skeletal muscle of obese adults shows alterations in key mitochondrial energy pathways (including branched chain amino acid (BCAA) dependent intermediary metabolism and carnitine palmitoyltransferase I (CPTI)). Myocytes differentiated from multipotent MSC populations have the potential to reveal early changes that may predispose infants to obesity. Here we test the hypothesis that umbilical cord-derived MSCs from infants of obese (Ob) vs normal weight (NW) mothers differentiated into myocytes (skMSC) manifest impaired mitochondrial metabolism when challenged with lipid exposure. Design/Methods: We used data from the Healthy Start Study. This pre-birth cohort study of maternal-infant pairs provides detailed maternal data, including pre-pregnant body mass index (BMI), and cord tissue samples. MSCs derived from full term infants born to NW (BMI<25, n=5) and Ob mothers (BMI>30, n=5) were grown and differentiated into myocytes under control and lipid exposed conditions. On day 21 of differentiation media from all cells was tested for acetyl-CoA concentrations using the Biochrom 30 Analyzer. Two-way Analysis of Variance compared the effects of maternal pre-pregnant BMI, treatment, and their interaction. Results: SkMSC from infants of Ob mothers had elevations in BCAA (leucine and isoleucine), Alanine, and Citrulline versus skMSC from infants of NW mothers (p=0.05, 0.04, 0.01, and 0.05). All of these increases in the conditions of Hypoglycemia (VMC), Molybdenum (MD), and Acid, an under-recognized methionine and threonine intermediary metabolism, decreased with lipid exposure in Ob but increased in NW samples (p= 0.02 interaction). Conclusions: Differential handling of amino acids suggest differences in mitochondrial energy metabolism in offspring of Ob mothers.

More data is needed to demonstrate use in intermediary metabolism rather than differences in cell transport or anabolism, however our data is consistent with amino acid findings in obese adult serum. This MSC model will further be developed in the targeted study of energy metabolomics, transcriptomics, and epigenomics in fetal programming and obesity.
3135S
Autophagy retards inflammatory mRNA decay and elicits a white phe-
notype during adipocyte maturation. J. Shan, A. Worschech, R. Thomas, L. Chouchane. Laboratory of Genetic Medicine and Immunology, Weill Cor-
nell Medical College in Qatar - Qatar Foundation, Doha, Qatar.

Background: Recently, the role of autophagy in glucose and lipid metabo-
lism has been emerging. Mice experiments showed that autophagy defi-
ciency could prevent diet-induced obesity, characterized by less fat and a
browning phenotype of white adipocyte (WAT). However, the underlying
molecular mechanism is not well explored and the data from human are
limited. Method: The mRNA sequencing data of undifferentiated and differ-
entiated human adipocyte cell lines, including two white adipocyte (WAT) and
one brown adipocyte (BAT) were included in our analysis. Gene expression
was reduced by RNA interference in human adipocyte and was enhanced
by glucocorticoid, respectively. Qualifying the lipid droplet content and quan-
tifying the adipolysis and differentiation marker expression were applied to
evaluate WAT differentiation. LC3 was used as a marker to examine autoph-
agy function of adipocyte. Result: We found a remarkable feature of adipoc-
yte differentiation that inflammation signaling was significantly strengthened
during WAT maturation, but not during BAT maturation. The alteration of Zinc
Finger protein 36 (ZFP36), which mediates the decay of mRNA transcripts of
inflammation molecules, obviously affected the phenotype of mature WAT:
silencing of ZFP36 gene resulted in a more whitening phenotype and induc-
ition of ZFP36 resulted in a browning phenotype. ZFP36 activity was associ-
ated with p38 MAPK signaling that was regulated by autophagy. Conclusion:
ZFP36 links autophagy to the determination of mature adipocyte phenotype.
Therefore, ZFP36 is a potential target to prevent obesity and improve glucose
and lipid metabolism.

3136S
RNA-Seq to identify novel markers for neural tissue differentiation. Y. Sun1, K. Giorda1, E. Frey2, M. Taylor1, T. Barron1, C. Davidson1, D. Piper2, G. Meredith3. 1) Thermo Fisher Scientific, South San Francisco, CA; 2) Thermo Fisher Scientific, Madison, WI.

Neural tissue differentiated and cultured from patient-derived stem cells is expected to revolutionize treatment of patients with brain and spinal injuries and diseases. Critical for these cellular therapies is accurate control and monitoring of differentiation but current methods for such cell typing are limited to qPCR and immunocytochemisty (ICC) which is not sufficient to
monitoring of differentiation but current methods for such cell typing are
limited to qPCR and immunochemistry (ICC) which is not sufficient to
discriminate between the numerous (likely >100,000) possible neural cell-
types. RNA-Seq profiling using next-generation sequencing systems permits
coloration and discovery of much-needed novel markers. RNA was isolated over a time course from human embryonic stem cells (H9) and
induced neural stem cells (NSCs) in triplicate. ICC was performed on the
putative NSC pools at d7 and d14 for markers of pluripotency (Oct4) and
neural differentiation (nestin, Sox1, and Pax6) and H9 cells were stained
on d14 for markers of pluripotency (Oct4 and SSEA4). Ion Torrent semi-
conductor sequencing libraries were created to profile expression of miRNAs
and whole transcriptomes for each of the 15 cell populations. We generated
≥1.5 million small RNA reads and ≥229 million whole transcriptome reads
per sample. Cluster analysis of the RNA-Seq profiles indicates that the cell
populations have characteristic molecular signatures. Among genes that
are decreased in induced cells are OCT4 (POU5F1), JARID2, NANOG, consistent with the differentiation of iPSCs into neurons. Among genes that
showed increased expressions are NTRK2, Pou5F2, and a number of HOX
family genes. We also find lincRNA are involved in cell differentiation. For Research Use Only. Not for use in diagnostic procedures.

3137S
A conserved role for IRF6 in neureulation. Y.A. Kousa1, H. Zhu2, W. Fakhouri2, Y. Lei2, A. Kinoshita2, R.R. Roushangar2, E.J. Leslie2, T.D. Busch1, J.T. Williams1, Y. Cha1, B.A. Amendt1, J.C. Murray1, G.M. Shaw1, A.G. Bassuk6, A. Ashley-Koch1, S. Gregory11, R.H. Finnell2, B.C. Schutte12. 1) Biochemistry and Molecular Biology Department, Michigan State University, 48824 East Lansing, Michigan, USA; 2) Dell Pediatric Research Institute, Department of Nutritional Sciences, University of Texas at Austin, 78723 Austin, Texas, USA; 3) Department of Diagnostic & Biomedical Sciences, School of Dentistry, University of Texas at Houston, 77054 Houston, Texas, USA; 4) Department of Human Genetics, Nagasaki University, Nagasaki, Japan; 5) Center for Craniofacial and Dental Genetics, Department of Oral Biology, University of Pittsburgh; 6) Department of Pedi-
atrics, University of Iowa, 52242 Iowa City, Iowa, USA; 7) Department of Craniofacial Biology, University of Colorado Denver at Anschutz Medical Campus, 80045 Aurora, Colorado, USA; 8) Center for Craniofacial Molecular Biology, Ostrw School of Dentistry, University of Southern California, 90033 Los Angeles, California, USA; 9) Department of Anatomy and Cell Biology, University of Iowa, 52242 Iowa City, Iowa, USA; 10) Department of Pedi-
atrics, Stanford University School of Medicine, 94305 Stanford, California, USA; 11) Duke Molecular Physiology Institute, Department of Medicine and Molecular Genetics and Microbiology, Duke University, 27701 Durham, NC, USA; 12) Department of Microbiology and Molecular Genetics, Michigan State University, 48824 East Lansing, Michigan, USA.

IRF6, TFAP2A and GHRH3 are part of a gene regulatory network that is
required for development of the lip and palate in humans and mice. However, in mice, Tlp2a and Ghrl3 are also required for neureulation. To test the hypothesis that Irf6 is also required for neureuration, we evaluated an allelic series of Irf6 mutant mice for neural tube defects. We found that over-expressing Irf6 led to anencephaly by directly repressing Tlp2a transcription.
Reducing Irf6 expression led to the mouse equivalent of spina bifida
by regulating Tlp2a and Ghrl3. In addition, we observed Irf6 expression
in neural and non-neural ectoderm and early migrating neural crest cells.
Moreover, we observed that MCS9.7, an enhancer for IRF6, was active in
pre-migratory neural crest during neureulation. Considering these data in
mice, we asked if rare or common IRF6 variants are associated with human
neural tube defects. We found that rare IRF6 variants are strongly associ-
ated with neural tube defects. In 92 individuals with spina bifida, we identified a rare nonsynonym-
ous substitution previously reported in dominantly inherited orofacial
claefing. In the 3’UTR, we identified a conserved region that regulated irf6
stability in a cell culture assay. Within this element, we found a common
SNP (rs17317411 T/C) whose derived allele (C) is predicted to create a
new miRNA-binding site. Genotyping in 735 trios with spina bifida showed
a significant association between the derived allele of rs17317411 and
lipomyelomeningocele (N = 68 trios; p value = 0.04), but was not statistically
significant for other types of spina bifida. To date, we have also sequenced
MC9.7 in 158 cases and 171 controls. We observed a significant associa-
tion between the derived allele of rs76145088 and spina bifida (p=0.016).
Interestingly, the derived allele at rs76145088 is predicted to abrogate a
highly conserved cis-regulatory motif for the TEAD/TEF family of transcrip-
tion factors. The TEAD family members mediate Hippo signaling in the
neural plate border and pre-migratory neural crest cells and regulate neural
progenitor cell number. In sum, we identified a conserved role for IRF6 in
neureulation between mouse and humans that is in addition to its role in
orofacial clefing. This is a rare example of two common variants, located in
the same enhancer element, that contribute risk for more than one common,
complex human disease.
3138S  
Analysis of CAPZB function in cleft pathogenesis and lower jaw exten- 
sion. K. Mukherjee1,2, M. Grimaldi3, M. Talkowski3, J. Gusella4, R. Macas3, C. Morton1,4, E. Liao1,4. 1) Center for Regenerative Medicine, Massachusetts General Hospital; 2) Center for Human Genetic Research, Massachusetts General Hospital; 3) Brigham and Women’s Hospital; 4) Harvard Medical School, Boston, MA.

Purpose: Orofacial clefts are among the most common congenital birth anomalies worldwide. Clefts and other craniofacial anomalies arise due to defects in craniofacial morphogenesis. In an ongoing gene discovery effort, the Developmental Genome Anatomy Project (DGAP) has developed whole genome sequencing strategies to characterize genes contributing to such human congenital anomalies. Through DGAP, we discovered a new candidate gene for cleft palate (CP), CAPZB that encodes an actin-capping protein. The isolated disruption of CAPZB was identified in a 6-month old female presenting cleft palate, micrognathia and hypothyroidism. We have exploited the zebrafish model to determine the function of capzb in cranio- 
facial morphogenesis. Methods: The zebrafish embryonic Meckel’s cartilage and ethmoid plate are analogous to the mammalian mandible and primary palate respectively, making it an ideal model system to study palate and lower jaw morphogenesis. The spatiotemporal gene expression of capzb was determined by whole mount in situ hybridization (WISH) during embryogenesis. Craniofacial cartilaginous structures and muscles were examined in the capzb mutant, identified from an insertional mutagenesis screen.

Results: WISH analysis shows that capzb is ubiquitously expressed, demonstra-ng its potential requirement in the function of many tissue types. Moreover, preliminary analysis of the capzb mutants show that the lower jaw elements are smaller and retrusive and the palate is only partially fused with a cleft in the anterior palate. The actin cytoskeleton is in disarray without capzb, with loss of cell morphology in the palate chondrocytes and highly disorganized myofibris, leading to atrophied muscles. Conclusion: We have identified CAPZB to be important in craniofacial and muscle morpho- 
genesis, disruption of which is pathologic for palate development. Preliminary results from characterization of the capzb mutants suggest that capzb plays a role in palate fusion and lower jaw extension, and affects tissue types where actin organization is critical to cell morphology and higher order tissue function. Gene expression analysis demonstrates fundamental cellular mechanisms where actin dynamics and cell signaling pathways intersect. This study illustrates how clinically based studies can uncover fundamental mechanisms that govern cell biology and tissue morphogenesis.

3139S  
A novel transcriptional regulatory pathway in cardiac and skeletal mus- 

Gene expression can be regulated at the level of transcript elongation, when RNA polymerase II pauses along the partially transcribed mRNA. Following the recruitment of transcription elongation factors (TEFs), RNA polymerase II resumes transcription allowing for gene expression. Our aim was to determine the role of transcriptional elongation in the development of cardiac and skeletal muscle. Using zebrafish as a model organism, we confirm the expression of a cardiac and skeletal muscle specific elongation factor in these tissues. Morpholino-mediated knockdown of this TEF leads to severe defects in development of the heart and somites, including aberrant looping, reduced circulation and detachment of somites from myosepta. Interestingly, these defects resemble those seen in zebrafish models of cardiac and skeletal muscle. Using zebrafish as a model organism, we were able to determine the role of transcription elongation factors in the development of atrioventricular septum and its associated valves. We conclude that a novel transcriptional pathway may play a role in the development of atrioventricular septum and its associated valves. This study illustrates how clinically based studies can uncover fundamental mechanisms that govern cell biology and tissue morphogenesis.
3142S
Haploinsufficiency of RERE contributes to the development of cleft palate in 1p36 deletion syndrome. B. Kim1, H.P. Zaveri2, V. Jordan1, D.A. Scott3, 1) Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX.

Oral clefts—cleft palate and cleft lip with or without cleft palate—affect 1:7,000 newborns and can lead to long-lasting adverse outcomes for health and social integration. Deletions of chromosome 1p36 are the most common telomeric deletions in humans and carry a high risk of oral clefts. The arginine-glutamic acid dipeptide (RE) repeats gene (RERE) is located in one of three non-overlapping critical regions for orofacial clefts. We have defined on chromosome 1p36. RERE encodes a nuclear receptor coregulator that positively regulates vitamin A/retinoic acid signaling in the developing embryo. Since cleft palate can be caused by abnormalities in retinoic acid signaling, we hypothesized that haploinsufficiency of RERE may contribute to the development of cleft palate in children with 1p36 deletions. To test this hypothesis, we generated an allelic series of RERE-deficient mice using an Rere null-allele and an Rere hypomorphic allele (eyes3) identified in our laboratory. On a C57BL/6 background, cleft palate was seen in 80% (4/5) of Rere+/-embryos at E15.5. When we ablated Rere in neural crest cells using a transgenic Wnt1-Cre, we found that Rere+/-:Wnt1-Cre embryos and mice also developed cleft palates in which the opposing palatal shelves elevated into the normal horizontal position but failed to make contact in the midline. Further studies revealed that RERE is expressed in the bend region of the developing palate; a region where retinoic acid signaling controls cell proliferation and the expression of key palatal genes including TBX1 that causes cleft palate associated with 22q11.2 microdeletions. We conclude that RERE plays a critical, cell-autonomous role in the neural crest cells during palatal development and that deletion of RERE contributes to the development of cleft palate in children with 1p36 deletions.

3143S
An animal model to investigate genetic variants in patients with 46,XY Disorders of Sex Development. H. Barsegian1, T. Baker2, A. Eskin1, S. Nelson1, E. Deliot1, E. Vilain1. Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, California.

Disorders of Sex Development (DSD) are defined as congenital conditions in which development of chromosomal, gonadal, or anatomic sex is atypical. These conditions have a frequency of 0.5-1% of live births and encompass a wide variety of urogenital abnormalities ranging from mild hypospadias to sex reversal. At present, a specific molecular diagnosis is identified in only a minority of DSD cases. We used whole exome sequencing (WES) with gene list approach specific to DSD to identify the underlying genetic etiology for 46,XY DSD cases with a range of phenotypes. In 35% of these cases, exome sequencing identified mutations in genes, classified as pathogenic or likely pathogenic, that could explain the given phenotype. In the unexplained 65% of cases, we identified a large number of variants of uncertain clinical significance (VUS). In order to investigate these variants, we utilized a powerful mouse model for studying undervirilization in 46,XY individuals. In this model, the presence of a Y chromosome originating from an M. domestica strain (YPOS) in C57BL/6J (B6) background results in XY sex reversal and severe hypovirilization. We hypothesized that abnormal gonadal expression of specific genes in the B6-YPOS mouse model will correlate with VUS in genes of 46,XY DSD patients identified through WES. Thus, we isolated gonadal tissue at embryonic day 11.5 and performed RNA sequencing in order to assess the differential gene expression levels in the bipotential gonads of wild type (WT) B6 and undervirilized B6-YPOS males. The obtained differential expression data from the mouse model was analyzed in conjunction with WES data. We identified 12 genes in which the differential expression between WT and B6-YPOS males was significantly lower in B6-YPOS gonads, and in which a missense variant with an alternate allele frequency of less than 1% was identified in 46,XY DSD cases. Among these genes, four (FBLN2, SMTNL2, ADAMTS16, COLSA2) were predicted to be damaging by in silico tools. Additional studies and validation techniques are necessary to investigate the role of these new candidate genes for 46,XY DSD.

3144S
An Allelic Series Reveals Novel Roles of Fgf Ligands in Skeletogenesis. I.H. Hung1,2, G.C. Schoenwolf1, M. Lewandoski2, D.M. Ornitz3. 1) University of Utah, Salt Lake City, UT; 2) National Cancer Institute, Frederick, MD; 3) Washington University School of Medicine, St. Louis, MO.

Chondrodysplasia and craniosynostosis syndromes are caused by activating mutations in fibroblast growth factor (FGF) receptors. Although the roles of the FGF/FGFRs in bone development have been relatively well-characterized, only two FGF ligands, FGF9 and FGF18, have been identified to regulate embryonic skeletogenesis. We have generated an Fgf allelic series to further elucidate the functions of these growth factors during skeletal development. Mice lacking both Fgf9 and Fgf18 exhibit severe defects in endochondral and intramembranous ossification, demonstrating partial functional redundancy between these ligands. We provide evidence that FGF signaling is required at early developmental stages to prevent an arrest in chondrocyte differentiation of the proximal limb and to promote calvarial osteogenesis.

3145S
Phenotypic and functional characterization of Bst+/- mouse retina. G. Sun1,2, H. Riazifar2,3, X. Wang1, F.N. Ross-Cisneros1, V. Carelli1, A.A. Sadun1, T. Huang1,2. 1) Cincinnati Children’s Hospital Medical Center, Cincinnati, OH; 2) Department of Pediatrics, Division of Human Genetics, University of California, Irvine, CA; 3) Doheny Eye Institute, Department of Ophthalmology, University of Southern California, Los Angeles, CA; 4) Dipartimento di Scienze Neurologiche, Università di Bologna, Bologna, Italy; 5) Baylor Contributor.

Current animal models for subretinal neovascularization (SRN) depend on using chemistry or physics methods to stimulate subretinal vessels grow. The animal model for retinal degenerative disease is also very limited. Genetic animal model for retinal disease is needed. The Belly spot and tail (Bst) mouse phenotype is caused by mutations of the ribosomal protein L24. Bst+/- mouse has striking ocular phenotypes, with the feature of delayed closure of choroid fissures, decreased ganglion cells, and subretinal vascularity. In seeking mouse model for stem cell therapy for retinal degenerative disease due to retinal ganglion cell (RGC) loss, we further characterized the Bst+/- mice and investigated the underlying molecular mechanisms. We found that although RGC was significantly reduced in retinal ganglion cell layer in Bst+/- mouse, melanopsin-positive RGC, also called ipRGCs seemed preserved. Pupillary light reflex is complete absent in Bst+/- mouse, while circadian rhythm is normal. In order to examine the pathological abnormalities in Bst+/- mice, we performed electronic microscope test and found mitochondria morphology was deformed with irregular borders and lacking cristae. The complex activities of mitochondrial electron transport chain were decreased significantly. Finally, for subretinal vascularization, the delay of angiogenesis was observed in Bst+/- mice associated with delayed haloid regression. Our characterization of Bst+/- mouse retina suggests that Bst+/- mouse could be a useful model for stem cell therapy.
Sox7 we also found that Sox17—a known regulator histology and evidence of developmental arrest. Using these mouse models, effusions, vascular hemorrhage, ventricular septal defects, aberrant liver failure of yolk sac vasculature remodeling. The same phenotype was around E10.5 with signs of cardiac failure including pericardial edema and 8p23.1 microdeletions. the development of cardiovascular malformations associated with recurrent 8p23.1 microdeletions.
A Tgds mutation in a novel mouse model of Pierre-Robin sequence-type clefting, B.C. Bjork1, I. Saad2, L. Plitnick3, S. Docksey3, D.R. Beier3. 1) Department of Biochemistry, Midwestern University, Downers Grove, IL; 2) Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City, KS; 3) Center for Developmental Biology and Regenerative Medicine, Seattle Children’s Research Institute, Seattle, WA. Non-syndromic orofacial clefting is one of the most common human birth defects with both multigenic and environmental influences that contribute to its etiology. Pierre Robin sequence (PRS)-type clefting is a “palate-extrinsic” mechanism of cleft secondary palate (CP) that results as an indirect consequence of small, mandibular, palatal processes. We identified the little chin (lc) mutant in a three-generation ENU-mutagenesis screen designed to identify autosomal recessive mouse mutants that model human congenital birth defects. lc mutants die shortly after birth due to respiratory distress andPRS-type CP, as evidenced by mandibular hypoplasia. lc mutant palate shelves fail to elevate above the tongue. Mutants also exhibit a shortened snout and a bifid yphoid process. We initially mapped the lc mutation to a 1.1 Mb critical region on chromosome 14 by positional cloning. The region contains six genes, at least two of which, Sox21 and Gpc6, represented attractive candidates for involvement in craniofacial development. However, initial sequencing of coding exons and highly conserved genomic sequences within the critical region failed to reveal any causative mutations.

Recently, we performed Next Generation whole-exome sequencing of genomic DNA from lc and wild type embryonic tissues and identified a homozygous point mutation in the Tgds gene that encodes the DTDP-glucose 4, 6-dehydrogenase enzyme. The C to T missense mutation creates a novel Valine (Val) non-consensus amino acid at residue 206 of Tgds, the same residue that is altered in a homologous mutation of Tgds that is predicted to be damaging by PolyPhen. This is the only homozygous damaging sequence variant identified from this study. The exact function of Tgds and its relationship to craniofacial development is unknown. It has primarily been studied for its role in cell wall integrity in plant species and Candida albicans. Tgds also plays an important role in NAD+/NADH shuttling, and the A26V mutation in lc mutants alters its highly conserved NAD binding motif (GGAGFG). We are validating the mechanism of Tgds mutation and the mechanism of clefting in lc mutants. Impaired Tgds function in lc mutants represents a novel mechanism to be explored as a contributing factor in the incidence of PRS-type clefting in humans.

The transcriptional co-regulator Jab1 is required for skeletogenesis. G. Zhou, L. Bashur, Z. Chen, S. Murakami. Orthopaedics, Case Western Reserve University, Cleveland, OH. Skeletogenesis is a finely-tuned process governed by a spatiotemporal-specific transcriptional circuit. RUNX2 is the master transcription factor required for BMP signaling-mediated chondrocyte maturation and osteoblast differentiation. RUNX2 haploinsufficiency results in cleidocranial dysplasia, while mutations in various BMP signaling components lead to a spectrum of human skeletal developmental defects. Through analyzing a novel mouse model of lethal chondrodysplasia (Jab1floxflox;Col2a1-Cre), we have recently shown that Jab1, a highly conserved transcriptional co-regulator, is a novel factor crucial for chondrocyte maturation, likely by inhibiting Runx2 expression. Jab1 is known to be a binding partner of Runx2. In our study, Jab1 deletion leads to reduced chondrocyte population, decreased expression of Runx2 and Col1a1, and reduced osteoblast differentiation. Thus, Jab1 is required for early osteoblast differentiation of osteochondral progenitor cells. Next, we generated Jab1floxflox;Oss-Cre conditional knockout (Oss-Cre/floxflox;Col2a1-Cre;Oss-Cre) mouse mutants and identified the stage-specific expression of Jab1 on Runx2 and BMP signaling during skeletogenesis. Our study will provide novel insights into the pathogenesis of skeletal disorders caused by dysregulated RUNX2 and BMP signaling.
The chromosome 3p22.3 region is a potential novel locus for complex heart disease, anorectal malformation and intellectual disability. G. Aklier1, RA. Donchy1, A. Angiulli2, S. Sharma3, MK Geiger2, L. Mehta1. 1) Icahn School of Medicine at Mount Sinai, New York, NY; 2) Elmhurst Hospital Center, Elmhurst, NY; 3) Quest Diagnostics Nichols Institute, San Juan Capistrano, CA.

Deletions in the 3p22.3 chromosomal region are rarely reported. Here we describe a female infant, currently 11 months old, with a de novo 1.5 Mb deletion encompassing 14 genes in the 3p22.3 region (no deletion of the 22q11 region detected). Significant phenotypic findings present included tetralogy of Fallot with absent pulmonary valve and a variant imperforate anus (rectoepithelial fistula). A small external previs was reported in early renal sonograms. The association of these anomalies in the otherwise non-dysmorphic child is suggestive of incomplete VACTERL association [MIM 192350]. Currently the patient is status-post repair of the heart defect, but has had motor delays, recurrent respiratory infections and failure to thrive prior to the repair. During embryologic development, aortopulmonary outflow tract septation, and down-growth of the urorectal septum with desintegration of the anal canal occurs between 5-8 weeks suggesting that these birth defects may result from factors disrupting development at this time. There are two previously reported interstitial deletions of this region: deletion 3p22.2-p24.2 in a patient with a small perimembranous ventricular septal defect, global developmental delay and short stature, and deletion 3p22.3p22.2 encompassing the ARRP21 [MIM 605488] and CLASP2 [MIM 605853] genes in a family with intellectual disability and dysmorphic facial features. The CLASP2 gene found in the overlapping region of all three cases may be a potential candidate gene for intellectual disability, given recent evidence of CLASP2 functioning as a key regulator of neuronal polarity and synaptic activity function. One of the other genes deleted in our patient, TRIM71 has been shown to be evolutionarily conserved and its regulation has been shown to play an important role in development. Recent literature has identified several CNVs and SNVs in patients with VACTERL/ VACTERL-like association. However, currently no consistent candidate regions or genes have emerged. This report adds to the literature on potential genetic etiologies for congenital heart defects and associated birth defects as in the VACTERL spectrum.

The sex-determining factor SRY is involved in numerous early events of testis differentiation including testis cord formation. Y. Lau1, Y. Li1, M. Zheng1. 1) Dept Medicine/ VAMC-111C5, Univ California, San Francisco, San Francisco, CA; 2) Dept Anesthesia, Stanford Univ, Palo Alto, CA.

The Y-located sex-determining region Y (Sry) gene encodes the testis-determining factor, which plays a critical role in male sex determination during embryogenesis. A current hypothesis suggests that SRY collaborates with steroidogenic factor 1 (SF1) and activates a related transcription factor gene, SRY-box 9 (Sox9), the essential mediator for testis differentiation. Sox9, in turn, interacts with SF1 and propagates its own expression and testis differentiation. Using the ChIP-CHIP strategy, we have identified large numbers of targets for both SRY and Sox9 in the fetal gonads of mouse embryos at the time of sex determination. Subsequent studies indicate that SRY or Sox9 bindings could be confirmed independently to about 90% of respective targets, which could be regulated by SRY or Sox9 in various reporter and transient gene transfection assays. SRY and Sox9 share about half of the respective targets, suggesting that these two transcription factors regulate a significant number of common target genes. Importantly, SRY, but not Sox9, binds to numerous ovarian differentiating genes and represses their activation by the WNT/β-catenin signaling. Among the SRY and Sox9 targets, numerous ones have been demonstrated to play critical roles in early events of testis differentiation. Notable ones include Ptgds (prostaglandin D2 synthase), critical in Sertoli cell recruitment; Cyp26b1 (cytochrome P450, family 26, subfamily B, polypeptide 1), important for determining male lineage of the germ cells; and Fgfb (fibroblast growth factor 9) and Gdmf (glial cell derived neurotrophic factor), vital for Sertoli cell proliferation and male germ cell niche development. Gene ontology analysis identifies the Sertoli cell-Sertoli cell junction signaling, important for testis cord formation, as the top canonical pathway among the SRY and Sox9 targets. Our findings challenge the current paradigm and suggest that SRY is involved in regulating genes critical in early events of testis differentiation. Therefore SRY determines the Sertoli cell fate by repressing ovarian and activating testicular differentiating genes, promotes Sertoli cell proliferation and recruitment, establishes the male germ cell lineage, and induces the early Sertoli cells to form testis cord. It then passes on its functions to Sox9, which regulates a set of common targets and activates its own gene regulatory program, beyond SRY action, in mammalian sex determination.

3156S

New insights in holoprosencephaly inheritance: Exome sequencing in families reveals new double-hit and recessive cases. C. MOUDEN6, C. DUBOURG1,2, S. ROSE1, G. VIOT4, B. HERON4, W. CARRE4, M. DE TAYRAC1,2, V. DUPE1, S. ODENT1,3, V. DAVID1,2,1. 1) Genetics of pathologies related to development, IGDR UMR 6290 CNRS, Rennes, France; 2) Laboratory of molecular genetics and genomics, CHU Pontchaillou, Rennes, France; 3) Service de Génétique médicale, CHU Hôpital Sud, Rennes, France; 4) Service de gynécologie-obstétrique 1 GHU Cochin - Saint Vincent de Paul, Paris, France; 5) Service de neuropédiatrie, Hôpital Jean Verdier, Bondy, France; 6) Génomique Fonctionnelle Intégrée et Biomarqueurs, IGDR, UMR6290 CNRS, Rennes, France.

Holoprosencephaly (HPE) is an anomaly of early development resulting from cleavage default of prosencephalon. It represents the most frequent cerebral malformation, with an occurrence of approximately 1 in 250 embryos and 1 in 10,000 births. About 60% of HPE cases are due to chromosomal defaults, to environmental causes like maternal alcoholism or diabetes, or are part of a polymalformative syndrome. A part of the remaining 40% of HPE cases, called genetic cases, are due to genetic alterations involving four main genes: SHH, ZIC2, SIX3 and TGIF. About 10 other minor genes are also implicated. But all these genetic alterations explain only 30% of HPE genetic cases. Although these genes are formally involved, their penetrance is usually incomplete and mutations are often inherited from a healthy parent. Thus the outbreak of HPE should be due to a second event, like the implication of another gene. In this context, two high-throughput sequencing strategies (NGS) were used to identify new HPE genes. On one hand, we sequenced the whole exome in 15 families, in which known mutations were inherited from one of the two parents asymptomatic or presenting a micro-malformative syndrome, or de novo mutation in a new candidate gene. Analysis of one family with a known SHH mutation, revealed a second mutation in DIP3P1 in the HPE fetus. The screening for these two mutations in the entire family showed a co-segregation with HPE. This suggests for the first time a double hit involving SHH and DIP3P1. On the other hand, recessive inheritance of HPE can also be suspected in consanguineous families with intrataminal recurrence. Homozygosity mapping was undertaken in eight families with history of consanguineous marriage, or recombination for regions that are identical by descent. Exome sequencing in these consanguineous families revealed a homozygous mutation in two HPE children in a new candidate gene, involved in centrosome duplication. The KO of this gene in mouse leads to HPE pheno-type. Further tests are in progress to validate the double hit role of this mutation in this good candidate gene. The discovery of new genes by these different strategies should permit to improve the prenatal diagnostic, and more generally the knowledge on early brain development.
Whole Exome Sequencing (WES) to Analyze the Genetic Basis of Non-Syndromic Cleft Lip and Palate. M. BASHA, M. QUENTRIC, R. HEL-AERS, N. REVENCU, B. BAYET. 1) Laboratory of Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Europe, Belgium; 2) Center for Human Genetics, Cliniques universitaires St Luc, Université catholique de Louvain, Brussels, Belgium; 3) Centre Labiopalatin, Division of Plastic Surgery, Cliniques universitaires St Luc, Université catholique de Louvain, Brussels, Belgium.

Cleft lip with or without cleft palate (CL/P) is the most common craniofacial birth defect with an incidence of ~1/700 live births, varying with ethnicity and cleft type. It is a debilitating condition requiring an expensive and lifelong treatment. We performed whole exome sequencing (WES) on 20 CL/P subjects, by using True Seq enrichment kit capture on Hiseq2000 (illumina) (38x mean coverage, 100bp 2x paired-end reads). Our pilot WES cohort predominantly consisted of familial cases (n=15) and a few isolated syndromic CL/P subjects (n=5). We WESed 1 affected subject per family. Their sub-phenotypes ranged from a full-blown bilateral CL/P to a subtler velopharyngeal insufficiency (VPI). After bioinformatic processing of raw data with an in-house developed pipeline, we analyzed our samples on Highlander (a software developed in our group by R. Helaers). Each sample had on average ≈50,000 detected variants. We retained variants for further analysis if they met the following criteria: (i) passed the GATK filter (≈49,000), (ii) reported allele count of ≤10 in the population from the 1000 genomes project (≈40,000), (iii) not reported in the population from the GO-NL (genome Netherlands) project (≈6,000), (iv) occurrence in our in-house WES variant list in CL/P samples that passed GATK and at most 2 subjects afflicted with a different condition than CL/P that did not pass GATK (≈957). As “likely pathogenic” variants, we considered those with a high impact (stops) or a moderate impact (non-synonymous-NS), as predicted by SnpEff software. On average, a sample harbored 12 stops and 78 NS variants. All NS variants were predicted as damaging by at least 3 softwares. To identify causative gene(s) shared by multiple subjects, we used the following combinatorial filtering: {a} ≥2 samples with high impact variants in the same gene: 2 variants in 1 gene retained; {b} ≥2 samples with high or moderate-impact variants: 37 variants in 16 genes retained; and {c} ≥2 samples with moderate-impact variants: 267 variants in 122 genes retained. Our results demonstrate important locus heterogeneity for familial CL/P. To be able to distill out the causal gene from the aforementioned data, we will WES additional affected subjects from multiplex families, for which continued collection is also ongoing.
3158T


Pallister-Killian syndrome (PKS) is a sporadic genetic disorder caused by the presence of a tissue-specific mosaicism for isochromosome 12p -i(12p) and is characterized by facial dysmorphism including coarse facies, upslanting palpebral fissures; bitemporal alopecia, pigmentary skin anomalies, and is characterized by facial dysmorphism including coarse facies, upslanting palpebral fissures; bitemporal alopecia, pigmentary skin anomalies, and is characterized by facial dysmorphism including coarse facies, upslanting palpebral fissures; bitemporal alopecia, pigmentary skin anomalies, and coarse facies, hypotelorism, congenital heart defects, mental retardation and CNS abnormalities. MLPA analysis of palmar sweat ducts, buccal smear or others tissues are essential for PKS diagnosis because cytogenetic analysis of peripheral lymphocytes usually fails to detect the mosaicism of isochromosome 12p -i(12p). We reviewed the medical records of patients with confirmed PKS followed in our service (since 1981 to 2014) cytogenetic analysis of peripheral lymphocytes usually fails to detect the mosaicism of isochromosome 12p -i(12p). We reviewed the medical records of patients confirmed PKS followed in our service (since 1981 to 2014) and performed MLPA (Multiplex Ligation-dependent Probe Amplification) for diagnosis and confirmation in two of them. In this study we propose a new methodology for PKS diagnosis: MLPA of buccal smear. We describe eight patients with PKS. Clinical features of all patients were consistent with those described in the literature. Age at diagnosis varied from prenatal to 3 years. In all patients, peripheral blood karyotypes were normal. Seven patients were diagnosed by a skin fibroblast karyotype with i(12p) in 50% to 100% of cells. Recently, it was performed in two patients MLPA technique of buccal smear, showing 4 copies of 12p. Two patients have been lost during follow up and one patient died from pneumonia at 14 years old. The current age of five patients ranges from 9 months to 18 years, all of them with severe mental retardation. A precocious PKS diagnosis is important to optimally manage the disease and to provide genetic counseling. MLPA on buccal smears is a good and non-invasive method to detect extra copies of 12p and should be considered at the first exam, before a skin biopsy for a fibroblast karyotype is performed.

3159T


Small supernumerary marker chromosomes (sSMCs) that cannot be identified with standard cytogenetic analysis. G-banding techniques are estimated to occur in approximately 0.04% of live births potentially causing congenital abnormalities if associated with copy number gains in regions of clinically significant genes. We present cytogenetic findings from two newborn patients referred for cytogenetic analysis. Patient 1: Clinical indications for patient 1 included respiratory distress and suspected sepsis. G-banded chromosome analysis of lymphocytes collected from this patient showed the variable presence of 4 to 6 marker chromosomes in 20 cells counted. Whole genome SNP microarray analysis detected a total copy number gain of 51.14 MB consisting of the pericentric regions of chromosomes 1, 6 and 9, as well as an interstitial region of chromosome 17, suggesting the possibility of at least one composite marker. Subsequent FISH studies confirmed that these genomic regions were present in three distinct sSMCs. Patient 2: This patient was born at 29 weeks with congenital anomalies including a ventricular septal defect (VSD), patent ductus arteriosus (PDA), low set ears and widely spaced nipples. Whole genome SNP microarray analysis detected a total copy number gain of 15.80 MB originating from the pericentromeric regions of chromosomes 6, 16 and X. A consistent percentage of mosaicism was detected for most of these regions, again suggesting the presence of several mitotically unstable composite marker chromosomes. Chromosome studies confirmed the variable presence of 3 to 6 supernumerary marker chromosomes in 30 cells counted. These results indicate that sSMCs can originate from complex genetic rearrangements and illustrate the necessity and efficacy of whole genome microarray analysis to precisely identify the overall genomic imbalance. Designation of the specific integration of the additional interstitial segments in each marker will entail numerous FISH studies which can be facilitated by multiprobe painting cocktails.

3160T


The incidence of multiple, apparently balanced translocations in the general population is rare. Our lab has identified five families with multiple translocations, two of which have been previously reported. This study details the cytogenetic findings in the remaining three families. Family 1: Chromosome studies on amniotic fluid and buccal smear failed to detect the presence of either one or both balanced translocations. Family 2: Karyotypic analysis of a product of conception (POC) specimen on an amniotic fluid, buccal smear or others tissues are essential for PKS diagnosis because cytogenetic analysis of peripheral lymphocytes usually fails to detect the mosaicism of isochromosome 12p -i(12p). We reviewed the medical records of patients confirmed PKS followed in our service (since 1981 to 2014) and performed MLPA (Multiplex Ligation-dependent Probe Amplification) for diagnosis and confirmation in two of them. In this study we propose a new methodology for PKS diagnosis: MLPA of buccal smear. We describe eight patients with PKS. Clinical features of all patients were consistent with those described in the literature. Age at diagnosis varied from prenatal to 3 years. In all patients, peripheral blood karyotypes were normal. Seven patients were diagnosed by a skin fibroblast karyotype with i(12p) in 50% to 100% of cells. Recently, it was performed in two patients MLPA technique of buccal smear, showing 4 copies of 12p. Two patients have been lost during follow up and one patient died from pneumonia at 14 years old. The current age of five patients ranges from 9 months to 18 years, all of them with severe mental retardation. A precocious PKS diagnosis is important to optimally manage the disease and to provide genetic counseling. MLPA on buccal smears is a good and non-invasive method to detect extra copies of 12p and should be considered at the first exam, before a skin biopsy for a fibroblast karyotype is performed.

Small supernumerary marker chromosomes (sSMCs) that cannot be identified with standard cytogenetic analysis. G-banding techniques are estimated to occur in approximately 0.04% of live births potentially causing congenital abnormalities if associated with copy number gains in regions of clinically significant genes. We present cytogenetic findings from two newborn patients referred for cytogenetic analysis. Patient 1: Clinical indications for patient 1 included respiratory distress and suspected sepsis. G-banded chromosome analysis of lymphocytes collected from this patient showed the variable presence of 4 to 6 marker chromosomes in 20 cells counted. Whole genome SNP microarray analysis detected a total copy number gain of 51.14 MB consisting of the pericentric regions of chromosomes 1, 6 and 9, as well as an interstitial region of chromosome 17, suggesting the possibility of at least one composite marker. Subsequent FISH studies confirmed that these genomic regions were present in three distinct sSMCs. Patient 2: This patient was born at 29 weeks with congenital anomalies including a ventricular septal defect (VSD), patent ductus arteriosus (PDA), low set ears and widely spaced nipples. Whole genome SNP microarray analysis detected a total copy number gain of 15.80 MB originating from the pericentromeric regions of chromosomes 6, 16 and X. A consistent percentage of mosaicism was detected for most of these regions, again suggesting the presence of several mitotically unstable composite marker chromosomes. Chromosome studies confirmed the variable presence of 3 to 6 supernumerary marker chromosomes in 30 cells counted. These results indicate that sSMCs can originate from complex genetic rearrangements and illustrate the necessity and efficacy of whole genome microarray analysis to precisely identify the overall genomic imbalance. Designation of the specific integration of the additional interstitial segments in each marker will entail numerous FISH studies which can be facilitated by multiprobe painting cocktails.

The incidence of multiple, apparently balanced translocations in the general population is rare. Our lab has identified five families with multiple translocations, two of which have been previously reported. This study details the cytogenetic findings in the remaining three families. Family 1: Chromosome studies on amniotic fluid and buccal smear failed to detect the presence of either one or both balanced translocations. Family 2: Karyotypic analysis of a product of conception (POC) specimen on an amniotic fluid, buccal smear or others tissues are essential for PKS diagnosis because cytogenetic analysis of peripheral lymphocytes usually fails to detect the mosaicism of isochromosome 12p -i(12p). We reviewed the medical records of patients confirmed PKS followed in our service (since 1981 to 2014) and performed MLPA (Multiplex Ligation-dependent Probe Amplification) for diagnosis and confirmation in two of them. In this study we propose a new methodology for PKS diagnosis: MLPA of buccal smear. We describe eight patients with PKS. Clinical features of all patients were consistent with those described in the literature. Age at diagnosis varied from prenatal to 3 years. In all patients, peripheral blood karyotypes were normal. Seven patients were diagnosed by a skin fibroblast karyotype with i(12p) in 50% to 100% of cells. Recently, it was performed in two patients MLPA technique of buccal smear, showing 4 copies of 12p. Two patients have been lost during follow up and one patient died from pneumonia at 14 years old. The current age of five patients ranges from 9 months to 18 years, all of them with severe mental retardation. A precocious PKS diagnosis is important to optimally manage the disease and to provide genetic counseling. MLPA on buccal smears is a good and non-invasive method to detect extra copies of 12p and should be considered at the first exam, before a skin biopsy for a fibroblast karyotype is performed.
3162T
Reversing differences in chromatin accessibility that distinguish homologous mitotic metaphase chromosomes. W.A. Khan,1 P.K. Rogers2,3 J.H.M. Knoll4,5 1Department of Pathology, London, ON; 2Department of Biochemistry, London, ON; 3Department of Computer Science, London, ON; 4Cytogenomix, London, ON Canada.

Treatment with agents that alter histone modifications, DNA conformation or its sequence are well established strategies for cancer chemotherapy. Apart from their effects on interphase chromatin (G1, S), little is known about how these compounds act during and preceding metaphase. We have previously demonstrated that most individual single copy (SC) FISH probes (1-5.5 kB) show fluorescence patterns that are independent of chromosome metaphase targets. A subset of these probes show reproducible, non-random hybridization intensity differences between homologous targets. Atomic and super-resolution microscopy analyses indicate that the differences correspond to differential accessibility (DA) at these loci. We investigated epigenetic characteristics of DA with agents that inhibit histone deacetylation, phosphorylation, cytosine methylation, and alter superhelicity in lymphoblastoid cell lines. Compared to controls, DA was impervious to chemical treatments targeting histone modifying enzymes or cytosine methylation from five distinct genomic regions (RG57, 1q43; CACNA12, 8q34; ADORA2B1851, 17p12; PMP221853, 17p12 and ACR, 22q13.33). We also hybridized etoposide-treated metaphase cells with probes to the same genomic regions that did not hybridize in untreated cells. DA was reversed at all of these loci at various physiological etoposide concentrations (p < 0.05; two-proportion z-test, n = 30-40 cells/region). Differences in fluorescent probe intensity ratios measured between homologs from etoposide-treated controls were also 4 to 5-fold lower relative to untreated controls. With 3-D structured illumination microscopy, we monitored the resolution differences in hybridized probe volume between allele targets significantly decreased (p < 0.05, two-tailed t-test) in treated cells (μA = 0.237µm³) compared to untreated controls in which one allele was less accessible (μA = 0.7525±μm³). Our results suggest that etoposide induces decreased DNA supercoiling and axial metaphase chromosome condensation. Our findings suggest that allele differences in metaphase chromosome accessibility represent a stable chromatin mark that is largely maintained on mitotic metaphase chromosomes, and that the local differences in condensation are due to topoisomerase activity, which result in DA.

3163T
Longitudinal shortening in telomere length as a biomarker for demenita status of adults with Down syndrome. E.C. Jenkins1, B. Ye1, S.J. Kirsksy-McHale1, W.B. Zigman1, N. Schupf1,2, W.P. Silverman3 1Department of Human Genetics, NYS Institute for Basic Research in Human Genetics, Staten Island, NY; 2Taub Institute for Research on Alzheimer’s Disease and the Aging Brain, Columbia University, New York, NY; 3Kennedy Krieger Institute and The Johns Hopkins University School of Medicine, Baltimore, MD.

We have previously reported shorter telomeres in T-lymphocytes from participants in a cross-sectional study of 11 people with Down syndrome (DS) and mild cognitive impairment (MCI) or dementia versus those with DS only, using PNA (peptide nucleic acid) probes for telomeres and the chromosome 1 centromere (cen (2)) PNA probe. The intensity/length for the cen (2) PNA probe can be used to generate an “internal standard” to normalize telomere measurements. With 3-D structured illumination microscopy, we demonstrated that the mean differences in telomere length of individuals with and without MCI revealed clear and consistent stable, while the light intensity/length of the telomeres is expected to shorten or dementia (N = 1) by the time a second sample was obtained, all had not have MCI/dementia at Time 1 but who had converted to MCI (N = 5) or dementia group (0.08 - 0.13). We are anxious to extend this study with additional prospective follow-up of a larger sample of individuals with Down syndrome who have not yet converted to MCI/dementia. Supported in part by NYS OPWDD, Alz. Assoc. IIRG-96-077, IIRG-07-60558; NIH PO1-HD035897, RO1-HD037425, RO1-AGO14673, and P30-HD024061.

3164T
The role of copy number variation in non-syndromic cleft lip and palate. L.A. Haney1,2, B.W. Darbro1,2, A. Long2, J. Standley1, J.C. Murray1,2, J.R. Manak1,2 1Department of Pediatrics, University of Iowa, Iowa City, IA; 2Department of Biogenetics, University of Iowa, Iowa City, IA; 3Interdisciplinary Genetics Program, University of Iowa, Iowa City, IA.

Clefts of the lip and/or palate (CLP) are common congenital abnormalities affecting approximately 1 in 700 live births. It is estimated that over 90% of CLP cases are non-syndromic (NS) which excludes individuals with CLP who exhibit cognitive or multiple congenital abnormalities. A variety of candidate genes, genome-wide association studies, and animal models have been used to identify genetic loci for clefting. Approximately half of copy number variation (CNV) in clefting. We performed array-based comparative genomic hybridization on a NSCLP cohort from the Philippines to identify CNVs associated with clefting. After quality controls minimized false-positives, we processed 84 NSCLP cases and a replication cohort of 854 NSCLP cases for further analysis. Specifically focusing on deletions for this study, we used an analysis pipeline which identified CNV losses that overlapped with exons of genes in regions sharing 50% or less overlap with segmental duplications and common CNVs. Analysis of CNVs in the 84 NSCLP cases identified 33 deletions containing 53 genes, 1 of which overlapped with a previously unidentified clefting locus CDH1 (MIM 192090). We also investigated the GWAS region significantly associated with NSCLP. Bq24 (MIM 126859). Analysis of Bq24 revealed three small deletions within a 1.4Mb region containing the most associated SNP rs987525 (MIM 611888) and a recently reported murine medianosal enhancer of MYC (MIM 190080). Characterization of these deletions is underway. Seventeen regions from the small cohort, containing 23 genes, replicated in the larger cohort. In addition, we found a 300kb deletion which was present in all NSCLP cases including six previously identified clefting loci XRPC3 (MIM 600675), JAG2 (MIM 602570), Bq24, NTN1 (MIM 601614), DLG1 (MIM 601014), and LPL (MIM 609708). To elucidate the role of these CNVs in CLP, we are performing experiments to determine the altered copy number regions, as well as altering their dose in zebrafish. In addition, we are conducting a trio study using the loci identified in the small cohort to determine if the CNVs are de novo or familial. Since this large dataset provides immense potential for the identification of novel CLP loci, we plan to extend the analysis to intronic and intergenic CNVs. It is our hope to define how CNVs contribute to NSCLP and identify novel causative variants for the disease.

3165T
Characterization of a fusion gene involving the leptin gene generated by a de novo 7q32.1 duplication associated with severe anorexia. J.Lévy1,2,3, E. Pipiras1,4,5, L. de Pontual2,6,7, E. Ghoulouzi1,3, N. de Roux2,4,5, A. Tabet2,4, B. Gressens1,3, B. Benzacken2,4,6,6, S. Lebon1, A. Delaty2,4,6,6,1, Jean Robert Debré, Paris, France; 2AP-HP, Service de Genétique et de Cytogénétique, Hôpital Robert Debré, Paris, France; 3Université Paris Diderot, Sorbonne Paris Cité, Paris, France; 4Université Paris 13, Sorbonne Paris Cité, UFR AP-HP, Service d’Histologie, Embryologie, et Cytogénétique, Hôpital Jean Verdier, Bondy, France; 5Université Paris 13, Sorbonne Paris Cité, UFR SMBH, Bobigny, France; 6AP-HP, Service de Pédiatrie, Hôpital Jean Verdier, Bondy, France; 7Inserm U781, Université Paris Descartes, Paris-Sorbonne Cité, Institut IMAGINE, Paris, France.

Leptin, a polyepitope comprised of several copies of a protein with the same size and the same cellular localization that is supposed to play a major role in the regulation of appetite and weight. Leptin action is mediated by a large family of leptin receptors, LeptinR and LEP receptor genes cause extreme obesity in mice and humans. Here we describe a de novo 7q32.1 duplication spanning the LEP gene discovered in a 2-year-old boy referred for unexplained precocious anorexia associated with postnatal ponderal growth retardation.

Leptin-sequence-gene sequencing revealed that the duplication was a direct tandem duplication. This duplication generated a fusion between a potential pseudogene paralogous to the TLR2 gene (thereafter referred as TLK2P3) and the LEP gene. Several fusion transcripts generated by alternative splicing of TLK2P3 were detected by RT-PCR on patient’s fibroblasts and lymphocytes. Thus, we performed 84 NSCLP cases and a replication cohort of 854 NSCLP cases for further analysis. Constructions of expression vectors encoding HA-tagged leptin, TLK2P3 and TLK2P3-LEP were performed to characterize the potential fusion proteins. The expression of TLK2P3-LEP, LEP and TLK2P3 were analyzed by immunoblots and western blot in patient fibroblasts and L6A cells.

The first results of the fusion gene expressed in HeLa cells showed the production of a protein with the same size and the same cellular localization as leptin. Additional alternative fusion transcripts expressions in HeLa cells were detected by RNAseq and the corresponding alternative RNA-seq alignments were generated. The characterization of the fusion gene, predicted by a chromatin state segmentation for each of nine human cell types. This highly active promoter may induce deregulated and ectopic expression of leptin.

Several hypotheses of the mechanisms by which this 7q32.1 duplication affect the phenotype of the patient will be discussed. Although our results are preliminary, this study suggests that fusion genes can be a mechanism by which CNVs contribute to Human diseases.
Detectable mosaic 13q14 deletions in non-hematologic cancers and healthy controls. M. Yeager1,2, M. Machiela1, W. Zhou1,2, M. Dean3, S.J. Chanock1. 1) Cancer Genomics Research Laboratory, Leidos Biomed, FNLCR, Frederick, MD; 2) Division of Cancer Epidemiology and Genetics, NCI, Bethesda, MD; 3) Center for Cancer Research, NCI, Frederick, MD.

Mosaic loss of 13q14.3 is a common chromosomal event found in approximately 50 percent of B-cell chronic lymphocytic leukemia (CLL) cases. Studies suggest 13q14.3 loss may be evident in somatic DNA of other non-hematologic tumor types, but to date no comprehensive study has been performed on the frequency of 13q14.3 mosaic loss in lymphocyte-derived DNA of solid tumor cases and healthy controls. Based on preliminary findings indicating an excess of 13q14.3 mosaic loss in our previous studies of detectable clonal mosaicism, we performed a comprehensive scan of chromosome 13 from 46,254 non-hematologic cancer cases and 36,229 controls using modified B-allele frequency and log R ratio based detection algorithms. We detected a total of 60 individuals with 13q14.3 mosaic loss, 1 occurrence of mosaic copy neutral uniparental disomy, and 13 individuals with stretches of homozygosity that spanned the 13q14.3 region. While detected 13q14.3 mosaic losses were variable in size, the minimally deleted region (MDR) in our analysis (chr13:49,590,000-49,983,100, NCBI36/hg18) closely resembled that seen in CLL and included transcripts of DLEU1 and DLEU2. A breakpoint analysis on 13q14.3 loss boundaries indicates a clustering of breakpoints with statistically significant enrichment for breakpoints around gene rich regions and areas of open chromatin when compared to random permutations of same-sized events spanning the chromosome 13 MDR. Our analysis indicated the frequency of 13q14.3 loss significantly increases with increasing age (p-value=0.028). The frequency of 13q14.3 mosaic loss did not significantly differ between non-hematologic cancer cases (0.084%, 95% CI: 0.051-0.111%) and controls (0.058%, 95% CI: 0.033-0.083) nor did frequency estimates differ from age and sex standardized SEER estimates. While 13q14.3 deletions may serve as early biomarkers for CLL, our study suggests 13q14.3 loss may be evident in somatic DNA of other non-hematologic cancers and healthy controls.
First case of homozygous deletion in the ABAT gene leading to GABA-T deficiency and severe neonatal neurologic disease. A. Mosca-Boi- dron1,2, G.S. Salomons3,4, E.A. Struyf5, C. Viazy-Saban5,6, L. Faivre7,8, M. Lefebvre5, N. Marie1,2, F. Feillet9, M. Payet1, C. Ragon1, F. Mugnere1, A. Masurel-Paulet1, J. Thevenon10,5, M.G. Mourot de Rougemont1, C. Thauvin-Robinet10,11, P. Callier1,2, S. El Chehadeh12,13,1, Laboratoire de cytoté- nique, Plateau technique de biologie, CHU de Dijon, France; 2) GAD, EA 4271, Génétique et Anomalies du Développement, Université de Bourgogne, Dijon, France; 3) Metabolic Unit, clinical chemistry, VUMc Medical center, de Boelelaan, Amsterdam, the Netherlands; 4) Service des maladies héréditaires du métabolisme et dépistage néonatal, Groupement hospitalier EST, Bron, France; 5) FHU TRANSALD, Centre de référence maladies rares « anomalies du développement et syndromes malformatifs » de l’Est, Centre de Génétique, CHU de Dijon, France; 6) Service de médecine infantile, Hôpital d’Enfants de Brabois, CHU de Nancy, Vandoeuvre-les-Nancy, France; 7) Service de radiologie, Hôpital d’Enfants, CHU de Dijon, France.

Background 4-gamma-aminobutyrate transaminase (GABA-T) catalyzes the conversion of gamma-aminobutyric acid (GABA), one of the most important inhibitory neurotransmitters in the central nervous system of mamma- lians, into succinic semialdehyde and is encoded by the ABAT gene. GABA- T deficiency is a very rare autosomal recessive disorder characterized by severe psychomotor retardation, hypotonia, hyperreflexia, seizures, associated with early infantile death. It has been reported in only 3 patients to date, 2 of whom were siblings. Electrocorticographic and brain magnetic resonance imaging (MRI) abnormalities including respectively burst suppression pattern, cerebellar hypoplasia, posterior fossa cyst, corpus callosum agenesis and abnormal gyration, have been previously described. We report here on a fourth male patient born to consanguineous healthy Moroccan parents who presented with neonatal epileptic encephalopathy including major axial hypotonia, lethargy with hypotonia and intractable seizures, without any dysmorphic features. He died at 15 days of life. Methods Investigations that were done in the first week after birth included array- CGH, brain MRI, and wide metabolic screening. Results Brain MRI showed atrophy of the cerebellar vermis and hemispheres with a cerebellar cyst aspect, reduced white matter, very thin corpus callosum and diffuse thicken- ing, the cortex showing no localization of abnormalities. MRI and metabolic screening did not show any abnormality. Chromatography of urinary aminoacids showed an increased level of glutaric and that they had consequences on neuronal development.

First case of homozygous deletion in the ABAT gene leading to GABA-T deficiency and severe neonatal neurologic disease. A. Mosca-Boidron1,2, G.S. Salomons3,4, E.A. Struyf5, C. Viazy-Saban5,6, L. Faivre7,8, M. Lefebvre5, N. Marie1,2, F. Feillet9, M. Payet1, C. Ragon1, F. Mugnere1, A. Masurel-Paulet1, J. Thevenon10,5, M.G. Mourot de Rougemont1, C. Thauvin-Robinet10,11, P. Callier1,2, S. El Chehadeh12,13,1, Laboratoire de cytoté- nique, Plateau technique de biologie, CHU de Dijon, France; 2) GAD, EA 4271, Génétique et Anomalies du Développement, Université de Bourgogne, Dijon, France; 3) Metabolic Unit, clinical chemistry, VUMc Medical center, de Boelelaan, Amsterdam, the Netherlands; 4) Service des maladies héréditaires du métabolisme et dépistage néonatal, Groupement hospitalier EST, Bron, France; 5) FHU TRANSALD, Centre de référence maladies rares « anomalies du développement et syndromes malformatifs » de l’Est, Centre de Génétique, CHU de Dijon, France; 6) Service de médecine infantile, Hôpital d’Enfants de Brabois, CHU de Nancy, Vandoeuvre-les-Nancy, France; 7) Service de radiologie, Hôpital d’Enfants, CHU de Dijon, France.

Background 4-gamma-aminobutyrate transaminase (GABA-T) catalyzes the conversion of gamma-aminobutyric acid (GABA), one of the most important inhibitory neurotransmitters in the central nervous system of mamma- lians, into succinic semialdehyde and is encoded by the ABAT gene. GABA- T deficiency is a very rare autosomal recessive disorder characterized by severe psychomotor retardation, hypotonia, hyperreflexia, seizures, associated with early infantile death. It has been reported in only 3 patients to date, 2 of whom were siblings. Electrocorticographic and brain magnetic resonance imaging (MRI) abnormalities including respectively burst suppression pattern, cerebellar hypoplasia, posterior fossa cyst, corpus callosum agenesis and abnormal gyration, have been previously described. We report here on a fourth male patient born to consanguineous healthy Moroccan parents who presented with neonatal epileptic encephalopathy including major axial hypotonia, lethargy with hypotonia and intractable seizures, without any dysmorphic features. He died at 15 days of life. Methods Investigations that were done in the first week after birth included array- CGH, brain MRI, and wide metabolic screening. Results Brain MRI showed atrophy of the cerebellar vermis and hemispheres with a cerebellar cyst aspect, reduced white matter, very thin corpus callosum and diffuse thicken- ing, the cortex showing no localization of abnormalities. MRI and metabolic screening did not show any abnormality. Chromatography of urinary aminoacids showed an increased level of glutaric and that they had consequences on neuronal development.

First case of homozygous deletion in the ABAT gene leading to GABA-T deficiency and severe neonatal neurologic disease. A. Mosca-Boidron1,2, G.S. Salomons3,4, E.A. Struyf5, C. Viazy-Saban5,6, L. Faivre7,8, M. Lefebvre5, N. Marie1,2, F. Feillet9, M. Payet1, C. Ragon1, F. Mugnere1, A. Masurel-Paulet1, J. Thevenon10,5, M.G. Mourot de Rougemont1, C. Thauvin-Robinet10,11, P. Callier1,2, S. El Chehadeh12,13,1, Laboratoire de cytoté- nique, Plateau technique de biologie, CHU de Dijon, France; 2) GAD, EA 4271, Génétique et Anomalies du Développement, Université de Bourgogne, Dijon, France; 3) Metabolic Unit, clinical chemistry, VUMc Medical center, de Boelelaan, Amsterdam, the Netherlands; 4) Service des maladies héréditaires du métabolisme et dépistage néonatal, Groupement hospitalier EST, Bron, France; 5) FHU TRANSALD, Centre de référence maladies rares « anomalies du développement et syndromes malformatifs » de l’Est, Centre de Génétique, CHU de Dijon, France; 6) Service de médecine infantile, Hôpital d’Enfants de Brabois, CHU de Nancy, Vandoeuvre-les-Nancy, France; 7) Service de radiologie, Hôpital d’Enfants, CHU de Dijon, France.

Background 4-gamma-aminobutyrate transaminase (GABA-T) catalyzes the conversion of gamma-aminobutyric acid (GABA), one of the most important inhibitory neurotransmitters in the central nervous system of mamma- lians, into succinic semialdehyde and is encoded by the ABAT gene. GABA- T deficiency is a very rare autosomal recessive disorder characterized by severe psychomotor retardation, hypotonia, hyperreflexia, seizures, associated with early infantile death. It has been reported in only 3 patients to date, 2 of whom were siblings. Electrocorticographic and brain magnetic resonance imaging (MRI) abnormalities including respectively burst suppression pattern, cerebellar hypoplasia, posterior fossa cyst, corpus callosum agenesis and abnormal gyration, have been previously described. We report here on a fourth male patient born to consanguineous healthy Moroccan parents who presented with neonatal epileptic encephalopathy including major axial hypotonia, lethargy with hypotonia and intractable seizures, without any dysmorphic features. He died at 15 days of life. Methods Investigations that were done in the first week after birth included array- CGH, brain MRI, and wide metabolic screening. Results Brain MRI showed atrophy of the cerebellar vermis and hemispheres with a cerebellar cyst aspect, reduced white matter, very thin corpus callosum and diffuse thicken- ing, the cortex showing no localization of abnormalities. MRI and metabolic screening did not show any abnormality. Chromatography of urinary aminoacids showed an increased level of glutaric and that they had consequences on neuronal development.

1) Informatics, Affymetrix Inc., Santa Clara, CA; 2) Cytogenetics Laboratory, Laboratory Corporation of America, Research Triangle Park, NC; 3) The Center for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada; 4) Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 5) Radboud University Medical Centre, Nijmegen, Netherlands; 6) Department of Clinical Genetics Skåne University Hospital - 221 85 Lund, Sweden; 7) Department of Immunogenetics, Genetics and Pathology, Rudbeck and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; 8) Grupo de Medicina Xenómica, UdC, CeGen- CIBERER, Santiago de Compostela, Spain.

A large cohort of 2,321 genomic DNA samples from phenotypically normal individuals was characterized for genomic copy number variations to understand the frequency and genomic distribution of non-pathogenic CNVs in the population. DNA samples were extracted from blood samples obtained at 6 sites in 6 countries, including Sweden, the Netherlands, Norway, Spain, Canada and the United States. Samples were processed with the Affymetrix CytoScan® HD Microarray Kit and the data analyzed for copy number calls using Affymetrix Power Tools. Copy number calls were extracted from the output files and compiled into a large dataset, including only CNVs that exceeded a minimal size threshold, 25 markers and 25 kbp for copy number losses or 50 markers and 50 kbp for copy number gains, using a composite profile from 100 blood and 280 cell line samples as a reference. The CNV data set was evaluated for the frequency of finding a loss or gain at every genomic location. Frequencies of adjacent genomic locations were smoothed using an exponential decay window-smoothing algorithm and the data were segmented to reconstruct common regions of CNVs. We found less than 5.5Mb of the human genome to be associated with CNV regions having population frequencies of 10% or more. These common CNV regions were analyzed for genomic structure, such as segmental duplications and chromosomal position to help understand the genomic attributes associated with regions with frequent CNV polymorphisms.

Familial transmission of 5p13.2 duplication due to maternal der(-X)ins(X;5): L.C. Walters-Sen, K. Windemuth, J. Nandhali, J.M. Milunsky. 1) Center for Human Genetics, Inc., Cambridge, MA; 2) Boston Medical Center, Boston, MA.

Submicroscopic duplications of 5p13 have been recently reported in a number of cases, warranting the description of a new clinical entity (chromosome 5p13 Duplication Syndrome, MIM 613174). These microduplications, while variable in size, all contain at least part of the NIPBL gene (MIM:608667). Patients with duplications in this region present with intellectual disability/developmental delay (ID/DD) and dysmorphic facies. In addition, skeletal and brain abnormalities have been variably reported, as well as propensity for obesity and hypotonia. We report a family with two affected sons, each carrying a duplication at 5p13.2 encompassing the 5’ portion of SLC1A3 and the 5’ portion of NIPBL. The proband was originally referred for a SNP microarray analysis, which identified the 341 kb 5p13.2 duplication, also found in his brother by a different laboratory. Both brothers had delays in speech development and diagnoses of possible autism and ADHD. The proband’s weight was in the 95-97th percentile, while that of his brother was in the 75-90th percentile. In addition, the brothers shared facial dysmorphism, including long eyelashes, long palpebral fisure, a broad nasal tip with anteverted nares, and a small chin. Other anomalies included a high palate, pes planus, and prominent breast tissue. Upon confirming the SNP microarray finding by FISH in the proband, it was discovered that the 5p13.2 duplication was located on the short arm of the X chromosome. Further FISH studies on the mother demonstrated that both she and the proband carried a derivative X chromosome with insertion of material from 5p13.2 into the intermediate region of Xp [der(X)ins(X;5) (p22.1;p13.2p13.2)]. As the mother did not share facial dysmorphism or ID/DD with her sons, X-inactivation studies were performed. The mother showed skewed inactivation (82:18), providing a mechanism for suppression of expression of the duplicated 5p13.2 material. To our knowledge, this is the first report of an inherited duplication of 5p13.2 with multiple affected family members. This family underscores the need to confirm array findings by FISH, both in the proband and family members, to discern implications for pathogenicity and more accurately define the recurrence risk.
Multiple homozygosity regions in a girl with unexplained intellectual disability. C.C. da Silva, A.V. Melo, M.R. Cunha, C.L. Ribeiro, T.C. Vieira, D.M. e Silva, A.D. da Cruz. 1) Pontifical Catholic University of Goias, Department of Biology, Replicon Research Group, Brazil; 2) Pontifical Catholic University of Goias, Genetics Master’s Program, Brazil; 3) Federal University of Goias, Biotechnology and Biodiversity PhD Program, Brazil; 4) Federal University of Goias, Genetics and Molecular Biology Master’s and PhD Programs, Brazil; 5) Human Cytogenetics and Molecular Genetics Laboratory, Secretary of Goias State for Public Health, Brazil; 6) State University of Goias, UnU Eseffez, Brazil.

Intellectual Disability (ID) is a complex disorder characterized by significant limitations in both intellectual functioning and adaptive behavior that begin before the age of 18 years. The detection of excessive long continuous stretches of homozygosity (LCSH) on multiple chromosomes may suggest consanguinity, and could be useful in determining candidate genes for further testing for autosomal recessive Mendelian disorders. Herein, we report a first case of a Brazilian girl who presented moderate intellectual disability and LCSH detected by Chromosomal Microarray Analysis (CMA). A 9 years old girl born of first-cousin parents, at 40 weeks gestation to a 27-year-old mother and 25-year-old father, her birth weight was 1210g and crown-heel length 42 cm, and she presented mild facial dysmorphism, highly arched eyebrows, cognitive impairment, and divergent strabismus. Affymetrix’s GeneChip Human SNP 6x6 v1.0A Genotyping array revealed several LCSH in 9 chromosomes (2p23.3; 2p13.2; 7q22.3; 8q24.23; 10q26.12; 11q24.1; 14q24.3; 14q32.2; 15q13.1; 21q11.2 and Xp21.1) and none pathogenic CNVs. The morbid genes included TMEM18, SNTG2, COL11A1, MYT1L, TSSC1, SOX11, RNF144B, DCTN1, PCGF1, LRRTM4, CUBN, TCTN2, KCNH2, TRPM1, MAP1A, GATM, UBR1, BTG3, TSPAN7, respectively. NCBI database analysis showed that mutations in COL11A1 gene are related to Malpuech-Michels-Mingarelli-Carnevale Syndrome (3MC Syndrome [MIM265050]) which has autosomal recessive inheritance. The results suggest that LCSH involving the COL11A1 gene may be responsible for the intellectual disability of this child. Although the several LCSH not diagnostic a specific condition, it indicate a causative gene located within of these regions. Thus, further studies require Exome Next Generation Sequencing to identify the mutations for candidate genes and confirm the pathogenicity of such changes. Homozygous regions identified using Exome Sequencing (ES) provides a powerful tool to identify recessive mutations in candidate genes.

Mate-pair sequencing analysis of karyotypically balanced chromosomal rearrangements associated with cryptic imbalances reveals additional structural variants and complex genomic reorganization typical of chromothripsis or replication mechanisms. A.C.S. Fonseca, A. Bonaldi, M. Bak, K.T. Abe, A.M. Vianna-Morgante, N. Tommervik. 1) Wilhelm Johannsen Centre for Functional Genome Research, Department of Cell Biology and Molecular Medicine, University of Copenhagen, Denmark; 2) Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil; 3) Sarah Network of Rehabilitation Hospitals, Brasilia, DF, Brazil. About 7% of karyotypically balanced chromosomal rearrangements (BCR) are associated with intellectual disability (ID), developmental delay (DD) and/ or congenital malformations (CM). Cryptic imbalances on the rearranged chromosomes appear as an important cause of the clinical phenotypes, being detected in about 23% of the BCRs analyzed by aCGH (Feenstra et al. Eur J Hum Genet 19:1152, 2011). Aiming at investigating if structural variants (SVs) undetected by aCGH co-occur with such imbalances, we applied mate-pair sequencing (MPS) to characterize eight sporadic BCRs (four intrachromosomal and four complex rearrangements) carried by patients with ID, DD and CM. Previously performed aCGH (180K, OGT) had revealed one or more imbalances (2.2kb to 5.8Mb) at or in cis to the BCR breakpoints. Mate-pair libraries were prepared using the Nextera kit (illumina), and paired-end sequenced (2×100 bp) on HiSeq2000. Filtering strategies and cluster analysis of discordant mate pairs were performed using an in-house script. SVDetect and Delly were used for SV prediction. In total 109 breakpoints (BPs; resolution <1kb) were identified, which resulted in 60 balanced and 29 unbalanced SVs. The number of BPs and SVs varied among the 8 BCRs (n1=6; n2=11; n3=10; n4=13; n5=11; n6=5; n7=12; n8=22). Among the 89 SVs identified by MPS, 25 unbalanced and 14 balanced SVs had been detected, respectively, by aCGH or karyotyping. The rate of SVs detected exclusively by MPS varied among the different BCRs, ranging from 3% to 41%. SVs detected using MPS included deletions (40/23), duplications (0/6), inversions (2/2), intrachromosomal direct (15/16) and inverted insertions (10/11), translocations (2/12), interchromosomal direct (7/8) and inverted insertions (10/11). All but one BCR had SVs identified only by MPS (n1=0; n2=10; n3=12; n4=4; n5=10; n6=1; n7=5; n8=22). One BCR (n4) harbored three SVs: deletions (4/23), duplications (0/6), inversions (2/2), intrachromosomal direct (15/16) and inverted insertions (10/11), translocations (2/12), interchromosomal direct (7/8) and inverted insertions (10/11). All but one BCR had SVs identified only by MPS (n1=0; n2=10; n3=12; n4=4; n5=10; n6=1; n7=5; n8=22). One BCR (n4) harbored three SVs: deletions (4/23), duplications (0/6), inversions (2/2), intrachromosomal direct (15/16) and inverted insertions (10/11), translocations (2/12), interchromosomal direct (7/8) and inverted insertions (10/11). All but one BCR had SVs identified only by MPS (n1=0; n2=10; n3=12; n4=4; n5=10; n6=1; n7=5; n8=22). One BCR (n4) harbored three SVs: deletions (4/23), duplications (0/6), inversions (2/2), intrachromosomal direct (15/16) and inverted insertions (10/11), translocations (2/12), interchromosomal direct (7/8) and inverted insertions (10/11). All but one BCR had SVs identified only by MPS (n1=0; n2=10; n3=12; n4=4; n5=10; n6=1; n7=5; n8=22). One BCR (n4) harbored three SVs: deletions (4/23), duplications (0/6), inversions (2/2), intrachromosomal direct (15/16) and inverted insertions (10/11), translocations (2/12), interchromosomal direct (7/8) and inverted insertions (10/11).

Novel H3K4me3 marks are enriched at human- and chimpanzee-specific cytogenetic structures. G. Giannuzzi, E. Migliavacca, A. Reymond. Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland.

Human and chimpanzee genomes are 98.8% identical within comparable sequence. They however differ structurally in nine pericentric inversions, one fusion that originated human chromosome 2 and content and localization of heterochromatin and lineage-specific segmental duplications. The possible functional consequences of these cytogenetic and structural differences are not fully understood and their possible involvement in speciation remains unclear. We show that subtelomeric regions—that have a species-specific organization, are more divergent in sequence, and are enriched in genes and recombination hotspots—are significantly enriched for species-specific histone modifications that decorate transcription start sites in different tissues in both human and chimpanzee. Human lineage-specific chromosome 2 fusion point and ancestral centromere locus as well as chromosome 1 and 18 pericentric inversion breakpoints showed enrichments of human-specific H3K4me3 peaks in prefrontal cortex. Our results reveal an association between plastic regions and potential novel regulatory elements.
Posters: Cytogenetics

813

3178T

3180T

Detection of copy number variation in single cells by next generation
sequencing. F. Kaper 1, C. April 1, X. Cai 1, H.Y. Chuang 1, T. Royce 1, H.
Joshi 2, C.E. Michel 2, P. Burns 2, J.B. Fan 1. 1) Illumina Inc, San Diego, CA;
2) Illumina Inc, Cambridge, UK.
The detection of copy number variations (CNVs) in single cells has many
applications, including cell lineage mapping, cancer heterogeneity research,
neuronal mosaicism research and preimplantation genetic screening (PGS).
Current protocols employ an upfront whole genome amplification (WGA)
step such as multiple displacement amplification (MDA), primer extension
preamplification PCR (PEP-PCR) or degenerate oligonucleotide-primed
PCR (DOP-PCR) followed by a next generation sequencing (NGS) library
preparation. Here we describe results obtained with a novel protocol that
combines NGS library preparation and WGA into one robust and rapid
protocol. Single, three and five cell samples were isolated through FACS
sorting or narrow-bore cell transfer pipetting from commercially available
human lymphoblastoid cell lines with well-described subchromosomal
abnormalities. NGS libraries were sequenced on a MiSeq® or a HiSeq®
2000 sequencing system. After alignment using Bowtie, only perfectly and
uniquely aligning reads were kept for tag counting. The reference genome
was divided into non-overlapping bins such that each bin is expected to
contain 100 uniquely mapping reads. Deviations from this expectation are
correlated with copy number. Unlike MDA-based methods, the protocol does
not result in overamplification pile ups, resulting in increased signal to noise.
Furthermore, the library fragments are generated in a random, uniform
fashion across the entire genome unlike the reproducible bias of PCR-based
WGA methods. Therefore genome coverage is correlated with increasing
cell number. Using 1 million reads per NGS library, confirmed CNVs of 4Mb
were reproducibly detectable at single, three and five cell input levels. The
entire work flow from the start of sample preparation to available sequencing
data takes 9 hours when using a MiSeq sequencing system. For the detection
of smaller events, NGS libraries were sequenced to greater depth on a
HiSeq 2000 system. In conclusion, this technique provides a simple and
rapid protocol to detect CNVs in single cells by NGS; experiments to further
validate the method are currently ongoing.

Array-based analysis reveals partial 11q14 duplication in a familial
case with intellectual disability, short stature and mild dysmorphic
features. R. Satomi 1, M. Ohta 1, K. Matsumoto 1, T. Miyashita 1, M. Hayata 1,
H. Shiraku 1, N. Matsuda 1, K. Kurosawa 2, K. Enomoto 1. 1) Department of
Pediatrics, JA Toride medical center, Toride, Ibaraki, Japan; 2) Division of
Medical Genetics, Kanagawa Children’s Medical Center, Yokohama, Japan.
Partial chromosomal duplications of 11q without additional imbalances are
rare. A total of seven cases associated with 11q14 duplication without any
other chromosomal involvement have been reported. The phenotypes in
these previous cases were characterized by a wide variety of clinical findings
and variable intellectual disability or psychomotor delay with facial dysmorphic features. This may be due in part to varying size of the duplicated
segments. Herein, we report a new case with a 9.6Mb duplication involving
11q14.1 to q14.3 defined by array-based comparative genomic hybridization
analysis: arr 11q14.1q14.3(79,131,430-88,758,610)x3 (hg19). The patient
is a 9-year-old boy who exhibits developmental delay, intellectual disability,
minor dysmorphic features, and short stature. He was born at 32-week
gestation. His birth weight and length are 1854g (25-50th percentile) and
45cm (75-90th percentile), respectively. He was healthy and uncomplicated
in infancy. Initial visit to our hospital was at age 9, complaining short stature.
His height was 119.5cm (below -2.5SD). Physical examinations found that
he also had mildly upslanting palpebral fissures and cryptorchidism. Subsequent examinations reveal that his intelligence quotient is 50 and ADHD
Rating Scale based on DSM IV criteria is 26. His brother also has intellectual
disability, learning disorder, and short stature. Furthermore, we have discovered that he has some relatives presenting similar features on his mother’s
side. Though, this family history strongly shows as if they have the X-linked
recessive inheritance, the copy-number change is even in the autosome. Our
observations delineate the phenotypic spectrum associated with a clearly
defined duplication of chromosome 11q14. Phenotype-genotype correlation
will be discussed in view of all the reported cases and our case’s family members.

3181T
3179T
Mosaicism in a Mosaic: Reduced capacity of female X chromosomes
to resist age-related structural erosion. M.J. Machiela 1, W. Zhou 1,2, M.
Dean 3, J.N. Sampson 1, N. Rothman 1, N.D. Freedman 1, S. Wacholder 1,
M.A. Tucker 1, F.X. Real 4,5, L.A. Perez-Jurado 5,6, M. Yeager 1,2, S.J. Chanock 1, NCI GWAS Mosaicism Consortium. 1) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA; 2) Cancer
Genomics Research Laboratory, Leidos Biomedical Research Inc., Frederick, Maryland, USA; 3) Laboratory of Experimental Immunology, Center for
Cancer Research, NCI-Frederick, Frederick, Maryland, USA; 4) Spanish
National Cancer Research Centre (CNIO), Madrid, Spain; 5) Departament de
Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona,
Spain; 6) Centro de Investigación Biomédica en Red de Enfermedades
Raras (CIBERER), Barcelona, Spain.
Previous studies characterizing large clonal mosaic events have not
focused on the X chromosome. Therefore, we analyzed deviations in log2
intensity ratio and B allele frequency on the X chromosome using existing
genome-wide association data of 38,581 female participants from cancer
association studies. We detected 131 mosaic events ≥2 Mb in size on the
X chromosome of 104 females (0.27%). The rate of mosaicism per basepair was approximately 5 fold higher for the X chromosome than the autosomes (1.10 vs 0.22 events per 100,000 Mb). Similar to autosomal events,
most detected X events were mosaic copy neutral losses of heterozygosity
(53%) followed by mosaic losses (34%) and mosaic gains (14%). Contrary
to autosomal observations, mosaic X events tend to be larger in size (mean
86.1 vs 44.6 Mb) and far more likely to involve the whole chromosome (49%
vs 5%). Interestingly, when compared to individuals with no X mosaicism,
individuals with detected X events were observed to have a 15 times
increased odds of additionally harboring an autosomal event (p-value=
2.20×10 -8). Frequency of X mosaicism ranged from 0.1% in 50 year old
women to 0.4% in women 75 years or older (trend p-value=0.0066). While no
overall cancer association was observed (p-value=0.30), subtype analyses
suggest an association of X mosaicism with increased lung cancer risk (pvalue=0.04), an association also observed for autosomal mosaicism, but
further studies are required to confirm this finding. Methylation arrays were
hybridized for a subset of 10 samples in an attempt to localize events to
either the active (Xa) or inactive (Xi) X chromosome. Initial results indicate
8 out of 10 events have a shift in methylation profile favoring more copies
of Xi and fewer copies of Xa; however, the observed trend was not significant
in our limited sample (p-value=0.11). Our results indicate a higher per base
pair rate of mosaicism on female X chromosomes than on autosomes and
show that age-related erosion of the genome, previously observed in studies
as mosaic events on the autosomes, also occurs on female X chromosomes.

15q11.2 duplication encompassing only the UBE3A gene is associated
with developmental delay and neuro-psychiatric phenotypes. A. Noor 1,
L. Dupuis 2, K. Mittal 3, T. Stockley 1,4, J.B. Vincent 3, R. Mendoza-Londono 2,
D.J. Stavropoulos 1,4. 1) The Hospital for Sick Children, Department of Paediatric Laboratory Medicine, Toronto, Ontario, Canada; 2) The Hospital for
Sick Children, Department of Pediatrics, Division of Clinical and Metabolic
Genetics, University of Toronto, Toronto, Ontario, Canada; 3) Molecular
Neuropsychiatry & Development Lab, Neurogenetics Section, The Campbell
Family Brain Research Institute, Centre for Addiction and Mental Health,
Toronto, Ontario, Canada; 4) Department of Laboratory Medicine and
Pathobiology, University of Toronto, Toronto, Ontario, Canada.
Due to the presence of low copy repeats (LCRs) in chromosome 15q11-q13
region is prone to recurrent structural rearrangements including deletions,
duplications and inversions. This region is also subject to genomic imprinting,
thus associated phenotypes show parent-of-origin specific effect. Interstitial
duplications of the 15q11-q13 region with maternal imprints are associated
with a wide spectrum of neuro-psychiatric disorders including, autism spectrum disorders (ASD), developmental delay, learning difficulties, schizophrenia and seizures. On the other hand, individuals with duplications of the
same region with paternal imprints are often phenotypically normal or in
rare instances mildly affected. These observations suggest that a dosage
sensitive imprinted gene or genes within this region that underline the risk
for neuro-psychiatric phenotypes. Recent studies have shown that UBE3A
is the only gene within this region that is solely expressed from the maternal
allele in mature neurons. To date, no case with duplication of only UBE3A
has been reported, therefore the contributions of this gene to the 15q11q13 duplication phenotypes remains unclear. We present a female patient
with developmental delay in whom we identified a 129 Kb duplication in
chromosome region 15q11.2 encompassing only the UBE3A gene. The
duplication was found to be maternally inherited in the proband. We further
tested the segregation of this duplication and it was found to be segregating
with a neuro-psychiatric phenotype in four generations. Expression analysis
in cultured fibroblast confirmed the over-expression of the UBE3A in the
proband compared to age-matched controls. Our study shows for the first
time the clinical features associated with over-expression of UBE3A in
humans, and underscores the significance of this gene in the phenotype of
individuals with 15q11-q13 duplication.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.


A duplication of the CDKL5 gene identified in a boy with developmental delay with autistic behavior, short stature and microcephaly. K. Takano1, T. Nishimura2, K. Waku3, S. Takahashi3, Y. Inaba2, T. Kosho2, Y. Fukushima1. 1) Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan; 2) Department of Pediatrics, Shinshu University School of Medicine, Matsumoto, Japan; 3) Department of Pediatrics, Asahikawa Medical University, Asahikawa, Japan.

Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5) gene [MIM#300203] cause neurological developmental disorders involving early infantile epileptic encephalopathy, atypical Rett syndrome, severe intellectual disability (ID), infantile spasms and epilepsy. Chromosomal microdeletions including CDKL5 also cause similar neurodevelopmental disorders. Although duplications of the short arm of X chromosome including CDKL5 have been reported in patients with ID, dysmorphic features and/or epilepsy, the effect of CDKL5 copy number gains on neurological development remains unclear because of those duplications encompassing other X-linked ID genes. Here, we present a boy with developmental delay and autistic behavior having a microduplication including the only gene CDKL5 that could affect developmental brain. The patient is a 3-year-old boy, born at 38 weeks of gestation with the weight as 2.226 g (2.1 SD), length as 46 cm (-1.0 SD), and OFC as 29 cm (-2.8 SD). His mother has normal intelligence and short stature with the height as 145 cm. He had surgical repair of inguinal hernia at the age of 9 month. His motor development was delayed with sitting alone at 7 months and walking independently at 1 years and 1 month. His DQ was 51 at 1 year and 8 months and he presented autistic behavior. He began to refuse food at 2 years and 8 months, and nasogastric tube feeding was started. At 3 years, he showed growth impairment with the height as 94.0 cm (-3.1 SD), weight as 11.5 kg (-1.6 SD), and OFC as 46.5 cm (-2.1 SD). G-banded chromosomes were 46,XY. Chromosomal microarray analysis revealed a 200-kb duplication in Xp22.13 spanning along the most of the genomic region of CDKL5 which was derived from his mother’s paternal duplication of CDKL5, implying neurodevelopmental effects of increased gene dosage of CDKL5.

CMIP, a Strong Candidate Gene Involved in the Autism Spectrum Disorder. R. Politi1, R. Burns2, R. Pasierb1, P. Papenhausen2. 1) Dynagene, Div Lab Corp America, 3701, Kirby Drive, Houston, TX 77098; 2) Labcorp, 1904 Alexander Drive, Research Triangle Park, NC 27709.

Although deletions in 22q13 deletions may be a comorbid disorder with ASD, the role of CMIP haploinsufficiency in ASD patients. Additionally, our study is the first report on a microduplication of CDKL5, implying neurodevelopmental effects of increased gene dosage of CDKL5.

Complex mosaic chromosome rearrangement in a patient with Phelan-McDermid Syndrome. C. Purnien1, W. Froehlich2, J. Bernstein3, R. Dolmetsch4, J. Hallmayer1, A.E. Urban1,2. 1) Department of Psychiatry & Behavioral Sciences, Stanford University, Palo Alto, CA; 2) Department of Genetics, Stanford University, Palo Alto, CA; 3) Department of Pediatrics, Stanford University, Palo Alto, CA; 4) Department of Neurobiology, Stanford University, Palo Alto, CA; 5) current address: Novartis Institutes for Biomedical Research, Cambridge, MA.

Phelan-McDermid Syndrome is caused by deletions or unbalanced translocations of chromosome band 22q13, for which few mosaic cases have been described. Here, we report a detailed genetic analysis of a patient with a 22q13 terminal deletion showing a complex rearrangement containing three different breakpoints. The affected patient exhibits global developmental delay, moderate intellectual disability, autism, and severely impaired speech. She was initially diagnosed with Phelan-McDermid Syndrome using multiplex ligation-dependent probe amplification (MLPA). Molecular karyotyping of a fibroblast sample using a high-density Illumina HumanOmniExome array revealed a complex rearrangement of a 1 Mb region missed by the MLPA array pointing to two or possibly three different cell populations in the sample. This finding was followed up by digital droplet PCR and microsatellite analysis. Subsequent analyses include Fluorescence in situ hybridization (FISH) as well as targeted chr22 capture followed by deep sequencing for fine-mapping of the breakpoints. The combined approach revealed a complex rearrangement consisting of two different breakpoints of the 22q13 terminal deletion (BP1 and BP2). Furthermore, the patient showed an additional rearrangement between BP2 and a third breakpoint (BP3) pointing to a possible third cell population. To the best of our knowledge, this case represents the first mosaic chromosome rearrangement of this kind in Phelan-McDermid Syndrome. This case also highlights the effective use of established and new laboratory technologies for studying complex chromosomal rearrangements.
3186T
Cytogenetic studies of the drug Methotrexate (MTX) on the blood lymphocytes of colon cancer patients. Z. MT Jafar. Biotechnology Center, Ministry of Science and Technology - Iraq, Baghdad, Baghdad, Iraq.

The aim of the study is to investigate the cytogenetic parameters in order to define the effect of methotrexate (MTX) on blood lymphocytes, manifested by blast genic index (BI), mitotic index (MI), replicative index (RI), and sister chromatid exchanges (SCEs). The collected blood samples were cultivated in RPMI 1640 supported with 10% fetal bovine serum, 0.3 ml of phytohaemagglutinin (PHA), antibiotics (penicillin and streptomycin) and different concentrations of MTX (0.2, 0.5, 0.1, 2, 4, 8 µg/ml). The results showed significant reduction in BI and MI and RI indices with concentrations (1, 2, 4, 8 µg/ml) when comparing with the control. While in 0.2, 0.5, 0.1 µg/ml of MTX gave no significant decrease when compared with the controls.

The results of BI shows significant reduction according to increasing of conc. Of drug, it gave 26% in 0(0) conc. and increase to 19.1% in 0.2 µg/ml and 11% at 2µg/ml and 4.5% in 8µg/ml. The replicative index also show reduction which is proportional with increasing conc. Of MTX, it gave 0.72% at zero conc. and 0.44% at conc. of 0.2 µg/ml of MTX and 0.22% at conc. of 0.5 µg/ml and 0.02% at conc. of 1 µg/ml of MTX. On the other hand the SCE frequency in colon cancer samples which gave 8.69 in zero conc. and 8.81 in 0.2 µg/ml, 10.66 at conc. of 0.5µg/ml. The SCEs in the other conc. of drug can not be detected due to the toxicity of drug in high concentrations. In control samples, the SCEs frequency gave 5.45 at zero conc., 5.66 at 0.2 µg/ml and 6.56 at conc. of 0.5 µg/ml. The mutation fraction increased in colon cancer patient which gave 8.93 in 0.2 µg/ml, and 15.72 in 0.5 µg/ml and 4.1 in 1ug/ml while in control gave 3.17 in 0.2 µg/ml and 0.21 in 0.5 µg/ml and 0.74 in 1µg/ml. The conclusion showed decrease in the (MI), (BI) and (RI) and increase in SCEs frequencies and mutation fraction in colon cancer patients when compared with control.

The presence of mutant cells in lymphocytes of colon cancer that resistant to MTX, give an idea about amplification in DHFR gene that regulate the activity of DHFR enzyme and also indicate the resistance of cancer cells to drugs.

3187T
Molecular cytogenetic characterization of a patient diagnosed with dimorphic anemia carrying de novo rare ring chromosome 7 along with t(7;9). S.K. Bhattacharya, V. Lal. Cytogenetics, Dr. Lal Path Labs. Pvt. Ltd., Newdelhi, Delhi, India.

Cytogenetic abnormalities are described in a few patients with typical dimorphic anemia (DA). The possible clonal nature of this disease is still controversial. We present a case of a derivative chromosome 7 formed by a ring chromosome 7 and t(7;9). The patient is phenotypically normal. This is a unique case report from India, on occurrence of r(7) along with t(7;9).

The patient is born to cytogenetically normal parents. We have performed conventional cytogenetic technique using GTG banding and molecular FISH technique using LSI D7S486 probe and WCP for chromosome 7 and 9. The conventional cytogenetic technique using LSI D7S486 probe and WCP for chromosome 7 and 9 has been studied by molecular techniques. To our knowledge, this is the first report of t(7) and t(7;9) found in a child with dimorphic anemia, characterized by molecular cytogenetic analysis.

3188T
Reciprocal microdeletions and microduplications of CNTN6 gene (3p26.3) in patients with intellectual disability. I. Lebedev1, A. Kashevarova1, N. Salyukova2, L. Nazarenko2. 1) Laboratory of Cytogenetics, Institute of Medical Genetics, Tomsk, Tomsk region, Russian Federation; 2) Laboratory of Hereditary Pathology, Institute of Medical Genetics, Tomsk, Tomsk region, Russian Federation.

To date several reciprocal microdeletion / microduplication syndromes or «genomic sister disorders» have been described. Due to application of high resolution microarray technologies new submicroscopic chromosomal rearrangements are being published. The reciprocal rearrangements are particularly valuable, since they allow to determine new dosage-sensitive pathogenic genes in human genome. We performed array-CGH analysis for 80 patients with idiopathic intellectual disability using CGH Microarray Kits 4x44K and 8x60K (Agilent Technologies, USA). Pathogenically significant cases were confirmed by qPCR. We present two patients with microdeletion (369 kb) and patient with microduplication (766 kb) at 3p26.3 containing the only gene - CNTN6. The microduplication was inherited from apparently healthy father and grandmother. The brother and sister with microdeletion were orphans. The microdeletion was the same size and localization in both sibs. Recently, the microduplication in 3p26.3 (containing CNTN6) has been shown to be associated with autism spectrum disorders. Contactin 6 is also suggested to play a neuroprotective role in ischemic injury and contribute to granule cell maturation and/or synaptic formation in the developing cerebellum. Considering the experiments with mice we found myotonic syndrome, late development of sit and walk ability in the anamnesis, current fine motor skills impairment and dysarthria in patient with dup3p26.3. Dysarthria was also observed in the boy with del3p26.3. Obviously, CNTN6 can be a novel pathogenic gene associated with autism spectrum disorder, intellectual disability, and motor functions impairment. This study was supported by Russian Scientific Foundation (grant N 14-15-00772).

3189T
Co-occurrence of non-mosaic trisomy 22 and inherited balanced t(4;6)(q33;q23.3) in a live-born female: Case report and review of the literature. J. Liu1, F. Kehinde2, C. Anderson3, J. McGowan4, R. Jethva5, M. Wahab6, A. Glick7, M. Stemer1, J. Fascasio1. 1) Dept Path & Lab Med., St Chris Hosp for Children, Drexel University College of Medicine, Philadelphia, PA; 2) Department of Neonatology, St. Christopher’s Hospital for Children, Drexel University College of Medicine, Philadelphia, PA 19134; 3) Department of Medical Genetics, St. Christopher’s Hospital for Children, Drexel University College of Medicine, Philadelphia, PA 19134.

Trisomy 22 is the third most common autosomal trisomy occurring in about 0.4% of all clinically recognized pregnancies. Complete non-mosaic trisomy 22 is extremely rare in live births, and most of the affected children died before one year of age. To date, only 29 cases have been reported in the literature, and none of them carried an additional genetic lesion. In this report, we describe clinical presentations, cytogenetics and cytogenomics findings in a live-born female with a complete non-mosaic trisomy 22 as well as a paternally inherited, balanced reciprocal chromosomal rearrangement t(4;6)(q33;q23.3). The proband manifested features commonly seen in individuals with non-mosaic trisomy 22 such as intrauterine growth retardation (IUGR), single umbilical artery, cranial abnormalities, short neck, cleft lip and palate, dysmorphic ears, congenital heart defects, dysplastic kidneys, genital anomalies, hypoplastic nipples, and digital malformations. In addition, she had lobar holoprosencephaly, aqueductal stenosis, limb and eye problems that have not been associated with complete trisomy 22. The proband died 35 days after birth due to complex heart disease and renal failure. We are hereby expanding cytogenetic and clinical spectrum of this rare chromosome disorder. Clinical features of live-born children with non-mosaic trisomy 22 are reviewed and compared to those in our proband. The impact of genomic content in relation with the survival of trisomies in humans is also discussed.
Identification and Characterization of Marker Chromosome in a Patient with Turner’s Syndrome. S. Sharma, P. Srivastava. Indraprastha Apollo hospital, New Delhi, New Delhi, India.

Introduction: We are describing a 14 year old female who was referred to us with severe growth retardation, primary amenorrhea, poor development of secondary sexual characteristics, web neck, increased FSH and LH. A provisional diagnosis of Turner syndrome was made clinically. On routine cytogenetic analysis she was found to be mosaic for 45,X and 46,X,+marker chromosome. The presence of a marker chromosome in Turner syndrome generally implicates a sex chromosome origin but it may also originate from a non-sex chromosome. If the marker chromosome originates from Y, the patient is at risk of developing Gonadoblastoma. Therefore, FISH test was performed for marker chromosome to rule out the presence of Y chromosome so as to delineate the risk of gonadoblastoma in the patient. Methods: Peripheral blood sample was collected from the proband and cytogenetic examination was carried out using standard techniques of 72hr phytohemagglutinin stimulated peripheral blood lymphocyte culture. FISH was done using centromere enumerating probe for X and Y chromosome. Results: The marker chromosome identified in the chromosomal analysis was found to be centromeric part of X chromosome on FISH analysis, thus ruling out the Y chromosome and the risk of gonadoblastoma. Conclusion: We strongly recommend that FISH should be done in all cases of Turner syndrome where a marker chromosome is identified on chromosomal analysis so as to rule out the risk of gonadoblastoma and to provide appropriate genetic counseling. Keywords: Turner’s Syndrome, Marker chromosome, Cytogenetics.


Chromosomal abnormalities account for approximately 50 to 70% of sporadically first trimester miscarriages with the remainder usually classified as idiopathic. A majority of the abnormalities are numerical with structural chromosomal abnormalities or other genetic mechanisms accounting for less than 10% of cases. Cytogenetic analysis of products of conception (POC) is essential to determine the cause of pregnancy loss, aid the prenatal diagnosis of subsequent pregnancies, calculations of recurrence risks, and delineation of the different kinds of abnormalities causing spontaneous abortions. We report our case series of over 5000 POC samples referred to the North Shore University Hospital and compare our data with the published data.

Haplo-insufficiency and triploid-insensitivity of the same 6p25.1p24.3 region in a family. Z. Qiu1, L. Jeng2, A. Slavotinek3, J. Yu1. 1) Dept. of Pathology, University of California San Francisco, San Francisco, CA; 2) Dept. of Medicine, Pediatrics and Pathology, Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD; 3) Dept. of Pediatrics, University of California San Francisco, San Francisco, CA.

Chromosome 6pter-p24 deletion syndrome is a clinically recognized chromosomal disorder (OMIM #612582). Of more than 40 reported cases of this disorder, only four showed interstitial deletions within the 6p25-3p23 region. To date, only two interstitial duplications involving the 6p25-3p23 region have been reported. We present a family carrying both deletion and duplication within the 6p25.1p24.3 region (4,745,144-10,384,769 bp, Build hg19). A 5.6 Mb interstitial deletion within the region was detected by array CGH in a 26-month-old female proband who presented learning disabilities affecting motor and speech skills, bilateral conductive hearing loss, dysmorphic features including hypertelorism, a left preauricular pit, downturned corners of the mouth and small teeth, a hemangioma of the right neck and a small pectus excavatum. Instead of deletion, her typically developing 5-year-old brother carried an interstitial duplication of the same region. High-resolution G-banding and FISH studies demonstrated that the deletion and duplication were apparently inherited from their 35-year-old mother who carries both cell lines of the deletion (~70%) and duplication (~10%), as well as a normal cell line (~20%) in her peripheral blood. The proband showed phenotypes reportedly associated with the chromosome 6pter-p24 deletion syndrome. Their mother had an embolic stroke at age of 26 years, a patent foremen ovale, and an arterial septal aneurysm. Our findings are consistent with the reported haplo-insufficiency of the 6p25.1p24.3 region, and suggest possible triploid-insensitivity of the same region. In addition, the cells with the duplication may compensate the phenotypic effect of the cells with the deletion as inferred by the maternal phenotype and karyotype that was most likely derived from a rare somatic inter-chromatid non-allelic recombination that occurred early in embryogenesis.
hit. We hypothesize that our patient had the initial TEL/AML1 fusion with second hit following the initial TEL/AML1 fusion, or is by itself the primary apy. Since our patient presented to us following BMT, we cannot ascertain late-onset relapses by its ability to evade complete elimination during ther-

Furthermore, the manifestation of relapse following remission clearly sup-

ing deletion of the TEL gene region on the non-rearranged TEL allele. Our modified method did not miss any abnormalities, produced clear results and increased confidence in setting tighter cutoff limits.

3196T Homozygous Deletion of TEL (ETV6) in Childhood Acute Lymphoblas-

The fusion of TEL/AML1 (ETV6/RUNX1) gene regions resulting from the t(12;21), is the most common genetic abnormality observed in childhood ALL. The neoplasm in these patients persists not by the initial fusion event, but in most cases by a secondary or subsequent leukemogenic event involv-

ing deletion of the TEL gene region on the non-rearranged TEL allele. Although variant patterns are seen by fluorescence in situ hybridization (FISH) including extra copies of AML1, e.g. fusions of TEL/AML1 and rearrangements of TEL, deletion of the second TEL allele is the most com-

mon and can have significant prognostic implications. In these cases, persis-
tence of the original pre-leukemic clone may indicate proliferative advantage and lead to prolonged remission followed by periods of relapse as the residual secondary clones emerge. We report a case on a 26 year old male with ALL harboring homozygous deletion of the TEL gene region. To our knowledge, this is the first reported case of a homozygous TEL deletion without a concomitant TEL/AML1 fusion. His initial presentation to our labora-
tory was in 2003 status post bone marrow transplant and showing normal result by chromosome analysis (CA). Nine years later (07/2012) the patient returned with relapse. CA showed a complex karyotype with multiple abnor-

malities. FISH analysis with ALL panels showed 3 copies of cMYC gene region, deletion of p16 gene region and homozygous deletion of TEL gene region with 6 copies of AML1. Follow-up chromosome and FISH studies up until 01/2014 were all normal. Homozygous deletion of the TEL gene region in our patient may suggest that absence of TEL gene region on both alleles behaves as a strong adverse factor to cases with the classic TEL deletion of only one allele. Furthermore, the manifestation of relapse following remission clearly sup-

ports TEL deletion being the principle causative factor for subsequent and late-onset relapses by its ability to evade complete elimination during ther-

apy. Since our patient presented to us following BMT, we cannot ascertain if the homozygous TEL deletion, which was never before reported, is the second hit following the initial TEL/AML1 fusion, or is by itself the primary hit. We hypothesize that our patient had the initial TELAML1 fusion with deletion in this TEL allele at diagnosis and this in turn may have resulted in the loss of second TEL allele similar to loss of heterozygosity in other can-
cers.
3199T

Two new cases of chromosome 7p22.1 microduplication detected by array CGH. R.G. Hutchinson 1, J. Nicholl 1, L. Montgomery 1, C. Barrett 1, K. Bourdony 1, S. Yu 1 1 Cytogenticists, SA Pathology, Adelaide, South Australia; 2) Clinical Genetics, SA Pathology at WCH, King William Rd, North Adelaide; 3) Neurology Department, TQEH, Woodville.

Array CGH has enhanced our ability to detect rare microduplications previously too small to be observed cytogenetically. Cytogenetically detectable aberrations are necessarily large and the clinical features associated with them have been variable. The use of microarray technology has enabled the phenotypic features more generally associated with duplication of a chromosomal region to be assigned to smaller chromosomal areas and even specific genes. Microarray also detects abnormalities where the pathogenicity of the change may be unclear. Interpretation of these rare abnormalities relies on the cooperation of institutions involved in genetic diagnosis in sharing rare cases where the anomaly is similar. Two cases of 7p22.1 microduplication (1.7Mb and 1Mb) detected by array have been reported in the literature. [1,2] We present two patients with 7p22.1 microduplication for comparison with the reported cases. Our cases include a 4Mb (case 1) and a 700kb (case 2) microduplication. The 700kb duplication is almost entirely contained within the 1mb duplication reported by Preiksaiten et.al. (2012) [2], while both published cases are encompassed within the larger duplication. Brief clinical descriptions of our cases are as follows. Case 1: A 4Mb duplication was detected in a 2 1/2 year old girl with global developmental delay particularly language, autistic features, dysmorphic facies and renal pelvis dilatation. Case 2. A 700kb duplication was detected in a 22 year old male reported to have intellectual disability with language delay and dysmorphic features including macrocephaly. Parental studies were not performed and a balanced rearrangement could not be detected by chromosome analysis. Our data shows that CMA provided a comprehensive cytogenetic diagnosis with no apparent decrease in sensitivity compared to a full chromosome analysis. In the two cases with balanced rearrangements, the CMA analysis was not only sufficient for interpretation of the phenotype, but also provided a comprehensive cytogenetic diagnosis with no apparent decrease in sensitivity compared to a full chromosome analysis.

4Mb (case 1) and a 700kb (case 2) microduplication. The 700kb duplication is almost entirely contained within the 1mb duplication reported by Preiksaiten et.al. (2012) [2], while both published cases are encompassed within the larger duplication. Brief clinical descriptions of our cases are as follows. Case 1: A 4Mb duplication was detected in a 2 1/2 year old girl with global developmental delay particularly language, autistic features, dysmorphic facies and renal pelvis dilatation. Case 2. A 700kb duplication was detected in a 22 year old male reported to have intellectual disability with language delay and dysmorphic features including macrocephaly. Parental studies were not performed and a balanced rearrangement could not be detected by chromosome analysis. Our data shows that CMA provided a comprehensive cytogenetic diagnosis with no apparent decrease in sensitivity compared to a full chromosome analysis. In the two cases with balanced rearrangements, the CMA analysis was not only sufficient for interpretation of the phenotype, but also provided a comprehensive cytogenetic diagnosis with no apparent decrease in sensitivity compared to a full chromosome analysis.

3202T

Partial Microduplication of PTEN in a girl with multiple congenital abnormalities including agenesis corpus callosum, C.P. Oliveira 1, 2, R.S.H.G Silva 3, R.O.A. Benicio 4, M.T.O. Cardoso 1, S.F Oliveira 1, A. PicTaylor 4, J.F. Mazzuè 2 1 Programa de Pós Graduação em Ciências Médicas, Faculdade de Medicina, Universidade de Brasília, Brasília, Brasil; 2) Núcleo de Genética (NUGEN), Hospital de Apoio (HAB), Secretaría de Saúde, Distrito Federal, Brasil; 3) Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasilia (UnB), Distrito Federal, Brasil; 4) Programa de Pós Graduação em Ciências Médicas e Biotecnologia, Universidade Católica de Brasília (UCB), Distrito Federal, Brasil.

The corpus callosum is the largest commissure connecting the cerebral hemispheres. Agenesis corpus callosum (ACC) is a result of 45,X and is one of the most common brain malformations with an incidence of 0.5-70 in 10,000, and can occur isolated or as part of many syndromes. Although ACC is predominantly genetic, few genes have as yet been identified and chromosome abnormalities have been described in a large number of patients. We report a patient with multiple congenital malformations including agenesis corpus callosum, asymmetrical face, microcephaly, hypertelorism, microphthalmia and short neck. She has two accessory nipples, asymmetrical limbs, sacral dimple and aural hemangioma. ECG showed interatrial communication. To screen for chromosomal imbalances high-resolution array Genomic Hybridization Assay was performed using the Affymetrix® CytoScan™ HD platform. Analysis revealed a 49 Kb of novel microduplication 10q23.31 (89,629,908-99,679,426 - hg19) including exon 2 of PTEN gene. The phosphorylated and activated PTEN (phosphotyrosine (PTEN) gene) on chromosome 10 is also upregulated as a tumor suppressor gene that is frequently mutated in human cancers. Further studies showed that PTEN plays an important role in brain development involved in cell migration, cell number and cell size regulation. Clinical manifestations of PTEN microduplication (10q23) in both males and females include Cowden, Bannayan-Riley-Ruvalcaba, Lhermitte-Duclos, and Proteus Syndromes. The patient here reported shows a complex phenotype that does not correspond to any of the syndromes listed above and includes ACC. Thickening of corpus callosum is one of the most common features besides those compatible with her Turner syndrome karyotype. The mosaic condition and the unidentified regions including in the extra chromosomal mosaicism. We have studied a cohort of 52 patients under these clinical criteria. We present a patient with 2 abnormal cell lines, one with 45,X and the other with a small supernumerary marker chromosome. The proband is a 2 year and 7 months old female, first child of healthy non-consanguineous parents. She has psychomotor delay and suffers from epileptic seizures. Height is 88 cm (PC<25), weight 12.6 kg (PC<25), and head circumference 55 cm (PC<25). She has macrocrania, high hair implantation, prominent forehead, hypertelorism, telecanthus, left dacryostenosis, nasal bridge with broad root, bulbous nose and anteverted nares. She presents hyperpigmented linear macules in arms and legs following Blaschko lines, and a dorsal inverted V macule. Cytogenetic analysis with GTG bands in lymphocytes revealed mos 47,XX,+mar[28]/45,X[6]/46,XX[16]. In fibroblasts from both types of skin, two cell lines were detected; light skin: mos 45,X[35]/47,XX,+mar[2]/46,XX[6], dark skin: mos 45,X[20]/ 46,XX[10] 45,X[6]/46,XX[13]. The origin of the marker chromosome was determined by FISH using DXZ1, p11.23-p11.22 (4968723-49885097) and 1q11.2-p11.22 probes. The marker chromosome comprised Xp11.22-p11.23 region. P. Pérez-Vera 1, V. Ulloa-Avilés 1, E. Lieberman-Hernández 1, C. Duran McKenzie 1, S. Gómez Carmona 1, R. Cruz-Alcivar 1, M. Hidalgo 1, A. Roque-León 1, V. Del Castillo-Ruiz 1, C. Salas-Labadía 1 1) Genética Humana, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico; 2) Dermatologia, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico.

3202T

Pigmentary mosaicism with 45,X and an extra marker containing the Xp11.22-p11.23 region. P. Pérez-Vera 1, 2, V. Ulloa-Avilés 1, E. Lieberman-Hernández 1, C. Duran McKenzie 2, S. Gómez Carmona 2, R. Cruz-Alcivar 1, M. Hidalgo 1, A. Roque-León 1, V. Del Castillo-Ruiz 1, C. Salas-Labadía 1 1) Genética Humana, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico; 2) Dermatologia, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico.

Pigmentary mosaicism characterized by hyperpigmented macules following Blaschko lines. These lesions are strongly associated with the chromosomal mosaicism. We have studied a cohort of 52 patients under these clinical criteria. We present a patient with 2 abnormal cell lines, one with 45,X and the other with a small supernumerary marker chromosome. The proband is a 2 year and 7 months old female, first child of healthy non-consanguineous parents. She has psychomotor delay and suffers from epileptic seizures. Height is 88 cm (PC<25), weight 12.6 kg (PC<25), and head circumference 55 cm (PC<25). She has macrocrania, high hair implantation, prominent forehead, hypertelorism, telecanthus, left dacryostenosis, nasal bridge with broad root, bulbous nose and anteverted nares. She presents hyperpigmented linear macules in arms and legs following Blaschko lines, and a dorsal inverted V macule. Cytogenetic analysis with GTG bands in lymphocytes revealed mos 47,XX,+mar[28]/45,X[6]/46,XX[16]. In fibroblasts from both types of skin, two cell lines were detected; light skin: mos 45,X[35]/47,XX,+mar[2]/46,XX[6], dark skin: mos 45,X[20]/46,XX[10] 45,X[6]/46,XX[13]. The origin of the marker chromosome was determined by FISH using DXZ1, p11.23-p11.22 (4968723-49885097) and 1q11.2-p11.22 probes. The marker chromosome comprised Xp11.22-p11.23 region. P. Pérez-Vera 1, V. Ulloa-Avilés 1, E. Lieberman-Hernández 1, C. Duran McKenzie 2, S. Gómez Carmona 2, R. Cruz-Alcivar 1, M. Hidalgo 1, A. Roque-León 1, V. Del Castillo-Ruiz 1, C. Salas-Labadía 1 1) Genética Humana, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico; 2) Dermatologia, Instituto Nacional de Pediatría, Mexico, Distrito Federal, Mexico.
The identification of ring 13 chromosome and breakpoint region at a Brazilian child requires karyotype, FISH and microarray analysis. I.P. Pinheiro1,2, L.B. Minasi1,2, R.V. Meloni3,2, J.G. Almeida3,2, D.M.C. Cunha1,2, C.L. Ribeiro1,2, G.P. Silva1,2, M.G. Brasil1,2, D.M. e Silva1,2, C.C. da Silva1,2,3,4, A.D. da Cruz1,2,5,6 1 Pontifical Catholic University of Goias, Genetics Master’s Program; 2 Pontifical Catholic University of Goias, Department of Biology, Replicon Research Group; 3 Federal University of Goias, Biotechnology and Biodiversity PhD Program; 4 Clinical Hospital of Federal University of Goias; 5 Federal University of Goias, Genetics and Molecular Biology Master’s and PhD Programs; 6 Human Cytogenetics and Molecular Genetics Laboratory of the Faculty of Health Sciences. Deletions of chromosome regions are generally the result of double strand chromosome breaks with loss of produced acentric fragments during the next cell division, and could lead to ring chromosome. Patients with (r13) presented different phenotypic abnormalities at different breakpoints. Here, we report the first postnatal diagnosis in Central Brazil of a girl with cytogenetic abnormalities involving chromosome 13 using 3 different laboratory methodologies. At the age of 8 months, she presented low tubular renal function in both kidneys, craniofacial dysmorphisms, hypotelorism, bilateral microphthalmia, atrophy of the optic nerves and chiasma, and epicanthal folds. Magnetic Nuclear Resonance (MNR) revealed decreased cerebral and axial tonsus, reduced fontanelle, lack of visualization of the septum pellicudum with persistent of the cavum septum pellicudum and cavum Vergae, atrophic hippocampus, absence of falx cerebi, partial fusion of the thalami with an appearance suggestive of holoprosencephaly. She also had hands with elongated fingers and disproportionate in size, and heart problems. At the age of 6 months, she showed the proportion of the proximal region severe delay in psychomotor development and intellectual disability. Cytogenetics analysis revealed a constitutive 46.XX,r(13)[77]/46.XX,-13/17[46.XX,iddic r(13)[6]. Fish analysis also showed the absence of 13pter and the presence of 13q14.3. Karyotyping of the ring chromosome it was also indicative of the presence of an isodicentric chromosome. Chromosome Microarray Analysis (CMA) with Affymetrix CytoScan™ HD Array detected a de novo 15.39Mb deletions at 13q32.3-q34. This deletion involved 44 morbid genes from OMIM. Both monosomy and isodicentric were confirmed with FISH. CMA was also useful to determine the deletion of the terminal end of the long arm of chromosome 13 in the ring formation. However, banding cytotyping and FISH were not able to precisely define the breakpoint of the terminal deletion. It was noted that the ring chromosome led to a partial deletion of the long arm of chromosome 13. Haploinsufficiency of the genes within the deleted region is the most probable cause of the proband phenotype. Thus, based on probe density, high-resolution CMA made it possible to refine the breakpoint region for the affected chromosome 13 and it was also useful to determine the gene content within the region.

3204T Proximal 3p deletions: phenotypic characterization and molecular delineation. J.M. Lee1, T. Schreiner1, A.F. Sandhu1, N. H. L. De Oliveira1, R. Leiva-Bauza2, E. Novak3, O. Somers4, T. P. Pina5, C. M. de Lima6, D. Hochedez1, A. T. Tricio1 1 Department of Laboratory Medicine, University of Ulsan College of Medicine and Asan Medical Center, Seoul, Korea; 2 Department of Pediatrics, University of Ulsan College of Medicine and Asan Medical Center, Seoul, Korea; 3 Asan Biomedical Sciences, Seoul, Korea; 4 Medical Genetics, Department of Pediatrics, University of Virginia, Charlottesville, VA, USA; 5 Clinical Hospital of Federal University of Goias, Genetics Master’s Program; 6 Human Cytogenetics and Molecular Genetics Laboratory of the Faculty of Health Sciences. Proximal interstitial deletion of chromosome 3p is a rare intrachromosomal rearrangement known to cause neurodevelopmental delay and intellectual disability, and associated with various clinical manifestations such as autistic features, hypotonia, heart and kidney abnormalities. The chromosome configuration affect the expression of genes in two copies in the patient’s genome, causing phenotypic alterations. Gene expression of eight genes was assessed using TaqMan assays, in peripheral blood from the patient and seven controls, matched by age and sex. qPCR experiments were run in triplicate and normalized by two endogenous genes (GAPDH and ACTB) using 2-∆∆Ct analysis method. PTGER2, RPS6KA5, RORC1 and VRK1 genes presented normal level of expression. ARID4A, GPHN, RORC1 and ARID4A2 genes were significantly downregulated in the patient compared to controls. The website GeneMania (www.genemania.org) used to verify gene interaction network showed that, except for GPHN gene, all the other genes studied are co-expressed. We propose that these downregulated genes may affect the expression of other genes explaining the patient’s phenotype. One mechanism proposed to explain ring syndrome phenotype is telomere effect position caused by the ring formation that changes the chromatin architecture and consequently the gene expression. The juxtaposition of the DNA from p-arm inactive, centromere and subtelomeric region close to the q-arm active euchromatin could have silenced the RORC1 gene and others in 14q32 due to inactivation spreading. However, the other downregulated genes (ARID4A, GPHN and ACTNB1) may have been silenced by other mechanisms as the repositioning of the ring chromosome in the nucleus that may affect the gene expression. The transcriptional network needs to be taken into account in the evaluation of the genetic consequences of complete rings with clinical phenotype. (Financial support, FAPESP, Brazil.)

Craniosynostosis is defined as a premature fusion of at least one cranial suture, which can be accompanied by other phenotypes. Of syndromic cases, 14-22% have been associated with chromosomal rearrangements. This report describes a Brazilian boy with syndromic craniosynostosis who also presented with mental retardation, microcephaly, frontal bossing, bitemporal narrowing, short neck, syndactyly and cardiac defects. Chromosome banding showed an apparently normal male karyotype. Subsequent chromosomal microarray analysis (CMA) using Affymetrix CytoScan 750K Array showed a duplication of 2.1 Mb on chromosome 17q and a deletion of 1.4 Mb on chromosome 20q. The data suggested an unbalanced translocation, which was confirmed by Fluorescence in situ hybridization analysis (FISH). While there are several reports in the literature of chromosome 17q duplication syndrome accompanied by partial monosomies of other chromosomes, this is the first case featuring partial monosomy of 20q. The patient’s phenotype is generally consistent with 17q duplication syndrome, however the craniosynostosis has rarely been associated with this chromosomal anomaly. Thus, CMA provides a sensitive, rapid and powerful technology for a “genotype-first” approach in conditions such as craniosynostosis that are difficult to associate with a specific genetic defect based on clinical features alone.

Meiotic I error in a Thai girl with tetrasomy 9p syndrome identified by SNP microarray. C. Charalsawadi1, S. Jaruratanasirikul1, J. Wiriyarattanarot1, V. Praphanphon2, S. Puangpech2, K. Jaruthamsophon2, P. Limprasert2. 1) Department of Pathology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand; 2) Department of Pediatrics, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand; 3) Rajanukul Institute, Bangkok, Thailand.

Tetrasomy 9p is a rare chromosomal disorder caused by the presence of an extra isochromosome consisting of the short arm of chromosome 9. The syndrome has a wide range of phenotypic features, ranging from stillborn babies with multiple malformations to normal healthy adults. We report the case of a 4-year-old Thai girl with mild developmental delay and no major physical malformations. The girl also presented with pteridactyly of the abdominal wall and skin pigmentary abnormalities. G-banding karyotyping of cultured lymphocytes revealed 47.XX,+mar in all analyzed cells. The Illumina HumanCytoSNP-12 v2.1 array was carried out on uncultured peripheral blood DNA and revealed an elevated logR ratio and an altered pattern of B allele frequency, consistent with tetrasomy of the short arm of chromosome 9. A FISH assay using probes specific to chromosome 9p and 9q subtelomeres confirmed the presence of 4 copies of chromosome 9p and 2 copies of chromosome 9q. Her karyotype was designated 47,XX,+i[9]q(12)[+] and 9p24.3p12(46,587,42,374,011)[+]x4. G-banding karyotyping of tissues from hypopigmented skin, normal skin and the pilomatricoma revealed 28.5%, 5%, and none of the aberrant chromosome, respectively. Tissue-limited mosaicism may explain mild developmental delay in the patient and the pilomatricoma presented in this case was likely unrelated to the syndrome. Haplotyping information obtained from a SNP microarray revealed meiosis I nondisjunction which determines the meiotic origin of tetrasomy 9p. The rearrangement thereafter which leading to duplication of the short arm and loss of the long arm of chromosome 9 was probably a mechanism of the isochromosome formation. Subsequent postzygotic tetrasomy rescue in certain tissues was a possible cause of tissue-limited mosaicism in our patient. To the best of our knowledge, we are the first to speculate a mechanism insight into a supernumerary isochromosome 9p formation using an extensive collection of SNP markers.
3210T

A de novo microduplication at 7q11.23 from Central Brazil detected by Chromosomal Microarray Analysis. L.B. Minasi1,2, I.P. Pinto1,2, A.V. Malheiro1,2, D.M.C. Cunha1,2, D.L. Ribeiro3, C.C. da Silva1,2,3, D.M. e Silva2,4, A.D. da Cruz1,2,3,5
1) Pontifical Catholic University of Goias, Genetics Master's Program; 2) Pontifical Catholic University of Goias, Department of Biology, Replication Research Group; 3) Federal University of Goias, Biotechnology and Biodiversity PhD Program; 4) Federal University of Goias, Genetics and Molecular Biology Master's and PhD Programs; 5) Human Cytogenetics and Molecular Genetics Laboratory, Secretary of Goias State for Public Health.

The Chromosome 7q11.23 Duplication Syndrome (MIM 609757) is a geno-mic disorder caused by the duplication of a common 1.5 Mb segment spanning 26 genes. This region is also associated with the Williams-Beuren Syndrome (WBS [MIM 194050]), and duplications of the WBS region should occur at the same frequency as deletions based on an inter- and nonallelic homologous recombination mechanism. The duplication at 7q11.23 shows variable clinical manifestations most common being speech delay, mild craniofacial anomalies, and neurocognitive and behavioral impairment, specifically, intellectual disability. Herein, we report a first case of 16 year old boy who presented moderate intellectual disability, mild facial dysmorphism with broad forehead, high- brow nose, short philtrum, straight eyebrows, and speech delay with a de novo 7q11.23 microduplication in Central Brazil. Chromosomal analysis by GTG banding using the software IKAROS® (Metasystems Corporation, Germany) showed 46,XY and Chromosomal Microarray Analysis (CMA) with Affymetrix’s CytoScan™ HD SNP array revealed a de novo 1.43 Mb microduplication at 7q11.23, encompassing 1.54q markers, and containing fewer described 26 genes (NSUN5, TRIM50, EKBP6, FZD9, BAI1, C17L7, TBL2, MLXIPL, VPS37D, DNAJC30, WBSCR22, STX1A, IRX4, AHSID1-AS1, AHSID1, CLDN3, CLDN4, WBSCR27, WBSCR28, ELN, LMK1, EIF4H, MKS90, LAT2, RFC2, CLIP2, GTF2R1, GTF2). The progenitor’s CMA confirmed de novo geno-mic imbalances in their child. Some genes included in that genomic region have been shown to be implicated in Williams Beuren Syndrome. However, deletion or duplication in that region produces different degrees of impairment with variable phenotypes. In addition, strategically CMA has proved to be a useful tool to carry out genetic diagnostic and it is especially useful to investigate children with developmental and neurobehavioral delays, reducing the time to reach a clinically relevant diagnoses in which genetic rearrangement display an important etiological role such as the deletion of 26 contiguous genes, including elastin (ELN [OMIM 130160]) on chromosome 7q11.23. The Williams syndrome phenotype is characterized by facial dysmorphism, cardiac abnormalities, behaviour characteristics, intellectual disability, elastin arteriopathy, short stature, connective tissue abnormalities, infantile hypercalcaemia, and a unique personality and cognitive profile. We present here three cases of Williams syndrome who were referred to paediatric cardiologist because of a systolic murmur in upper sternal border. Case 1 aged 1 year had neonatal coarctation and aortic stenosis with diffuse narrowing of ascending aorta. She underwent surgical correction of coarctation followed by balloon angioplasty for recoarcta-tion. Her surgical findings were markedly thickened aortic wall. No aortic biopsies were done. Case 2 aged 9 months had mild supravalvar aortic stenosis (SVAS) and was on regular medical follow up. Case 3 aged 2 years was initially referred to have large tricuspid septal defect (VSD), but on further postnatal cardiac echocardiogram, mild SVAS had surgical fenestrated VSD patch closure, branch pulmonary artery pericardial patch repair with DORV’s repair of SVAS. With a triad of dysmorphic facies, cognitive disorder and congenital heart defect, a diagnosis of Williams syndrome was made clinically. Karyotyping and fluorescent in situ hybridization (FISH) techniques employing the elastin gene probe were performed. Karyotyping was normal in all the three patients whereas FISH result was positive for 7q11.23 microduplication. Patients were screened for microdeletion by FISH and were found normal. In this paper we suggest a defined protocol with more attention while evaluating heart murmur in childhood period, especially when the patient has facial dysmorphism or developmental delay.
3214T Target-specific synthetic oligonucleotide libraries for use in Fluorescent In Situ Hybridization. K.C. Semrau1, Y.E. Murgha1, E. Robinson2, F.A. Ray3, C. Proudfit3, B. Hao4, J. Skok5,6, E. Gulan6,7, J.M. Rouillard8,9,10. 1) Mycroarray, Ann Arbor, MI; 2) KromaTID Inc., Fort Collins, CO; 3) Department of Pathology, New York University School of Medicine, New York, NY; 4) NYU Cancer Institute, New York University, New York, NY; 5) Department of Chemical Engineering, University of Michigan, Ann Arbor, MI.

Human cytogenetic applications rely on detecting the presence, position, and location of specific chromosomal regions within the nucleus as well as chromosomal abnormalities, mainly through the use of Fluorescent In Situ Hybridization (FISH). Conventionally, FISH probes have been generated from PCR amplification of genomic regions or BACs followed by nick translation to incorporate fluorophore(s). These probes are accompanied by limitations associated with the specificity of probes. We have developed a method to produce large quantities of customizable FISH probes from synthetic oligonucleotides that show improved specificity and coverage over traditional genome-derived probes. Here we present an algorithm to design complex probe libraries to target specific regions of the human genome (or any sequenced genome) optimized for reproducible hybridizations. The resulting probe sequences are synthesized on a microarray, cleaved, and linearly amplified by in vitro transcription before being converted to single-stranded probe sequences. Through this IVT-RT procedure, we have shown improvement in the yield of labeled probes over previously used methods. Here we present highly-specific detection of genomic regions ranging from a few kilobases to several megabases in size using probe densities ranging from 1 to several probes per kilobase. These libraries have been successfully used for FISH in mammalian, insect, and plant cells as well as for 3D FISH.

3215T Mapping breakpoints of a familial chromosome insertion (18;7) (q22.1; q36.2q21.1). W. Fan1, L. Li2,3, H. Chen1, J. Skok1,4, J. Wang1, S. Zheng1, J. Zhang2. 1) Molecular genetics, Hebei University, Baoding, Hebei, China; 2) Institute of Medical Genetics, Linyi People's Hospital, Shandong 276003, China; 3) BGI, 11-2 Building, Northern Industry District, Shenzhen 518083, China.

It is widely accepted that the incidence of chromosomal aberration is 10-15.2% in the azoospermic male; however, the exact genetic damages are currently unknown for more than 40% of azoospermia. To elucidate the causative gene defects, we used the next generation sequencing (NGS) to map the breakpoints of a chromosome insertion from an azoospermic male who carries an balanced, maternally inherited karyotype 46, XY, inv ins (18;7) (q22.1; q36.2q21.11). The analysis revealed that the breakage in chromosome 7 disrupts two genes, dipeptidyl aminopeptidase-like protein 6 (DPP6) and CACNA2D1 genes in an azospermic male. W. Fan1, L. Li2,3, H. Chen1, J. Skok1,4, J. Wang1, S. Zheng1, J. Zhang2. 1) Molecular genetics, Hebei University, Baoding, Hebei, China; 2) Institute of Medical Genetics, Linyi People's Hospital, Shandong 276003, China; 3) BGI, 11-2 Building, Northern Industry District, Shenzhen 518083, China.

3216T Refining 16p11.12 microdeletion region for Intellectual Disability/Developmental Delay (ID/DD). P.S. Lai1, Y. Rong1, K.M. Eu1, P.S. Low1, E.C. Zhang1,2. 1) Dept. Pediatrics, National Univ Singapore, Singapore, Singapore; 2) KK Hospital, Singapore, Singapore.

Copy number variations on chromosome 16p11.2 have been associated with extensive phenotypic variability and diverse phenotypes such as autism spectrum disorders (ASDs), intellectual and developmental disabilities (ID/DD), schizophrenia, etc. Within this region, a proximal 16q11.2 microdeletion region (~593 kb spanning Chr16: 29.5 to 30.1 Mb) has been commonly associated with developmental delay, autism spectrum disorder, epilepsy and obesity. An atypical and less frequently reported microdeletion region (~820 kb spanning Chr16:28.74 to 28.95 Mb) adjacent and distal to this has been reported in patients with developmental delays, behaviour problems and unusual dysmorphic features. We report a seven-year-old child from non-consanguineous marriage, who presented with ID and DD. The child was diagnosed with severe learning impairment and psychomotor developmental delay. No dysmorphic features were observed with the exception of a bifid thumb. Array CGH analysis using the 180K oligo chip (Agilent) identified a de novo 1.8 Mb deletion at 16p11.2 (UCSC HG18, Chr16: 28,285094-30,089069). Further qPCR dosage analysis delineated the genomic deletion to a narrower region between 28.39 to 27.9 Mb with the breakpoints spanning EIF3C1 gene to ZG16 genes. As this deletion spans the proximal typical 16p11.2 microdeletion syndrome and distal atypical 16p11.2 microdeletion syndrome, genotype-phenotype investigations can aid in delineating the critical regions underlying the clinical phenotypes. Our patient’s phenotype most closely corresponds to that of proximal 16p11.2 microdeletion syndrome, with developmental delay as the predominant shared symptom, and without facial dysmorphism. This deletion is most closely associated with the distal 16p11.2 microdeletion syndrome. The deletion region in our patient overlaps with the microdeletion regions of both syndromes, harboring 19 annotated protein-coding genes, namely ATXN2L, TURM, SNOB1, ATP2A1, HABEP2, CD19, NFATC2IP, SPSN1, LAT, BOLA2, BOLA2B, SLX1B, SLX1A, SULT1A3, SULT1A4, SPN, QPTR, C16orf54 and ZG16. Recent gene dosage studies in animal models suggest that genes with deletion dosage sensor properties may contribute towards abnormal brain development and function leading to developmental and cognitive impairments. The clinical phenotype in our patient could most likely arise from the haplinsufficiency of one or more of the above genes, and contributes towards refining the critical 16p11.2 region underlying ID/DD.

3217T A Five Year Retrospective Analysis of the Utility of Family Segregation Analysis in the Evaluation of the Clinical Significance of Variants of Uncertain Significance Detected by Chromosomal Microarray: The Greenwood Genetic Center Experience. P.S. Lai1, Y. Rong1, K.M. Eu1, P.S. Low1, E.C. Zhang1,2. 1) Dept. Pediatrics, National Univ Singapore, Singapore, Singapore; 2) KK Hospital, Singapore, Singapore.

Chromosomal microarray (CMA) and massively parallel sequencing (MPS) technologies have been transformative to the practice of laboratory medical genetics over the past decade. These technologies allow the highest resolution interogation of the entire genome in such a cost-effective manner that these techniques have become widely adopted by laboratories. In cytogenetics laboratories CMA has become a first tier test in the evaluation of nonsyndromic pediatric patients presenting with intellectual disability, autism or multiple congenital anomalies. CMA offers dramatic increases in the ability to detect genomic copy imbalances, but evaluation of the clinical significance of these imbalances presents a challenge for clinical diagnostic laboratories. Family segregation analysis (FSA) is among the best methods available to assist a clinical diagnostic laboratory in the evaluation of variants of uncertain significance (VUS) detected by whole-genome assays such as CMA and MPS. In order to evaluate the utility of FSA in attributing clinical significance to CMA detected VUS, we conducted a 5 year retrospective study of FSA performed for CMA detected VUS across all referral reasons at a single clinical genetics laboratory. In total 2224 reported copy imbalances from 1682 consultands have been evaluated by quantitative PCR (qPCR) based FSA. 85% of VUS were found to be de novo in origin, with a 2.125:1 ratio of deletions to duplications. Deletions ranged from 50kb to 23Mb in size, with average and median of 2.4Mb, and 1.4Mb respectively. Duplications ranged in size from 95kb to 17.5Mb with average and median of 2.5Mb and 825kb, respectively. Among inherited cases, a 1.5:1 ratio of maternal to paternal transmission was detected. A clear transmission pattern was only established for 64% of cases, largely due to the inability to obtain parental samples from one or both parents in 22.4% of cases. Analysis of de novo hot spots, and comparison of FSA results by referral reason will also be presented. Our results add to the value added by FSA in the evaluation of VUS, and highlight the inability to obtain relevant family members in a significant number of cases as a significant limitation of this approach.
3218T Interstitial duplications of 19p13.3. H. Risheg1, R. Pasien2, S. Schwartzz3, E. Prijoles3, E.A. Keilges1. 1) Laboratory Corporation of America/Dynacare, Department of Cytogenetics, Seattle, WA; 2) Laboratory Corporation of America, Center for Molecular Biology and Pathology, Department of Cytogenetics, Research Triangle Park, NC; 3) Greenwood Genetic Center, Columbia, SC.

Interstitial deletions involving 19p13.3 characterized by microarray analysis have recently been reported in the literature. However, there is little molecular cytogenetic information reported on interstitial copy number gains (i.e. duplications and triplications) of 19p13.3. We present 4 patients with overlapping copy number gains (3 duplications and 1 triplication) of 19p13.3 identified by high resolution SNP microarray analysis. Three of the 4 copy number gains with parental studies performed confirmed a de novo origin. Features described in these patients include intellectual disability, microcephaly, motor and speech delay, and short stature. Copy number gain sizes ranged from 0.613 Mb-2.45 Mb and included a common region arr[hg19]19p13.3(2,885,504-3,501,271). The patient phenotypes were compared with 7 additional patients with overlapping interstitial 19p13.3 duplications from DECIIPHER and ISCA databases. All 7 patients from both ISCA and Decipher databases were reported as de novo 19p13.3 duplications. The common region of overlap of all 11 cases further narrowed the region of overlap to arr[hg19]19p13.3 (3,098,056-3,451,152) and includes 5 OMIM genes (GNA11, GNA15, STIP1, CELF5, NFIC). The most common clinical features reported among all 11 were intellectual disability, microcephaly and short stature. Less common features were intrauterine growth retardation and fine motor and speech delays. Our results combined with those reported in the databases provide support for copy number gains within 19p13.3 as a cause of a constellation of findings, including intellectual disability/developmental delay, microcephaly and poor growth. Identifying additional patients with imbalances within this region is important to further recognize associated clinical features.

3219T Prenatal Chromosome Rearrangements and Markers: Normal SNP Microarray Analysis Associated with Favorable Pregnancy Outcome. J.H. Tepperberg1, R.M. Pasien1, I. Gadi1, R.D. Bumsdie2, L. Kline1, B. Willford1, K. Phillips3, E. Keilges2, H. Risheg1, A. Penton1, J. Schievel1, S. Schwartzz1, P. Papenhausen1. 1) Laboratory Corporation of America, Research Triangle Park, NC; 2) DynaCare/Laboratory Corporation of America, Seattle WA.

Prenatal microarray is effective in delineating copy number changes not detectable by cytogenetic analysis in and around chromosome rearrangements. The goals of this study were: 1. investigate how many prenatal cases with either G-banded chromosome rearrangement or marker chromosome analysis by high resolution array were reported as "balanced," within the resolution of the 2,695,000 SNP/oligo microarray. 2. how many apparently balanced translocations were “balanced” by the array; 3. how many de novo balanced rearrangements had an unbalanced result on array. In order to understand the significance of chromosome rearrangements ascertained at the time of amniocentesis/CVS, we reviewed ~2675 prenatal cases from Oct 2011 to June 2012 by a SNP array. Of those referred for SNP microarray testing, 127 (4.75%) had a translocation, inversion, insertion or marker chromosome. Fifty of these (39%) were duplicated in the translocation, 26 had an inversion, 3 had an insertion and 46 had a supernumerary marker. Of the 50 simple translocations, 19 cases (38%) were de novo. Of these, 17 (89.4%) had a normal SNP analysis and apparently normal ultrasound and/or pregnancy outcome. Two (10.5%) cases had copy number findings at the reported breaksite - one significant and one variant of unknown significance (VOUS). Two reported balanced translocations had pathogenic copy number changes identified by array that were presumed to be de novo. Of the 26 inversion cases, 12 (46.4%) had normal signalization. Two were VOUS findings within the inversion breakpoint and another two had secondary abnormal microarray findings unrelated to the inversion breakpoints. Of the 46 cases with supernumerary chromosomes, 16 (34%) were normal by array. In conclusion, the majority of apparently balanced chromosome rearrangements appeared to be balanced following microarray analysis. Including the two pathogenic cases that were presumed to be de novo, 19% (4/21) had a gain or loss at the reported rearrangement breaksite. Of the 10 de novo rearrangements and normal microarray, only 1 of the cases was reported to have an abnormal copy number. 34% of prenatal cases with supernumerary chromosome had a normal array result and only three of these had a reported abnormal ultrasound. One of these was UPD. While additional follow-up is pending, this preliminary data suggests that a normal microarray result after normal chromosome analysis appears to reduce the risk for an abnormal pregnancy outcome.

3220T Validation of an Ion AmpliSeq™ RNA Lung Fusion Panel, workflow, and analysis solution: an OncoNetwork collaborative research study. J.G. Oienlueg3, K. Bramlett1, C.P. Vaughn13, L. Lacroix14, R. Petraroli15, M. Budagyan15, F. Hyland15, R. Gottumukkala16, O. Shells6, B. Tops2, D. Le Corre8, H. Kurth8, H. Blons2, E. Amato8, A. Mafficini9, A.M. Rachiglio2, A. Heilmann3, C. Noppen3, C. Aimail16, P. Laurent-Puig8, R. Franco8, H. Feilottor10, P. Park1, J. Schageman1, J. Creet1, J.L. Costa11, M. Lightenberg8, A. Scarpa6, J.C. Machado17, K. Nishio18 OncoNetwork Consortia. 1) Life Sciences Solution Group, Thermo Fisher Scientific, Austin, TX; 2) Trinity College, Dublin, Ireland; 3) Radboud University Medical Center, Nijmegen, Netherlands; 4) University Paris Descartes, Paris, France; 5) Viollet AG, Basel, Switzerland; 6) ARC-NET University of Verona, Italy; 7) Centro Ricerche Oncologiche Mercogliano, Italy; 8) Warwick Medical School, United Kingdom; 9) Surgical Pathology, Instituto Nazionale Tumori “Fondazione Istituto di Patologia”, Napoli, Italy; 10) Queen’s University, ON, Canada; 11) J.ATIMUP, University of Porto, Portugal; 12) Kinki University Faculty of Medicine, Osaka, Japan; 13) ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT; 14) Institut Gustave Roussy (IGR), Paris, France; 15) Thermo Fisher Scientific, Carlstadt, CA.

Fusion transcripts resulting from translocation events in the oncogenic driver genes ALK, RET, ROS1, and NTRK play an important role in lung adenocarcinoma. There is a need to detect these fusion transcripts with up to date technologies as they may serve as viable therapeutic targets. We have utilized a targeted sequencing approach and developed an Ion AmpliSeq™ RNA Lung Fusion panel, a workflow, and an Ion Reporter analysis solution to detect these known fusion events. The panel detects transcripts from 37 ALK, 9 RET, 15 ROS1, and 11 NTRK fusion variants along with 5 housekeeping genes to serve as internal controls. The workflow is FPPE compatible requiring an input of only 10 ng of total RNA with the capacity to multiplex up to 16 libraries on a single Ion 318™ chip. The panel was initially validated using 10ng of total RNA from a cocktail of 3 cell lines containing known ALK and/or ROS1 fusions (H2228 - EML4-ALK and/or ALK 3a and 3b, HCC78 - SLC34A2-ROS1 and LC-2/ad - CCDC6-RET). The library was sequenced using the Ion PGM™ system and analyzed with the AmpliSeq™ RNA Lung Fusion workflow in the Ion Reporter. Analysis showed that the workflow amplified all expectant fusions and house-keeping genes and reported zero false positives fusions. This multiplexed fusion transcript targeted sequencing solution is currently being validated by all members of the OncoNetwork Consortium who will test lung cancer tissue samples from 2011 and report all inconsistencies, with no ALK and/or NTRK fusion or signal. Initial results from OncoNetwork Consortium members reveal 100% concordance between the AmpliSeq™ RNA Lung Fusion panel and FISH in 25 lung tissue samples.

3221T Variation in the Zinc Finger Binding Domain of PRDM9 is Associated with the Absence of Recombination on 21q. T. Oliver1, C. Middlebrooks2, A. Harden1, B. Johnson1, C. Wilkerson1, S. Saffold2, N. Scott1, S. Sherman1. 1) Department of Biology, Spelman College, Atlanta, GA; 2) Laboratory of Translational Genomics, The National Institutes of Health, Bethesda, Maryland; 3) Department of Human Genetics, Emory University, Atlanta, GA.

Proline Rich Domain Containing 9 (PRDM9) is a major determinant of meiotic recombination. Variation in the zinc finger-binding domain (ZFBD) of PRDM9 is linked to altered placement of recombination in the human genome. As altered recombination (both the absence and altered placement of recombination) is also observed among chromosomes 21 that nondisjoin, we generated the PRDM9 ZFBD among mothers of children with trisomy 21 to examine the relationship between variation in the ZFBD of PRDM9 and chromosome 21 nondisjunction. In our approach, PCR was used to amplify the ZFBD of PRDM9 and these PCR products were then subjected to competitive real-time Sanger sequencing. In order to identify samples with PRDM9 minor alleles, DNA sequencing reads were aligned and compared to the sequence of the PRDM9 major allele previously identified by Berg et al. 2010. Chi-Square analysis was then used to compare the distribution of major and minor alleles between our cases (N=228, mothers of children with trisomy 21) and controls (N=94). Among individuals exhibiting no recombination on 21q, the PRDM9 major allele was observed at a significantly lower frequency among our cases (0.63) when compared to that of controls (0.86). This suggests that variation in the PRDM9 ZFBD plays a novel role in the regulation of meiotic recombination on chromosome 21 which is a major risk factor for its nondisjunction and trisomy 21.
Importance of cytogenetic and molecular characterization of patients with pigmentary mosaicism. C. Salas-Labada¹, R. Cruz-Alcivar², V. Uloa-Aviles¹, A. Reyes-Leon¹, M. P. Navarrete-Meneeses¹, S. Gomez-Carmona¹, C. Durán-McKinster¹, E. Lieberman-Hernández¹, V. Del Castillo-Ruiz¹, P. Pérez-Vera¹, D. E. Cervantes-Barragán¹,², ¹Instituto Nacional de Pediatria, Mexico City; ²Hospital Central Sur de Alta Especialidad, PEMEX-Mexico City.

Pigmentary mosaicism (PM) describes a heterogeneous group of skin anomalies and in association with multisystem involvement (central nervous, musculoskeletal and ocular), could be related with chromosomal abnormalities. One of the most well described that chromosome analysis on stimulated cultures may not become evident a structural alteration not observed by GTG bands. It is detected by cytogenetics. In 1/4 patients of group 4, the molecular analysis was able to identify a cell line present in low-level mosaicism (<5%), not detected by cytogenetics. In 1/4 patients of group 4, the molecular analysis delineated the structural alteration; in one case, it was able to identify a cell line present in low-level mosaicism (<5%), not detected by cytogenetics. In 1/4 patients of group 4, the molecular analysis became evident a structural alteration not observed by GTG bands. It is well described that chromosome analysis on stimulated cultures may not detect somatic chromosome mosaic in individuals with PM and the number of metaphases counted may not be sufficient to detect it. Here, we screened a large number of metaphases from three different tissues applying in some cases SNP array/FISH. This plan of analysis could contribute to a better characterization of anomalies in patients with PM. The molecular tools identified cases with chromosomal mosaicism, allowing the delineation of the abnormality, to perform better associations with the phenotype. CONACYT 2012-01-1882277.

3223T
Chromosome Therapy: Correction of Large Chromosomal Aberrations by Inducing Ring Chromosomes in Induced Pluripotent Stem Cells (iPSCs).

T. Kim¹, M. Bershteyn¹, A. Wynshaw-Boris¹,². ¹Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH; ²Institute for Human Genetics and Department of Pediatrics, University of California, San Francisco, CA, 94143.

Approximately 1 in 500 newborns are born with chromosomal abnormalities that include trisomies, translocations, large deletions and duplications. There is currently no therapeutic approach for correcting such chromosomal aberrations in vivo or in vitro. Recently, we attempted to produce induced pluripotent stem cell (iPSC) models from patients that contained ring chromosomes. We hypothesized that a chromosome from the aberrant chromosome in iPSCs, which would then be eliminated and replaced by a normal chromosome. We are testing this hypothesis by inducing ring formation in patients with large deletions of chromosome 17 via a Cre/loxP approach. LoxP sites will be inserted by CRISPR/Cas9 mediated gene editing at sites near each end of chromosome 17. Once this is accomplished, we will infect these cells with Cre recombinase to induce the formation of a ring chromosome. To visualize cells that form a ring chromosome, we designed and cloned the partial tandem dimer Tomato (tdTomato) along with these loxP sites such that the cell will express tdTomato upon formation of a ring chromosome due to Cre-mediated recombination. If successful, we will have created a generalized system of “chromosome therapy” for the correction of large chromosomal aberrations by the induction of ring chromosomes through genome editing followed by duplication of the normal chromosome.
Targeted RNA sequencing of breast cancer genes using a genetic capture approach: cBROCA. S. Casadei1, S. Gulsunen1, A.M. Thornton1, J.B. Mandell1, M.K. Lee1, M.C. King1,2, T. Walch1, 1Medical Genetics, University of Washington, Seattle, WA; 2) Genome Sciences, University of Washington, Seattle, WA.

RNA sequencing provides a powerful method for measuring expression of transcripts of interest. We have adapted BRCQA, our targeted construction and multiplex sequencing approach for genomic DNA, for RNA sequencing. We call this approach cBROCA. The goals of cBROCA analysis are to detect two classes of mutations: those that alter transcript length, as the result of a point mutation, and those that alter expression of the protein in a constitutive and alternate transcripts, and to measure different proportions of alternate transcripts. We will next apply cBROCA to genomically characterized samples for which a causal mutation has not been identified.

Identification of Large Intergenic Non-coding genes as Candidate Targets for Prostate Cancer risk-SNPs Utilizing a Normal Prostate Tissue eQTL Dataset. Y. Zhang1, S. McDonnell1, 2, J. Oliver1, M.N. Timofeeva1, 3, V. Gaborieau1, 4, M. Johansson2, A. Chabrier1, M.B. Wozniak1, D. Brenner1, 4, G. Bynes1, P. Brennan3, J.D. McKay1, ARCADE, CEE, IARC MC oral cancer, Latin America, Polish, ACTREC, Rome&Japan Head&Neck cancer studies. 1) Genetic Cancer Susceptibility group (GCS), International Agency for Research on Cancer (IARC), Lyon, France; 2) Genetic Epidemiology group (GEP), International Agency for Research on Cancer (IARC), Lyon, France; 3) Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh and Medical Research Council (MRC) Human Genetics Unit, Edinburgh, UK; 4) Biostatistics, International Agency for Research on Cancer (IARC), Lyon, France.

Recent meta-analysis of lung cancer GWAS have identified 13q13.1 and 12p13.33 regions, which encompass the BRCA2 and RAD52 genes respectively, as lung cancer susceptibility loci, and particularly lung squamous cell carcinomas (LUSC). Here, we report detailed investigation of two variants rs11571833 and rs10849605, located within the 13q13.1 and 12p13.33 regions respectively, for association with upper aerodigestive tract cancer squamous cell carcinoma (UADT). Using Sanger UADT cancer sequencing controls from 9 different studies, we demonstrate that rs11571833, a rare truncating variant in BRCA2 (K3326X), is also associated with risk of UADT (OR=2.53, 95% CI: 1.89-3.38, p=3x10^-10). Similarly rs10849605, a common intron variant in RAD52, is associated with UADT (OR=1.09, 95% CI: 1.04-1.15, p=6x10^-4). There was little evidence for association between rs11571833 carrier status and the loss of wild type allele of the BRCA2 gene or a cis-expression quantitative trait locus (eQTL) effect in the Cancer Genome Atlas (TCGA) data. However, we identified rs10849605 as a RAD52 cis-eQTL in TCGA and LUSC, with the risk allele correlated with increased RAD52 tumor expression levels. The 12p13.33 locus, encompassing rs10849605/RAD52, was identified as a significant genetic risk allele copy number amplification in UADT (n=324, OR=2.53, 95% CI: 1.89-3.38, p=3x10^-10). RAD52 is a DNA repair gene and the risk allele correlated with higher RAD52 tumor expression levels (p=6x10^-4 and p=3x10^-9 in UADT and LUSC respectively). Within this 12p13.33 region, rs10849605 was a cis-eQTL for RAD52 only, making it the most plausible candidate gene at this locus for RAD52. rs10849605 also correlated with the 13q13.1/BRC2 loci, linked by their roles in homologous recombination based DNA repair, are associated with genetic susceptibility of the UADT.
3228M  
**BRCA1 and BRCA2 mutational screening in 223 hereditary breast cancer patients in Chile: genotype-phenotype correlations.**  
C. Alvarez1, T. Tapia1, E. Perez2, D. Wiener3, C. Ruiz2, M. Rios3, I. Ayendurako4, C. Rodriguez5, C. Cortés6, L. Matamala7, A. Cruz8, C. Missarelli9, M. Camus9, P. Carvallo10.  
1) Departamento de Biología Celular y Molecular, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Santiago, Chile; 2) Facultad de Medicina, Pontificia Universidad Católica de Chile, Santiago, Chile; 3) Hospital Base de Valdivia, Valdivia, Chile; 4) Hospital Barros Luco Trudeau, Santiago, Chile; 5) Hospital Regional de Antofagasta, Antofagasta, Chile.

**BRCA1 and BRCA2** are the only genes described so far which mutations confer a high risk for breast/ovarian cancer. Frequency of families carrying a mutation in these genes varies across populations and in Chile account for almost 20% of all hereditary cases. Usually, patients are selected for **BRCA1** and **BRCA2** genetic screening by defined criteria related to family history of breast and/or ovarian cancer. More recently, other patients with no family history are being screened for these mutations. These patients may present bilateral breast cancer, age of onset for breast cancer before 40 years old, and breast and ovarian cancer in the same woman. Taking all these criteria in consideration we selected 223 Chilean patients for screening of **BRCA1** and **BRCA2** mutations. Screening was performed through Sanger and Next Generation Sequencing. Considering familial cases 34/179 (19%) carried a mutation, 13 in **BRCA1** and 21 in **BRCA2**. Among patients without family history, 6/44 carried a mutation (13.6%) all in **BRCA1**, being those diagnosed before age of 40 or with bilateral breast cancer the most prone to carry a mutation. Those six patients correspond to the 31.6% of **BRCA1** mutation carriers. In contrast, all **BRCA2** mutation carriers had family history. Average age of onset of index patients was 35 years old for **BRCA1** mutation carriers vs 44.4 years old for **BRCA2** mutation carriers. Considering all patients, average age of onset of mutation carriers was 42 years old vs 44 years old in patients without mutation. It has been described that several types of cancer are recurrent in **BRCA1** and **BRCA2** families. In our group of families, ovarian cancer was more frequent in **BRCA1** families (54%) than **BRCA2** (29%). Other types of cancer found were prostate and gastric, either in **BRCA1** or **BRCA2** carrier families. Interestingly uterine/ cervix cancer was present in 26% of **BRCA1** mutation families, and absent in **BRCA2** carrier families. These observations support different and specific cancer susceptibilities among **BRCA1** and **BRCA2** mutation carriers. Acknowledgments to FONDEF ca1210152 and FONIS sa1212299.

3229T  
**Identification of new familial breast cancer susceptibility genes: are we there yet?**  
I. Campbell1, E. Thompson1, M. Doyle1, P. James2, A. Trainer3, R. Scott4, G. Mitchell5, Lifepool.  
1) Research Div, Peter MacCallum Cancer Ctr, East Melbourne, Australia; 2) Familial Cancer Centre, Peter MacCallum Cancer Ctr, East Melbourne, Australia; 3) Hunter Area Pathology Service, Newcastle, Australia.

The genetic cause of the majority of multiple-case breast cancer families remains unresolved. Next generation sequencing has emerged as an efficient strategy for identifying predisposing mutations in individuals with inherited cancer. We are conducting whole exome sequence analysis of germline DNA from multiple affected relatives from breast cancer families, with the aim of identifying rare protein truncating and non-synonymous variants that are likely to include novel cancer predisposing mutations. Data from more than 200 exomes show that on average each individual carries 30-50 protein truncating mutations and 300-400 rare non-synonymous variants. Heterogeneity among our exome data strongly suggest that numerous moderate penetrance genes remain to be discovered, with each gene individually accounting for only a small fraction of families (~0.5%). This scenario marks a rate-limiting step in resolving the missing heritability of breast cancer. The aim of this study is to screen genes that are recurrently mutated among our exome data in a larger cohort of cases and controls to assess the prevalence of inactivating mutations that may be associated with breast cancer risk. We are using the Agilent HaloPlex Target Enrichment System to screen the coding regions of 168 genes in 1,000 **BRCA1/2** mutation-negative familial breast cancer cases and 1,000 cancer-naive controls. To date, our interim analysis has identified 21 genes which carry an excess of truncating mutations in multiple breast cancer families versus controls. Establishing breast cancer susceptibility gene **PALB2** is the most frequently mutated gene (13/998 cases versus 0/1009 controls), but other interesting candidates include NPSR1, GSN, POL2D and TOX3. These and other genes are being validated in a second cohort of 1,000 cases and controls. Our experience to date on the prevalence of mutations that are recurrent in the remaining breast cancer predisposition genes is likely to be very low making definitive validation exceptionally challenging.

3230S  
**Significant evidence for linkage of cutaneous malignant melanoma to 1q41.**  
L. Cannon-Albright1, J.M. Farnham1, K.K. Thai1, J.J. Zone1, C.C. Teeng2.  
1) Division of Genetic Epidemiology, Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT; 2) Dermatology, University of Utah School of Medicine, Salt Lake City, Utah.

A genome-wide linkage analysis using various illumina SNP platforms was performed in 46 high risk extended Utah melanoma pedigrees including 237 genotyped cutaneous malignant melanoma cases. For linkage analysis a set of 25,437 genomewide SNPs with no LD was selected from the intersection of the 3 illumina high density SNP platforms (550k, 610k, 720K). Analyses of distantly related carrier cases in these 2 linked pedigrees is underway.

3231M  
**Expression and insertion of MMTV/HMTV env gene sequences in human breast cancer.**  

Breast cancer is the leading cause of cancer death for women in Mexico. Recently, it was demonstrated that MMTV/HMTV retrovirus contributes to breast cancer progression. The aim of this study was to identify env gene expression and insertion sites of MMTV/HMTV retrovirus in breast tumours in Mexican women. DNA and RNA were extracted from a total of 73 tumours and normal tissues. In order to assess DNA quality, a 700bp GAPDH fragment was amplified for all samples. MMTV detection was performed by nested PCR for two fragments of env gene (660bp and 250bp), as a positive control MMTV env gene (3CH4) in pBR322 was used. Presence of the retrovirus was validated by qPCR with TaqMan probes and quantification of the number of copies using the env gene inserted into pBR322. Expression was evaluated by RT-qPCR with DNase and then a RT-qPCR with TaqMan probes to detect env gene. MMTV / HMTV retrovirus insertion sites were evaluated by RNA treatment with DNase and then a RT-qPCR with TaqMan probes for two fragments of env gene (660bp and 250bp) as a positive control. MMTV env gene detected in melanoma cell lines was 10% in 73 tumours. The data was confirmed by qPCR with TaqMan probes designed for a specific region of the env gene. Sequence analysis shows that env gene fragments present an identity of 95% with MMTV and 97% with HMTV. The number MMTV / HMTV viral copies is variable for each sample (228-7729 copies). env gene expression was determined in 50% of the positive samples. Preliminary analysis reveals retroviral insertions throughout the genome of infected breast tumours. We are working to establish the participation of MMTV or its homolog, HMTV in tumour progression and breast cancer development.
3232T
Association of P2RX7 gene polymorphisms and cervical squamous cell carcinoma risk. T. Chang1, Y. Yang2,3, Y. Yeh4,5, T. Chen2, W. Lin3, S. Chang1. 1) Medical Research Department, Mackay Memorial Hospital, New Taipei, Taiwan; 2) Department of Gynecology and Obstetrics, Mackay Memorial Hospital, Taipei, Taiwan; 3) Department of Gynecology and Obstetrics, Taipei Medical University, Taipei, Taiwan; 4) Pediatrics, Mackay Memorial Hospital, Taipei, Taiwan; 5) Pediatrics, Taipei Medical University, Taipei, Taiwan.

Cervical cancer is a multifactorial disease and increasing evidence suggests that host immunogenetic background may contribute to its pathogenesis. P2RX7 receptor has been implicated in the regulation of immune response and associated with certain cancer development. The aim of this study is to investigate associations between the P2RX7 gene single nucleotide polymorphisms (SNPs) and cervical cancer susceptibility. We genotyped 4 functional SNPs (rs17525809 C/T, rs208294 C/T, rs1718119 A/G, and rs3751143 A/C) in 507 cervical squamous cell carcinoma (CSCC) patients and 430 age/sex matched healthy controls by using the Pre-Developed TaqMan Allelic Discrimination Assay. The presence and genotypes of HPV in CSCC patients were determined by PCR. We found no significant associations between the polymorphisms or haplotypes and CSCC. Stratified by the positivity of HPV-16 infection also did not find marked association. Our findings provide no support for the hypothesis that P2RX7 gene polymorphisms are associated with increased risk for CSCC in the Taiwanese population.

3232S
Functional Variants at The 21q22.3 Locus Involved in Breast Cancer Progression Identified by Screening of Genome-Wide Estrogen Response Elements. H. Chu1,2, C. Huang1,2, Y. Huang1, W. Chou1, L. Hu1, H. Hsu1, P. Wu1,2, M. Hou1, J. Yu1, C. Shen1,2,5,1) Institute of Biomedical Sciences, Academia Sinica, Taipei, 115, Taiwan; 2) Taiwan Biobank, Academia Sinica, Taipei, 115, Taiwan; 3) Department of Surgery, Tri-Service General Hospital, Taipei, 114, Taiwan; 4) Cancer Center and Department of Surgery, Kaohsiung Medical University Chung-Ho Memorial Hospital, Kaohsiung, 804, Taiwan; 5) College of Public Health, China Medical University, Taichung, 404, Taiwan.

Estrogen forms a complex with the estrogen receptor (ER) that binds to estrogen response elements (EREs) in the regulatory region of estrogen-responsive genes, and regulates their transcription. Since sequence variants in the regulatory regions have the potential to affect the transcription factor-regulatory sequence interaction, resulting in altered expression of target genes, this study explored the association between single-nucleotide polymorphisms (SNPs) within the ERE-associated sequences and breast cancer progression. The ERE-associated sequences throughout the whole genome, demonstrated to bind ERα in vivo, were blasted against online information from SNP datasets, and 54 SNPs located adjacent to estrogen-responsive genes were selected for genotyping in two independent cohorts of breast cancer patients, 779 in the initial screening stage and another 888 in the validation stage. With death from breast cancer or recurrence of breast cancer being defined as the respective event of interest, the SNPs at 21q22.3 were significantly associated with overall survival and disease-free survival of patients. Furthermore, these SNPs (rs2839494 and rs1078272) could affect the binding of this ERE-associated sequence to ERα or Rad21 (an ERα coactivator), respectively, resulting in a difference in ERα-activated expression of the reporter gene. These findings support the idea that functional variants in the ERα-regulating sequence at 21q22.3 is important in determining breast cancer progression.

3234M
Post-GWAS functional characterization of the 12p11.23 renal cancer susceptibility locus. L.M. Colet1, P. Bigot1,2, L. Jessop1, M. Machiela1, T. Myers2, S. Chanock1,1) Division of Cancer Epidemiology and Genetics, NCI - NIH, Gaithersburg, MD; 2) University Hospital of Angers, Angers, France.

In previous GWAS, rs718314 and rs1049380 at 12p11.23 were associated with renal cell carcinoma (RCC). The aim of our study is to perform a functional analysis of the 12p11.23 region in relation to RCC risk. We performed an imputation analysis within 1 Mb of rs718314 in three different previously published RCC GWAS studies (4197 cases and 8527 controls). The genotyped and high-quality imputed SNPs were tested for association with RCC. After meta-analysis, 44 SNPs demonstrated nominally significant association with RCC risk (p<5×10-5). The two initial GWAS SNPs, rs718314 and rs1049380, were strongly associated with RCC (Padj = 3.44×10-6 and Padj = 5.27×10-6). All nominally significant SNPs were in a non-coding region which contains the 3′-UTR of ITPR2. Six of the 44 variants were in regions enriched for H3K4me1 and H3K27ac, chromatin marks found in enhancers. Only rs7132434, which is highly correlated with the initial GWAS signal (rs718314, r2=1), showed allele specific regulatory activity in luciferase assays and allele specific differences in protein binding by EMSA.

The RCC-associated variants were examined for an effect on nearby gene expression (ITPR2, SSPN, SHARP1), using the TCGA database. Five SNPs associated with RCC were present in the TCGA database and all of them were associated with SHARP1 expression in tumor tissues (p<0.05). The most significant association was found with rs12814794 (p=1×10-8) which is in high linkage disequilibrium with rs718314 (r2=0.956) and rs7132434 (r2=0.956). There was no association between these SNPs and SSPN or ITPR2 expression. Recently, SHARP1 has been shown to be involved in the HIF pathway in breast cancer. Also, SHARP1 regulates adipogenic differentiation and could be a link between RCC and obesity, a known risk factor for RCC. Our results suggest rs7132434 is the functional SNP responsible for the GWAS signal and that this locus could act as an enhancer of SHARP1. Further investigations will be necessary to confirm the link between rs7132434 and SHARP1 and to understand the role of SHARP1 in renal carcinogenesis.
Targeted Gene Sequencing in Familial Colorectal Cancer Type X. J. Cunningham,1, A. French,1, M. DeRycke,1, S. Riska,2, S. McDonnell,3, S. Gunawardena1, Z. Fogarty1, S. Middha2, S. Baheti2, D. Schaid2, Y. Zhang2, S. Gallinger2, M. Cottlechio2, R. Haile5, G. Casey6, J. Jenkins6, D. Schaid2, M. Wood6, L. Le Marchand1, J. Potter1, P. Newcomb1, D. Duggan1, M. Cleidenning1, D. Buchanan1, N. Lindor4, E. Gooder1, S. Thibodeau1 on behalf of the Colon Cancer Family Registry. 1) Dept Lab Med & Pathology, Mayo Clinic & Fndn, Rochester, MN; 2) Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 7) Centre for Epidemiology & Biostatistics, Melbourne School of Population & Global Health, University of Melbourne, Melbourne, Victoria, Australia; 8) Centre for Molecular, Environmental, Genetic & Analytic Epidemiology, School of Population Health, The University of Melbourne, Melbourne, Australia; 9) Discipline of Genomics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada; 10) Epimediology Program, University of Hawaii Cancer Center, University of Hawaii, Honolulu, HI; 11) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 12) Translational Genomics Research Institute (TGen), Phoenix, AZ; 13) Oncogenomics Group, Genetic Epidemiology, Department of Pathology, University of Melbourne, Australia; 14) Department of Health Science Research, Mayo Clinic, Scottsdale, AZ.

Much of the heritability of colorectal cancer (CRC) remains unexplained. Type X CRC cases are those that meet Amsterdam criteria for hereditary non-polyposis CRC but have microsatellite unstable tumors and lack mutations in mismatch repair (MMR) genes. In our recent germline next generation sequencing custom capture project, 144 Type X CRC cases were sequenced. The capture covered 2.9Mb and included known hereditary colon cancer genes (HCC, n=18) and suspected HCC genes (SHCC, n=18). A minimum read depth of 20 and a genotype quality score of 30 were required for sample analyses. Regions containing pseudogenes and variants with < 90% total successful call rate were excluded. Excluding synonymous SNVs, 64 variants in HCC genes and 64 variants in SHCC genes were detected in Type X CRC cases. Three were known pathogenic variants (STK11 F354L, MLH3 V741F and APC E1229Q). Three novel variants were splice variants (one each in PALB2, NUDT1, RECQL5), one was in an initiation codon (NUDT1), and 29 were predicted to alter splicing in silico tools (AXIN1, CTNNB1, FLCN, GALNT12, PALB2, RECQL5, APC, BMP4, CHEK2, MLH1, MSH2, MUTYH, PMS1, PTEN, SMAD4). The occurrence of potentially deleterious MMR gene variants suggests possible misclassification of MMR status. Notably, among these 144 Type X CRC cases, 35 (24.3%) carried likely deleterious mutations in known CRC and suspected CRC genes, suggesting the value of a broad cancer gene panel. To uncover the remaining heritable factors in Type X CRC, however, more work remains. This work was supported by grant UM1 CA167551 from the National Cancer Institute.
Targeted Germline Sequencing of Young Onset, Proficient Mismatch Repair Colorectal Cancer Genes. M.S. DeRycke, S.R. Riska, S.K. McDonnell, A.J. French, Z.C. Fogarty, S. Midha, S. Baheti, Sunaw ardena, D.J. Schaid, Y. Zhang, D. Buchanan, M. Clendenning, G. Casey, M. Cotterchio, D.J. Duggan, S. Gallinger, R.W. Halle, J.L. Hopper, M.A. Jenkins, L. Le Marchand, J.D. Potter, P.A. Newcomb, M.O. Woods, N.M. Lindor, E.L. Goode, S.N. Thibodeau on behalf of the Colon Cancer Family Registry. 1) Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Oncogenomics Group, Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Melbourne, Australia; 3) Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, CA; 4) Prevention and Cancer Control, Cancer Care Ontario, Ontario, Canada; 5) Translational Genomics Research Institute (TGen), Phoenix, AZ; 6) Lunenfeld Tannenbaum Research Institute, Toronto, Ontario, Canada; 7) Department of Medicine and Stanford Cancer Institute, Stanford, CA; 8) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Victoria, Australia; 9) Centre for Epidemiology & Biostatistics, Melbourne School of Population & Global Health, University of Melbourne, Melbourne, Victoria, Australia; 10) Epidemiology Program, University of Hawai’i Cancer Center, Honolulu, HI; 11) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 12) Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John’s, Newfoundland, Canada; 13) Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ.

Although the overall incidence of colorectal cancer (CRC) is decreasing, incidence in individuals diagnosed before the age of 50 has increased. While familial syndromes account for approximately 20% of young onset cases, the genetic understanding remains incomplete for the remaining cases. We studied 370 young-onset (YO, <50 years) proficient mismatch repair (pMMR) cases using targeted capture and sequencing to search for genes contributing to the susceptibility of YO pMMR CRC. Agilent SureSelect custom capture targeted 2.9 Mb of coding DNA for sequencing genes on an Illumina HiSeq2000, including known and suspected hereditary CRC genes (HCC, n=36), several miRNAs (n=118), and genes identified by linkage (n=185) or whole exome sequencing (n=363) for the remaining cases. We identified multiple TF binding sites located on the OLFML3 promoter. Multiple TF binding sites located on the 3'UTR in humans and has been validated in animal lung as well as many other cancers. One miRNA, miR-155, was predicted to be significantly under-expressed in LC tissues. Additionally, we explored the SNPs located on OLFML3 and validated this prediction in LC tissues.}

Characterization of OLFML3 mutations in non-small cell lung cancer. C. Drennan, M. Orloff. Department of Epidemiology, University of Arkansas for Medical Sciences, 4301 W. Markham Street, #820 Little Rock, Arkansas 72205.

In Arkansas and the U.S., lung cancer (LC) accounts for more deaths each year than breast, colon, and prostate cancers combined. To date, the phenotypic heterogeneity associated with LC, particularly non-small cell LC (NSCLC), has been the main obstacle in the development of effective prevention and treatment. Recently, an integrative ‘omics’ analysis of significant regions of tracts of homozygosity combined with genome-wide expression array data was used to map and shortlist NSCLC-related genes (p-values <0.0001). Of these 9 significant genes, OLFML3[MIM 610088], which has been implicated in apoptosis, tumor growth, and cell cycle regulation in cancer cells, was reported to be significantly under-expressed in LC cases who were ever and never smokers (p-values <0.0001 and <0.006, respectively). As a follow-up, we postulate that mutations located in OLFML3 exons, splice sites, and promoter regions contribute to either a risk or protective effect in NSCLC patients.

To address our hypothesis, NSCLC tumor and matched adjacent normal tissues were retrospectively obtained from the University of Arkansas for Medical Sciences (UAMS) Tissue Procurement Facility. From these tissues, we simultaneously extracted DNA, RNA, and protein. We screened for mutations using the Sanger sequencing method and in patients who were positive for OLFML3 mutations, the transcript levels of OLFML3 are being assessed. We explored predicted transcription factor (TF) and microRNA (miRNA) binding sites located on OLFML3. Our search identified multiple TFs (MZF1, C/EBPb, SRY, Oct-1, Sp1, AML-1a, delta E, HNF-3b, v-Myb, XFD1, Nkx-2.1, CdxA) that have been previously described as important regulators of carcinogenesis. Additionally, we explored the SNPs located on OLFML3 to identify predicted TF binding patterns that could be affected and thus potentially change the expression levels. On OLFML3 exon 3, rs2055542, resides 8 base pairs from where transcription factor SRY was predicted to bind at 100% threshold. SRY transcription factors have been shown to regulate lung as well as many other cancers. One miRNA, miR-155, was predicted to bind to OLFML3 3' UTR in humans and has been validated in animal studies. miR-155 is known to target OLFML3, miR-155 has also been linked to angiogenesis and tumor proliferation. Our continued investigation and characterization of OLFML3 mutations combined with OLFML3-specific functional analyses may unravel a biologically plausible explanation to NSCLC.
Hox pattern expression and non coding transcripts in the HOX locus are associated with adult medulloblastoma subtype. A. M. Fontes\textsuperscript{1}, K. J. Abraham\textsuperscript{1}, D. B. Graha\textsuperscript{1}, D. T. Coven\textsuperscript{2}, B. Veiga\textsuperscript{3}, R. M. R. Trappe\textsuperscript{4}, F. S. Ramahlo\textsuperscript{5}, A. A. Cardoso\textsuperscript{6}, S. L. Gerson\textsuperscript{7}, C. G. Carlotto\textsuperscript{1}.\textsuperscript{1} Department of Genetics, School of Medicine of Ribeirão Preto, University of Sao Paulo, Brazil; \textsuperscript{2} Department of Pediatrics, School of Medicine of Ribeirão Preto, University of Sao Paulo, Brazil; \textsuperscript{3} Department of Sao Paulo, Brazil; \textsuperscript{4} Department of Surgery and Anatomy, School of Medicine of Ribeirão Preto, University of Sao Paulo, Brazil; \textsuperscript{5} Department of Pathology and Legal Medicine, School of Medicine of Ribeirão Preto, University of Sao Paulo, Brazil; \textsuperscript{7} Hermann B. Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, USA; \textsuperscript{8} Case Comprehensive Cancer Center, School of Medicine, CWRU, USA.

Introduction: Medulloblastoma (MB) represents a heterogeneous group of neuroepithelial primary tumors arising from the cerebellum. MB can be classified at the molecular level into at least four major subtypes: WNT, SHH, Group C and Group D. Although the current management of MB patients involves specific treatment approaches to their molecular subtype, a significant proportion of patients do not respond to treatment suggesting heterogeneity within these subtypes. We investigate for differences in Hox gene expression and their regulators among MB subtypes to provide additional biomarkers for MBs and molecular mechanism of HOX regulation involved in tumor progression. Methodology: In this study we analyzed 5 frozen tumor samples from 5 adult MB patients, 3 control cerebellar tissues, 4 MB cell lines and 3 control cerebellar primary cells. Two color microarray-based gene expression profiling was performed according to the manufacturer’s instructions. The microarray analysis could be related to HOX genes, since the cerebellum Hox genes are downregulated or not expressed, with the exception of CNS. We found that: HoxD1, D3, D4, D8 and D9, HoxC4, HoxA9 and A10 are overexpressed in WNT MB; HoxA2, A3, A4, A5, A6, A7, A8, A10, A11, HoxC8, C10 and A1 are overexpressed in SHH and Group CMB and Group D MB. Since HOTAIR is located a distance of 5,726 bp from HoxC11, we look for 5'-ends and 3'-ends of long ncRNAs located 1bp - 10 kbp apart from each 39 HOX genes. The analysis shows the presence of 5 Hox-related long ncRNAs: HOTSIR-35 mapped at HoxD9, HOTSIR-24 mapped at HoxD3, HOTSIR-49, HOTSIR-69 and HOTSIR-27 mapped at HoxC domain. HOTAIR is not expressed in MB. These long ncRNAs are differentially expressed among MB subtypes. Validation analyses and genetic studies are ongoing to elucidate the molecular mechanisms of HOX regulation according to MB subtype, which can provide useful insights for biomarkers and therapeutic intervention.

Ptprj-interacting susceptibility genes for colorectal cancer. M. Gerber, M. Cianciolo, A. Tolan. Department of Molecular Virology, Immunology, & Medical Genetics, The Ohio State University, Columbus, OH.

Colorectal cancer (CRC) causes over 50,000 deaths in the United States each year and is the third leading cause of cancer deaths. The identification of genetic risk factors underlying CRC will have immense value as a tool to identify individuals with increased predisposition to this cancer and to highlight new potential therapeutic targets. In the mouse, the gene Ptprij maps to a region of the genome linked to colon cancer susceptibility in mice. The Ptprij locus interacts synergistically with two other CRC susceptibility loci mapped in the mouse (Scs5 and Scs13) to further increase CRC risk. The Scs5 locus also independently interacts in a reciprocal manner with Scs0 to augment susceptibility. Importantly, the susceptibility genes at Scs4, Scs5, and Scs13 have not yet been identified. Our goals are (1) to identify which of the multiple genes present at these Scs loci are responsible for modifying CRC risk, and (2) to understand how the combination of the susceptibility genes at these regions increases CRC risk. To achieve these goals, RNA-Seq was used to identify genes with genetic variants or expression levels that differ between the CRC-resistant and CRC-susceptible mouse strains used to map the Scs loci. Next, SNPs of the human orthologs of these candidate susceptibility genes were tested in 194 pairs of normal and colon tumor DNA from human CRC patients for evidence of allele-specific gains or losses. SNPs were identified in 5 of these genes in which candidate susceptibility genes were tested in 194 pairs of normal and colon tumor DNA samples compared to matched healthy colon DNA samples (p-values<0.05). From these studies, Epas1 (Scs4), Csnk1a1 (Scs5), and Prdm5 (Scs13) emerged as leading candidates at the loci of interest. Preliminary studies to uncover interactions among these candidate genes using in vitro models have thus far shown that the transcription factor Epas1 may regulate expression of Csnk1a1 and Ptprij. The epigenetic regulator Prdm5 may also exert regulatory effects at these loci. Furthermore, Epas1 may promote Wnt signaling in colon epithelial cells and therefore function in an oncogenic capacity. Future work will delve deeper into the exploration of this complex network of gene-gene interactions by manipulating the expression levels of these genes and identifying effects on cell growth, death, and other cancer-relevant phenotypes.
3243M
A Novel Risk Variant at the 8q24 Cancer Susceptibility Locus in Men of African Ancestry


S100A4 is one of the members of the S100 family that characterized as a small calcium-binding protein, and frequent overexpression as well as positive association between overexpression and metastasis were reported. However, detailed mechanisms for such characteristics are not well understood. We analyzed S100A4 in pancreatic and lung cancers and observed frequently overexpression in both tumor types, irrespective of histological subtype. Methylation status in the CpG sites in intron 1 associated with expression of S100A4. In some tumors, however, not methylation at the CpG sites but acetylation of histone regulates expression. Then we performed knockdown and forced expression of S100A4 in pancreatic and lung cancers. Results of both tumor types shared high similarities; specific knockdown of S100A4 effectively suppressed cell proliferation, mainly by induction of apoptosis, only in cancer cells with S100A4-overexpression, and forced expression of S100A4 accelerated cell motility only in S100A4 low-expressing cancer cells. Furthermore, S100A4 expression was significantly correlated with perineural invasion (P = 0.029) and invasion pattern (P = 0.001) in primary pancreatic cancer tissues. Microarray analyses using pancreatic cancer were performed both after specific knockdown and forced expression of S100A4, and we identified several candidate genes such as PRDM2 and VASH1 after knockdown, and IFI27 and NOV after forced expression. It is notable that positive association between IFI27 expression and perineural invasion (P = 0.0023) was observed. Our results suggest that S100A4 plays an important role in pancreatic and lung carcinogenesis, and that these two tumor types may share the same or similar pathway in cell proliferation and motility.

3243T
Molecular characterization of oncogenic properties of S100A4 in pancreatic and lung cancers and identification and characterization of candidate downstream genes.


S100A4 is one of the members of the S100 family that characterized as a small calcium-binding protein, and frequent overexpression as well as positive association between overexpression and metastasis were reported. However, detailed mechanisms for such characteristics are not well understood. We analyzed S100A4 in pancreatic and lung cancers and observed frequently overexpression in both tumor types, irrespective of histological subtype. Methylation status in the CpG sites in intron 1 associated with expression of S100A4. In some tumors, however, not methylation at the CpG sites but acetylation of histone regulates expression. Then we performed knockdown and forced expression of S100A4 in pancreatic and lung cancers. Results of both tumor types shared high similarities; specific knockdown of S100A4 effectively suppressed cell proliferation, mainly by induction of apoptosis, only in cancer cells with S100A4-overexpression, and forced expression of S100A4 accelerated cell motility only in S100A4 low-expressing cancer cells. Furthermore, S100A4 expression was significantly correlated with perineural invasion (P = 0.029) and invasion pattern (P = 0.001) in primary pancreatic cancer tissues. Microarray analyses using pancreatic cancer were performed both after specific knockdown and forced expression of S100A4, and we identified several candidate genes such as PRDM2 and VASH1 after knockdown, and IFI27 and NOV after forced expression. It is notable that positive association between IFI27 expression and perineural invasion (P = 0.0023) was observed. Our results suggest that S100A4 plays an important role in pancreatic and lung carcinogenesis, and that these two tumor types may share the same or similar pathway in cell proliferation and motility.

3245S
Estrogen Receptor Gene Polymorphisms and Lung Adenocarcinoma Risk in Never-smoking Women. CF. Hsiao1, KY. Chen2, GC. Chang3, YH. Tsa4, WC. Su5, YM. Chen6, MS. Huang7, C. Hsing8, CJ. Chen9, PC. Yang1.

1) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; 2) Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan; 3) Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung; 4) Division of Pulmonary Medicine, National Yang-Ming University, Taipei, Taiwan; 5) Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan; 6) Division of Hematology/Oncology, Department of Internal Medicine, National Cheng Kung University, Tainan, Taiwan; 7) Division of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 8) Genomics Research Center, Academia Sinica; and Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan.


S100A4 is one of the members of the S100 family that characterized as a small calcium-binding protein, and frequent overexpression as well as positive association between overexpression and metastasis were reported. However, detailed mechanisms for such characteristics are not well understood. We analyzed S100A4 in pancreatic and lung cancers and observed frequently overexpression in both tumor types, irrespective of histological subtype. Methylation status in the CpG sites in intron 1 associated with expression of S100A4. In some tumors, however, not methylation at the CpG sites but acetylation of histone regulates expression. Then we performed knockdown and forced expression of S100A4 in pancreatic and lung cancers. Results of both tumor types shared high similarities; specific knockdown of S100A4 effectively suppressed cell proliferation, mainly by induction of apoptosis, only in cancer cells with S100A4-overexpression, and forced expression of S100A4 accelerated cell motility only in S100A4 low-expressing cancer cells. Furthermore, S100A4 expression was significantly correlated with perineural invasion (P = 0.029) and invasion pattern (P = 0.001) in primary pancreatic cancer tissues. Microarray analyses using pancreatic cancer were performed both after specific knockdown and forced expression of S100A4, and we identified several candidate genes such as PRDM2 and VASH1 after knockdown, and IFI27 and NOV after forced expression. It is notable that positive association between IFI27 expression and perineural invasion (P = 0.0023) was observed. Our results suggest that S100A4 plays an important role in pancreatic and lung carcinogenesis, and that these two tumor types may share the same or similar pathway in cell proliferation and motility.

3245S
Estrogen Receptor Gene Polymorphisms and Lung Adenocarcinoma Risk in Never-smoking Women. CF. Hsiao1, KY. Chen2, GC. Chang3, YH. Tsa4, WC. Su5, YM. Chen6, MS. Huang7, C. Hsing8, CJ. Chen9, PC. Yang1.

1) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; 2) Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan; 3) Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung; 4) Division of Pulmonary Medicine, National Yang-Ming University, Taipei, Taiwan; 5) Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan; 6) Division of Hematology/Oncology, Department of Internal Medicine, National Cheng Kung University, Tainan, Taiwan; 7) Division of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan; 8) Genomics Research Center, Academia Sinica; and Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan.

1.168 - 1.950; rs985192: OR: 1.390, 95% CI: 1.001 - 1.712). Therefore, we assessed the associations between lung adenocarcinoma risk and SNPs in ER. The case-control study included 532 never-smoking female patients with lung adenocarcinoma and 532 healthy controls. The ESPR1, a total of 98 SNPs were retrieved and analyzed. Among the retrieved SNPs, 7 tagged SNP associated with lung adenocarcinoma risk was identified. A stepwise forward selection logistic regression approach was performed, with adjustment for hormone replacement therapy (HRT), education level, passive smoking, cooking fume exposure, and other lifestyle factors. Two SNPs, rs7753153 and rs985192, were significantly associated with lung adenocarcinoma risk. (rs7753153: OR: 1.509, 95% CI: 1.168 - 1.950; rs985192: OR: 1.390, 95% CI: 1.001 - 1.712). Therefore, we may conclude that estrogen receptor gene SNPs may be associated with lung adenocarcinoma in never-smoking women.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
PALB2 mutations can help to identify to be determined. Our analysis will contribute to the improvement in estimation of the disease-associated risk in the PALB2 mutation carriers. Moreover, unravelling the families with germline PALB2 mutations can help to identify high-risk individuals who may benefit from early detection screening tests. Supported by grants: IGA MZCR NT 14006-3/2013, PRVOUK-P27/FL1/1, and SVV-UK 2682-2014.

**3247T**


1) Hematology, oncology, transfusion medicine center, Vilnius University Hospital Santariskiu Clinics, Vilnius, VNO, Lithuania; 2) State Research Institute, Innovative Medicine Center, Vilnius, VNO, Lithuania; 3) Department of Gynecology, Centre of Women’s Physiology and Pathology, Vilnius University Hospital Santariskiu Clinics, Vilnius, VNO, Lithuania; 4) Institute, Innovative Medicine Center, Vilnius, VNO, Lithuania

**BACKGROUND.** Genetic predisposition to breast and/or ovarian cancer is largely confined to mutations in BRCA1/2 genes, although rarer mutations in other known genes are also important. Massively parallel (or next-gener-ation, NGS) sequencing technology is used for identification of mutations that predisposing mutations in selected known genes (panels) and discover new associations.

**METHODS.** We aimed to better characterise cancer predisposing landscape in clinically selected 127 breast (BC) and 88 epithelial to ovarian (EOC) cancer cases from Lithuania (with family history or early age at diagnosis and negative for previously tested BRCA1/2 genes mutations) by performing NGS based targeted analysis (genomic DNA) of genes previously associated with both common (e.g., breast, ovarian) and rare cancers. Custom made TruSight Cancer Nextera Custom hybridization-based target enrichment method was used for the targeted hybridization and preparation of genomic libraries, which were sequenced on MiSeq (2x150-cycles; Illumina). VariantStudio software was used for annotation and filtering of genetic variants. Further analysis filter encompassed 32 genes previously associat-ed with BC/OC predisposition and other mechanistically implicated cancer predisposition genes. RESULTS. Germline loss-of-function protein truncating mutations (PTM) were identified in 13.4% of BC (17/127) and 3.4% of OC (3/88) BRCA1/2 negative samples. Seven PTM were implicated in 6 genes previously tested with BC/OC predisposition, and were further investigation. Other PTM were in 6 known BC/OC genes (BRIP1, CHEK2, PALB2, RAD51C, MLH1, TP53). RESULTS. Potentially pathogenic missense mutations (MM) were identified in 10.2% of BC (13/127) and 3.4% of OC (3/88) BRCA1/2 negative samples. Several PTM were located in 5 genes (ATM, CDH1, BRIP1, NBN, MUTYH) previously associated with BC predisposition. One compound HOXB13 p.V64E + NBN p.I171V and one MUTYH p.C690D + p.S910L were identified in BC cancer patient with identified. Two PTM were in 2 known BC/OC susceptibility genes (PALB2, TP53) with CDH1 (1 MM) was not previously associated with predisposition to OC. Further functional characterisation of selected MM variants is currently ongoing. CONCLUSION. NGS panel based resequencing is effective way for better characterising of cancer predisposition landscape.

**3247S**

Identification of germline mutations in hereditary prostate cancer families satisfying clinical testing criteria for hereditary breast and ovarian cancer (HBOC). The University of Michigan Prostate Cancer Genetics Project (PCGP) selectively enrolls families with 1) a living affected relative pair diagnosed with Ca at any age or 2) an individual diagnosed with Ca before age 56 regardless of family history. For this study, we restricted our sample to PCGP families with at least 1 male BrCa case or a female BrCa or ovarian cancer case in a first-degree relationship to a PCa case. With the exception of the families with male BrCa, we further enriched our sample set to increase our expectation of finding a mutation in BRCA1/2 by requiring these families to also have at least 1 family history factor described by the U.S. Preventive Services Task Force that would qualify as a suspicious woman for clinical BRCA1/2 testing (N=50). The PCa cases with the earliest onset disease (Avg age dx = 54.9 years; range 41-77), with some preference given for clinically significant PCa, was selected for sequencing of BRCA1/2. Due to the lack of available HBOC genotyping in these PCGP families, we performed NGS based targeted analysis (genomic DNA) of genes previously identified by the Cancer Genome Project as high-risk genes for Ca. We identified 30 potentially pathogenic mutations in 30/50 (60%) PCGP PCa families.

**3248S**

Fine-mapping of 67 prostate cancer GWAS regions identifies better and multiple association signals. Z. Kote-Jarai1, A. AminAl Olama2, T. Dadaev1, D. Leongamornlert1, D.J. Hazelett3, E. Saunders1, D. Easton2, G.A. Coetzee2, D. Conti4, R. Eeles1, The PRACTICAL Consortium. 1) Oncogenetics, Institute of Human Genetics, University of Oxford, Sutton, Surrey, United Kingdom; 2) Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, United Kingdom; 3) Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

Germ line common prostate cancer (PrCa) susceptibility loci have been reported in various genome wide association studies. However, the published associated SNPs are most often only tagging the genomic regions that include the possible causative variants and it remains challenging to identify these variants. We aimed to develop a powerful new statistical analysis tool to evaluate the association between genetic variation and risk across 67 PrCa regions. The genotyping data from a custom Illumina iSelect array, iCOGS, from PRACTICAL, a large multi-national collaboration and two UK GWAS datasets were used for imputation using 1000 Genomes. Further analysis filter encompassed 32 genes previously associated with Not yet determined if these the association did not reach genome wide significance. Four regions were identified as having evidence of stratifying population risk using polygenic risk score after including the best signal and selected SNPs significant at P<10^-5 level for the second SLR. After this only one SNP remained in the model in 21 regions and 19 regions only one SNP remained in the model . In 9 of these the association did not reach genome wide significance, four regions were identified as having evidence of stratifying population risk.
3250T

Localization and Expression Level of p16 Correlate with Patient’s Survival and Human Papilloma Virus Status in Oropharyngeal Squamous Cell Carcinoma. S. Lai1,2,3, V. Sandulache1,4, J. Zevallos1,2,1. 1) Department of Pathology, Michael E. DeBakey VA Medical Center, Houston, TX, USA; 2) Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA; 3) Department of Otolaryngology, Baylor College of Medicine, Houston, TX, USA; 4) Department of Otolaryngology, Michael E. DeBakey VA Medical Center, Houston, TX, USA.

p16 is a tumor suppressor gene. Its gene product has been reported to be overexpressed in many head and neck squamous cell carcinoma (SCC) during the last decades. The association between human papilloma virus (HPV) and head and neck SCC has also been recognized. HPV is thought to cause p16 overexpression in E7-induced inactivation of the retinoblastoma protein. p16 immunoreactivity is often considered as an excellent surrogate marker for cancer risk and outcome. However, the differential expression of p16 overexpression in nuclear and cytoplasm and correlation with HPV status of oropharyngeal SCC of veteran male patients has not been studied. Tumors in oropharyngeal area are difficult to be resected completely, often unresectable due to complexity of location. Study of the pattern of p16 overexpression, HPV infection, and patient’s clinical outcomes plays an important role for therapeutic implication. We studied p16 expression by immunohistochemistry in 135 oropharyngeal SCC of veteran male patients. P16 expression pattern is divided to the five groups: low nuclear/low cytoplasm, high nuclear/low cytoplasm, low nuclear/high cytoplasm, high nuclear/high cytoplasm and no expression. Five year disease-free survival is 28.4%, 74.7%, 0%, 93.1% and 13.7% in each group, respectively. Five year overall survival is 23.5%, 74.2%, 0%, 88.7% and 2.4%, respectively. Nuclear and cytoplasmic expression of p16 correlated with HPV status (p<0.05). We observed the strongest when cases with a 1st and 2nd degree carriers (carrier frequencies: 1.55% vs. 0.51%). The association was stronger when cases with a 1st and 2nd degree carriers (carrier frequencies: 1.55% vs. 0.51%). There was no association of Hoxb13 G84E with PrCa risk and prognosis using data from the UK Genetic Prostate Cancer Study Collaborators. 1,2,3 A.C. Antoniou1, F. Lhota1, E.J. Sawyer1,2,3, C. Mikropoulos1, T. Daday1, M. Tymrakiewicz1, E.J. Saunders1, S. Jogumart-Little1, K. Govindasami1, M. Guy1, R.A. Wilkinson1, E.J. Sawyer1, A. Morgan1, D. Neal2, F. Hamdy1, J. Donovan1, A.C. Antoniou1, Z. Kote-Jarai1, R.A. Eeles2, The UK Genetic Prostate Cancer Study Collaborators. 1) Cancer Genetics, The Institute of Cancer Research, Sutton, Surrey, United Kingdom; 2) Surgical Oncology, University of Cambridge, Addenbrooke’s Hospital, Cambridge CB2 0QQ, UK; 3) Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge CB2 2RE, UK; 4) Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK; 5) School of Social and Community Medicine, University of Bristol, Bristol BS8 2PS, UK; 6) Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge CB1 8RN, UK; 7) The Royal Marsden NHS Foundation Trust, London SM2 5PT, UK.

3252M

Identification of hereditary alterations predisposing to breast cancer using Next-Gen Sequencing. F. Lhota1, V. Stranecky2, P. Boudova1, J. Soukopova1, H. Hartmannova2, K. Hodanova2, P. Kleiblova1, M. Janatova1, P. Palheilich1, S. Knoc2, Z. Kleibal1. 1) Inst. of Biochemistry and Exp. Oncology, 1st Faculty of Medicine, Charles University in Prague, Prague, Czech Republic; 2) Inst. of Inherited Metabolic Disorders, 1st Faculty of Medicine, Charles University in Prague, Prague, Czech Republic.

The most common cancer not only in the Czech Republic is represented by Breast cancer (BC). Previously, we have tested 1035 high-risk breast/ovarian cancer patients and identified 269 carriers of pathogenic variants in eight BC susceptibility genes (BRCA1, BRCA2, p53, PALB2, CHEK2, ATM, NBN, and PPM1D). The most frequent alterations in identified mutation carriers were pathogenic variants in BRCA1 (118%). Identification of some pathogenic variant failed in 74% of analyzed patients from high-risk families. With the aim to identify underlying BC susceptibility variants within hereditary breast cancer (HBC) patient subgroup (N=853), we launched next-gen sequencing (NGS) project targeting 594 genes that code for proteins involved in DNA repair or influencing BC pathogenesis using custom sequence capture panel and SOLID sequencing. The NGS analysis in 314 high-risk patients and 101 non-cancer controls has been performed so far. We identified 84 (66 unique) protein truncating variants in 58 genes; 35 of these mutations were found in DNA repair genes in 33 out of 314 patients (10%). Totally, 387 missense variants (in 186 genes) were predicted as deleterious in 15% of patients. Moreover, we found also 19 splicing site alterations. All truncating mutations were confirmed by Sanger sequencing. These variants affect several interesting genes including genes coding for Fanconi anemia proteins and proteins involved in DNA-repair pathways (e.g. ATM, NBS1, NBN). These findings allow to analyze free and overall population is a rational, powerful, and economic strategy surpassing the classical strategies of mutation analysis. However, the implementation of exome-wide or genome-wide NGS approaches will be indispensable for the detection of population-specific rare variants or pathogenic variations in individual high-risk patients. By identification of truncating variants in potentially “actionable” genes, we are reducing the percentage of uncertainty in characterization of BC-susceptibility genes in high-risk patients from 75% down to 50 %. Supported by grants: IGA MZCR NT14054, IGA MZCR NT14006-3, PRVOUK P27, and SVV-UK 23622014.
Germline copy number variant analysis as a mechanism to identify novel high-risk endometrial cancer gene mutations. F. Lose1, G.L. Morris2, J.F. Pearson3, M. Bowman4, R.I. Scott5, M. McEvoy6, J. Attia5, E. Holliday5,6, P.D. Pharoah7,10, A.M. Dunning7, D.J. Thompson6, D.F. Easton8,10, L.C. Walker9, A.B. Spurdle1, HCS, SEARCH, ANECS. 1: Molecular Cancer Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia; 2: Mackenzie Cancer Research Group, Department of Pathology, University of Otago, Christchurch, New Zealand; 3: Biostatistics Unit, University of Otago, Christchurch, New Zealand; 4: Centre for Information Based Medicine and the School of Biomedical Science and Pharmacy, University of Newcastle, Newcastle, New South Wales, Australia; 5: Hunter Medical Research Institute, Hunter Area Pathology Service, John Hunter Hospital, Newcastle, New South Wales, Australia; 6: Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia; 7: Department of General Medicine, Hunter Medical Research Institute, John Hunter Hospital, Newcastle, New South Wales, Australia; 8: Centre for Information Based Medicine and the School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia; 9: Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK; 10: Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK.

The current findings from studies demonstrate CNVs have potential to drive discovery of novel cancer genes, and there have been no CNV studies in endometrial cancer to date. We have utilised our published endometrial cancer Genome-wide association studies (GWAS) and large scale replication studies have identified more than 80 common loci associated with breast cancer. However, for the majority of these loci the underlying causal genes and variants are unknown. We attempted to define the causal variants underlying one such locus, on 1q24, using dense genotyping of variants on an ~200K custom array (IGCS), conducted by the Breast Cancer Association Consortium (BCAC). The region contains no RefSeq genes, and is flanked by TBX3, a gene known to be somatically mutated in breast tumours, and MD13L.

Across a 530kb region, 589 variants were genotyped on 48,155 cases and 43,612 controls of European ancestry from 41 studies and 6,269 cases and 6,624 controls of Asian ancestry from 9 studies. Genotypes for a further 4,799 variants were imputed using the 1000 genomes project data as a reference.

We identified 58 variants associated with breast cancer risk at P<5×10−8 in Europeans, the most strongly associated variant being rs1391721 (per-minor allele OR and 95% CI 0.92(0.90-0.94), P=3.7×10−17). rs1391721 is closely correlated with the original GWAS hit, rs129201. Forward stepwise regression analysis in the region revealed two additional independent variants contributing to the breast cancer risk, at P<5×10−6: rs476780 which lies around 400kb centromeric to the original variant (adjusted OR and 95%CI 0.96(0.94-0.98), P=8.2×10−6), and a 4 base insertion rs200803242 (adjusted OR and 95%CI 1.06(1.03-1.08), P=3.3×10−5). rs1391721 was associated primarily with ER-positive disease (case-only P=0.00013), whereas rs476780 and rs200803242 were associated with both ER-positive and ER-negative disease. Analyses in the Asian samples revealed a significant association for rs1391721 (OR and 95%CI 0.89(0.84-0.94), P=9.7×10−8) but not the other two variants.

For the primary signal rs1391721, we defined a set of 8 variants, spanning a 1.1kb interval, that could not be excluded as causal (iCHAV1): based on a likelihood ratio of 100:1. Of these variants, only rs1391721 lies in a predicted regulatory region, predicted to be binding site for multiple transcription factors including CTCF and FOXA1 in MCF7 cells based on ENCODE data. These results provide a basis for further analyses to define the causal gene(s) underlying this association.

1) INSTITUTO DE GENÉTICA HUMANA, UNIVERSIDAD DE GUADALAJARA, GUADALAJARA, JALISCO, MEXICO; 2) TRAYECTORIA DE GENOMICA ALIMENTARIA, UNIVERSIDAD DE LA CIENEGA DEL ESTADO DE MICHOCAN, SAHUAYO MICHOACAN MEXICO; 3) HOSPITAL CIVIL DE GUADALAJARA “DR. JUAN I. MENCHACA” GUADALAJARA, JALISCO MEXICO; 4) DEPARTAMENTO DE BIOLOGÍA MOLECULAR Y GENOMICA, GUADALAJARA JALISCO MEXICO; 5) GI CANCER RESEARCH LABORATORY, BAYLOR UNIVERSITY MEDICAL CENTER AT DALLAS, TEXAS USA; 6) DEPARTAMENTO DE CLINICAS, CUAJITOS, UNIVERSIDAD DE GUADALAJARA, TEPATITLAN JALISCO MEXICO.

BACKGROUND. Lynch Syndrome (LS) is characterized by germline mutations in the DNA mismatch repair (MMR) genes: MLH1, MSH2, MSH6 or PMS2. This syndrome is inherited in an autosomal dominant pattern and is characterized by early onset colorectal cancer (CRC) and extra-colonic tumors. The aim of this study was to identify mutations in MMR genes in four Mexican patients with LS. MATERIAL AND METHODS. The study comprised tumor and blood samples obtained from four patients diagnosed clinically with Lynch syndrome at the “Dr. Juan I Menchaca” Civil Hospital of Guadalajara, Jalisco, Mexico. All the patients signed an informed consent. The protein expression was evaluated by immunohistochemical analyses in tumoral tissues. Genomic DNA was extracted from peripheral blood samples for mutation analysis using dhPLC, sequencing and MLPA analysis.

RESULTS. Immunohistochemical analyses: LS-23 and LS-41 samples showed the absence of nuclear staining for MLH1, and samples LS-3 and LS-52 showed loss of nuclear staining for MSH2. Germline mutation analysis: One was a substitution at splice donor site, located in first base of intron 18 of MLH1, and another was a deletion/insertion of two bases, c.1852_1853delinsGC (K618A) which resulted in a frameshift deletion in MSH2. Germline mutation analysis: Altered or may introduce potential intronic splice sites, that are now under investigation. Rearrangement analysis was first validated on control patients with known rearrangements. We confirmed these results and could reliably detect the exact breakpoints of an inversion in MLH1, a duplication in MSH2 and four deletions in MLH1, PMS2 and MSH6. Rearrangement analysis of the patient cohort with the validated pipeline is currently being performed.

Chemotherapy-induced peripheral neuropathy (CIPN) affects up to 60% of breast cancer patients undergoing chemotherapy. However, work remains to elucidate the etiology of CIPN, including the role that germline mutations may play in CIPN symptom (CIPN-sx) onset and severity. Leandro-Garcia et al., 2013, found that SNPs in Ephrin A receptor genes EPHA4, EPHA5, and EPHA6 were associated with CIPN in cancer patients treated with taxanes. In addition to replicating this finding, we investigated the association of significant SNPs with different types of CIPN-sx to further clarify the underlying biological mechanism and functional implications. This study utilized data from female breast cancer patients treated with (Ctx+, n=26) and within the (Ctx−, n=26) taxanes, assessed at baseline (post-surgery, pre-treatment) and one month post chemotherapy completion. CIPN-sx type and severity was collected with a validated 11-item self-report subscale (FACT/GOG-Ntx). In addition to total score, items are divided into sensory (numbness and tingling), motor, hearing, and functional domains. Candidate SNPs were selected by meta-analysis using the most significant SNP for each gene. Total CIPN-sx increased in Ctx+ compared to Ctx− (p=0.001); multivariate analysis showed that this signal was driven by significant increases in sensory (p=0.001) and motor (p=0.024) domains. Analysis of 18 Ctx+ white British Hispanic index individuals indicated that the deletion of exon 22, producing the incorporation of 11 frameshifted residues and the loss of 132 amino acids from the FANCM C-terminus (p.Gly1906Alafs12*). We also showed that this exon skipping is most likely due to the creation of a 6-bp insertion site for the splicing factor CIPN1. This is the first report providing robust evidence of a heterozygous loss-of-function mutation of the FANCM gene acting as a risk factor for familial breast cancer.

Identification of familial Wilms tumor predisposition genes using whole genome sequencing (WGS) in a family with bilateral nephronseismhesis. D. Papadopoulou1,2, H. Bhattacharya1, K. McGovern2, T. Bouchard1, M. Garcia1, E. Calva2, J. Smith1, D.J. West3, B. Schneider1,2, V.L. Champion2,4, A.J. Saykin1,5, A. Nyul5, 1) Department of Medical and Molecular Genetics, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 2) School of Nursing, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 3) Department of Radiology and Imaging Science, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 4) Training In Research for Behavioral Oncology and Cancer Control Program, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 5) Simon Cancer Center, Indiana University-Purdue University Indianapolis, Indianapolis, IN.

Whole genome sequencing of two members from Family 2 revealed the mutations that cosegregate in 2/47 families (G803R, R800fs). Peripheral blood samples from members of 47 WT families revealed novel germline DICER1 mutations in axonal guidance and repair following injury, and the potential impact of sensory and motor symptoms. Given the function of Ephrin receptor genes in EC1M 6BQ.

The CDH1 gene as a susceptibility locus for lobular breast carcinoma. C. Petridis1,2, I. Shinomiya1, S. Nowinski1, I. Tomlins1,2, R. Roylance1, M.A. Simpson2, E.J. Sawyer1. 1) Research Oncology, Division of Cancer Studies, King’s College London, Guy’s Hospital, Great Maze Pond, London, SE1 9RT; 2) Medical and Molecular Genetics, King’s College London, Guy’s Hospital, Great Maze Pond, London, SE1 9RT; 3) The Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN; 4) Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary, University of London, Whitechapel, London E1 1BQ.

Introduction: Both Invasive lobular breast cancer (ILC) and lobular carcinoma in situ (LCIS) are characterised by loss of E-cadherin expression, an adhesive molecule encoded by the CDH1 gene. Germline CDH1 mutations have been described in cases of familial ILC associated with diffuse gastric carcinoma (DGC), they are rare in sporadic lobular breast cancer and have not been described in women with LCIS. Common variation at this locus has been genotyped in breast cancer and have not been described in women with LCIS. Common variation at this locus has been genotyped in breast cancer patients and have not been described in women with LCIS. Common variation at this locus has been genotyped in breast cancer patients and have not been described in women with LCIS.

Anemia pathway, in 8,635 familial cases and 6,625 controls from different populations and found association with breast cancer risk (OR = 3.93, 95%CI = 1.28-12.11, P = 0.017). Following information theory-based prediction of significant SNPs with different types of CIPN-sx to further clarify the underlying biological mechanism and functional implications. This study utilized data from female breast cancer patients treated with (Ctx+, n=26) and within the (Ctx−, n=26) taxanes, assessed at baseline (post-surgery, pre-treatment) and one month post chemotherapy completion. CIPN-sx type and severity was collected with a validated 11-item self-report subscale (FACT/GOG-Ntx). In addition to total score, items are divided into sensory (numbness and tingling), motor, hearing, and functional domains. Candidate SNPs were selected by meta-analysis using the most significant SNP for each gene. One of these genes may play in CIPN symptom (CIPN-sx) onset and severity. Leandro-Garcia et al., 2013, found that SNPs in Ephrin A receptor genes EPHA4, EPHA5, and EPHA6 were associated with CIPN in cancer patients treated with taxanes. In addition to replicating this finding, we investigated the association of significant SNPs with different types of CIPN-sx to further clarify the underlying biological mechanism and functional implications. This study utilized data from female breast cancer patients treated with (Ctx+, n=26) and within the (Ctx−, n=26) taxanes, assessed at baseline (post-surgery, pre-treatment) and one month post chemotherapy completion. CIPN-sx type and severity was collected with a validated 11-item self-report subscale (FACT/GOG-Ntx). In addition to total score, items are divided into sensory (numbness and tingling), motor, hearing, and functional domains. Candidate SNPs were selected by meta-analysis using the most significant SNP for each gene. Total CIPN-sx increased in Ctx+ compared to Ctx− (p=0.001); multivariate analysis showed that this signal was driven by significant increases in sensory (p=0.001) and motor (p=0.024) domains. Analysis of 18 Ctx+ white British Hispanic index individuals indicated that the deletion of exon 22, producing the incorporation of 11 frameshifted residues and the loss of 132 amino acids from the FANCM C-terminus (p.Gly1906Alafs12*). We also showed that this exon skipping is most likely due to the creation of a 6-bp insertion site for the splicing factor CIPN1. This is the first report providing robust evidence of a heterozygous loss-of-function mutation of the FANCM gene acting as a risk factor for familial breast cancer.

Identification of familial Wilms tumor predisposition genes using whole genome sequencing (WGS) in a family with bilateral nephronseismhesis. D. Papadopoulou1,2, H. Bhattacharya1, K. McGovern2, T. Bouchard1, M. Garcia1, E. Calva2, J. Smith1, D.J. West3, B. Schneider1,2, V.L. Champion2,4, A.J. Saykin1,5, A. Nyul5, 1) Department of Medical and Molecular Genetics, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 2) School of Nursing, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 3) Department of Radiology and Imaging Science, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 4) Training In Research for Behavioral Oncology and Cancer Control Program, Indiana University-Purdue University Indianapolis, Indianapolis, IN; 5) Simon Cancer Center, Indiana University-Purdue University Indianapolis, Indianapolis, IN.

Despite ongoing research, metastatic melanoma five-year survival rates remain low and treatment options limited. Researchers can access a rapidly growing amount of molecular and clinical information about melanoma that may be critical to understanding this disease and making clinically relevant treatment decisions. However, this information is becoming difficult to collate and clinically interpret due to its dispersed nature. Presented here is Melanoma Profiler, a new cloud-based web application for clinically relevant metastatic melanoma genomic research. It performs a tumor/normal comparative variant analysis, comparing proband modified genes and pathways to those in a curated set of characterized melanoma tumor samples. To provide clinical decision support, modified genes and pathways are cross-referenced with clinical and molecular data. These data are incorporated from publicly available sources, including: associations between gene expression and patient survival, data concerning drug targets, biomarkers, and druggability as well as past clinical trials. The resultant MelanomaDB cross-reference database integrates data from cutting edge research on a continuous basis. The interface is presented in two tiers: a front-panel summary tier and an evidential detail tier. For the busy clinician, immediately presented in the summary view is a dynamic report clearly highlighting the relevant actionable information, potential drug targets, and a summary of affected genes and pathways. For researchers that wish to dig deeper, they may browse annotated pathway data, and view a comprehensive comparative summary of affected genes in waterfall and heatmap reports. Melanoma Profiler is a free research tool hosted by Biomatiers, produced in collaboration with the University of Auckland, New Zealand. A recent Frontiers in Oncology methods article describing the research behind the application can be found at http://bit.ly/1bUilqS. The Melanoma Profiler web application may be found at https://apps.biomatters.com/melanoma-profiler/.


Breast and Ovarian Cancer (BC, OC) are by far the most common cancers affecting women. Approximately 5-7% of BC and 11-15% of OC are estimated to be caused by germline DNA mutations mainly in the BRCA1 and BRCA2 genes, but mutations in additional genes have been also demonstrated to account for significant cancer risk. Knowing the underlying molecular defect can be very valuable for diagnosis, guiding treatment and estimating recurrence risks. Next generation sequencing (NGS) enables development of high-quality clinical tests for disease diagnosis with fast turnaround times. We developed a comprehensive NGS diagnostic panel including 38 genes relevant for breast and ovarian cancer. In order to increase the diagnostic yield we divided these genes into two subpanels according to both phenotypes. Our NGS diagnostic criteria require that sequencing reaches a medium quality of more than 30 per sequence cycle (detection precision of 99.9%) and that coverage of more than 20 sequences per base pair in at least 98% of the analyzed regions is assured. In addition, for BRCA1 and BRCA2 we achieve the highest level of exactitude for NGS that a lab can offer at the current stage (i.e. 100% sequencing of the coding region and flanking intronic sequences with more than 20 sequences per base pair, and deletion/duplication analysis with MLPA). The NGS platform at M2G has been validated for diagnostic routing in terms of accuracy, analytical sensitivity, analytical specificity and precision. Up to 24 samples of patients with suspected hereditary breast and/or ovarian cancer are processed weekly on an automated workflow which combines capture workflow enrichment sample preparation (TruSightTM Cancer Panel) with massive parallel sequencing (Illumina MiSeq). Data analysis is performed with a bioinformatics pipeline consisting of BWA, SAMtools, and mpEff. Our preliminary results show that 7.4% of the patients had BRCA1/2 patogenic mutations and 92.6% were non-patogenic or inconclusive. Further analysis of these patients revealed 13.9% of the cases patogenic variants (CDH1, CHEK2 and RAD51D genes) and 25% presented inconclusive VOUS (ATM, BRIP1, FANCI, MSH2, PALB2, PMS2, PTEN, RECQL4 and RHBD2 genes). Still 61.1% of the cases remained unsolved. Some of the genes were identified as causative only once, emphasizing the advantage of diagnostic panels for breast and ovarian cancer testing.

ABRAXAS (FAM175A) and breast cancer susceptibility: no evidence of association in the Breast Cancer Family Registry. A. Renault, F. Lesueur, P. Soucy, Y. Hamdi, Y. Coulombe, S. Gobeil, F. Le Calvez-Kelm, M. Vallée, I.L. Hopper, I.L. Andruis, M.C. Southey, E.M. John, Y.J. Masson, S.V. Tavtigian, J. Simard. The Breast Cancer Family Registry, 1) Curie Institute, Paris, France; 2) Cancer Genomics Laboratory, CHUQ Research Center, Quebec City, Canada; 3) Genome Stability Laboratory, Laval University Cancer Research Center, Quebec City, Canada; 4) CHUQ Research Center, Faculty of Medicine, Laval University, Quebec City, Canada; 5) Genetic Cancer Susceptibility, International Agency for Research on Cancer, Lyon, France; 6) Center for Molecular Environmental, Genetic and Analytical Epidemiology, School of Population Health, EGA The University of Melbourne, Victoria, Australia; 7) Department of Molecular Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada; 8) Genetic Epidemiology Laboratory, The University of Melbourne, Victoria, Australia; 9) Cancer Prevention Institute of California, Fremont, USA; 10) Stanford University School of Medicine and Stanford Cancer Institute, Stanford, USA; 11) Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, USA.

Background: Currently, less than 20% of the familial breast cancer can be explained by highly penetrant mutations in the BRCA1 and BRCA2 genes. Other breast cancer susceptibility genes conferring a moderate increase in risk have been recently identified. However, approximately 50% of the heritability is still unexplained, suggesting that other susceptibility genes remain to be discovered. DNA double-strand breaks are one of the most damaging events occurring in a cell, as they can disrupt the integrity and stability of the genome. The majority of breast cancer susceptibility genes already identified, including BRCA1, BRCA2, CHEK2, ATM, BRIP1 and PALB2, are involved in the DNA double-strand breaks repair pathway. ABRAXAS is involved in this pathway as a member of the “A complex” which leads BRCA1 to DNA damage site during homologous recombination repair. Mutations in ABRAXAS impair BRCA1 recruitment to DNA damage foci and increase cell sensitivity to ionizing radiation. Moreover, a recurrent germline mutation has been reported in Finnish high-risk breast cancer families. To determine if ABRAXAS could be a breast cancer susceptibility gene in other populations, we conducted a population-based case-control mutation screening study in the Breast Cancer Family Registry. Methodology/Principal finding: The nine coding exons of ABRAXAS were screened in 1,123 controls and 339 breast cancer cases. Sixteen distinct rare variants were identified, 15 in cases and 11 in controls: of these one was an in-frame deletion, eight were non-synonymous, four were synonymous, two were intronic and one was located in the 5’UTR of the gene. The two variants p.Thr141Ile (found in cases and controls) and p.Gly39Val (found in one case), were predicted to disrupt the protein function, and we confirmed experimentally that both of them diminish phosphorylation of gamma-H2AX, an important DNA damage signaling event. Conclusion: Overall, likely damaging or neutral variants were evenly represented among cases and controls suggesting that rare variants in ABRAXAS may explain at most a small proportion of hereditary breast cancer.
Association between rare and common variants in DNA repair genes and prostate cancer using the iCOGS genotyping array. A. Amin Al Olama1, D. Leongamornlert1, M. Tymrakiewicz1, S. Jugurnauth-Little1, A. Amin Al Olama2, S. Benlloch2, R. Eeles1, Z. Kote-Jarai1, The PRACTICAL Consortium. 1) Oncogenetics, Institute of Cancer Research, Sutton, Surrey, SM2 5NG, United Kingdom; 2) Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, Cambridge, CB1 8RN, United Kingdom.

Prostate cancer (PrCa [MIM 176807]) is the most frequently diagnosed male cancer in developed countries and amongst the leading causes of cancer related death. It is known to have a strong heritable component and so far 77 common, low penetrance susceptibility variants have been identified by genome-wide association studies (GWAS). In addition, a small number of rare variants have been found to give rise to greater risk; the majority of which are in DNA repair genes. We previously genotyped 211,155 SNPs on a custom chip (iCOGS) in blood DNA from 21,780 PrCa cases and 21,727 controls of European ancestry from the international PRACTICAL consortium. These SNPs were selected predominantly to evaluate previous suggestive low penetrance associations, to finemap validated associations, or to evaluate candidate genes and pathways. In this study we have analysed the association between common and rarer variants in DNA repair genes and PrCa. The iCOGS chip contained 9,192 SNPs within 10kb of a widely recognised DNA repair gene and at 7 of the 77 known PrCa susceptibility loci these were also significantly associated with risk (Chr3 - RUVBL1; Chr5 - FGF10; Chr9 - RAD23B; Chr11 - CCND1; Chr14 - RAD51B; Chr20 - RTEL1; Chr22 - BIK). These may therefore represent potential candidate genes for the aetiology of prostate cancer risk. No statistically significant associations with risk were observed for any of the rare variants (predominantly BRCA1/2 missense SNPs) submitted on the iCOGS chip. We also investigated association between DNA repair gene variants on the iCOGS array and phenotypic characteristics of PrCa; to examine whether candidate genes for the aetiology of prostate cancer risk. No statistically significant associations with risk were observed for any of the rare variants (predominantly BRCA1/2 missense SNPs) submitted on the iCOGS chip.

CONCLUSION

( NB and RMS). 

INTRO

Outside familial syndromes, the pathogenesis of rhabdomyosarcoma (RMS) and neuroblastoma (NB) remains obscure. We sought to identify rare deleterious germline variants that may play a role in the initiation of sporadic RMS and NB.

METHODS/RESULTS

We investigated two cohorts consisting of RMS (n=133) and NB (n=222) patients; the NB data is from the TARGET initiative. We first called high-quality protein-coding changing single nucleotide variants (SNVs) in both paired germline and tumor DNA from the illumina, SOLID and CGI platforms (exome and whole-genome sequencing). We excluded variants with minor allele frequency (MAF) > 0.01 in 1000G or NHLBI ESP to yield 56,336 germline variants. To find the most deleterious variants, we filtered for a C-score >= 20 (top 1%) using CADD, a method to integrate many diverse bioinformatics predictions (SIFT, PolyPhen, conservation) into a single score. We then binned variants in 3 categories: 1) variants in known cancer genes and select human syndromes (n = 177), 2) variants reported in >= 3 tumors from TCGA and our somatic NB and RMS data (n = 116), and 3) variants not known to be cancer-associated, have ESP MAF = 0 and a germline variant allele count of 2-8 in our data (n = 159). Of the resulting 446 SNVs (6 variants overlapped categories), we picked the 20 most-frequent and 198 random SNVs for Sanger verification sequencing. Out of the 305 verification sequencing reactions, no DNA was available for 121 sequencing reactions. Of the 184 successfully sequenced reactions, 2 failed PCR, and the verification rate equaled 97%; none of the TARGET NB samples had a false positive. Of the 446 SNVs, there were 2/133 (1.5%) germline variants in 2 known RMS-associated genes (TP53, CTCP) and 10/222 (4.5%) germline variants in 4 known NB-associated genes (ALK, NFI, PINK1, BARD). We also observed 3/133 (2%) germline ALK variants in 3 patients with RMS and 1/222 (each) germline TP53 and BUB1B variants in 2 NB patients. Known reported somatic mutations in NB with germline variants in NB patients included NRAS (1/222, 0.5%) and VANGL1 (4/222, 2%). Known reported somatic mutations in RMS with germline variants in RMS patients included FGFR4 (1/133, 1%) and BCOR (1/133, 1%). We also observed predicted deleterious germline variants in SDHB (NB), DICE1 (NB), ERCC2/3/4 (NB and RMS) and TSC1/2 (NB and RMS).

CONCLUSION

This high-quality dataset using stringent filtering, we found multiple novel germline variants associated with NB and RMS.
### 3266S

A recurrent germline mutation in the splicing factor SRRM2 gene is implicated in papillary thyroid cancer predisposition. J. Tomisco1, H. He5, K. Akagi1, S. Livanarchos1, Q. Pan1, B. Bertani1, R. Nagy2, D. Syme1,4,3, B. Blencowe2,2, A. de la Chapelle1,1. Department of Molecular Virology, Immunology and Medical Genetics, The Ohio State University, Columbus, OH, USA; 2) Banting and Best Department of Medical Research, University of Toronto, Toronto, Canada; 3) Department of Internal Medicine, The Ohio State University, Columbus, OH, USA; 4) Department of Biomedical Informatics, The Ohio State University, Columbus, OH, USA; 5) Department of Molecular Genetics, University of Toronto, Toronto, Canada.

Purpose: The role of chromosome 7 is thought to act as tumor-related genes, little is known about the details of them. Williams-Beuren syndrome (WBS) is one of the most well-known diseases related to chromosome 7 microdeletions and thought to be a tumor risk of multiple tumor types. We analyzed deletions in candidate genes, such as BCL7B. However, there is a little knowledge regarding the specific functional roles of BCL7B.

Methods: We analyzed the functional significance of BCL7B in KATOIII human gastric cancer cells and its homolog in Caenorhabditis elegans (e. c. elegans). We generated a bcl-7 deletion mutant, observed and analyzed its phenotype. As a result, we found that bcl-7 is required for the asymmetric differentiation of neuronal precursors of the seam and the gonad. This result suggested that BCL-7 functions as a negative regulator of the WNT pathway in c. elegans. In addition, bcl-7 deletion mutants exhibited nuclear enlargement, which was reminiscent of the anaplastic features of malignant cells. Second, to analyze the function of BCL7B in human cells, we used KATOIII cells, which were derived from gastric signet-ring cell cancer and expressed only BCL7B of the BCL7 family members. As a result, BCL7B knockdown induced nuclear enlargement, as observed in c. elegans, and promoted the multilobular phenotype. Moreover, we showed that Pat-1, a protein that is involved in WNT signaling in c. elegans, was also increased in the cell with enlarged nuclei. BCL7B knockdown also suppressed cell death by inhibiting the apoptotic pathway.

Results: In contrast, the function of the bcl-7 gene in C. elegans was also involved in WNT signaling in c. elegans and in human gastric cancer cells. Additionally, human BCL7B may function as a negative regulator of WNT signaling and a positive regulator of apoptosis in KATOIII cells. Furthermore, this study revealed that BCL7B is also involved in nuclear enlargement, which is an important signature of malignancy. Collectively, our data suggest that the members of the BCL7 family and their homologous proteins may function as tumor suppressors by affecting multiple pathways.

### 3268T

Investigating the genetic basis of multiple primary tumors. Clinical and gene panel analyses. J. Whitworth1,2, J. Hoffman1, AB. Skytte5, ER. Mahesh1,2, J. Boland1,2, 1) Clinical Genetics, Birmingham Women’s Hospital, Birmingham, United Kingdom; 2) Medical Molecular Genetics, University of Birmingham, Birmingham, United Kingdom; 3) Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark; 4) Medical Genetics, University of Cambridge, Cambridge, United Kingdom.

Multiple primary malignant tumors (MPMT) are frequently taken as an indicator of potential inherited cancer susceptibility and occur at appreciable frequency both among unselected cancer patients and referrals to cancer genetics services. Analysis of a referral based series of 212 MPMT cases showed that only around 40% of those who underwent genetic testing and 20% of referrals overall were identified as having a pathogenic germline variant conferring predisposition to malignancy. Comparison of individuals who tested positive and negative revealed considerable overlap between the two groups with respect to clinical characteristics indicative of an inherited cancer syndrome, suggesting that many of the latter group also have a genetic basis. Analysis of PTEN and TP53 by Sanger sequencing, however, did not reveal any significant variants. Failure to detect a causative mutation may result from mosaicism for a mutation in a known inherited cancer gene, or an unusual phenotype that leads to the relevant gene being overlooked or mutation in a novel inherited cancer gene. To address these possible explanations, next generation sequencing techniques are being applied to blood samples from an expanded series of mutation negative MPMT cases (>2 cancers diagnosed before 60 years) ascertained through clinical genetics services. Initial analysis is being performed using a panel of 94 known inherited cancer genes. Results from the first 63 samples produced 309 variant calls, of which 127 were categorized as likely pathogenic in the heterozygous state based on literature review and predicted effect. These 10 variants were identified among 8 cases (12.7%) and all but one was associated with a characteristic tumor. One case harboured three separate mutations in SDMC, FLCN and NF1, all of which were penetrant. In silico predictions revealed multiple further variants as potentially causative of the patient’s phenotype. Individuals not identified with explanation variants through panel analysis are being entered into a study conducting whole genome sequencing to identify potentially novel inherited cancer genes with subsequent functional validation. A sequencing dataset from large heterogeneous series such as this should provide opportunity for analysis of multiple phenotypic subsets and allow for potential pleiotropic effects of unidentified genes relevant to cancer predisposition.
3270M
Exome sequencing identified potential causative candidate genes for hyperplastic polyposis syndrome. S. Arezzi,2 C. Trueck,3 J. Altmueller4,2, S. Horpaopan1, P. Hoffmann1,1,4, H. Thiele1,3, J. Spier1,1 Institute of Human Genetics, University Hospital Bonn, Bonn, Germany; 2) Cologne Center for Genomics, University of Cologne, Germany; 3) Institute of Human Genetics, University of Cologne, Germany; 4) Department of Genomics, Life & Brain Center, University of Bonn, Germany; 5) Division of Medical Genetics, University Hospital Basel; Department of Biomedicine, University of Basel, Basel, Switzerland.

Purpose: Hyperplastic polyposis syndrome (HPS), also known as serrated polyposis syndrome (SPS), is a yet poorly defined colorectal cancer (CRC) predisposition characterized by the occurrence of multiple and/or large serrated lesions throughout the colon. A serrated poly-CRC sequence (serrated pathway) has been postulated, however, to date, only few molecular signatures of serrated neoplasia were described and neither the etiology of the syndrome nor the distinct genetic alterations during tumorigenesis have been identified. Methods: To uncover predisposing causative genes, the exomes of 11 unrelated and clinically well characterized HPS patients with sporadic appearance were sequenced (Illumina HiSeq platform). The variants were filtered for rare truncating germline mutations assuming a monogenic disease model. For data analysis and variant filtering the GATK software and in-house tools (VARBANK pipeline) were applied. Results: Allogether, 261 rare truncating germline variants were found. After stringent filtering steps including quality scores, the comparison with large databases from population-based controls, detailed manual investigations of the variants and data mining according to functions and pathways, 139 unique variants in 136 genes remained. Of those, six genes were affected by biallelic variants (recessive model) in at least one patient and 19 genes by heterozygous variants (dominant model) in at least two patients. The majority of these genes is supposed to be associated with cancer or is involved in molecular and cellular functions related to tumorigenesis such as DNA repair, apoptosis, or cell proliferation. Another 53 genes, which are affected by heterozygous variants in only one of the patients, are regarded as interesting candidates according to functional scores and known somatic mutations in cancer. Conclusions: Using exome sequencing we identified several potentially causal genes. Further investigations are needed to identify causative mutations of HPS. Our findings may uncover genetic changes specific to this population.

3272S
Deleterious mutations in multiple cancer-risk genes in individuals from a high-risk cancer genetics clinic. C.M. Laukaïtis1,2, C. Mauz1, M. Uruojj2, A. Chaudhury3, K. Maher1, J. Jeter1,2. 1) Department of Medicine, University of Arizona College of Medicine, Tucson, AZ; 2) University of Arizona Cancer Center, Tucson, AZ; 3) Departamento de Medicina, Universidad de Sonora, Hermosillo, Sonora, Mexico.

Neoplasms of serrated origin, a mutually exclusive, and in some cases paralleled next generation sequencing (NGS) technology has led to surprising findings. We are using NGS to identify inherited genetic changes predisposing to cancer in people cared for by the University of Arizona Cancer Center high-risk cancer genetics clinic. Under an IRB-approved research protocol, we sequenced a custom panel of 104 cancer-related genes in blood samples from 48 individuals. This was a heterogeneous population containing both patients who were diagnosed with cancer and people without a personal history of cancer but with a strong cancer family history. Surprisingly, 6 individuals (12.5 percent) carried clearly deleterious mutations in multiple genes linked to increased cancer risk. While 35 percent of the cohort has been diagnosed with cancer, five of six people with multiple deleterious mutations have had cancer (p = 0.017). Individual mutations are being validated by Sanger sequencing. Additional clinical characteristics (participant age, function of mutated genes, type of cancer, family history of cancer, type of cancer treatment) are being evaluated to identify associated characteristics in order to determine whether this evidence of genomic instability is the cause or an effect of the cancer diagnosis.

3273M

1) Department of Medicine, Division of Hematologic Oncology, University of Pennsylvania, Philadelphia, PA; 2) Clinical Genetics Research Lab, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY; 3) Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN; 4) Department of Medicine Division of Translational Medicine and Genetics, University of Pennsylvania, Philadelphia, PA; 5) Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY; 6) Beckman Research Institute of City of Hope, Duarte, CA; 7) National Institutes of Health.

Over 50% of the familial relative risk of breast cancer is unexplained by currently identified common low risk polymorphisms and rare to high risk alleles in genes such as BRCA1/2. Early identification of the individuals with an inherited susceptibility to breast cancer may lead to enhanced screening and prevention strategies and the opportunity for targeted therapeutics of associated cancers. In order to identify novel genes which confer an inherited risk to breast cancer, whole exome sequencing was undertaken in 333 BRCA1/2 negative individuals with at least one primary breast and ovarian cancer or with high-risk familial breast cancer, defined as a proband and at least two first or second degree relatives with breast cancer. Samples included 256 individuals from 109 families and 77 unrelated individuals, for a total of 186 independent cases. Data were first analyzed for identification of rare mutations in 49 cancer susceptibility genes. Eight independent cases (4.3%) were found to carry deleterious mutations in high penetrance cancer susceptibility genes, namely TP53, CDH1, MSH6, and CDKN2A. In addition, sixteen of 186 independent cases (12.3%) were found to have deleterious mutations in other breast and/or ovarian cancer predisposition genes, namely ATM, CHEK2, PALB2, BARD1, RAD51D, RAD50, and MRE11A. For novel cancer susceptibility gene discovery, whole exome data was harmonized from three separate sites into a single vcf file for variant filtering and analysis. Two major methodologies were used to select genes for further study, a candidate gene approach and an agnostic gene ranking approach. In the candidate gene approach, genes matching either specific gene ontologies or those identified by published somatic tumor studies were queried for the presence of rare deleterious mutations in the dataset. In the agnostic gene ranking approach, genes were ranked based on the frequency of types of variants in the dataset versus population estimates and the level of segregation of these rare variants in families. These methodologies identified 636 candidate cancer susceptibility genes in the discovery exome set. These genes will be evaluated by a targeted sequencing approach in a validation panel of over 200 individuals with familial breast cancer. Overall, our data demonstrate a significant level of heterogeneity in the genetic basis of familial breast cancer in BRCA1/2 negative individuals.
3274T
Germline epigenetic inactivation of BAP1 in a subset of patients with uveal melanoma. R. Pilarski1, G. Bonu2, O.H. Sagr6, K. Rai3, J.B. Massengill2, C.M. Cebulla1,2, M.H. Abdel-Rahman1,2,4,5. 1Dept of Internal Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX; 2Dept of Ophthalmology, Duke University School of Medicine, Durham, NC; 3Dept of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, TX; 4Department of Pediatrics, Texas Children's Cancer Center, Baylor College of Medicine, Houston, TX; 5Department of Pathology, Menoufiya Univ, Egypt.

Objective: To identify the frequency of germline deletions and/or epigenetic inactivation in BAP1 in uveal melanoma (UM) patients with high-risk for BAP1 tumor predisposition syndrome (TPDS) but no detectable mutation.

Methods: Twenty-three UM patients with high-risk for BAP1 TPDS, including 12 UM patients with family history of UM, one UM patient with bilateral disease and 10 UM patients with personal or family history of renal cell carcinoma (RCC), were studied. Constitutional decrease in BAP1 mRNA expression was assessed in an additional 17 UM patients with available non-tumor choroidal tissue. Germline copy number variation was studied by Multiplex Ligation-dependent Probe Amplification (MLPA). Methylation specific PCR and pyrosequencing was used to assess germline promoter hypermethylation. The expression of BAP1 mRNA was carried out utilizing quantitative RT-PCR. Results: No germline CNV or significant promoter methylation was detected in any of the 23 patients with high-risk for BAP1 TPDS. Two out of the 17 non-tumor choroidal tissues showed significant decreases in BAP1 expression compared to normal controls, but no germline CNV or aberrant promoter methylation were identified. Conclusion: Epigenetic inactivation of BAP1 through mechanisms other than promoter hypermethylation, could be responsible for germline constitutional inactivation of BAP1 in a subset of UM patients. Copy number variation is not a major mechanism for germline inactivation of BAP1. Our results have important implications for designing clinical assays for detection of BAP1 germline inactivation in patients at risk for the BAP1 TPDS.

3275S
Integrating Whole Genome and Exome Sequencing with Structural Variation Analysis to Identify Potential Causative Mutations in Patients with Cancer Phenotypes Suggestive of Li-Fraumeni Syndrome. D.I. Ritter1,2, B. Powell1, J. Bojadzieva2, D.A. Wheeler1,2, R. Gibbs3,4, L.C. Strong2, S.E. Plon2,4. 1) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX; 2) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 3) Department of Cancer Genetics, University of Texas MD Anderson Cancer Center, Houston, TX; 4) Department of Pediatrics, Texas Children’s Cancer Center, Baylor College of Medicine, Houston, TX.

While mutations in TP53 explain ~70% of the inherited cancer condition Li-Fraumeni Syndrome (LFS), a sizeable fraction of families or individuals with cancer phenotypes suggestive of TP53 mutations remains unexplained. To identify additional candidate genes underlying cancer phenotypes suggestive of LFS, we applied an integrative analysis of whole exome rare coding variants with whole genome structural variants (SV) and copy number variation (CNV) for 27 subjects who were negative for mutations in TP53, including family members when available (four kindreds). Proband demonstrated (1) childhood sarcoma and at least one additional primary malignancy by age 40 or (2) a rare childhood cancer with at least one first-degree relative with a rare cancer. Whole exome sequence data (VCRome2.1) were analyzed for rare heterozygous single nucleotide variants (SNVs) and insertions/deletions (INDELs), encoding missense, nonsense, frameshift and splice variants. We eliminated variants >1.5% in 1079 unaffected normal samples, and >1% dbSNP minor allele frequency. Variants were limited to those shared among affected family members, when available. We applied a proportions test to variants in genes and exons against 100 random controls. To identify structural variants, we applied complementary algorithms of insert-size (Breakdancer) and soft-clip stacking (CREST). We used an empirical annotation filtering strategy against normal, unaffected whole genomes. Structural variants were progressively clustered to identify recurrent and CNVs. We prioritized SV events in genes harboring rare variants, interrupting coding regions, introns or <=10kb of genes. We identified 110 unique genes harboring rare coding variants that affect >=4 independent families. The top genes of interest are TER1 (gene: p=3.6E-3, exon 30: p=0.02) and ATR (gene: p=0.02, exon38p=4.9E-4). We found no rare coding SNVs or INDELS within TP53, nor in CHEK2, a gene known to be mutated in a subset of LFS phenotypes. We found no evidence for SV within 1MB of either of these genes. We have identified multiple additional regions of rare SVs and CNVs affecting alternative genes, and are combining these results with exome variant data. With this effort, we aim to identify additional genetic causes for pediatric inherited cancer syndromes with phenotypes suggestive of Li-Fraumeni. Supported by RP10189 from CPRIT and R01-CA138836 to SEP, and K12GM084897 from IRACDA to DIR.

3276M
Causative novel POLE mutations in hereditary colorectal cancer syndromes. A.M. Rohlin1, T. Zagoras1, F. Eiengård2, Y. Engwall2, S. Nilsson2, U.T. Lundstam3, J. Björk4, G. Karlsson2, M. Nordling2, 1) Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden; 2) Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden; 3) Department of Surgery, Sahlgrenska Academy at University of Gothenburg, Sahlgrenska University Hospital/Ostra, Gothenburg, Sweden; 4) The Swedish Predisposso Registry, Department of Medicine, Karolinska Institute, Stockholm, Sweden.; 5) The Swedish NMR-centre, University of Gothenburg, Gothenburg, Sweden.

Background: In Familial Adenomatous Polyposis (FAP) it is today possible to find almost all of the disease-causing mutations. However, in patients with a more attenuated phenotype (AFAP) with less than 100 polyps only a fraction of the disease-causing mutations can be identified. The low detection rate implicates the probable presence of additional disease-causing genes still to be identified. FAP is caused by autosomal dominantly inherited mutations in the APC (Adenomatous polyposis coli) gene. In AFAP families 20%-30% of the patients have a germline APC mutation. Recently a new CRC syndrome, polymerase proof reading associated polyposis (PPAP) was described. This syndrome is characterized by a dominantly inherited predisposition to the development of a variable number of colorectal adenomas and carcinomas [1] The aim of this study was to sequence the exon-nuclease domain of POLE in 88 index patients with a familial history of polyposis or non-polyposis and/or early onset CRC that had previously tested negative for mutations in APC, MÜTYH and/or mismatch-repair genes MSH2, MLH1, MSH6 and PM2S. Method: In one large family exome sequencing was performed on four family members and for the remaining 84 index patients a semi-automated targeted enrichment strategy using Sanger sequencing of POLE exonuclease domain (ex 3-14) was conducted. Results and Conclusion: We have identified two novel mutations in the exonuclease domain of POLE. The first mutation was identified from exome sequencing in a large Swedish family with CRC. The POLE: c.1089C>A, p.Asn363Lys mutation is directly involved in DNA binding. Family members carrying this mutation demonstrate a high penetrant predisposition not only to CRC but also to extra-intestinal tumours such as ovarian, endometrial and brain tumours [2]. The second mutation identified in a patient with early onset CRC. Theoretical prediction of the amino acid substitution suggests a profound effect of the substrate binding capability and a severe impairment of the catalytic activity for both these mutations, which strongly suggest a functional impact of these mutations. Screening the proofreading domains of POLE should be considered in routine genetic diagnostics in families with hereditary CRC. 1.Palles C, Nature genetics 2013, 45(2):136-144. 2.Rohlin A, International journal of oncology 2014, 45(1):77-81.
3277T
Parental Inheritance and WT1 Abnormality Types May Affect the Penetration Rate of Hereditary Wilms Tumor. 1. Y. Kaneko1, H. Okita2, M. Haruta1, Y. Arai3, T. Oue4, T. Koshinaga5, M. Fukuzawa6, Japan Wilms Tumor Study Group. 1) Research Institute for Clinical Oncology, Saitama Cancer Center, Ina, Saitama; 36-0806 Japan; 2) National Institute for Child Health and Development; Hematology/Oncology; 2-10-1, Ohkura, Setagaya-ku, Tokyo; 157-0074; 3) National Cancer Center Research Institute; 5-1-1 Tsukiji, Chuo-ku, Tokyo; 104-0045; 4) Osaka University Graduate School of Medicine; 2-15 Yamadaoka,Suita, Osaka; 565-0871; 5) Nihon University; 30-1 Ohyaguchi-Kamimachi, Itabashi-ku, Tokyo 173-8610; 6) Osaka Medical Center and Research Institute for Maternal and Child Health; 840 Murodo-machi, Izumi, Osaka 594-1101.

Background: Wilms tumor (WT) is a genetically heterogeneous disease which arises from developmental kidney. The frequencies of hereditary WT are unknown, and only some hereditary WTs are thought to be caused by a germline mutation in WT1. IGF2 is an imprinted gene expressed by the paternal allele. Both WT1 and IGF2 genes are located on the short arm of chromosome 11 (11p), and uniparental disomy (UPD) on 11p was found in 30-40% of WT, and regularly accompanied by maternal allele loss and paternal allele duplication. The inheritance of WT1 mutations has been poorly studied in familial WTs, although carriers with WT1 mutations are now increasing because multidisciplinary therapies have improved the survival rates of patients with bilateral WTs and those with a unilateral WT with a germline WT1 mutation. We examined the statuses of WT1 and IGF2 in bilateral and familial WTs in Japan, and summarized the present and previous findings on the penetrance rate for children who inherited various types of WT1 abnormalities from their fathers or mothers, or had de novo WT1 abnormalities that occurred in the paternal or maternal germ cell. Results: We detected WT1 abnormalities in 25 (81%) of 31 patients with bilateral WTs and 2 of 2 families with hereditary WT. Of 35 WTs from the 25 patients, 31 had duplications of a small WT1 mutation and paternal IGF2 caused by 11p UPD, while 4 had large deletions in one WT1 allele and small mutations or deletions in the other with the retention of 11p heterozygosity. The high incidence of WT1 abnormalities in bilateral WTs in Japan sharply contrasts with the lower incidence in America (27%). The penetrance rate was shown to be 100% if children inherited small WT1 mutations from their fathers and 67% if inherited the small mutations from their mothers, or inherited 11p13 deletions or had de novo 11p13 deletions irrespective of parental origin (P=0.057). The duplication was identified in most WTs with small WT1 mutations of paternal origin, but not in WTs with large 11p13 deletions irrespective of parental origin. Conclusion: We hypothesize that individuals who inherited small WT1 mutations from their fathers may be more likely to develop WT due to the expressing IGF2 than those who inherited small mutations from their mothers or inherited large 11p13 deletions or had de novo large deletions irrespective of parental origin. These findings may be useful for the genetic counseling of individuals who may inherit WT1 mutations.

3278S

1) Department of Genetics and Computational Biology, QIMR Berghofer, Herston, QLD, Australia; 2) University of Southern California, Los Angeles, CA; USA; 3) University of Cambridge, Straingeways Research Laboratory, Cambridge, UK; 4) Dana-Farber Cancer Institute, Boston, MA, USA; 5) Centre Hospitalier Universitaire de Québec (CHUQ), Québec, Canada; 6) Mayo Clinic, Rochester, MN, USA.

Multiple independent genome-wide association studies (GWASs) have identified variation at the 19p13 locus that predisposes to estrogen receptor-negative breast cancer and high-grade serous ovarian cancer (HGSOC). Fine-mapping of the region using the iCOGS customized Illumina Infinium genotyping array, followed by imputation, limited the risk associated region to a 20 kb window encompassing two protein coding genes - ANKLE1, which may be involved in DNA repair, and ABHD8 of unknown function - and just distal to the BRCA1-interacting gene BABAM1. The strongest association signal is represented by 13 highly correlated single nucleotide polymorphisms (P=10−22). By integrating genotyping data with epigenetic data generated from normal ovarian and breast epithelial cells we identified risk SNPs intersecting known or putative enhancers, and ELF1, ELK4, GABP and GATA3 transcription factor binding sites. Expression quantitative trait locus (eQTL) analyses identified significant genotype-gene expression associations for ANKLE1 in normal ovarian epithelial cells (P=0.02), and for ABHD8 in both HGSOCs (P=3.0×10−5) and normal breast tissues (P=2.8×10−5), but not for BABAM1. Risk associated SNPs were also significantly associated with allele specific expression of ABHD8 in breast cancer (P=5.8×10−5). ABHD8 (but not ANKLE1 or BABAM1) was prioritized in a co-expression network analysis based on expression data in primary HGSOCs. Genes in the 19p13 region are frequently overexpressed in both breast cancers and HGSOCs; however, stable overexpression of neither BABAM1 nor ANKLE1 was observed across breast cancers and HGSOCs. ABHD8 expression was however clearly upregulated in normal breast and ovarian epithelial cells. Using chromosome-conformation-capture (3C) assays in these same cell types, we identified interactions between a putative regulatory region and the promoter of ABHD8 as well as in breast and ovarian cancer cell lines. Taken together these results suggest that the same functional mechanism at this locus underlies breast and ovarian cancer development, and that ABHD8 may represent a novel breast/ovarian cancer susceptibility gene.
3279M
Identification of germline mutations in **TEP1** among familial and sporadic pediatric acute myelogenous leukemia cohorts. N.R. Oak1, D.I. Ritter2, B.C. Powell1, H.C. Cheung3, M.M. Gramatges2, D.A. Wheeler1, S.E. Plon1,2,3. 1) Molecular and Human Genetics, Baylor College of Medicine, Houston, TX; 2) Department of Pediatrics, Baylor College of Medicine, Houston, TX; 3) Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.

Identification of cancer susceptibility gene mutations can impact surveillance for secondary primary malignancies and risk in family members. Germline mutations in genes related to telomere maintenance are associated with bone marrow failure, as well as hematologic malignancies. Given their predisposition to pediatric cancers, we performed whole exome sequencing of constitutive DNA from childhood cancer patients including seven kindreds with familial leukemia/lymphoma in at least two first degree relatives. We filtered for rare variants (<1% in dbSNP or in another non-cancer database. We limited our analysis to truncating (nonsense and frameshift) mutations. In one of the families, we found 7 frameshift and 4 nonsense mutations. The most interesting candidate among these was the novel variant p.R314X in **TEP1**. Loss-of-function mutations in **TEP1** are very rare in sequence databases including the NHLBI exome sequencing project. This rare mutation was shared by a parent-child pair affected by acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML) respectively. We further explored whether germline mutations in **TEP1** are found in larger cohort of apparently sporadic AML patients. AmpliSeq™ sequencing of 43 genes related to telomere maintenance and DNA repair pathways was performed from DNA from pediatric patients with AML. Thirteen patients revealed 10 rare missense germline mutations in **TEP1**. Four of these 10 mutations predicted to have damaging effect on protein function according to Combined Annotation Dependent Depletion (CADD) tool (CADD scaled score of 15 or more). The **TEP1** familial p.R314X truncating mutation in addition to germline **TEP1** missense mutations in sporadic AML patients warrant further functional studies. These studies will include assessing the role of these **TEP1** mutations on the telomere length phenotype of cells from the patients carrying the mutation. The results of these studies are designed to determine whether **TEP1** represents another member of the telomere maintenance genes to be associated with AML predisposition. This work was supported by research grant RP10189 from the Cancer Prevention and Research Institute of Texas and grant R01-CA198836 to Plon, SE and training grant T32 GM007526 to Powell, BC.

3280T
**Hdac9** Intronic Enhancer Variants as Candidates for Skin Cancer Risk. A. Toland1,2,3, T. Siekmann1,2,3, H. Foulkes2,3,1. 1) Human and Medical Genetics, 2) Biomedical Sciences Program; 3) The Ohio State University, Columbus, OH.

The susceptibility to skin tumorigenesis 5 (Skit5) locus on mouse chromosome 12 was mapped through linkage analysis of skin tumor susceptible mice with susceptibility to the skin tumor resistant Mus musculus (NIH/Ola) and skin tumor resistant Mus musculus (SPRET/Ei) mice. Expression and sequence analysis of genes at Skts5 led to the identification of **Hdac9** as a potential candidate for Skts5; **Hdac9** contains both amino acid variations and differential expression in skin between the two strains. Further, we found that variants in human **Hdac9** show allele-specific imbalance in human cutaneous squamous cell carcinomas (cSCC), suggesting a role for this gene in human cSCC. Interestingly, studies by others identified an exonic/intronic enhancer in **HDAC9** that impacted expression of **Twist1**. In this study, we observed that the mouse **Hdac9** might also contain an enhancer element and that variants in this region might contribute to differential expression of the oncogene, **Twist1**. To test this hypothesis, we performed sequencing analysis and identified 45 sequence variants between NIH/Ola and SPRET/Ei from the shared orthologous region of the human **HDAC9** enhancer. We subcloned this region into nine segments; two of these segments differentially impacted luciferase expression in vitro. NIH/Ola clones showed 2-fold increased luciferase expression relative to vector alone or the similarly subcloned Mus musculus (SPRET/Ei) mice. Furthermore, cells transfected with this segment of the NIH/Ola intron 17 led to a 2.2 fold increase in **Twist1** expression, but the same region in SPRET/Ei resulted in no up-regulation of **Twist1**. In silico transcription factor analyses identified a number of transcription factors that were predicted to potentially bind NIH/Ola and SPRET/Ei variants. Chromatin immunoprecipitation studies of two transcription factors, Gata3 and Oct1, demonstrated differential binding between NIH/Ola and SPRET/Ei DNA that fit the in silico predictions. Together these studies show evidence that the mouse **Hdac9** enhancer to a human **HDAC9** enhancer and acts as an enhancer for **Twist1**. As **Hdac9** intron 17 sequence variants between NIH/Ola and SPRET/Ei differentially impacted luciferase expression, **Twist1** expression and Gata3 and Oct1 binding, they are candidates for differences in skin tumor susceptibility locus Skts5.

3281S
Allelic imbalance in gene expression as a mechanism in breast cancer development. I. Pulyakhina1, M.P.G. Vreeswijk1, J.F.J. Laros1,2, C.M. Meijers1,2, J.J. den Dunnen1,2, P. Devilee1, P.A.C. ’t Hooren1,1. 1) Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands; 2) Leiden Genome Technology Center, Leiden University Medical Center, Leiden, the Netherlands.

Differential allelic expression (DAE) refers to differences in expression levels between the two alleles of a gene. DAE can be caused by genetic variation and is known to affect a considerable fraction of genes in the human genome. DAE has been suggested to play an important role in human phenotypic variability, including complex traits and diseases. We hypothesize that DAE contributes to the pathogenesis of breast cancer. Loss of the highest expressed allele of a tumor suppressor gene would lead to substantially lower expression levels of the gene than loss of the lowest expressed allele. This may promote tumorigenesis. We performed whole transcriptome sequencing of paired normal-tumor samples from breast cancer patients and analyzed the allelic ratios for heterozygous, coding SNPs. We focused specifically on SNPs that showed DAE in the normal samples and a switch or loss of DAE in the tumor. We separated genomic variants from RNA editing events and investigated the behavior of allelic imbalance and the underlying molecular mechanism in a subset of these SNPs. 102 SNPs in 60 genes demonstrated significant changes in allelic ratios between normal and tumor tissue. A decrease in allelic imbalance in tumor was most frequently observed. 20% of SNPs showed a complete switch in ratios, where the allele with higher expression in normal cells was lower expressed in tumors and vice versa. Multiple coding SNPs in the same gene and the same sample showed consistent changes in allelic imbalance. Genes outside the breast cancer region showed no significant differences in allelic imbalance across multiple samples, although sometimes measured by different SNPs. Some of the affected genes have been shown to play a role in cancer development, but contribution of their allelic imbalance has not been described so far. We propose that DAE in a larger cohort and investigation of a possible role for DAE in other cancer types is ongoing.

3282M
Germline Mutations in Men with Multiple Primary Malignancies from a Hereditary Prostate Cancer Cohort. P.G. Pilie, K. Zuhlke, A. Johnson, K. Cooney. Internal Medicine, University of Michigan, Ann Arbor, MI.

There exists a variety of well-known familial cancer syndromes originating from germline mutations that lead to multiple primary malignancies in a single individual. Cases of men with early-onset and/or familial prostate cancer in addition to other primary cancers, even including other urogenital cancers, are relatively rare. Furthermore, the molecular underpinnings are not well understood. High-throughput DNA sequencing encompassing large cancer exomes from familial prostate cancer patients was performed to identify the most interesting candidates among these patients. The rarest mutation was shared by a parent-child pair affected by prostate cancer at age 57 as well as significant family history of prostate cancer. Interestingly, the other parent was discovered in a man with a personal history of thyroid cancer diagnosed at age 44, prostate cancer diagnosed at age 55, renal cell cancer diagnosed at age 57 and significant family history of prostate cancer. Loss-of-function mutations in **FGFR3** which likely contribute to early-onset prostate cancer and multiple other primary cancers led to the identification of a novel germline mutation (frameshift) mutations. In one of the families, we found 7 frameshift and 4 nonsense mutations in 25% from the prostate cancer cohort with multiple primary cancers harboring deleterious germline mutations, one a frameshift mutation in **FGFR3** gene and the other a truncating variant in ATM gene. The **FGFR3** mutation was discovered in a man with a personal history of thyroid cancer diagnosed at age 44, prostate cancer diagnosed at age 55, renal cell cancer diagnosed at age 57 as well as significant family history of prostate cancer. Interestingly, multiple first and second degree relatives with only prostate cancer tested negative for **FGFR3** mutation. In summary, our strategy of selecting men with prostate cancer and multiple other primary cancers led to the identification of a novel deleterious mutations in **ATM** and **FGFR3** which likely contribute to cancer risk in these individuals and their families. Whereas somatic **FGFR3** mutations have been described in a number of cancers, this is the first report of a **FGFR3** germline mutation that may increase cancer risk and studies are ongoing to understand the biologic implications.

3.26.2014 The American Society of Human Genetics. All rights reserved.
3283T

Genetic analysis of the chromosome 15q25.1 region identifies IREB2 variants associated with lung cancer. C. Amos1, I. Gotov1, Y. Han1, L. LeMarchand2, X. Ji1, D. Christian1, M. Frazier1, C. Wei1, J. McKay1, P. Brennan3, J. Field4, Y. Li5, R. Hung1, 1) Dartmouth, Lebanon, NH; 2) University of Hawaii Cancer Center, Honolulu, HI; 3) Harvard University, Boston, MA; 4) U.T. M.D. Anderson Cancer Center, Houston, TX; 5) International Agency for Research on Cancer, Lyon, France; 6) University of Liverpool, Liverpool, UK; 7) University of Toronto, Toronto, CA.

Genome-wide association studies of lung cancer identified the region of chromosome 15q25.1 that includes a nicotinic acetylcholine receptor cluster as being the most strongly associated with lung cancer risk. To characterize the impact that specific functional variants in this region have upon risk for lung cancer development, we performed fine mapping selecting all currently known SNPs influencing lung cancer risk along with all coding SNPs in the 200 megabase region surrounding CHRNA5, a gene known to influence smoking behavior in this region. SNPs were initially identified by Sanger sequencing of 96 individuals for CHRNA5, CHRNA3, CHRNA4, and PSMA4. Additional markers were selected from dbSNP. Markers that were selected for genotyping were based upon following the known functional effect on activity of genes in the region, validation in African or European populations, position across the region, predicted effect on function, r-square value for LD with respect to other markers being less than 80%, and 268 markers had p-values less than 0.05 and 101 SNPs met the multiple testing corrected threshold (p<3.5×10^-5). Across this region, 268 markers had p-values less than 0.05 and 101 SNPs met the multiple testing corrected threshold (p<3.5×10^-5).

The most significant SNPs lie in a region including of IREB2, with the most significant associated variant being rs17483686 (OR=1.26, p=8.93×10^-9). The previously well characterized SNP in CHRNA5, rs16969968, which significantly associated with lung cancer risk beyond the previously identified SNP rs11835340 in CHRNA5, rs16969968 (OR=1.25; p-value: 7.6×10^-9) which may have implications for disease prevention.

3284S

Genome-wide analyses identify gene interaction between SMAD7 and body mass index with risk of colorectal cancer. P.T. Campbell1, C. Haynes2, J. Yang2, J. Song2, W.J. Gauderman3, P. Schwartz2, B. Bernier1, H. Wenk3, G. Casey4, A.T. Chan5,6,7, J. Chang-Claude10, C. Edlund13, J. Figueiredo6, G. Gielen11, L. Le Marchand14,5, M. Lemire15, L. Li14, P. Newcomb16,17, F. Schumacher6, M. Stiattetry6, D. Thomas3, E. White18,19, M.O. Woods17, S. Gruber20, U. Peters21,22, V. Moreno20, L. Hsu23,1 Department of Biostatistics, University of Washington, Seattle, WA, USA; 2) Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany; 3) Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, OH, USA; 4) Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, USA; 5) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA; 6) Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ) Heidelberg, Germany; 7) German Cancer Consortium (DKTK), Heidelberg, Germany; 8) Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; 9) Channing Division of Network Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA; 10) Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany; 11) Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia; 12) Epidemiology Program, University of California, Los Angeles, CA, USA; 13) Case Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; 14) Catalan Institute of Oncology and Research, Toronto, ON, Canada; 15) Department of Biostatistics, University of Washington, Seattle, WA, USA.

Body Mass Index (BMI) is a complex phenotype that is consistently associated with risk of colorectal cancer, although associations are often stronger in men and women. Genetic variants interact with BMI to modify colorectal cancer risk. We tested interactions between approximately 2.7M single nucleotide polymorphisms (SNPs) and BMI with colorectal cancer risk among 17,122 colorectal cancer cases and 17,689 controls. We used comprehensive statistical methods with individual-level data to evaluate sex-specific multiplicative interactions between SNP and BMI with colorectal cancer risk. BMI per 5 kg/m^2 was associated with higher risk of colorectal cancer (men, odds ratio (OR): 1.26; p-value: 2.1×10^-8; women, OR: 1.14; p-value: 5.4×10^-6). Among men, from traditional case-control logistic regression models, we identified a suggestive interaction between BMI and a locus in SMAD7 (T/T, OR: 0.77; p-value: 9.2×10^-16; G/T, OR: 1.13; p-value: 5.4×10^-6; G/G, OR: 1.26; p-value: 2.4×10^-10). Among women, we identified a statistically significant interaction between BMI and a locus in SMAD7. Specifically, when using the Cocktail method or the EDGxE method, which each involves two-step screening and testing methods and corrects for multiple testing using weighted hypothesis testing, an interaction was detected between BMI and rs4939827 in SMAD7 (p-observed: 0.0009; p-threshold: 0.005). A statistically significant interaction was also detected between BMI and rs4939827 in SMAD7 using a 2-d.f. joint test (p-value: 2.4×10^-16). ORs for the association between BMI per 5 kg/m^2 and colorectal cancer risk differed according to strata of rs4939827 genotype T/T, OR: 1.08; p-value: 0.02; G/T, OR: 1.13; p-value: 5.4×10^-6; G/G, OR: 1.25; p-value: 7.6×10^-9). rs4939827 in SMAD7 was previously identified in GWA studies as a colorectal cancer susceptibility locus. Herein, we showed that a common susceptibility locus in SMAD7 modifies the association between BMI and colorectal cancer risk for women. Further functional analyses are needed to understand this interaction. Future work with larger study samples is warranted to explore the borderline interaction between ADIPOQ and BMI in men, especially given the strong biological plausibility for a connection between obesity, circulating adiponectin and colorectal cancer. Our results identify a novel gene-BMI interaction for colorectal cancer risk, which may have implications for disease prevention.
3285M

Estimation of Whole Genome Variations of Hepatocellular Carcinoma among Chronic Hepatitis C Patients. Y. Chang1, M. Lee2, C. Liu3, H. Yang4, H. Chen5, C. Chen2. 1) Institute of Statistical Science, Academia Sinica, Taipei, Taiwan; 2) Institute of Clinical Medicine, National Yang Meng University, Taipei, Taiwan; 3) Genomics Research Center, Academia Sinica, Taipei, Taiwan.

Background & Aims: Host genetic susceptibility may be associated with the occurrence of hepatocellular carcinoma (HCC) among patients with chronic infection of hepatitis C virus (HCV). This study aimed to discover genomic variations associated with hepatocellular carcinoma risk through the genome-wide association study (GWAS) and imputation analysis. Methods: There were 472 HCC cases and 806 unaffected controls. All study subjects were adults seropositive for antibodies against HCV and seronegative for HBsAg. The demographic characteristics and serum markers for liver functions were also evaluated. High quality human genomic DNA was extracted from each blood sample to perform genotyping by AxionTM Genome-Wide CHB Array. Then, imputation algorithm was applied to get whole genome variations in patients based on genotyping data from SNP microarrays. The reference genomes of imputation were Han Chinese population in 1000 genome project. The logistic regression was used to evaluate association between disease and genotype based on four different genetic models (allelic, dominant, additive and recessive). P values <10^-8 were considered significant. Results: A total of 36,175,342 SNPs were obtained in the samples after imputation. There were 765, 134, 612, and 855 SNPs significantly associated with HCC based on allelic, recessive, dominant, and allelic genetic models, respectively. Among the SNPs discovered in our analysis, less than 4% could be detected by microarray chips. In addition, 18 nonsynonymous SNPs were found to be significantly associated with HCC under different genetic models. Only 3 out of 18 nonsynonymous SNPs could be detected in microarray chips. Interestingly, 7 nonsynonymous SNPs clustered on the human leukocyte antigen (HLA) complex region. It indicated that HLA region play an important role in host susceptibility to hepatocellular carcinoma within chronic hepatitis C patients. Conclusion: The SNPs associated with HCV-related HCC were identified in this study. It provide insight for identification of high risk population in HCV after clinical validation.

3286T

Functional characterization of PARP1 melanoma-associated locus. J. Choi1, M. Makowski1, M. Xu1, T. Zhang1, M. Law1, W. Kim1, M. Kovacs1, H. Parkh1, L. Aoude2, M. Gartsider3, H. Yin4, J. Trent4, S. Macgregor5, N. Hayward6, K. Brown6. 1) Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD; 2) Queensland Institute of Medical Research, Brisbane, QLD, Australia; 3) Translational Genomics Research Institute, Phoenix, AZ.

Recent genome wide association studies (GWAS) identified several new loci for melanoma susceptibility. While these results highlight potential pathways predisposing to melanoma, functional risks in these regions as well as the mechanism by which they influence risk have yet to be elucidated. To nominate functional variants we performed expression quantitative trait loci (eQTL) analysis in 62 melanoma cell lines. Transcript levels were measured using Affymetrix U133Plus2 expression microarray. SNPs (>700K) were then typed on Illumina OmniExpress arrays and the ones in GWAS loci were further imputed using 1000 genomes (1KG) data. Among 16 GWAS loci tested two chromosome 1 loci exhibited significant cis-eQTL Subsequent validation using Taqman quantitative PCR (qPCR) demonstrated that Poly[ADP-ribose] polymerase 1 (PARP1) expression is significantly altered with the lead SNP (p=0.03, genomic copy number adjusted). Namely, the risk allele is correlated with an increased PARP1 transcript levels after adjusting for genomic copy number (p=0.03). We then further interrogated the genotype-expression correlation by Taqman allele discrimination qPCR in 21 melanoma cell lines heterozygous for the GWAS lead SNP. The results demonstrated significantly higher proportion for the risk allele in PARP1 transcripts (p=0.0001). To identify functional risk variants mediating these effects we annotated the PARP1 locus using ENCODE database. Among 56 SNPs of strong linkage disequilibrium (LD) with the lead SNP (r2>0.5 using 1KG genotypes), six exhibited strong evidence as potential transcriptional enhancers in melanoma relevant cell types. One of them is a six-base pair indel in GC-rich region poorly covered by 1KG. Genotype reassessment of this indel using gel-based fragment analysis in 745 healthy Europeans from DCEG imputation reference set resulted in remarkably enhanced LD with the lead SNP (r2=0.94 from 0.67) supporting direct link to melanoma susceptibility. Subsequent Electro Mobility Shift Assays and luciferase assays for this indel demonstrated allele-specific protein binding and differential transcriptional activities in melanoma cell lines. Chromatin immunoprecipitation indicated that enhancer-binding proteins are enriched in melanoma-associated deletion allele in melanoma cell lines. Identification of proteins directly mediating this function will further elucidate PARP1 contribution to melanoma susceptibility.

3287S

Heritable missense variant rs3731249 underlies the CDKN2A association with childhood ALL and is preferentially retained by tumors harboring somatic CDKN2A loss. A.J. de Smith1, K.M. Walsh2, H.M. Hansen3, L.F. Barcellos1, A.P. Chakellangam3, R.B. Jenkins4, M.R. Wrensch5, J.K. Wiercnick6, C. Metayer7, J.L. Wiemels1. 1) Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA; 2) Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA; 3) Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, CA; 4) Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN.

Little is known about the interaction between constitutive and somatic genetic variation in carcinogenesis. Genome-wide association studies (GWAS) of childhood acute lymphoblastic leukemia (ALL [MIM 613065]) have identified 6 genetic associations, including rs3731217, located between the CDKN2A [MIM 600160] and CDKN2B [MIM 600431] genes. Somatic deletions of this 9p21.3 region are common in cancer, and occur in ~30% of childhood ALL. We replicated the rs3731217 association with childhood ALL in Hispanics (p=0.021) via a GWAS of 323 childhood ALL cases and 454 controls, using the illumina OmniExpress SNP array. SNP imputation analysis across 9p21.3 revealed a stronger association at rs3731249 (p=6.4x10^-4; OR = 2.77), a low-frequency missense variant in exon 2 (p.Ala148Thr) of CDKN2A (MAF_control=2.1%). These associations were replicated in an additional 378 ALL cases and 536 controls using Taqman assays, where the missense variant again had a greater magnitude of effect (P=3.7x10^-3; OR rs3731249 = 2.00 versus Prs3731217=0.029; OR rs3731217 = 1.38). We assessed the effect of rs3731249 on risk of other cancers with known chr9p21.3 GWAS associations, but observed no association suggesting the ALL-specific role for rs3731249. We hypothesized that cases heterozygous for rs3731249 may show allelic imbalance, with preferential loss of the wildtype (WT) allele when hemizygous chr9p21.3 deletions occur in leukemia. In this study, we developed a method termed "SMART-ddPCR" (Somatic Mutation Allelic Ratio Test using Droplet Digital PCR). Concentration of risk and WT alleles was measured by ddPCR using the rs3731249 Taqman assay in diagnostic bone marrow (i.e. tumor) DNA from 378 ALL cases and 378 normal bone marrow (i.e. normal) DNA. Hemizygous (WT=0, risk=1) CDKN2A alterations was identified in 17 cases, of which 14 had higher risk:WT allele ratio (p=0.006, 1-sided binomial significance test), thus demonstrating preferential loss of WT rs3731249 and retention of the missense allele. Subclonal heterogeneity of the CDKN2A locus was evident, suggesting that this is a late event in leukemogenesis. Preferential allelic imbalance towards the rs3731249 risk allele in ALL tumor DNA provides strong evidence of an important role for this missense variant in leukemogenesis and shows, for the first time, a direct relationship between the constitutive and somatic genetic variants underlying ALL development.
Genetic determinants of Breslow tumor thickness and their impact on melanoma progression. S. Fang\(^1\), Y. Wang\(^1\), D. Deng\(^1\), Q. Liu\(^1\), R. Feng\(^1\), K. Xu\(^1\), H. Liu\(^1\), M. Ross\(^1\), J. Gershenson\(^1\), J. Cormier\(^1\), R. Royal\(^1\), A. Lucchi\(^1\), J. Wargo\(^1\), C. Schachterer\(^1\), J. Reveille\(^1\), D. Sui\(^1\), R. Bassett Jr\(^1\), F. Demenais\(^2\), L. Wang\(^1\), M. Lu\(^1\), Q. Wei\(^1\), C. Amos\(^1\), J. Lee\(^1\), GenoMEL group. 1) University of Texas MD Anderson Cancer Center, Houston, TX; 2) The fifth affiliated Hospital of Zhengzhou University, Zhengzhou, Henan; 3) The University of Texas Health Science Center at Houston, Houston, TX; 4) Université Paris Diderot Paris; 5) Duke University School of Medicine, Durham, NC; 6) Dartmouth College, Lebanon, NH.

Breslow thickness of the primary melanoma tumor is one of the most important clinicopathologic factors predicting risk of melanoma progression. Genetic mechanisms determining Breslow thickness are not yet known. We performed a genome-association study of Breslow thickness using 1547 patients in MD Anderson (MDA) in discovery analysis and further evaluated their association using GenoMEL (UK and European) and French melanoma datasets. None of the SNPs identified in 4 candidate gene regions (MUC2/MUC5AC/CALR/RAD23A) in the MDA dataset was replicated in the GenoMEL or French datasets. However, gene-based VEGAS testing confirmed that two MDA candidate genes, MUC2 and MUC5AC, were significantly related to tumor thickness in the MDA dataset (both P-values<2.8x10^-6). Additionally, both CALR and RAD23A genes were associated with melanoma disease-free and overall survival (P-value<0.05) in the MDA dataset. Finally, immunohistochemistry of primary melanoma tumors demonstrated that both MUC2 and CALR expression levels increased with tumor thickness; expression of MUC2 was also elevated in regional nodal metastases from melanoma patients. These data suggest that MUC2 and CALR genes may determine melanoma tumor thickness and influence melanoma disease severity.
3292T
Regulatory polymorphisms in lymphoma and chronic lymphocytic leukemia risk. J. Hayes et al.

3292M
GWAS meta-analysis identifies three novel risk loci for melanoma at 6p22, 7q21 and 9q31. M. Law et al.

3293S
Role of polymorphic fibroblast growth factor receptor (FGFR) Gene and Breast Cancer Risk. M. Hosseini et al.

3294M
GWAS meta-analysis identifies three novel risk loci for melanoma at 6p22, 7q21 and 9q31. M. Law et al.

Posters: Cancer Genetics
3295T Detection of trans and cis splicing QTLs through large scale cancer genome analysis, K. Lehmann1, A. Kähäri2, C. Randolf1, W. Lee3, N. Schultz4, O. Stegle5, G. Rätsch1, 1) Memorial Sloan-Kettering Cancer Center, Computational Biology, New York, NY; 2) European Bioinformatics Institute; Hinxton; Cambridge; CB10 1SD; United Kingdom.

The comprehensive survey of molecular characteristics provided by The Cancer Genome Atlas (TCGA) enables large scale analyses across multiple cancers. However sophisticated tools for the joint analysis of the thousands of samples and cancer specific challenges are needed. In an effort to enable joint analysis, we have re-aligned and re-analyzed RNA and whole exome sequencing data of ~4,000 individuals across 12 cancer types in a uniform manner. We used a newly developed open source SpAddle pipeline to count gene expression as well as annotate and quantify a comprehensive set of alternative splicing events. We identified threefold more high confidence alternative splicing events than annotated in the GENCODE annotation which reflect cancer-specific and tissue-specific splicing variation. Comparisons to matching tissue normal samples confirm a ~20% increase of splicing complexity in tumor samples. We have identified 22 genes with splicing changes that recurrently occur in tumor samples (>10%) but are virtually never observed in normal samples or ENCODE cell lines (<0.5%) and could be possible targets for new drugs. While population structure is one of the most severe confounding factors in QTL analysis, tumor samples open up many new additional challenges. Tumor specific somatic mutations and recurrence patterns as well as sample heterogeneity can lead to spurious associations. Thus, we have developed a new strategy to perform a common variant association study using mixed models on tumor samples enabling us to account for tumor specific genotypic and phenotypic heterogeneity in addition to population structure. Due to sample size constraints, many previously analyzed mixed QTL studies have been limited to the analysis of cis-associated variants. The large sample size available from TCGA enables us to overcome this limitation and discover trans-associated variants as well. Preliminary data demonstrates that we find cis associations for ~10% of the analyzed genes, of which a large fraction replicates across tissue and cancer types. We also confirm a recently reported trans-association in the splice factor U2AF1 and detect several additional trans-associations with effect sizes >20%. Current efforts including including rare variant test for splicing QTLs in order to identify recurrently affected genes and networks leading to aberrant splicing.

3296S Genome-wide association study of breast cancer in Japanese population. S. Low1, A. Takahashi1, M. Kubo2, T. Katagiri2, 1) Lab. for Statistical Analysis, Ctr. for I.M.S., RIKEN, Yokohama, Kanagawa, Japan; 2) Lab. for Genotyping Development, Ctr. for I.M.S., RIKEN, Yokohama, Kanagawa, Japan; 3) Division of Genome Medicine, Inst. for Genome Res., The Univ. of Tokushima, Tokushima, Japan.

Breast cancer is the most common malignancy among women in worldwide including Japan. The breast cancer genome-wide association study (GWAS) has performed to identify genetic variants that are associated with the risk of breast cancer. Owing to the complex linkage disequilibrium structure and various environmental exposures among different populations, it is of important to investigate associated genetic variants with breast cancer in a specific population. In this study, we conducted a genome-wide association study as well as whole-genome imputation with 2,642 cases and 2,099 unaffected female controls collected from the Biobank Japan. We further examined 13 suggestive loci (P<1.0×10^-10, OR>1.21; 95% CI=1.15-1.28) and rs12292068 (KDR: OR=2.89, 95% CI=1.77-4.73) on chromosome 3q26.2 (FGFR2); rs3803662 (combined P-value of 2.79×10^-11, OR=1.51, 95% CI=1.43-1.60) and rs9319401 (combined P-value of 3.97×10^-10, OR=1.23, 95% CI=1.15-1.31) on chromosome 16q12 (TOX3-LOC643714). Weighted genetic risk score by using three significantly associated variants and two previously reported breast cancer associated loci in East Asian population revealed that individuals who carry the most risk alleles have 2.2 times higher risk of developing breast cancer in the Japanese population compare to those who carry the least risk alleles. In addition, pathway analysis suggested that variants within aryl hydrocarbon receptor signaling pathway from Ingenuity database are associated with the risk of breast cancer. Although we could not identify additional loci associated with breast cancer risk in East Asian population, pathway analysis suggested that variants within aryl hydrocarbon receptor signaling pathway in Japanese population compare to those who carry the least risk alleles. In future studies, we will further analyze the association between variants across the risk loci in Japanese population to identify additional genetic variants that could lead to a better, accurate prediction of breast cancer.

3297M New insights into ovarian cancer from the investigation of overall genetic sharing. Y. Lu1, A.B. Spurdle2, G.W. Montgomery2, K.T. Zonderman1, P.D. Pharoah1, G. Chenevix-Trench1, S. Maciejewski1, Ovarian Cancer Association Consortium. 1) Statistical Genetics, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia; 2) Molecular Cancer Epidemiology, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia.

To date, most genetic studies have focused on the common subtype, serous carcinoma. Studying rare histological subtypes of EOC (clear-cell, endometrioid, or mucinous) has been challenging because a large sample size is often required but not readily available. Epidemiological studies observed that certain subtypes of EOC are associated with other women’s diseases, e.g. clear cell and endometrioid carcinoma occur more frequently in women with endometriosis. But it remains to be elucidated whether there is a genetic overlap underlying the association observed from epidemiological studies. We investigated this question in the largest ovarian cancer data set to date from Ovarian Cancer Association Consortium, explicitly modelling the polygenic effect shared between EOC subtypes and other women’s diseases. We used the polygenic risk score approach that identifies polygenic cis- and trans-variants that associate with breast cancer risk across HapMap populations CEU and CHB/JPT. At the SNP-level, in CEU we saw an enrichment of functional regulatory variants across all cancers as well as whole-genome imputation with 2,642 cases and 2,099 unaffected female controls collected from the Biobank Japan. We further examined 13 suggestive loci (P<1.0×10^-10, OR>1.21; 95% CI=1.15-1.28) and rs12292068 (KDR: OR=2.89, 95% CI=1.77-4.73) on chromosome 3q26.2 (FGFR2); rs3803662 (combined P-value of 2.79×10^-11, OR=1.51, 95% CI=1.43-1.60) and rs9319401 (combined P-value of 3.97×10^-10, OR=1.23, 95% CI=1.15-1.31) on chromosome 16q12 (TOX3-LOC643714). Weighted genetic risk score by using three significantly associated variants and two previously reported breast cancer associated loci in East Asian population revealed that individuals who carry the most risk alleles have 2.2 times higher risk of developing breast cancer in the Japanese population compare to those who carry the least risk alleles. In addition, pathway analysis suggested that variants within aryl hydrocarbon receptor signaling pathway from Ingenuity database are associated with the risk of breast cancer. Although we could not identify additional loci associated with breast cancer risk in East Asian population, pathway analysis suggested that variants within aryl hydrocarbon receptor signaling pathway in Japanese population compare to those who carry the least risk alleles. In future studies, we will further analyze the association between variants across the risk loci in Japanese population to identify additional genetic variants that could lead to a better, accurate prediction of breast cancer.


Genome Wide Association Studies (GWAS) have identified large numbers of single nucleotide polymorphisms (SNPs) that reproducibly influence disease risk (Manolio et al. 2009). Such SNPs are more likely to be expression quantitative trait loci (eQTL) than random SNPs (Nicolaie et al. 2010), suggesting that they may exert their phenotypic effect by altering gene regulation rather than modifying the product of the implicated gene.

To investigate this phenomenon, we examined whether there is an overlap between cancer risk SNPs and eQTLs with cancer risk SNPs from all GWAS reported in the NHGRI catalogue, and whether these regulatory SNPs display ‘variant pleiotropy’ (i.e. when multiple variants affect the same gene/pathway). To test for enrichment at the SNP-level we first created random sets of SNPs by conditioning on minor allele frequency (MAF), genotyping array, function, and distance to transcription start site of cancer-associated SNPs. All proxy SNPs in LD at a threshold of r^2 > 0.8 and max-distance of 500kb were generated for the random sets, and Regulatory SNPs and eQTLs were aligned with these proxy SNPs. The number of proxy SNPs in each scoring category was used to create an empirical distribution of functional SNPs in LD with the random sets to calculate the enrichment of the cancer SNPs for each scoring category. The cancer SNPs were pruned (r^2 > 0.2) while the random sets were in order to obtain a conservative estimate of enrichment. To determine if functional variants affect similar genes/pathways we used DAVID (Huang et al. 2007) and GREAT (McLean et al. 2010).

We tested for functional variant enrichment and ‘variant pleiotropy’ in the HapMap populations of both CEU and CHB/JPT. At the SNP-level, in CEU we saw an enrichment of functional regulatory variants across all cancers as well as in bladder and prostate cancer. On the other hand, CHB/JPT was not enriched across all cancers but was in hepatocellular carcinoma, lung, and prostate cancer. Genes and pathways identified were enriched for cancer related gene families and pathways such as the MAPK pathway, TGF-Beta pathway, Basic-Helix-Loop-Helix genes, and Zinc Finger genes. Our findings support the hypothesis that regulatory variants play a role in disrupting the normal function of cancer associated genes and pathways.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3302S
Lynch syndrome [LS] is a hereditary disorder characterized by an increased risk of developing colorectal cancer and other specific cancers. LS follows an autosomal dominant pattern of inheritance and is caused by genetic variants that decrease DNA mismatch repair capacity. Individuals with LS have a 50-80% lifetime risk of developing colorectal cancer, with an average age-of-onset of 45 years. In LS families, offspring developing cancer approximately 10 years earlier than their parents, but the cause of anticipation in LS is not yet understood. Here, our goal is to determine whether anticipation in LS results from an excess of single nucleotide and small indel de novo mutations due to decreased DNA mismatch repair capacity. To test this hypothesis, we sequenced 39 individuals from 3 control and 7 LS families to greater than 30x coverage on two separate whole genome sequencing (WGS) platforms, Illumina and Complete Genomics. To identify single nucleotide and small indel de novo mutation candidates, we analyzed the Illumina WGS data with four pedigree-aware variant callers (GATK PhaseByTransmission, Denovogear, Famseq and Polymut) as well as two general-purpose variant callers (GATK HaplotypeCaller and GATK UnifiedGenotyper). We report the estimated single nucleotide and small indel de novo mutation rates in each family for candidates identified from the Illumina data and validated by the Complete Genomics data. With this approach, the expected number of false positive mutations is less than 1 per family. We also report the accuracy of the six variant callers as measured by the specificity, sensitivity and receiver operating curves (ROC) of each method.

3303M
Excess Prevalence of Gastric Cancer Family History Among Hispanic Breast Cancer Patients. I. Solomon, A. Sunga, K. Yang, J. Weltzel. City of Hope National Medical Center, Duarte, CA.
Purpose: As the 2nd most common cause of cancer death and 4th most common incident cancer, gastric cancer (GC) accounts for approximately 10% of all cancer deaths. Environmental and genetic risk factors have been established to play a role in the etiology of GC. However, outside of the few known inherited syndromes that cause GC, there is a lack of data exploring other contributing genetic factors, specifically family-based studies. In addition, there is a dearth of information regarding familial GC in the background of breast cancer aside from Hereditary Diffuse Gastric Cancer and Peutz-Jeghers syndromes. This report aims to characterize a population with a family history of GC from the City of Hope Clinical Cancer Genetics Community Research Network (CCGCRN) and explore the presence of familial GC in individuals with a personal history of breast cancer. We aim to contribute to the bridge-base of familial GC and illuminate potential associations. Methods: The CCGCRN’s Institutional Review Board-approved registry was queried (Progeny® v9.1) for the keywords “family history of GC” obtained from patients’ individual data spreadsheets. Review of information was entered into a database for analysis. Individual pedigree evaluations were completed for families with more than one family member afflicted with GC. Descriptive analysis using SPSS® v19 was conducted to create frequency tables. Results: There were 1116 (9.4%) families with a history of GC. Notably, 33% (365) reported Hispanic ancestry as compared to 26% Hispanic representation in the registry overall. Most families had one member with GC (919), 143 had 2 family members with GC and 54 had ≥ 3. A majority (657) of these families were based on a breast cancer index case. 58% (314) of these families carried a known mutation, with the majority (99) carrying an HBOC gene mutation. The most common mutation was the Ashkenazi Jewish Founder mutation BRCA1 185delAG (11), then 3492insT (7) and 6174delTT (5) in BRCA2. Two of the populations have been previously shown to be common Hispanic mutations in the U.S. Gastric cancer has been previously reported in association with HBOC. Our findings raise the question of an under appreciated association of GC in HBOC. However, a larger sample is needed to further define the phenotypic variant GC in HBOC families. In addition, this data prompts the idea of population-dependent genetic modifiers that may play a role in Hispanic HBOC families.
3304T Association of polymorphism in GSTM1 null with obesity in breast cancer patients triple negative. O. Soto-Quintana1,2, B. Sánchez2, R. Ramírez1,3, A. Ramos1,2, D. Carrillo1,3, J. A. Gutiérrez1,3, A. R. Rincon4, A. M. Puebla5, MP. Galettos1. 1) División de Genética Molecular, CIBO, IMSS, Guadalajara, Jal., Mex; 2) Doctorado en Farmacología, CUCS, U de G; 3) Doctorado en Genética Humana, CUCS, U de G; 4) Coordinación del Doctorado en Farmacología, Departamento de Fisiología, CUCS, U de G; 5) Laboratorio de Inmunofarmacología, Departamento de farmacología, CUCEI, U de G; 6) Consulta Externa de Oncología, UMAE Hospital de Ginecología, Facultad de Medicina, Universidad del Desarrollo, Santiago, Chile.

Introduction: Bladder Cancer (BC) is a heterogeneous disease with a variable natural history. It is currently the fourth most common malignancy in NPC cases (P<0.00087). All these pathways are critical for cell survival, growth and metastasis. Another set of NPC families (218 exomes, 110 families with 2-3 cases) will be exome sequenced to confirm our results and expand the search for NPC susceptibility genes.

3305S Genetic Variants related to presence of Bladder Cancer in a high risk, arsenic-exposed population in Northern Chile (Antofagasta). C. Vial1, K. Espinoza1, E. Chapparo2, G. Repeito3, M. Fernández1,3, 1) Centro de Genética y Genómica, Facultad de Medicina, Universidad del Desarrollo Clínica Alemana, Santiago, Chile; 2) Servicio de Urología, Hospital Regional de Antofagasta, Chile; 3) Departamento de Urología, Clínica Alemana.

Introduction: Bladder Cancer (BC) is a heterogeneous disease with a variable natural history. It is currently the fourth most common malignancy among men in the western world, following prostate, lung, and colon cancers. Incidence in an arsenic-exposed city in Northern Chile (Antofagasta 24.8/100,000) is 4 to 5 times higher than in the rest of the country. Concentrations in drinking-water in Antofagasta increased significantly from 90 to 800-900 µg/L when new drinking-water sources were introduced in 1958. Following this, nearly all of the population in Antofagasta was exposed to arsenic levels up to seventeen times over the WHO recommendation between 1958 and 1971. In the present study we aim to perform a Genome Wide Association Study (GWAS) on people exposed to Arsenic comparing cases with controls. Methodology: Individuals were invited to participate after signing an informed consent. A blood sample was obtained and a questionnaire with epidemiological and clinical information was applied. DNA samples were analyzed using Affymetrix Genome-Wide SNP Array 6.0. After filtering by missingness per individual, missingness per marker allele frequency and Hardy Weinberg Equilibrium we obtained 788,705 SNPs to be analyzed. Results: 42 BC patients and 36 control subjects have been enrolled so far in the study. All of them were exposed to arsenic at some point between 1955 and 1971. The epidemiological data showed that males are predominant among cases and controls (64.3% and 83.3%, respectively) and there is no significant difference concerning mean age, familial history of BC, occupational exposure or smoking status between groups. It is interesting to note that smoking prevalence is high and similar among cases and controls (59.5% and 58.3%, respectively; p=0.55). The sample for population stratification was analyzed using principal component analysis (PCA), clustering the different patients with the identity by state and found them to be a homogeneous population. An association test comparing cases and controls was performed and found two regions with a significant association: (a) rs4838646 in chromosome 10 (p=3.8E-06) and (b) rs12371702 in chromosome 12 (p=5.8E-06). Previous studies have linked polymorphisms in the former region to BC susceptibility. Conclusion: Initial results of a BC genomic case-control study in an arsenic exposed population are presented. Further analysis is warranted after completion of recruitment. Fondecy1120987.

Objective: A longer female reproductive lifespan increases exposure to endogenous estrogen number of ovulations, which can be a risk factor for breast and ovarian cancers. We investigated the effects of BRCA1 and BRCA2 on female reproductive traits and lifespan to understand the differences in risks for breast and ovarian cancers in BRCA1 compared with BRCA2 mutation carriers. Further, we explored the potential of other genes that influence reproductive traits/lifespan and are functionally independent of BRCA1 and BRCA2 mutations. Methods: We studied a total of 1058 women from the Creighton University Hereditary Cancer Registry. The women who tested positive or negative for their family’s deleterious mutation in BRCA1 and BRCA2, and provided age at menarche, age at natural menopause (ANM), or both were included in the study. The linear mixed-effects model was used to explore the effects of BRCA1 and BRCA2 on female reproductive traits and lifespan. Heritability estimates were calculated using variance-components methods in Statistical Oligogenic Linkage Analysis Routines (SOLAR) to explore the potential of other genes with respect to female reproductive traits and lifespan. Results: The reproductive lifespan of the women with a BRCA1 mutation was 3.12 years shorter (p-value = 0.048, n=38) than the non-carrier relatives. The significant difference was not observed in the BRCA2 families. Age at menarche was significantly heritable after adjusting for the presence of a BRCA1 or BRCA2 mutation (heritability [h^2] = 0.57 and 0.34, and n=510 and 416, respectively; p-values <0.01). Conclusions: Women with a BRCA1 mutation may have a shorter reproductive lifespan as compared to non-carrier relatives, causing a shorter period of exposure to estrogen and lower number of ovulations. The high heritability after adjusting for the presence of a BRCA1 or BRCA2 mutation underscores the importance of further work to identify other specific genes that contribute significantly to the variation in age at menarche in BRCA families.


Genome-wide association studies (GWAS) have identified more than 60 loci associated with prostate cancer (PCa) in men of European ancestry, including 23 loci recently identified in a large study from the PRACTICAL consortium. We assessed the evidence for these 23 new PCa loci and the aggregate predictive value of these 23 SNPs plus 40 established SNPs with respect to early-onset (EO) PCa using 931 unrelated men diagnosed with PCa prior to age 56 years from the University of Michigan Prostate Cancer Genetics Project and 1126 male controls. Ten of the 23 new SNPs demonstrated evidence (p < 0.05) for association with EO disease. On average, EO PCa cases carried one more total risk alleles compared to controls (21.61 vs. 20.69, p=2.0×10^-12 and 3.5 more risk alleles across all 63 SNPs (58.02 vs. 54.49, p=8.9×10^-59). We constructed receiver operating characteristic curves and calculated the corresponding area under the curves (AUC) for weighted aggregate risk allele counts for the 23 new PCa SNPs (AUC=0.59), 40 established PCa SNPs (AUC=0.69) and the set of 63 total PCa SNPs (AUC=0.71). The defined risk alleles and their associated weights were based on previous reports. We found that the aggregate burden of common risk alleles across previously associated PCa SNPs is highly predictive of EO PCa for men with values of total risk burden in the top 25% of the total risk allele distribution. The odds of a man having EO PCa given a burden of risk alleles in the upper, or lower, 5% of the combined case-control distribution of risk alleles across all 63 SNPs was estimated to be 11.11 [95% CI (5.91,20.90)] or 0.24 [95% CI (0.14,0.40)], respectively. While attenuated, odds remained strong for men in the upper or lower 25% of the total risk allele distribution (OR=3.81 [3.08,4.72] or OR=0.27 [0.21,0.33], respectively). Inclusion of results from the 23 new variants only modestly improved disease prediction over the 40 established stronger effect SNPs. Our results provide the first formal replication for many of the 23 new variants recently reported to be associated with PCa and firmly establish the importance of common variants, identified to be associated with PCa in GWAS, in men with early-onset disease.
3311S
Hispanic MMR Mutations: A Multi-Institutional Report from Southwestern United States and Puerto Rico. A. Sunga, C. Ricker, C. Espenschied, J. Herzog, S. Brennan, M. Cruz Correa, P. Lynch, S. Gruber, J. Weitzel. 1) Clinical Cancer Genetics, City of Hope, Duarte, CA; 2) Clinical Cancer Genetics, USC Norris Comprehensive Cancer Center, USC Los Angeles, CA; 3) Ambry Genetics, Aliso Viejo, CA; 4) Clinical Cancer Genetics, MD Anderson Cancer Center, Houston, TX; 5) Clinical Cancer Genetics, University of Puerto Rico Comprehensive Cancer Center, Rio Piedras, Puerto Rico.

Background Knowledge of founder mutations enable efficient and cost-effective strategies for genetic testing, a potential benefit for populations with limited access to services. We have shown several BRCA mutations, most of Spanish origin, to be founder mutations in Hispanic populations. We hypothesized that the same population factors may be operative in Lynch syndrome. There is limited literature on the spectrum of mutations in mismatch repair (MMR) genes (MLH1, MSH2, MSH6, and PMS2) in Hispanic populations. Our goal was to identify recurrent MMR mutations in Hispanic patients and explore potential ancestral origins of identified mutations.

Methods Subjects: Hispanics patients seen for genetic risk assessment at City of Hope Clinical Cancer Genetics Community Research Network collaborating institutions, MD Anderson Cancer Center, USC Norris Cancer Center and the University Of Puerto Rico Cancer Center. A total of 387 patients underwent evaluation for Lynch syndrome by genotyping of one or more of the MMR and/or tumor analysis by IHC. A comprehensive literature review and of MMR variant databases was conducted for all mutations.

Results Seventy one (18%) unrelated Hispanic families were found to have at least one MMR disease predisposing gene mutation. Most (80%) of the mutations were in the MLH1 and MSH2 genes and more than half (55%) were frameshift or nonsense type. Nine mutations were observed two or more times: MLH1 350C>T, 1790delAinsT, 2041G>A, 1852del3, 1024del16, and IVS7+5G>A; and MSH2 1216C>T, 425C>G, and 1765delGA. Most (6/9) recurrent mutations were detected in multiple institutions. MSH2 1216C>T and MLH1 1852del3 were seen 5 and 4 times respectively and have been reported multiple times in European populations. MLH1 350C>T, 2041G>A, 332C>T, and 676C>T and MSH2 1216C>T, exon 4-8 del, and exon 8 del were all being reported previously in Spanish MMR genes. MSH2 exon 4-8 del was seen in one Mexican family and has been reported as a Spanish founder mutation.

Conclusion Our finding that 3/9 recurrent mutations and the MSH2 exon 8 del identified in our cohort were previously reported in Spain supports the hypothesis that they may be part of a shared ancestral heritage on MMR gene mutations in Hispanic populations. While this is the largest reported cohort of Hispanic patients with MMR mutations in N America, a larger sample and haplotype analyses are needed to better define the spectrum and origin of MMR mutations in Hispanic populations.

3312M
Pleiotropy between Hodgkin lymphoma and other immunological diseases. W. Cozen, P. Khankhanian, Y. Kong, D. Himmelstein, R. Jarrett, J. McKay, A. van den Berg, F. Gilliland, K. Oneal, S. Baranzini, W.J. Gauderman, J. Oksenberg, H. Hjalgrim, EVE Consortium. 1) Preventive Medicine and Pathology, USC Keck School of Medicine, Norris Comprehensive Cancer Center, Los Angeles, CA; 2) Department of Neurology, University of California San Francisco 675 Nelson Rising Lane Suite 215 San Francisco, CA; 3) MRC - University of Glasgow Centre for Virology Research Institute of Infection, Immunity and Inflammation University of Glasgow Garscube Estate Glasgow G61 1QF; 4) Genetic Cancer Susceptibility group, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France; 5) Department of Pathology & Medical Biology University Medical Center Groningen University of Groningen Hanzeplein 1 P.O. Box 30.001 9700 RB Groningen The Netherlands; 6) Section of Hematology/Oncology The University of Chicago 900 East 57th Street, Rm 5140 MC 4060 Chicago, IL 60637; 7) Consultant Department of Epidemiology Research Statenum Serum Institut S Amagerskovvej 1 DK-2300 Copenhagen S Denmark.

Hodgkin lymphoma (HL) is a highly heritable B-cell lymphoma that primarily affects young adults. Its epidemiology and pathology are unique and more closely aligned with immunological diseases than with other solid tumors. To test this hypothesis, we conducted studies examining the genetic overlap between HL and multiple sclerosis (MS) and asthma. We performed a meta-analyses of HL and MS consisting of 1,816 HL patients, 9,772 MS patients, and 25,255 controls using 464,424 single nucleotide polymorphisms. Genetic overlap was observed in the HLA region and genome-wide, including modest genome-wide significant associations at two loci that did not reach genome-wide significance in either disease alone: IL12B rs2546890, ORHL-asthma = 1.1, p = 3.37 x 10^-9, a gene encoding a subunit of a cytokine involved in T-helper cell Type 1 (Th1) immunity, and NCOA5 (rs2425752, ORHL-asthma = 1.11, p = 2.96 x 10^-8), a gene included in an estrogen-dependent oncogenic pathway. We conducted a separate analysis of HL and asthma combining GWAS meta-analyses of 1,816 HL cases and 7,877 HL controls with a GWAS meta-analysis of 2,088 asthma cases and 2,743 asthma controls examining 904,634 SNPs in common. We found associations with two linked SNPs in the T-helper cell Type 2 (Th2) transcription factor gene GATA3 (rs422628, ORHL-asthma = 1.247, p = 3.36 x 10^-9) and rs444929, ORHL-asthma = 1.264, p = 2.09 x 10^-6, both of which were significant at the p < 0.05 level in the disease-specific GWAS, but reached genome-wide significance only in the HL-asthma meta-analysis. The effect was stronger in the subset of nodular sclerosis HL (rs422628, ORHL-asthma = 1.32, p = 2.33 x 10^-9 and rs444929, ORHL-asthma = 1.32, p = 4.28 x 10^-10). Another genome-wide significant SNP near KIAA1279 reached genome-wide significance only in the nodular sclerosis subtype subset (rs6864, ORHL-asthma = 0.81, p = 4.65 x 10^-8). In a genetic diseasome analysis, HL was more closely aligned to both MS and asthma and other immunological diseases than to solid cancers. Recognition of overlap in genetic predisposition to HL and immune diseases sheds light on the complex etiology of HL and may enable novel diagnostic and therapeutic approaches.
Prospectively Identified Incident Testicular Cancer Risk in a Familial Testicular Cancer Cohort. A. Pathak1, C. Adams2, J. Loud3, K. Nicholas2, M. Greene1, Y. D. Stewart1. 1) Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; 2) Westat, Inc., Rockville, MD, USA. Testicular germ cell tumors (TGCT) are the most common form of cancer in men aged 15-35 years. Approximately 8,850 new cases and 380 TGCT deaths are projected for 2014. Human testicular germ cell tumors (TGCT) have a strong genetic component and a high familial relative risk. TGCT has an estimated heritability that ranks as the 3rd highest for all cancers. Comorbidity considerations - which have not been high risk for men with TGCT - 1.5 to 2.5-fold - sons and brothers of affected men with TGCT have a 4 to 6-fold and 8 to 10-fold increased risk of TGCT versus the general population. These risks increase to 37-fold and 76.5-fold in dizygotic and monzygotic twins, respectively. However, linkage analysis has failed to identify any rare, highly-penetrant familial TGCT (FTGT) susceptibility loci. Currently, multiple low-penetrance genes are hypothesized to underlie the familial multiple-case phenotype. Despite the apparent heritability of TGCT, families with more than two affected members are unusual, unlike other hereditary cancer syndromes in which a single multifocal penetrance pedigree can harbor many affected individuals. Importantly, the prospective TGCT risk in this context is unknown. Thus, we performed a prospective quantitative analysis of TGCT incidence in a cohort of multiple affected persons and sporadic bilateral TGCT case families, 1,280 men from 140 families (10,207 person-years of follow-up) met our inclusion criteria. Age, gender, and calendar time-specific standardized incidence ratios (SIR) for TGCT relative to the general population were calculated using SEER*Stat. Eight incident TGCTs occurred during the study period (1.67 expected cases; 95% CI: 0.09-5.76). We calculated the incidence rate of TGCT is much greater among bloodline male relatives in multiple-case testicular cancer families than expected in the general population, a pattern characteristic of adult-onset Mendelian cancer susceptibility disorders. Remarkably, two incident TGCTs occurred in healthy relatives of sporadic bilateral cases (0.15 expected; SIR=13.4; 95% CI=1.6-48.6). Our data are the first to indicate that, despite relatively low numbers of affected TGCT individuals per family, families with multifocal affected persons with colorectal cancer (CRC) and breast cancer (BC) cases comprised high-risk groups for incident testicular cancer, which might benefit from tailored risk stratification and surveillance strategies.

Posters: Cancer Genetics

DEPTH: A Novel Algorithm for Feature Ranking with Application to Genome-Wide Association Studies Identifies that Variation in the CHEK2 Gene Region is Associated with Risk of Breast and Colorectal Cancer.

J. D. Schmidt1, E. Makalic1, M. Kapuscinski1, W. Chen2, P. M. Lynch3, C. I. Amos3, J. L. Hopper4, M. Pandi2. 1) School of Population and Global Health, The University of Melbourne, Victoria, Australia; 2) Department of Clinical Applications and Support, MD Anderson Cancer Center, TX, USA; 3) Department of Gastroenterology - Research, MD Anderson Cancer Center, TX, USA; 4) Division of Epidemiology and Community Health, School of Population and Global Health, The University of Minnesota, Minneapolis, Minnesota, USA. A conventional Genome-Wide Association Study (GWAS) involves selecting single nucleotide polymorphisms (SNPs), from among a very large set of correlated SNPs, that discriminate between two groups. The ultimate aim of a GWAS is to identify susceptibility genes, or gene regions; the dominant paradigm to date has been to do this by identifying individual SNPs associated with risk of disease. We propose a new algorithm, called DEPendency of association on the number of correlated SNPs, that discriminate between two groups. The ultimate aim of DEPTH is to aim more efficiently by using bootstrap statistics and stability selection and considering contiguous SNPs as a group. DEPTH: (i) exploits information from the correlation structure of the predictors without specifying an underlying data-generating process; (ii) (global), and Cox regression, so can be used to study continuous, binary and survival data, and (iii) can be run on a commodity computer but has also been implemented to run much faster in a parallel computing environment on the IBM BlueGene/Q supercomputer. In the context of a GWAS, the algorithm can consider SNPs across the whole genome, or subsets of SNPs (e.g. in a region or pathway, or of a particular ‘type’). We have found, using simulated data, that the algorithm shows good statistical performance when compared to several established procedures. We applied DEPTH to a subset of GWAS data from the Breast and Colon Cancer Family Registry (BCFCR) to try to enrich for SNPs associated with susceptibility to breast and/or colorectal cancer. Cases (985) were women with breast cancer and a family history of colorectal cancer, or persons with colorectal cancer with a family history of breast cancer. Controls (1,974) were women without breast and colorectal cancer. DEPTH identified that germline variation in the region of the checkpoint kinase 2 (CHEK2) gene is associated with risk of breast and/or colorectal cancer, a finding not evident from applying the conventional ‘genome-wide significant’ threshold. Lending support to our findings, several germline mutations in CHEK2 (e.g., 1100delC, IVS2+1G>A and 1577T) are known to be associated with the risk of breast, colorectal, ovarian, lung and other cancers. Our findings from studying a modestly sized GWAS dataset suggest that the CHEK2 gene region (using family history across diseases) and analysis using DEPTH).
Integrated pathway and gene-gene interaction analysis reveals novel candidate genes for melanoma. M. Brossard1,2, S. Fang4, A. Vayssé1,2, Q. Wei3, H. Mohamadi1,2, W.Y. Chen3, N. Lavieille1,2, E. Maubec1,2, M.-F. Avril1, M. Lathrop5, J.E. Lee5, C.I. Amos1, F. Demenais1,2, MELARISK and MDACC Melanoma Study Groups. 1) INSERM UMR-946, Paris, France; 2) Université Paris Diderot, Paris, France; 3) MD Anderson Cancer Center, Houston, Texas, USA; 4) Department of Medicine, Duke University School of Medicine, Durham, USA; 5) Hôpital Cochin, Université Paris Descartes, Paris, France; 6) Genome Quebec Innovation Centre, McGill University, Montreal, Canada; 7) Geisel College of Medicine, Dartmouth College, New Hampshire, USA.

Genome-wide association studies (GWAS) have identified 17 loci associated with melanoma risk. However, these loci account for a small part of melanoma susceptibility. These GWAS, which have focused on the analysis of individual SNPs, are underpowered to detect SNPs with small marginal effect. To increase the yield from GWAS, alternative approaches that use biological knowledge and allow testing for association on the basis of functional units (pathways) have been proposed. To identify new candidate genes for melanoma, we proposed an analysis strategy that combines pathway analysis and tests of gene-gene interactions (GxG) within melanoma-associated pathways. Pathway analysis was conducted using the gene-set enrichment analysis (GSEA) approach, which searches for gene sets, defined by gene ontology (GO) classes, enriched in genes associated with melanoma. GSEA was applied to single-SNP statistics obtained from GWAS of the French MELARISK study (3,976 subjects) and the North-American MDACC study (2,827 subjects). To identify enriched GO categories, we computed the false discovery rate (FDR), using 100,000 permutations of SNP statistics. We tested all cross-gene SNP-SNP interactions within each identified GO using INTERSNP. One million Hapmap3-imputed SNPs were assigned to 22,000 genes, that were assigned to 316 Level4-GO classes. Five GO categories were found to be significantly enriched in genes associated with melanoma (FDR<5% in both studies): response to light stimulus, regulation of mitotic cell cycle, induction of programmed cell death, cytokine activity and oxidative phosphorylation. A total of 101 genes were driving the enrichment signals in these pathways. Five of these genes (TP63, IL6, IL15, MAPK1, NDUFA2) were found to occur frequently with melanoma-related expression of four of these genes (TP63, IL6, IL15, MAPK1) has biological relevance. One million Hapmap3 imputed SNPs were assigned to 21,917 genes using NCBI Build 37.1. These genes were assigned to 317 level 4-GO categories. Three GO categories were found to be significantly enriched in genes associated with BT (FDR<5% in both studies): Hormone activity, Cytokine activity and Myeloid cell differentiation. A total of 61 genes were driving these pathways in both studies. Interestingly, expression of four of these genes (CXCL12, VEGFA, CDC42) was reported to be associated with melanoma progression in tumors. Analysis of cross-gene SNP-SNP interactions within each identified pathway showed evidence for interaction for three SNP pairs (P<10^-5 in MELARISK and replication at P<0.05 in MDACC). One of these gene pairs (SCINxCDC42, combined P over the 2 samples=2x10^-5) has biological relevance. SCIN and CDC42 proteins are both involved in actin cytoskeleton dynamics, that plays a major role in cell migration, and have opposite roles: CDC42 stimulates actin assembly while SCIN severs the actin filaments, thus preventing their growth. This study outlines the importance of integrating various approaches for gene identification. Funding: INCa_5982.
A Population-based survey of excess cancers observed in NF1 cases and in their first- and second-degree relatives. D. Abbott, D. Viskochil2, D.A. Stevenson1,2, L.A. Cannon-Albright1,2,1 1) Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT; 2) Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT; 3) George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT.

Background: Neurofibromatosis type I (NF1) is an autosomal dominant disorder that affects approximately 1 in 3000 individuals and is caused by a mutation in NF1, a tumor suppressor gene that encodes the protein neurofibromin. While NF1 individuals often have benign peripheral nerve sheath tumors (neurofibromas) and intracranial low-grade astrocytomas (gliomas), they are also at increased risk for malignant peripheral nerve sheath tumors (MPNST) and high-grade astrocytomas. Current research purports that NF1 individuals have increased risk for other cancers. We analyzed a unique population-based resource to characterize the increased risk of various cancers in NF1 individuals and their relatives in an unbiased fashion. Methods: The Utah Population Database (UPDB), a population-based Utah genealogy linked to both the Utah Cancer Registry and to hospital data from the University of Utah Health Sciences Center, was used to calculate relative risks (RR) of cancer in NF1 cases and in their first- and second-degree relatives. RR was calculated using age-, birth-state-, and birth-year-specific cancer rates. Results: We identified 237 individuals with NF1 who had ancestors in the UPDB. These cases had 1055 first-degree and 2770 second-degree relatives in the UPDB. The RRs for 6 cancers were significantly higher in NF1 cases: Malignant Peripheral Nerve Sheath Tumor-MPNST (RR=6.469.9; 95%CI: 2.60.7, 13.22.8), brain (4.30.20.03, 83.15), central nervous system (51.08-165.94, 1192.66), lymphoma (6.04-1.65, 15.47), skin melanoma (4.97-1.02, 14.52), and breast cancer (3.61.01, 9.23). The RR of seven cancers were significantly higher in first-degree relatives; MPNST (RR=118.03.47.45, 243.18), brain (5.48-2.20, 11.29); lymphoma (6.26-1.29, 18.29), acute lymphocytic leukemia (1.74-1.01, 2.82), and small intestine (7.07-1.01, 25.54). Lastly, in second-degree relatives there was a significantly increased risk of two cancer types: breast cancer (1.74-1.01, 2.82) and malignant melanoma (1.29-1.29, 18.29). These results support an increased risk of distinct cancers in individuals with NF1 and also in their first- and second-degree relatives. Relative risks for various cancers are much higher for NF1 cases themselves, but first- and second-degree relatives still display a significant excess of some cancers.
Chromosomal Mosaicism in Patients with Familial Chronic Lymphocytic Leukemia

L.R. Goldin1, W. Zhu2, M. Yeager3, M. Dean4, M.J. Machiela4, M.L. McMaster1, J. Boland5, M. Cullen6, N.E. Caporaso1, S.J. Chanock2, 1) Genetic Epidemiology Branch, DCEG, NCI, Bethesda, MD; 2) Cancer Genomics Research Lab, DCEG, NCI, Bethesda, MD; 3) Laboratory of Experimental Immunology, Center for Cancer Research, NCI; 4) Laboratory of Translational Genomics, DCEG, NCI, Bethesda, MD.

Chronic lymphocytic leukemia (CLL) is a malignancy characterized by the accumulation of small, mature-appearing lymphocytes in the bone marrow, blood, and lymphoid tissues accounting for about one third of the leukemias in US adults. About one third of patients have an indolent disease course.

Studies have shown significant familial aggregation of CLL and its precursor trait, monoclonal B-cell lymphocytosis (MBL). CLL tumors are characterized by several cytogenetic abnormalities, the most common being deletions of 13q14, 11q, 17p, 6q, and trisomy 12. We have used Illumina Omni Express arrays to detect these and other abnormalities in DNA derived from blood. Our analytic method can identify abnormalities present in 7% of cells or higher, allowing us to detect abnormalities in both normal tumor and normal cells. We genotyped 130 CLL patients from high-risk families (including some patients with longitudinal samples), as well as 24 first degree relatives with MBL. We use B-allele frequency and log2 intensity ratio to score copy number gains and losses, as well as copy neutral loss of heterozygosity. Mosaic abnormalities were found in 55% of CLL patients, but in none of the relatives with MBL. These abnormalities included those commonly seen in CLL patients as well as more recently reported trisomies of chromosomes 18 and 19. We also found evidence of chromothripsis in some patients. The rate of mosaic abnormalities in familial CLL patients is consistent with that seen in sporadic CLL patients. Among 15 patients studied on two or more time points, changes in the pattern of abnormalities could be seen over time. We are correlating the pattern of abnormalities with disease progression and treatment in this study with more patients, relatives, and time points. Future analyses will be conducted on purified tumor cells compared to normal cells and whole genome sequencing will be performed in selected patients to better define abnormalities. Together, these studies will allow us to identify the earliest genetic changes present in CLL, highlight those associated with progression, and aid in identifying genomic alterations in relatives at increased risk for developing CLL.
A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.

A t(1;19) translocation involving TCF3/PBX1 fusion within a context of a hyperdiploid karyotype in adult B-ALL.
3331T  
The dilution dilemma; a method to accurately estimate tumor fractions in complex tumor/normal DNA dilutions.  
Z. M. Weber1, S. Vattathil2, Y. Hu1, P. C. Yang1, E. A. Dull3, E. A. Dull3, J. A. Avera5, Institute for Human Genetics, Sioux Falls, SD, USA; 2) The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA.

The objective of this project was to create a series of highly accurate dilutions of matched tumor/normal DNA to test computational methods for detecting low-frequency chromosomal aberrations, such as allelic imbalance and copy number changes, in a sample of DNA obtained from heterogeneous cells. Such methods (e.g. hapLOH, GAP, BAFsegmentation, ASCAT) offer the potential to tumor cell fraction and cell line composition. However, the usefulness of these sections that contain normal tissue margins and to identify tumor sub-populations in samples that possess heterogeneous genetic profiles. Qubit is a fluorescence-based method used to quantify double stranded DNA. Using this method of quantitation, we created 32 tumor/normal mixtures (16 dilutions in 2 cancer cell lines) with tumor fractions ranging from 0.98 to 0.005. Dilutions containing ultra-low fractions of tumor DNA required a method of highly accurate quantitation. Breast cancer (CRL-2363) and lung cancer (HTB-172) tumor cells were grown in culture with their respective matched normal cell line lymphocytes, CRL-2343 and CRL-5998, respectively. DNA was extracted from the four cell lines using the QIAamp DNA Blood Midi protocol. The DNA from each cell line pair was normalized and quantitated in five replicates using Qubit High Sensitivity assays. The standard deviation (SD) and coefficient of variation (CV) between the five replicates was calculated for each DNA sample. Next, a series of tumor DNA fractions was prepared for both the breast and lung cancer DNA sets. Two sets of 16 dilutions were run on Illumina HumanOmni2.5-Beadchips. The SNP genotyping data was processed using the FALCON algorithm. Low frequency calls were used for downstream analyses with hapLOH (for example). Low SD values were successfully achieved between replicate sets (0.69 - 1.45 ng/ul) along with low CV values (0.013 - 0.028). Using a mathematical algorithm which incorporates the volume of DNA used in the dilution and the mean concentrations, we were able to estimate the amount of tumor DNA in the complex mixture. The mean GenCall score across all samples was p10 GC = 0.41 and p50 GC = 0.78 following the genotyping experiment. Call rates were all > 99% and error rates between the three tumor fractions at 0.0%.

3333M  
The actual impact of Fluorescence in Situ Hybridization (FISH) in the diagnosis and follow up of Acute Lymphoblastic Leukemia (ALL).  
H. Akita, K. Karam, B. Durnam, A. Durnam, A. Aydin, O. Cogull, University Medical School, The Department of Medical Genetics, Izmir, Turkey.

Acute lymphoblastic leukemia (ALL) is a disease characterized by abnormal clonal proliferation of lymphoid precursors which lose their ability to differentiate. It is the most common malignancy in childhood, accounting for 80% of all leukemia. Although it can occur at any age, its incidence is highest among children from 2 to 5 years at rate of about 70%, decreasing among adolescents and young adults. Cytogenetics and FISH have been the gold standard for defining genetic abnormalities and facilitating therapeutic stratification of pediatric ALL cases. In this study, it is aimed to determine the cytogenetic and molecular cytogenetic profile of submitting patients with ALL during the period from January 1st, 2013 to December 31st, 2013 to Ege University Medical Faculty, Medical Genetics Department. In ALL, the most important molecular markers are BCR/ABL, ALL translocations with their respective partner genes, and c-MYC aberrations that have to be detected at diagnosis and for minimal residual disease (MRD ) studies. Furthermore, for more than 90% of ALL patients specific markers can be detected by IGH mutations and serve as the most important MRD marker. We used a standardized a FISH panel for ALL including, BCR/ABL, cMYC, TCF/PBX, IGH, DLL, p16 and TEL/AML. We performed 150 FISH analyses for 77 patients. The most frequent aberration rate was 10.67% for IGH. The second most frequent aberration was detected for delp16 as 5.33%. They were all commented as diagnosis and follow up markers for ALL. The vast majority of this study group are of pediatric age group (85.33%). The repetitive analyses by the aim of follow up and cytogenetic results were also evaluated. Protocol evolution, followed by FISH results, and the way of translation of results to patient management and the follow up request for ALL patients. A more reliable approach would be to standardize wider FISH panels in order to detect more anomalies and using an algorithm to reach effective diagnose and follow up in ALL patients.

3333S  
Age-related mosaic loss of chromosome Y is associated with cancer in cohort studies.  
W. Zhou1, M. Machiela1, M. Dean2, J. Sampson2, N. Rothman3, N. Freedman2, S. Wacholder1, K. Jacobs4, M. Tucker2, F. Real5, L. Perez-Jurado3, K. Yanger4, S. Gowda1, 1) Cancer Genomics Research Laboratory, National Cancer Institute, Division of Cancer Epidermiology and Genetics, Leidos Biomedical Research Inc., Frederick, MD; 2) Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, Md., U.S.A.; 3) Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain; 4) Centro de Investigación Biomedica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain; 5) Bioinformed, LLC, Gaithersburg, MD; 6) Epithelial Carcinogenesis Group, Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain.

Prior studies indicate that loss of the Y chromosome occurs during the aging process and is associated positively with hematologic cancer, including acute myelogenous leukemia (AML), myelodysplastic syndrome (MDS), and pre-leukemia. To investigate the relationship between mosaic loss of chromosome Y and non-hematologic cancer, we utilized existing genome-wide association study intensity data from 14,099 males (8,545 non-hematologic cancer cases and 5,554 cancer-free controls) drawn from the San Antonio Heart Study. In this prospective cohort studies: Alpha-Tocopherol and Beta-Carotene Cancer Prevention Study (ATBC), Prostate, Lung, Colorectal, Ovarian Cancer Screening Trial (PLCO), and Cancer Prevention Study-II (CPSII). Subjects were examined for deviations from expected log2 intensity ratio for presence of loss of the male specific region of chromosome Y (MSY) between 6,671,468-22,919,969 (hg18build36). DNA derived from blood or buccal samples was genotyped on one of four Illumina SNP BeadChips with sufficient coverage of MSY (Human610, Human1M, HumanOmni1M, and HumanOmni2.5M). A total of 0.9% (95%CI 0.85-0.93%) individuals have detectable mosaicism of the MSY with the proportion of affected cells ranging from 9.9% to 61.7%; this is substantially higher than what has been observed for autosomal events of size >=2MB, approximately 1% overall. Mosaic chromosome Y loss is associated with cellular aging at DNA collection, with frequency of 1.09% in individuals less than 60 and 18.63% in those 80 years or older (OR 0.11, 95% CI = 1.10-1.13; P = 5.34E-69) and ever-smoking (OR = 1.43, 95% CI = 1.21-1.71; P = 8.16E-05). In addition, mosaic chromosome Y loss is more common in individuals developing cancer than in cancer-free controls (5.37%) (OR = 1.57, 95% CI = 1.36-1.81; P = 1.02E-10). Mosaic Y loss was associated with bladder cancer (OR = 1.47, 95% CI = 1.19-1.84; P = 7.15E-04) and prostate cancer (OR = 1.29, 95% CI = 1.08-1.56; P = 7.82E-03) but not lung cancer. Further studies are required to determine the role of mosaic Y loss in cancer.
3336M
A rare transformation case report: from Chronic Lymphocytic Leukemia to Multiple Myeloma. C. Hangul1, C. Aydin1, Ö.K. Yucel2, O. Salim2, S. Berkner Karazum2. 1) Department of Medical Biology and Genetics, School of Medicine, Akdeniz University, Antalya, Turkey; 2) Department of Pathology, School of Medicine, Akdeniz University, Antalya, Turkey.

Chronic Lymphocytic Leukemia (CLL) and Multiple Myeloma (MM) are closely related B-cell malignancies. The genetic characterization of CLL has made significant progress in the recent years. Genetic aberrations are detected in over 80% CLL cases with FISH methods. In the other hand, 30-50% cases of multiple myeloma shows random genetic rearrangements. Transformation from CLL to MM is seen in very rare cases. In this report, we present 62 years old patient who was first diagnosed with Chronic Lymphocytic Leukemia in 2009. Marked lymphoid population (CD5+, CD19+, CD23+) was detected by flow cytometric immunophenotyping and bone marrow. In the first diagnosis in 2009, his karyotype revealed from bone marrow sample was designed as 45, X, -Y[4]/46, XY[10] and in FISH analysis, an 11q22.3 deletion was found of 55% in this case as a sole abnormality. At the second biopsy, cytoplasmic kappa light chain positivity was seen. Lambda and CD20 staining were found negative in 2012. Bone marrow immunophenotyping revealed 30% of clonal plasma cells (CD38+, CD56+) and 0.9% of abnormal B lymphocytes (CD5+, CD23+, CD19+) with these findings the patient was diagnosed as Multiple Myeloma. In the second bone marrow sample, his karyotype was found to be: 46,XY,del(6)(p11.12)[11],+q11[23],t(11;14)(q13.3;q32), dup(17)(q23.2)[5]/17[46,XY][2]. Additionally two hundred nuclei were analysed for each probe from MM FISH panel and t(11;14) (q13.1;q32) translocation and deletion of 6q23 was found of 48% in interphase nuclei. All these results supported the diagnosis of MM and by the same time to transformation from CLL to MM. Clinical and laboratory findings will be discussed in detail with the company of literature.

Ependymoma (EP) is a slow-growing tumor, originates from the ependymal cell lines of the ventricular system, plexus choroid and central canal, located in regions within the Central Nervous System, principally developed at the fourth ventricle and less frequently in the brain parenchyma as a result of migration of ependymal cells during embryogenesis. It is known that the EP can proceed of cells derived from the transformation of neural stem cells and their progenitors to contribute to the initiation and the development of brain tumors. In recent studies in brain tumors, it has been found that the methylation affects genes responsible for DNA repair, regulation, migration and cellular apoptosis. Methylation is associated with the development of cellular malignancy by oncogene activation, promotion of genomic instability and loss of genomic imprinting. Additionally, the chromosomal aberrations are important for the development of EP, as amplifications in 7p, 9q, 15q and losses in 14q, 6q; it is known that changes in gene expression in oncogenic processes are responsible for large cell changes; so it is interesting to analyze the methylation pattern to relate with the chromosomal aberration in ependymomas. Ependymal tumors were collected surgically in pediatric patients, genomic DNA was extracted by organic method of phenol-chloroform-isoamyl, subsequently were performed microarrays methylation and aCGH with the platform of Agilent microarrays following the manufacturer’s recommended protocol. In this work was found amplification in 14q32.33 (100%), 2p11.2 (83%), 8p22 (83%), and losses in 11q11 (83%), 15q11.2 (66%), 16q12.2 (66%), with respect to the methylation pattern, found hypomethylated a region 14q32.33 showed hyp and hypermethylated of genes related to growth and neuronal development, apoptosis, transcription factors and system immune. It is suggested to the region 14q32.33 as alteration of great importance for the Mexican population, because of the genes located in this chromosomal region.


Pancreatic cancer is the fourth leading cause of cancer-related death in the United States, with an average 5-year survival rate of less than 5 percent. Cytogenetic analysis has allowed for the elucidation of clinically significant chromosomal abnormalities in both solid tumor and hematological malignancies, but pancreatic cancer still does not have any FDA-approved cytogenetic assays despite its known intratumor cytogenetic heterogeneity. In the present study, conventional cytogenetic and molecular genetic analysis was performed on the following 9 pancreatic cancer cell lines: AsPC-1, MIA PaCa-2, BxPC-3, Capan-2, CFPA-1, HPAF-II, Panc-3, Panc-10.05, and Panc-2.13. Giemsa-banding (G-banding), targeted fluorescence in situ hybridization (FISH), and chromosomal microarray (CMA) analysis was completed for each cell line, and results were pooled to elucidate commonly aberrant regions identified by all three methods. Recurrent numerical and structural abnormalities were observed on chromosomes 7, 8, 9, 18, and 20, and these findings were further corroborated by interphase and multicolor FISH studies. Abnormalities included deletion of SMAD4 (18q21.2), which has been associated with increased metastasis and consequent poor prognosis in pancreatic cancer, deletion of DCC (18q21.1), and deletion of p16 (9p21). Ultimately, phenotypic correlation of cytogenetic abnormalities using a larger population of patients can provide a framework for the development of diagnostic and prognostic cytogenetic assays as well as targeted therapies to combat pancreatic malignancies.

3342M The ZNF384 gene in pediatric acute lymphoblastic leukemia: multiple partner genes, immature (CD10 negative) immunophenotype, and potential good outcome. M. Shago1,2, O. Alba2,4, J. Hitzler2,4, S. Weitzmann1, M. Abdelhaleem2,3. 1) Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada; 2) Department of Paediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada; 3) Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; 4) Department of Pediatrics, University of Toronto, Toronto, Ontario, Canada.

The ZNF384 (CIZ) gene, located distal to the ETv6 (TEL) gene at 12p13.31, is a putative zinc finger transcription factor which has been reported to be rearranged in acute leukemia. Rearrangement of the ZNF384 gene results in attachment of various 5' partner gene sequences to form ZNF384 fusion genes. The ZNF384 gene has three known partners: TAF15 at 17q12 (19 cases), EWSR1 at 22q12 (4 cases), and TCF3 (E2A) at 19p13 (3 cases). The proteins encoded by the fusion genes consist of the transactivation domains of the 5' partner genes fused to the entire coding region of ZNF384. We present detailed karyotypic, genomic copy number, and gene expression analyses of seven new pediatric pre-B ALL patients with ZNF384 gene rearrangement. Our patients had lymphoblasts with a CD10-negative immunophenotype, similar to the immunophenotypic profile seen in ALL gene-rearranged ALLs. Follow up on the patients ranges from 5 to 7 years, and none of the patients have relapsed. Identification of the rearrangements was facilitated using dual-colour breakapart probes for the ZNF384, TCF3, and EWSR1 loci. Four of the patients had TCF3-ZNF384 gene rearrangement, and four had EWSR1-ZNF384 gene rearrangement. Two patients had novel ZNF384 gene translocations involving regions on chromosomes 6 and 22, identified as the ARID1B and EP300 genes respectively by RNA sequencing analysis. Similar to the previously identified ZNF384 partners, these novel partner genes, which function in chromatin remodeling, contribute 5' gene sequences to the ZNF384 gene fusion. Based on the total number of ALL patients seen at our centre during the study period, our data suggests that, collectively, ZNF384 gene rearrangement may have an incidence of ~3% in pediatric pre-B ALL, with an incidence of at least 18% in CD10-negative pre-B ALL. ZNF384 gene rearrangement may be associated with a more favorable prognosis than ALL gene rearrangement. Both good and poor outcomes have been proposed in the literature, however, the majority of data are based on TAF15-ZNF384 rearrangement in adult patients. The ZNF384 gene rearrangements described above are cryptic and not easily identified by G-banding. Inclusion of ZNF384 FISH analysis in patients with CD10-negative ALL, after ALL gene rearrangement has been excluded, will assist with the determination of the prognostic significance of ZNF384 gene rearrangement in pediatric ALL.

Mantle cell lymphoma (MCL) comprises 3-10% of all NHL. The t(11;14)(q13;32) is present in virtually all cases of MCL, independent of their morphologic or immunophenotypic presentation. We describe an 80-year-old female patient with splenomegaly, thrombocytopenia, elevated absolute lymphocyte count, and anemia. A bone marrow biopsy demonstrated a hypercellular marrow for age showing 50% cellularity with 30-40% of the total cells being mature immunoblasts. Flow cytometry showing co-expression of CD20 and BCL-1 in the aggregates of small B-cells. Flow cytometry of the bone marrow aspirate revealed a monotypic, kappa-restricted B-cell population representing 37% of total cells, positive for CD5, CD19, CD20, CD22, Flt3-7, and CD38; and negative for CD23 and CD10. These results were consistent with a diagnosis of mantle cell lymphoma. Conventional cytogenetics performed on the bone marrow showed a complex abnormal female karyotype with numerical and structural abnormalities but no evidence of a (11;14) translocation in 6/20 metaphase cells examined. These results were described as: 44,XX,del(6)(q25),add(8)(p21),add(10)(p15),add(11)(q13);add(11)(q13),-14, +mar(cp6)/46,XX[14]. However, FISH on interphase nuclei using the CCND1-IGH@ (CCND1-IGH@) dual-color, dual-fusion probe from Abbott showed evidence of CCND1-IGH@ fusion signals in 23% of the interphase cells studied. These findings were described as: nuc ish (CCND1,IGH@) x3(CCND1 conj IGH@ x2)/46,XX[200]. FISH on previously G-banded metaphases showed a CCND1-IGH@ fusion signal on the derivate chromosome 11 which was consistent with a balanced t(11;14). Cytogenetic and another fusion signal on one of the abnormal copies of chromosome 11. Therefore the karyotype was reinterpreted as: 44,XX,del(6)(q25),add(8)(p21),der(10)(11q25->11q23.3;4q23->4p15->10pter),add(11)(q13),+mar(cp6)/46,XX[14]. This particular pattern with a complex karyotype with three or more numerical and structural changes along with 6q-, t(11;14) and a marker chromosome is common findings in mantle cell lymphoma of the blastoid subtype and are usually associated with an unfavorable prognosis. This case also highlights the importance of FISH on previously G-banded metaphases to elucidate cryptic rearrangements.

Identification of Semi-Cryptic and Variant Translocation Partners of RUNX1 gene in Acute Myeloid Leukemia (AML). A. Yenamandra1, F.C. Wheeler2, A. Hollis3, D. Zalepa1, M. Kapp1, N.C. Richardson4, S.C. Borsenstein5. 1) Department of Pathology, Immunology and Microbiology, Vanderbilt University Medical Center, Nashville, TN; 2) Department of Pathology, Immunology and Microbiology, Vanderbilt University Medical Center, Nashville, TN.

Rearrangements of RUNX1 locus at 21q22 include translocation of t(8;21), t(12;21), t(1;21), t(1;21), t(3;21), i(3;21), i(5;21), t(17;21) and t(16;21) in myeloid and lymphoid malignancy. We report here three cases of AML with variant translocation partners of RUNX1: a pediatric AML with t(7;21), and two adult secondary AML (t(10;21) and t(16;21)) respectively. Semi-quantitative metaphase FISH and Cytogenetics were performed in all cases to identify the partner chromosome in the RUNX1 rearrangements. Case #1. A 9 year old male was referred for fever, fatigue, and pancytopenia. Bone marrow analysis using semi-quantitative cytogenetics with a mixture of myeloblasts and immature monocytic elements in 90% of cells as determined by both morphology and flow cytometry. Molecular analysis for mutations in FLT3, NPM1 and C-KIT were negative, although FISH revealed rearrangement of RUNX1 locus in the chromosome 7p22 region. Karyotype analysis confirmed the diagnosis, revealing a unique translocation: 46,XY,t(7;21)(p22;q22)[22]/46,XX[20]. Three adults and one child have been reported in the literature with AML or high grade MDS with t(7;21) involving RUNX1-ubiquitin-specific protease gene (USP42)-fusion, however, to date no consistent prognostic information has emerged from these cases. Our patient had induction chemotherapy and is currently in remission, awaiting stem cell transplant. Case #2. A 58 year old male was diagnosed with treatment related AML following urorheal cell carcinoma. In this case, cytogenetics was normal (46,XY) at diagnosis, but follow up bone marrow analysis four and eleven months later revealed progressive abnormalities (46,XY,del(9)(q13q22)[2])/46,XY[18] and t(10;21)(p13;q22)[12]/46,XY[9], respectively). The latter BMA was FLT3+ also demonstrated RUNX1 rearrangement. The gene at the 10p13 region is unknown. Case #3. A 73 year old female was diagnosed with a myeloproliferative disorder. Karyotype revealed 46,XX,t(16;21)(q24;q22)[16]/46,XX[4] with FLT3 positivity. FISH analysis revealed rearrangement of RUNX1 with MTG16 gene locus at 16q24. The majority of the cases with t(16;21) (MTG16; RUNX1) translocation could not be reported to occur in the context of MDS/ANLL and the prognosis is poor. We present here three cases with semi-cryptic RUNX1 translocation in pediatric primary and adult secondary AML with unique RUNX1 rearrangement indicating the importance of the regulatory role of RUNX1 gene in hematopoiesis.
**3347S**

RAS driver mutations are present in 36% of acute lymphoblastic leukemia cases in children with Down syndrome and are mutually exclusive with JAK2 mutations. D. Nizetic,2,3 O. Nikolaev,4 M. Garieri,5 F. Santon,5 E. Falconnet,5 P. Ribaux,6 M. Guipponi,6 A. Murray,6 J. Groel,6 E. Gianni,6 G. Basso,6 S.E. Antonarakis,6,7 1) The Blizard Institute, Barts and The London School of Medicine, Queen Mary Univ, London, Switzerland; 2) London Structural Genomics Centre (LSG), The Wellcome Trust, United Kingdom; 3) Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. Unit 04-11, Proteos Building, 61 Biopolis Drive, 138673 Singapore; 4) Department of Genetic Medicine and Development, University of Geneva Medical School. 1 rue Michel Servet, 1211 Geneva 4, Switzerland; 5) Geneva University Hospitals - HUG. Service of Genetic Medicine. 4 Rue Gabrielle Perret-Gentil, 1211 Geneva 4, Switzerland; 6) Dipartimento di Salute della Donna e del Bambino, University of Padua, IRP - Istituto di Ricerca Pediatrica - Fondazione Città della Speranza, Padua, Italy; 7) IGE3 Institute of Genetics and Genomics of Geneva. 1 rue Michel Servet, 1211 Geneva 4, Switzerland.

Children with Down syndrome (DS) and acute lymphoblastic leukemia (ALL) have poorer survival and more relapses, than non-DS children with ALL, highlighting an urgent need for deeper mechanistic understanding of DS-ALL. Using full exome or cancer genes-targeted sequencing of 42 ALL samples from 39 DS patients, we uncover driver mutations in RAS (KRAS and NRAS) recurring to a similar extent (15/42 as JAK2 (12/42) mutations or P2RY8-CRLF2 fusions (14/42)). RAS mutations were almost completely mutually exclusive with JAK2 mutations (p=0.016), driving a combined total of two thirds of analysed cases. Clonal architecture analysis revealed that both RAS and JAK2 drove sub-clonal expansions primarily initiated by CRLF2 rearrangements, and/or mutations in chromatin remodellers and lymphocyte differentiation factors. Remarkably, in 2/3 relapsed cases there was a switch from a primary JAK2 or PTPN11 mutated sub-clone to a RAS-mutated sub-clone in relapse. These results provide important new insights informing the patient stratification strategies for targeted therapeutic approaches for DS-ALL.

**3348M**

Integrated analysis of transcriptome and exome in cancer samples improves interpretation and reveals additional therapeutic insights. S.M. Boyle1, M.J. Clark1, E. Helman1, D.M. Church3, S. Lou1, S. Kirk1, P. Strapkadeevong1, M. Karbelashvili3, M. Pratt1, M. Snyder1, R. Altman1, R. Chen1, J. West1, 1) Personalis, Menlo Park, CA; 2) Department of Genetics, Stanford University, Stanford, CA.

Next-generation sequencing is increasingly applied as a method for cancer analysis in both research and clinical settings. Many research studies used DNA or transcriptome sequencing to assess the genetic basis of tumor progression and, on some occasions, guide therapeutic decision-making. While gene panel, exome and whole genome sequencing are widely used for clinical guidance and translational research, transcriptome sequencing has yet to be widely adopted in the clinical environment. However, analyzing the transcriptome allows for improved detection of expression changes, gene fusions, and alternative splicing events. To assess these unique features of RNA, we performed whole transcriptome sequencing along with ACE exome sequencing of cancer samples. Our sample set included commonly used cell lines, cell lines modified to contain specific mutations, patient-derived xenografts (PDx), and primary tumors with genetic alterations and limits of detection through analysis of gene expression levels, gene fusion events, and RNA-derived variant calls. These calls were then filtered and annotated with our cancer gene database. There are a growing number of therapeutically relevant gene fusions, which are critically important for accurate tumor mutational analysis. We identified important gene fusion events at the expression level and cross-referenced these with exome findings, where present. Many fusion events would have been missed by exome or gene panel analysis alone. Additionally, transcriptome sequencing improves upon assessment of the impact of CNVs on expression. We performed gene expression analyses to accurately quantify differential regulation, something that can only be assumed when analyzing CNVs from DNA alone. By combining expression and DNA-based CNV analyses, we were able to generate more accurate CNV interpretations. A third area transcriptome gives us additional information on is the expression of mutant alleles. To that end, we called variants directly from the RNA, and identified variants that are truly expressed in the tumor. We observed variants in important genes such as those involved in cellular signaling and cell adhesion functions. Taken together, these joint approaches demonstrate how dual interpretation substantially bolsters accuracy and leads to changes in both research results and clinical decisions.

**3349T**


Somatic genome variation is a key underlying the pronounced cellular and clinical heterogeneity observed across a wide range of cancers. The ability to detect and accurately quantify this variation in tumors, particularly as it evolves over time, may be useful for predicting the effectiveness of therapy and monitoring subsequent tumor burden, the emergence of pharmacoresistance, the occurrence of metastasis and, ultimately, project- ing a patient’s prognosis and survival. If functionally relevant variants can be detected at a low molecular frequency during monitoring, the time and options available for clinical intervention would be greater. For a tumor genome analysis method to be effective for this application, but also practical for widespread implementation, it must be sensitive, accurate, reproducible, fast, low-cost, uncomplicated, amenable to automation and have robust data analysis and reporting outputs. For primary tumor analysis, and for utility in noninvasive screening or monitoring applications, efficacy with low sample input is also an essential element. Targeted next-generation sequencing (NGS) approaches are a good fit for many of these requirements. We developed a NGS method based on optimized, multiplexed, molecular inversion probes (MIPs) for the identification of low frequency somatic variation in cancer. A panel of ~700 MIPs was synthesized to interrogate mutational hot spots within a set of 55 cancer genes. The probes targeted both DNA strands and were designed to facilitate biobiformatic error correction. They also incorporated unique molecule identifiers (UIDs) to facilitate tagging of PCR duplicates and the accurate assessment of molecular complexity free of amplification bias. A performance evaluation was executed using a series of normal and cancer mutation reference samples. Analysis of experimental results using ≤10ng of input sample DNA and <400pf of raw sequencing (paired-end Illuma MiSeq) indicated that >98% of the target region was covered to at least 50x coverage depth, with a duplicate read rate less than 20%, and with >96% of the probes exhibiting ≥20% of the mean panel coverage. Initial evaluation of sensitivity and specificity in demonstrating the presence of mutations in the samples could be reliably detected when present at frequencies down to below 1%, with low false-positive rates. We conclude that optimized MIP panels are a promising solution for important applications in cancer genomics.

**3350S**

Exome sequencing identifies novel cancer-predisposing genes in familial thyroid cancer. A. Chaudhuri1, R. Gupta1, T. Boben1, P. Kumar2, R. Patnaik3, S. Santhosh1, S.E. Antonarakis1, M. Guipponi1, A. Murray1, M. Pratt1, R. Altman1, R. Chen1, J. West1, 1) Personalis, Menlo Park, CA; 2) Kerila Institute of Medical Sciences, Kerala, India.

Thyroid cancer is the most common endocrine malignancy accounting for ~1.7% of all cancer incidences in the world. In India, incidence of thyroid cancer ranges from 2-3% of all cancer incidences. However, certain regions in India - in particular, the state of Kerala, have a significantly higher incidence (3-4-fold higher) of thyroid cancer than the rest of the country. There is no clear mechanistic explanation for this observation. We have applied next generation sequencing (NGS) to identify novel cancer predisposing genes in this population. To this end, we have sequenced exomes of four affected members (3 females + 1 male) of a family diagnosed with thyroid cancer, along with a normal female member from the same family. Three of the four affected were diagnosed with papillary thyroid carcinoma (PTC) and one with follicular thyroid carcinoma (FTC) leading to the identification of variants specific to each cancer type. Our analysis revealed 209,135 variants in genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that genes, which included 36,084 coding region variants. By selecting variants with minor allele frequency of ≤5% from 1000genome, and those that
As a rare malignant tumor chordomas arises from notochordal remnants and frequently occurs in skull base of patients. miRNA dysregulation has been observed in chordomas when comparing whole genome array data in chordoma tissues with skeletal muscle, notochord, or nucleus pulposus. To give a more comprehensive profiling of miRNA in chordomas, we sequenced both miRNA and mRNA in two chordoma tissues and two fetal notochords on HiSeq 2000. In both notochords and chordomas, there are over 600 miRNA expressed. The top five most abundantly expressed miRNAs accounted for more than 50% of total miRNAs in the whole expression library of chordomas or notochords, and among these top miRNAs, two tissues had 3 in common. A total of 45 significantly differentially expressed miRNAs were identified in chordomas, including 16 up-regulated and 29 down-regulated with changed transcriptional activities of their targeted mRNAs accordingly. The functions of these genes were enriched in retinoic acid receptor activation, DNA methylation, transforming growth factor beta (TGFβ) pathway, etc. The chordoma up-regulated miR-29b-3p and its down-regulated target gene TGFβ3 were further validated in additional 8 chordoma tissues and 8 notochords. Compared to notochords, miR29b-3p was over expressed in all chordoma tissues together with TGFβ3 down-expressed as revealed by qPCR. As shown by analysis of TGFβ pathway in KEGG, we further found the genes involved in apoptosis were dysregulated in chordoma in comparison with notochords, suggesting the missing regulation of TGFβ induced cell death had important contribution to chordoma. To identify some variations in DNA level that cause the dysregulation of miR-29b-3p and TGFβ3, we further compared the copy number of their corresponding genes between chordoma tissues and their paired blood samples. Among the 8 chordoma patients, 5 had somatic copy number gain at miR-29b1 and miR-29b2 loci, while 4 had copy number loss in TGFβ3 locus, suggesting the dysregulation of TGFβ pathway may play a critical role in the formation of chordoma.
iCAGES: integrated CAncer GEnome Score for understanding personal cancer genomes. C. Dong1,2, H. Yang1, X. Lu1,2, K. Wang1,2,3, 1) Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA; 2) Biostatistics Division, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA; 3) Department of Psychiatry, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; 4) Human Genetics Center, The University of Texas Health Science Center at Houston, Houston, Texas, 77030, USA; 5) Division of Epidemiology, The University of Texas Health Science Center at Houston, Houston, Texas, 77030, USA. Cancer is a disease caused by accumulation of somatic mutations. Integration of personal genomic information with prior biological knowledge allows better understanding of cancer on a personal and systematic level. Here, we propose iCAGES (integrated CAncer GEnome Score), a statistical model that integrates multiple sources of information, from personal genomic information, to prior knowledge of its genes and networks and finally to drug-gene interactions. For each patient, iCAGES takes as input the genomic somatic mutations and calculates a prioritized list of cancer driver genes and candidate drugs, using a three-layer estimation procedure. The first layer summarizes personal genomic information using radial SVM (Support Vector Machine) algorithm based on the observation of a novel non-linear relationship between evolutionary conservation and function alterations of potential cancer driver mutations retrieved from COSMIC database, modeled with nine different deleteriousness prediction scores (PolyPhen-2, SIFT, MutationAssessor, FATHMM, LRT, GERP+, Siph and PhyloP, retrieved from ANNOVAR website and dbNSFP database) as its features. The second layer is based on previous knowledge on the association of each gene with cancer, modeled using Phenolyzer. Natural language processing schema was used for summarizing related biological research publications compiled from various databases, such as OMIM, ClinVar and GWAS catalog. Overall conditional probability for each gene on cancer was calculated and then integrated with protein-protein interaction and other network information from sources, such as HPRD protein interaction, NCBI’s Biosystem and HGNC gene family, to adjust for known protein-protein interactions. The third layer is based on gene-drug interaction, currently modeled using machine learning algorithms from machine learning algorithms. For each patient, iCAGES takes as input the genomic somatic mutations and calculates a prioritized list of cancer driver genes and candidate drugs, using a list of nsSNP as input, consistent with the original publication. In summary, we developed a statistical model that leverages personal genomic information with prior biological knowledge, shedding light into cancer driver genes identification, personalized drug discovery and cancer treatment.

The Activation of LINE-1 Retrotransposition in Barrett’s Esophagus and Esophageal Carcinoma. T.T. Doucet1,2, N. Rodic2, J. Jungbin Choi, C. Dong1,2. 1) Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD; 2) Pre-doctoral training program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA; 3) Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America; 4) Division of Gastroenterology and Hepatology, Department of Medicine and Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America.

One cause of genomic instability is the insertion of human retrotransposons that copy and paste themselves into new sites. We investigated how retrotransposition is involved in genomic instability in esophageal carcinogenesis. Using the L1-seq technique, we enriched for young active LINE-1 (L1) insertions and utilized next generation sequencing coupled with our analysis pipeline to isolate L1 insertions unique to the tumors of twenty-one esophageal adenocarcinoma (EAC) cases and squamous cell carcinoma (SCC) patients, divided into three cohorts. Our first cohort contained 1 EAC case and 4 SCC cases; the second cohort had eleven EAC cases. The third cohort consisted of five cases of EAC with matched DNA from the precursor lesion, Barrett’s esophagus (BE). Between the three cohorts, we confirmed 86 somatic insertions many of which occurred in introns of genes known to be involved in carcinogenesis. To investigate whether these confirmed insertions were clonal, we validated a subset of them in BE as well as tumor DNA. After validation, eight insertions were present only in BE, while two occurred in both BE and tumor from the same individual. In a separate cohort of squamous cell carcinoma cases for which we had tissue, at least two of the 42 high-stringency insertions were present in multiple regions of each tumor, suggesting that these insertions were clonal and occurred early in tumorigenesis. By immunohistochemistry, we also observed high levels of ORF1p, one of two L1 proteins, in 6 of 9 patient samples evaluated. In the individuals in whom we observed ORF1p expression, protein concentration increased linearly with progression from dysplasia to frank cancer. We are also investigating germline insertions within introns of genes involved in carcinogenesis of the esophagus to determine whether or not certain insertions predispose individuals to develop esophageal carcinoma. 174 germline insertions, including the genes PAD6, AKT3, and THADA (not polymorphic in the general population), were recurrent in multiple cohorts of patients. Furthermore, there was one case with an insertion predicted to add an ORF to the LINE-1 gene in two cohorts. We are currently elucidating effects of these insertions on gene expression. In light of these data, we conclude that the process of retrotransposition is a potential cancer driver in at least a subset of EAC patients. Using multiple approaches, we now seek to understand how retrotransposition contributes to genomic instability in tumorigenesis.

Inflammatory bowel diseases (IBDs) are chronic inflammatory disorders in the gastrointestinal tract, which arise as a result of the interaction of environmental and genetic factors. The risk of developing colorectal cancer (CRC) is increased for a person with IBD and correlates with the degree of colonic inflammation and disease duration. Patients with colitis-associated cancer (CAC) are younger than patients with sporadic colorectal cancer and show a poorer survival in the advanced stage. Both cancer types are multistep processes, but differ in certain mutation types. Compared to the more frequent sporadic CRC, the genomic signatures of CAC are far less understood.

In order to investigate the genetic background of colitis-associated carcinoma we chose a well-established inflammation-associated colorectal carcinogenesis model in mice (AOM/DSS) and compared the results from early and late stage cancer with inflamed tissue from DSS-treated and with healthy samples from untreated mice. Three experimental settings with different DSS concentrations and treatment durations were investigated. After tissue collection, the whole exome as well as the whole transcriptome of tumor and all proximal colon non-tumor samples were sequenced on the Illumina HiSeq 2000, followed by bioinformatic analyses. Beside known CAC-associated single nucleotide variants (SNVs) in e.g. Ctnnb1, we were able to identify a multitude of novel potentially damaging mutations. Furthermore, we observed regional clustering of somatic SNVs in several tumor samples and could show that the cancer stage is associated with the mutation rate and the mutation pattern. Additionally, we found a tumor specific gene expression profile. This also included the investigation of novel transcriptionally active regions, fusion transcripts and splice variants. Based on exome and transcriptome data, relevant genes and pathways were filtered out using different methods.

We performed one of the first studies analyzing the genetic background of CAC with next-generation sequencing technologies and demonstrated that a combined application of whole exome and strand-specific whole transcriptome sequencing can reveal new insights into the development of CAC.

The mutational landscape of peritoneal malignant mesothelioma. O. Harlinsmendez, H. Alakus, S.E. Yost, C.-P. Cheng, D.V. Jagiush, E.S. Moses, R.P. French, B. Woo, K. Jepsen, A. Frazer, A.K. Lowy, L. Miotke, S. Greer, L. Zadeh, J. Bell, S. Grimes, H. Ji, 1) UC San Diego Moores Cancer Ctr, La Jolla, CA; 2) UCSD Department of Surgery, La Jolla CA; 3) UCSD Institute for Genomic Medicine, La Jolla CA; 4) UCSD Clinical and Translational Science Institute, La Jolla CA; 5) Department of General, Visceral and Cancer Surgery, University of Cologne, Cologne, Germany; 6) UCSD Bioinformatics Graduate Program, La Jolla CA; 8) National Cheng Kung University, Tainan, Taiwan.

Malignant mesothelioma (MM) arises from mesothelial cells that line the pleural, peritoneal and pericardial surfaces. The majority of MM are pleural and have been associated with asbestos exposure. Previously, pleural MM have been genetically characterized by the loss of BAP1, either somatic or inherited, as well as loss of NF2 and CDKN2A. The rare peritoneal form of MM (PeMM) occurs in ~10% cases. With only ~300 cases diagnosed in the US per year, its link to asbestos exposure is not clear and its mutational landscape is unknown. We analyzed the somatic mutational landscape of epithelial PeMM using whole exome sequencing (N=7) and copy number analysis (N=5). In total, we identified 87 non-silent mutations in 83 genes with a median of 8 mutated genes per tumor, therefore resulting in a relatively low mutation rate (~1.6 to 6). BAP1 is the most recurrently mutated gene in PeMM, affecting 6/9 samples. BAP1 deletions occurred as focal events in 4/9 cases, in the context of a loss of chromosome 3p21 and are frequently accompanied by somatic mutations of the remaining allele. In one additional case, the loss of the entire chromosome 3 (encoding for BAP1) leaves a non-functional copy of BAP1 carrying a rare nonsense inherited variant, thus suggesting a potential genetic predisposition in this patient. The PeMM mutational landscape, dominated by the loss of chromosome 3p21, is reminiscent of the one observed in clear-cell renal carcinoma. In contrast to pleural MM, we do not observe deletions of NF2 and CDKN2A in PeMM, which suggests genetic differences between the two diseases. These findings support the use of molecularly guided clinical trials and therapies in PeMM.


Fallopian tube (FT) carcinoma is a rare gynecological malignancy, whose incidence is most likely underestimated due to misdiagnoses as ovarian cancer. Recent advances in pathology with more thorough serial sectioning techniques have helped clinicians to better recognize FT cancer. We undertook to define germline and somatic mutations in 98 FT carcinoma cases enrolled at diagnosis in IRB-approved studies at the University of Washington, excluding cases with carcinoma identified at the time of risk-reducing surgery. We sequenced DNA from blood and neoplastic tissue using BROCA, a targeted capture and massively parallel genomic sequencing approach that detects all classes of mutation in 64 genes. The fraction of all primary ovarian, peritoneal and FT carcinoma defined as a FT primary has increased over time: FT diagnoses and has increased from <1% to >2% in the previous 10 years (p=0.002). Of 98 subjects with FT carcinoma, 32 (33%) had loss of function germline mutations in 9 genes. Germline mutations were identified in BRCA1 in 18 (19%), and in BRCA2 in 4 (4%) patients. Germline mutations were also identified in other known or candidate breast and/or ovarian cancer susceptibility genes including 2 in BRI1 (2%), 2 in CHEK2 (2%), and each in BARD1, BLM, MRE11A, and ATM (1%). One mutation was found in FANCL. Mosaic mutations were found in TP53 (1%) and PPM1D (1%). Together, germline and mosaic mutations were identified in 34 (35%) cases. Neoplastic DNA from 30 FT advanced stage carcinoma sequenced with BROCA revealed 27 somatic mutations in TP53 (90%), 1 in PTEN (3%), and 1 in CHEK2 (3%). The germline mutation rate was higher in FT carcinoma compared to our previous studies in ovarian carcinoma (35% vs. 22%), and the profile of mutant genes was similar. Recently identified ovarian cancer susceptibility genes such as BRIP1, BARD1, and PPM1D are also associated with hereditary fallopian tube carcinoma. Known breast cancer genes (i.e. CHEK2, ATM) and other cancer susceptibility genes (BLM, MRE11A) have not been associated with FT or ovarian cancer. Increased recognition by pathologists of the FT as a primary site of disease is changing the relative distribution of FT, ovarian, and peritoneal carcinoma as primary diagnoses, making the profiling of FT cancers increasingly important.
3362S A rare somatic mutation in the TEL peptide of telomere protein TPP1 acts as a driver of childhood acute lymphoblastic leukemia. J. Healy1, J.F. Spinella1, P. Cassart1, N. Garnier1, C. Drulion1, R. Vidal1, V. Salliot1, C. Richer1, M. Ouimet1, S. Buscher2, B. Ge3, T. Pastinen3, D. Sinnett1,2,4 1) Division of Hematology-Oncology, Sainte-Justine UHC Research Center, Université de Montréal, Montréal, QC, Canada; 2) Department of Human Genetics, McGill University, Montreal QC, Canada; 3) McGill University and Genome Quebec Innovation Centre, Montréal, QC, Canada; 4) Department of Pediatrics, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada.

Childhood acute lymphoblastic leukemia is a heterogeneous disease both biologically and clinically, and is the leading cause of cancer-related deaths among children. Thorough investigation of pediatric ALL genomes, and identification of not only recurrent but also rare/private somatic driver mutations, is crucial in order to understand the underlying genomic complexity of this disease and thus better diagnose and treat it. Here we performed deep re-sequencing of a hyperdiploid pre-B childhood ALL patient and used a strict filtering strategy to identify somatic events with high driver potential. Among the candidate mutations, we identified a rare non-synonymous mutation (p.G223V) in the oligonucleotide-oligosaccharide-binding(OB)-folds of the TEL peptide of telomere protein TPP1/ACD, member of the shelterin complex that protects chromosome ends and regulates telomerase activity. Using in vitro cytotoxicity assays in ALL cells, we demonstrated a strong protective effect of this mutation against apoptosis associated with telomere maintenance. Thus, for the first time we identified a somatic mutation in TPP1 as a novel driver of childhood ALL and highlighting the importance of rare/private somatic mutations in understanding disease etiology.

3363M Mutational spectrum of RET Proto-oncogene in Iranian Patients with Medullary Thyroid Carcinoma. M. Hedayati1, M. Zarif Yeganeh1, S. Sheikholeslami2, A. Azzizz. 1) Cellular and Molecular Research Center, Research Institute for Endocrine Science, Shahid Beheshti University of Medical Sciences, Tehran, Iran; 2) Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Introduction: Thyroid cancer is the most common endocrine malignancy. It divides into four groups, including papillary (80%), follicular (10-15%), medullary (5-10%), and anaplastic thyroid cancer. Medullary Thyroid Carcinoma (MTC) is one of the most aggressive thyroid cancers which occur in both hereditary (25%) and sporadic (75%) forms. Mutations of the RET proto-oncogene in MTC development have been well demonstrated. The aim of the study was to investigate the mutational spectrum of exons 3, 5, 8, and 10-18 of RET proto-oncogene in MTC patients. Material and Methods: This retrospective study has been started since 2001 in Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, and Tehran, Iran. 370 participants, including 224 patients (168sMTC, 39sMTC, 8MEN2A, 3MEN2B, Sipple syndrome, and sporadic MTC), and 146 relatives (80 Female, 66 Male) were included in the study. Genomic DNA was extracted by the standard Salting Out/ProteinaseK method and mutation detection was performed through direct DNA sequencing. Sequence analysis was performed by Chromas Software version 2.3. Results: Totally, in 224 patients (124 Female, 100 Male) and 146 relatives (80 Female, 66 Male), 79 mutations were identified in RET main exons, including exons 10, 11, 13-16. Furthermore, 282 Single Nucleotide Polymorphism (SNP) were found in exons 3, 13, and 14. Interestingly, SNPs G691S and S904S were 100% in linkage disequilibrium in 125 patients and 70 relatives. The most common mutation in our population was C634Y (4%).

3364T Analysis of RNA-Seq Data Reveals Association of JAK-STAT Pathway with NK/T-Cell Lymphoma. JH. Hwang1, HY. Park1, SB. Lee2, JS. Choi1, YW. Ko1, WS. Kim1, SJ. Kim1,4, HY. Yoo2, JH. Kim1,2,3 1) Seoul National Univ, College of Medicine, Seoul, South Korea; 2) Genomic Medicine Institute, Medical Research Center, Seoul National University, Seoul, Republic of Korea; 3) Department of Pathology, Samsung Medical Center; 4) Department of Medicine, Samsung Medical Center; 5) Sungkyunkwan University School of Medicine, Seoul, Republic of Korea.

NK/T-cell lymphoma is a rare and aggressive subtype of non-Hodgkin lymphoma. Although it has been identified that lymphoma is associated with JAK-STAT pathway, there are few studies of NK/T-cell lymphoma using RNA-sequencing. We sequenced the transcriptomes of three tissues, seven cell lines of NK/T-cell lymphoma, and three normal NK/T cell lines to identify a diverse transcriptional spectrum such as gene fusions, differentially expressed genes, and transcription factors that affect differentially expressed genes. In this study, we discovered 24 in-frame fusion genes in two tissues and five cell lines. Among these, we found SND1-LEP fusion in a single cell line, and the LEP gene in this sample was expressed 20 times more than normal samples. It has been reported that leptin is involved in the tumorigenesis by activation of cellular signal transduction of JAK-STAT pathway. The genes showing abnormal expression patterns were also shown to be enriched in the JAK-STAT pathway (p-value: 1.9x10-4). Based on the expression data, we tested enrichment of predicted transcriptional factor binding sites for genes specifically activated in each sample. The binding sites of KLF4, MZF1, and SP1, which were known to be associated with JAK-STAT pathway and lymphoma, were enriched in both tissues and cell lines. Taken together, these results show the association between NK/T-cell lymphoma and JAK-STAT pathway again at the transcriptional level.
3366M Detection of mutation hotspots through mutation set enrichment analysis. P. Jia1,2, Q. Wang3, Q. Chen1,3, K.E. Hutchinson4, W. Pao4, Z. Zhao1,2,3,4 1) Biomedical Informatics, Vanderbilt University, Nashville, TN; 2) Center for Quantitative Sciences, Vanderbilt University, Nashville, TN; 3) Department of Biostatistics, Vanderbilt University, Nashville, TN; 4) Department of Cancer Biology, Vanderbilt University, Nashville, TN; 5) Department of Medicine/Division of Hematology-Oncology, Vanderbilt University, Nashville, TN; 6) Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN.

Although cancer genomes are extremely heterogeneous both within and among tumors, the identification of mutation hotspots is an integral part of many cancer genes. Many of these mutation hotspots disrupt functional domains or active sites, leading to gain- or loss-of-function of the protein product. We leveraged these observations to identify and prioritize candidate cancer genes and formulated a mutation set enrichment analysis (MSEA). MSEA assesses whether somatic mutations within a gene have a tendency to cluster within certain regions of the gene’s coding sequence. We proposed two complementary approaches, MSEA-clust and MSEA-domain. MSEA-clust simulates a walk through the sequences and renders a quantitative measurement of the location and extent to which mutations cluster. MSEA-clust is hypothesis-free because the convergent regions to be discovered are independent of a priori annotations of domains or functional sites. In contrast, MSEA-domain explicitly detects whether mutations occur in a pre-defined functional domain more frequently than random regions. MSEA-domain implements a regression-based test and can only assess a priori domains; thus, it is hypothesis-driven. We assessed the two MSEA approaches using somatic mutation data from a comprehensive public resource of cancer somatic mutations (COSMIC), and found that many known cancer genes have mutation hotspots. We then applied these approaches to available somatic mutation data in eight cancer types from The Cancer Genome Atlas (TCGA) project. With false discovery rate (FDR) being less than 0.2, both MSEA-clust and MSEA-domain were effective in detecting well-studied cancer genes and new candidates that have not been intensively studied in cancer. Collectively, these analyses demonstrate that MSEA methods are well suited to detect cancer genes based on mutation hotspots and provide valuable tools for future cancer gene studies.

3367T Advanced qualification and quantification of amplifiable genomic DNA (gDNA) for PCR-based targeted enrichment prior to next-generation sequencing. Q. Jiang, Q. Peng, J. DiCarlo, Y. Wang, R. Samara, V. Devgan, E. Lader. Biological Research Content, QIAGEN Sciences, LLC, Frederick, MD, USA.

Formalin-fixed paraffin-embedded (FFPE) tissue archives are an invaluable source for the molecular characterization of disease using next-generation sequencing (NGS). Unfortunately, heavy damage and fragmentation from the preservation process often results in FFPE DNA samples that cannot be amplified by traditional methods. Commonly used methods, such as fiber amplification and spectrometry, are inadequate in assessing the amount of PCR-amplifiable DNA present in a sample. To overcome these challenges, we describe the development of a novel qPCR-based method for quantifying and qualifying FFPE gDNA for PCR-based targeted enrichment. This method utilizes two qPCR assays to target multi-copy loci mapping to regions distributed across the genome. 329 assays in 85 distinct genomic regions were screened for qPCR performance. Assays with outstanding performance and consistency across all major human populations were selected. Plaque amplified FFPE samples were quantified by the PGMI® MiSeq®, or HiSeq®. All (high- and low-quality) FFPE samples quantified (and guided for input amount) by this strategy achieved good library yields, high specificity (>90%), and fairly even coverage on all sequencing platforms. Qualitatively, NGS results from these samples indicated that the calculated gDNA was in good agreement with predicted false positive rates. Quantitatively, the PGMI® MiSeq®, or HiSeq® platform provided limited usefulness in qualifying or quantifying these samples, which frequently failed to yield satisfactory NGS results. These samples produced varied library yields, mostly of low quantity, and low-sensitivity NGS reads. This approach is useful for sample QC prior to NGS to assess both the quality and quantity of gDNA samples. This is critical for identifying whether a particular gDNA sample is suitable for NGS analysis and determining the appropriate DNA input for successful targeted enrichment via multiplex PCR. The applications presented here are for research use only. Not for use in diagnostic procedures.

3368S Genetic Alterations and Evolutionary Behavior in Liver Metastatic Colorectal Cancer. Z. Lim1, J. Kim2,3. 1) Medical Genomics Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIIB); 2) Department of Medicine/Division of Gastroenterology, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea; 3) Institute of Innovative Cancer Research and Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea.

Colorectal cancer (CRC) develops through well-established pathogenesis steps accompanying with specific mutations. However, mutations associated with CRC liver metastases (LMs) and their evolutionary behaviors are not fully understood. To investigate somatic mutations found in CRC LMs, we performed exome- and RNA-sequencing using 101 tissues from 41 CRC patients, comprising 22 non-metastatic and 19 liver-metastasized patients. Global genomic patterns, including variant allele frequency (VAF), base substitution spectrum and chromosomal aberration, of liver-metastatic primary tumors (LMPTs) were distinct from those of non-metastatic primary tumors (NMPTs) as well as LMs. Particularly, PIK3CA mutations were more frequently detected in LMPTs than NMPTs and significantly associated with clinical characteristics of advanced cancer. We observed 4 categories of evolutionary behavior of mutations (from Class 1 to 4) based on dynamic patterns of VAFs from LMPTs to LMs. The Class 1, which probably has low fitness to metastasis, was significantly associated with clinical characteristics of advanced cancer and frequently co-occurred with chromosomal aberrations in LMs compared with Class 2, which probably has low fitness to metastasis. We also observed discrepant VAFs between exome and transcriptome. Importantly, VAFs of mutant alleles from TP53 and APC at the transcriptic level were significantly correlated with expression level of transcriptional target genes, revealing the functional implications of selective expression of mutant alleles. Especially, TP53 mutant alleles tended to be positively selected through chromosomal aberration and allele specific expression in LMs. The examination of mutations and evolutionary behaviors at the genomic and transcriptic level is required to apply precision medicine for liver-metastatic CRC patients.

3369M Whole exome sequencing reveals that DNA repair and apoptosis pathways are affected in hereditary breast cancer cases. J. L. D. Mentoor1,2, E. J. van Rijn1,2, J. L. D. Mentoor1,2, E. J. van Rijn1,2, J. L. D. Mentoor1,2, E. J. van Rijn1,2, J. L. D. Mentoor1,2, E. J. van Rijn1,2. 1) Department of Genetics, University of Pretoria, Pretoria, South Africa; 2) Bioinformatics and Computational Biology Unit, Department of Biochemistry.

Family history serves as the most significant risk factor for breast cancer (BC). Variants in the BRCA1&2 breast cancer susceptibility genes account for ~30-60% of hereditary BC. Moderate risk genes (e.g., RAD51C, PALB2 & ATM etc.) account for ~3-5% of familial cases. A considerable number of familial BC cases are not due to mutations in high/moderate susceptibility genes; instead, they are identified thus far i.e. BRCA1&2 negative. Mutant alleles in novel risk genes may account for the missing heritability. The aim of our research is to discover variants in novel high-penetrant genes that may increase susceptibility for BC. Nine cases (BRCA1&2 negative) were selected from 6 families with a history of breast and/or ovarian cancer (≥3 cases). Paired-end exome sequencing was completed by the Beijing Genomics Institute (Illumina). The Genome Analysis Toolkit 2.4.9 was used for variant calling and annotation. Pathway analysis of key gene expression in LMs. The examination of mutations and evolutionary behaviors are not fully understood. To investigate somatic mutations found in CRC LMs, we performed exome- and RNA-sequencing using 101 tissues from 41 CRC patients, comprising 22 non-metastatic and 19 liver-metastasized patients. Global genomic patterns, including variant allele frequency (VAF), base substitution spectrum and chromosomal aberration, of liver-metastatic primary tumors (LMPTs) were distinct from those of non-metastatic primary tumors (NMPTs) as well as LMs. Particularly, PIK3CA mutations were more frequently detected in LMPTs than NMPTs and significantly associated with clinical characteristics of advanced cancer. We observed 4 categories of evolutionary behavior of mutations (from Class 1 to 4) based on dynamic patterns of VAFs from LMPTs to LMs. The Class 1, which probably has low fitness to metastasis, was significantly associated with clinical characteristics of advanced cancer and frequently co-occurred with chromosomal aberrations in LMs compared with Class 2, which probably has low fitness to metastasis. We also observed discrepant VAFs between exome and transcriptome. Importantly, VAFs of mutant alleles from TP53 and APC at the transcriptic level were significantly correlated with expression level of transcriptional target genes, revealing the functional implications of selective expression of mutant alleles. Especially, TP53 mutant alleles tended to be positively selected through chromosomal aberration and allele specific expression in LMs. The examination of mutations and evolutionary behaviors at the genomic and transcriptic level is required to apply precision medicine for liver-metastatic CRC patients.

Whole exome sequencing reveals that DNA repair and apoptosis pathways are affected in hereditary breast cancer cases. J. L. D. Mentoor1,2, E. J. van Rijn1,2, J. L. D. Mentoor1,2, E. J. van Rijn1,2. 1) Department of Genetics, University of Pretoria, Pretoria, South Africa; 2) Bioinformatics and Computational Biology Unit, Department of Biochemistry.

Family history serves as the most significant risk factor for breast cancer (BC). Variants in the BRCA1&2 breast cancer susceptibility genes account for ~30-60% of hereditary BC. Moderate risk genes (e.g., RAD51C, PALB2 & ATM etc.) account for ~3-5% of familial cases. A considerable number of familial BC cases are not due to mutations in high/moderate susceptibility genes; instead, they are identified thus far i.e. BRCA1&2 negative. Mutant alleles in novel risk genes may account for the missing heritability. The aim of our research is to discover variants in novel high-penetrant genes that may increase susceptibility for BC. Nine cases (BRCA1&2 negative) were selected from 6 families with a history of breast and/or ovarian cancer (≥3 cases). Paired-end exome sequencing was completed by the Beijing Genomics Institute (Illumina). The Genome Analysis Toolkit 2.4.9 was used for variant calling and annotation. Pathway analysis of key gene expression in LMs. The examination of mutations and evolutionary behaviors are not fully understood. To investigate somatic mutations found in CRC LMs, we performed exome- and RNA-sequencing using 101 tissues from 41 CRC patients, comprising 22 non-metastatic and 19 liver-metastasized patients. Global genomic patterns, including variant allele frequency (VAF), base substitution spectrum and chromosomal aberration, of liver-metastatic primary tumors (LMPTs) were distinct from those of non-metastatic primary tumors (NMPTs) as well as LMs. Particularly, PIK3CA mutations were more frequently detected in LMPTs than NMPTs and significantly associated with clinical characteristics of advanced cancer. We observed 4 categories of evolutionary behavior of mutations (from Class 1 to 4) based on dynamic patterns of VAFs from LMPTs to LMs. The Class 1, which probably has low fitness to metastasis, was significantly associated with clinical characteristics of advanced cancer and frequently co-occurred with chromosomal aberrations in LMs compared with Class 2, which probably has low fitness to metastasis. We also observed discrepant VAFs between exome and transcriptome. Importantly, VAFs of mutant alleles from TP53 and APC at the transcriptic level were significantly correlated with expression level of transcriptional target genes, revealing the functional implications of selective expression of mutant alleles. Especially, TP53 mutant alleles tended to be positively selected through chromosomal aberration and allele specific expression in LMs. The examination of mutations and evolutionary behaviors at the genomic and transcriptic level is required to apply precision medicine for liver-metastatic CRC patients.

Next generation sequencing of CDNA prepared from total RNA (RNA-seq) has become the method of choice for transcript profiling, and discovery. The standard whole-transcriptome approach faces a significant challenge as the vast majority of reads map to ribosomal RNA (rRNA). One solution is to enrich the RNA sample for polyadenylated transcripts using oligo (dT)-based affinity matrices; however, this also eliminates other biologically relevant RNA species, such as noncoding RNAs, and relies on having a high quality and quantity RNA sample. Here, we present a method to eliminate abundant RNAs from total RNA based on hybridization of probes to the targeted abundant RNA, followed by subsequent enzymatic degradation. We applied this method to remove cytoplasmic and mitochondrial rRNA from different eukaryotic total RNA samples (human, mouse and rat). We optimized this method for RNA with different degradation levels, from intact RNA to highly degraded formalin-fixed paraffin-embedded (FFPE) samples.

The standard whole-transcriptome approach faces a significant challenge as the vast majority of reads map to ribosomal RNA (rRNA). One solution is to enrich the RNA sample for polyadenylated transcripts using oligo (dT)-based affinity matrices; however, this also eliminates other biologically relevant RNA species, such as noncoding RNAs, and relies on having a high quality and quantity RNA sample. Here, we present a method to eliminate abundant RNAs from total RNA based on hybridization of probes to the targeted abundant RNA, followed by subsequent enzymatic degradation. We applied this method to remove cytoplasmic and mitochondrial rRNA from different eukaryotic total RNA samples (human, mouse and rat). We optimized this method for RNA with different degradation levels, from intact RNA to highly degraded formalin-fixed paraffin-embedded (FFPE) samples. We applied this method to remove ribosomal RNA from FFPE RNA extracted from breast cancer biopsy with an archive age of 1 and 10 year-old. We evaluated the depletion efficiency and off target effect of this method using strand specific RNA high-throughput sequencing. Ribosomal RNA depletion resulted in a minimal percentage of total reads mapping to rRNA sequences (<1% total reads), regardless of the species, input amount (1μg or 100 nanograms), or degradation level (intact or FFPE RNA). One such sample was used for an independent scrutiny experiment of the method using a low input amount of FFPE RNA. We also observed good expression (FPKM) correlation (R>0.93) between the material and FFPE RNA depleted and non-depleted libraries. This indicates that the method is not altering the expression of other genes. This method offers a robust and simple solution for transcriptome analysis of a variety of samples, including low quality and low quantity clinical samples such as FFPE RNA. Moreover, it is amenable to high-throughput sample preparation and robotic automation. This method is sensitive, specific, and produces increased coverage of less abundant, non-targeted transcripts in RNA-Seq studies.

Shared driver genes of familial and sporadic pancreatic cancer may explain the similar age of onset. A.L. Norris, N.J. Roberts, R.H. Hubbard, A.P. Klein, J.R. Eshleman.

Familial pancreatic cancer (FPC) accounts for approximately 10% of pancreatic ductal adenocarcinoma (PDAC) cases. While an earlier age of onset is a hallmark of most familial cancer syndromes, FPC does not present at an obviously earlier age. In this study, we collated the age of onset in FPC and sporadic PDAC reported in the literature. FPC cohorts had a reported mean or median age of 52 to 69 years old at diagnosis, while the range for sporadic PDAC cohorts was 60-74 years old. Our own small cohort, the median age of onset was 66 years old for both the FPC and PDAC patients. From our analysis of the literature, we show that any difference in the age of onset in FPC and sporadic PDAC cohorts is likely a result of ascertainment bias. The molecular progression of sporadic PDAC is well documented and the high prevalence driver genes are KRAS (>90%), CDKN2A/p16 (95%), TP53 (50-75%), and SMAD4/DPC4 (55%). We determined the prevalence of alterations in these PDAC driver genes in our cohort of FPC cell lines using an integrated genomics approach of whole exome sequencing (WES), whole genome sequencing (WGS), RNA-Seq, and high density SNP microarrays. We found that our FPC cohort had high frequencies of alterations in the 4 PDAC driver genes. Activating KRAS mutations were identified in all (12/12) FPC cases, predominately at the codon 12 hotspot (11/12, 92%). Inactivation of CDKN2A/p16 was identified in all (12/12) FPC cases, predominately by homozygous deletion (9/12, 75%). TP53 mutations were identified in 92% (11/12), predominately by single base substitutions with loss of heterozygosity (LOH) (9/11, 82%). Inactivation of SMAD4/DPC4 was identified in 75% (9/12) of FPC cases, predominately by homozygous deletion (5/9, 56%). Our collation of the literature confirms this observation and has a similar age of onset as its sporadic counterpart, which may also shed light on the need for more standardized controls to evaluate and impart confidence to NGS testing.

The Hotspot Frequency Ladder (HFL) is a proprietary genomic/synthetic DNA blended material. This material contains over 400 commonly sequenced COSMIC mutations. All variants were confirmed by Sanger sequencing and the genomic DNA was characterized by the Genome in a Bottle Consortium. A material was developed to contain ~400 variants at six finely tuned frequencies, with one frequency level per tube. Variants were quantified by digital PCR, enabling the materials to be manufactured consistently at a range of frequencies including near the limit of detection for both NGS or PCR platforms. The frequencies constructed were 48%, 29%, 18%, 11%, 5%, and 3%, and samples were tested using the Ion AmpliSeq™ Cancer Hotspot Panel v2 (CHPv2) with the Ion Torrent Personal Genome Machine® and Illumina® TruSeq® Amplicon Cancer Panel and TruSight® Tumor Panel on the MiSeq® instrument. At each frequency, the variants observed on each platform were compared to the expected engineered mutations. The limit of detection was then determined for ~400 variants. At the 48% level, the Ion AmpliSeq™ and TruSeq™ tests detected greater than 99% of all SNPs (383/385 for CHP2 and 351/351 for TruSeq®). Detection dropped to 78% (302/385 for CHP2 and 279/351 for TruSeq®) at the 3% level. Indel detection was 79% (23/29) for CHP2 and 74% (25/34) for the TruSeq® test. At the 3% level, indel detection ranged from 28% (8/29) for CHP2 and 44% (15/34) for TruSeq® test, respectively. To determine whether the material performs similarly to tumor tissue samples, tumor samples were diluted with matched normal samples to mimic a range of frequencies. Linearity and limit of detection between the material and diluted tumor tissue samples were compared. Overall, highly multiplex controls with tunable frequencies allow for much more extensive, yet streamlined, assay evaluation and facilitate implementation and impart confidence to NGS testing. The Hotspot Frequency Ladder is under development.


Increasing adoption of next-generation sequencing (NGS) technology has shed light on the need for more standardized controls to evaluate and optimize system performance. However, samples containing mutations of interest are difficult to source and cell line pooling experiments to determine the limit of detection require significant investments of time and money. To simultaneously evaluate variant calling performance in ~200 unique amplicons across 4 different genes targeted by NGS tests, AcroMetrix® has developed a proprietary genomic/synthetic DNA blended material. This material contains over 400 commonly sequenced COSMIC mutations. All variants were confirmed by Sanger sequencing and the genomic DNA was characterized by the Genome in a Bottle Consortium. A material was developed to contain ~400 variants at six finely tuned frequencies, with one frequency level per tube. Variants were quantified by digital PCR, enabling the materials to be manufactured consistently at a range of frequencies including near the limit of detection for both NGS or PCR platforms. The frequencies constructed were 48%, 29%, 18%, 11%, 5%, and 3%, and samples were tested using the Ion AmpliSeq™ Cancer Hotspot Panel v2 (CHPv2) with the Ion Torrent Personal Genome Machine® and Illumina® TruSeq® Amplicon Cancer Panel and TruSight® Tumor Panel on the MiSeq® instrument. At each frequency, the variants observed on each platform were compared to the expected engineered mutations. The limit of detection was then determined for ~400 variants. At the 48% level, the Ion AmpliSeq™ and TruSeq™ tests detected greater than 99% of all SNPs (383/385 for CHP2 and 351/351 for TruSeq®). Detection dropped to 78% (302/385 for CHP2 and 279/351 for TruSeq®) at the 3% level. Indel detection was 79% (23/29) for CHP2 and 74% (25/34) for the TruSeq® test. At the 3% level, indel detection ranged from 28% (8/29) for CHP2 and 44% (15/34) for TruSeq® test, respectively. To determine whether the material performs similarly to tumor tissue samples, tumor samples were diluted with matched normal samples to mimic a range of frequencies. Linearity and limit of detection between the material and diluted tumor tissue samples were compared. Overall, highly multiplex controls with tunable frequencies allow for much more extensive, yet streamlined, assay evaluation and facilitate implementation and impart confidence to NGS testing. The Hotspot Frequency Ladder is under development.
3373T
1) Center for Human Genetics Research, Vanderbilt University School of Medicine, Nashville, TN; 2) Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN; 3) Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN; 4) Department of Cancer Biology, Massachusetts General Hospital, Boston, MA; 5) Department of Pathology Services, Massachusetts General Hospital, Boston, MA; 6) Department of Medicine-Service, Division of Hematology and Oncology, Massachusetts General Hospital, Boston, MA; 7) Department of Pathology, Massachusetts General Hospital, Boston, MA; 8) Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN; 9) Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN; 10) Department of Cancer Research, Vanderbilt University School of Medicine, Nashville, TN.

Whole exome sequencing (WES) and RNA sequencing (RNA-Seq) are two widely applied next-generation sequencing (NGS) technologies used for detecting somatic mutations in cancer. In this study, we used 27 matched tumor-normal samples that have both WES and RNA-Seq data to systematically compare single nucleotide variants (SNVs) called from WES and RNA-Seq of the same patients. WES reads were mapped to the human reference genome (hg19) using bwa. Post-processing of the initial mapping included steps to mark duplicate reads using Picard and perform local realignment using GATK. RNA-Seq reads were mapped using TopHat 2. We used MuTect to call SNVs for both WES and RNA-Seq, VarScan 2 to determine sequencing reads of SNVs, and Cufflinks to compute gene-based expression levels (FPKM) of RNA-Seq data. Interestingly, we found a low overlap rate (on average 14%) of SNVs called from WES and RNA-Seq. Among WES-unique SNVs, 38 - 73% were missed in RNA-Seq due to no coverage (<1 read) and 10 - 23% due to low coverage (<8 reads). Interestingly, 9 - 53% of WES-unique SNVs were highly covered with >8 reads in RNA-Seq but were still not called. We next explored WES-unique SNVs’ gene expression levels in RNA-Seq. As expected, 55% of WES-unique SNVs were expressed in normal tissues, 45% in highly expressed regions (FPKM >20) yet remained undetected. We further examined WES-unique SNVs with available cDNA information, and found on average 51% were located on the non-transcribed strand causing them to be undetected in RNA-Seq. We also explored exon skipping as a factor for WES-unique SNVs being missed in RNA-Seq, and preliminary work discovered at least 4 mutations which result in potentially skipped exons. Among RNA-Seq-unique SNVs, 22 - 89% were in positions not covered by the WES kit. For SNVs covered by the WES kit, 81 - 97% had a callable coverage (≥ 8 reads) in WES. These SNVs were not missed due to technical issues. We therefore analyzed their allele frequency and determined only 3% (0 - 12%) of the alternate alleles occurred with a frequency ≥29% in WES. We also analyzed the mutation patterns of all RNA-Seq-unique SNVs. 54% of them displayed a T:A → C:G pattern, which is a signature of potential adenomas which cause RNA-Seq calling using two main NGS platforms.

3374S
Patterns of somatic mutations in hepatitis B virus-associated hepatocellular carcinomas. Q. Pan1, W. Yin1, S. Tong1, D. Li2, X. Li3, D. Zhao4, M. Li5, H. Hui6, H. Ren7, K. Ding1
1) Key Laboratory of Molecular Biology for Infectious Diseases, Ministry of Education, Chongqing, China; 2) Department of Hepatobiliary Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, P. R. China; 3) Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China; 4) Department of Pathology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China; 5) Department of Radiotherapy, Chongqing Cancer Hospital, Chongqing, China; 6) Department of Gastroenterology, Sichuan Province, P. R. China.

Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. Most hepatocellular carcinomas develop in the background of the advanced liver fibrosis and cirrhosis through a stepwise accumulation of various genetic alterations from cirrhotic nodule to HCC. The patterns of somatic mutations in the liver cancer and its matched cirrhotic tissues were not fully characterized. We performed whole exome sequencing on liver cancer tissues, its matched cirrhotic liver tissues and peripheral blood leukocytes (PBL) in 13 patients of hepatitis B virus-associated HCC. Diagnoses of HCC cases and the matched liver cirrhosis were histologically confirmed by two pathologists. Only HCCs with the percentage of tumor cells more than 70% were used for the analysis. Whole exome sequencing resulted in approximately 106 million raw reads with an average coverage of 110× in each sample. A bioinformatics pipeline was used to analyze sequencing reads and identify somatic mutations. The prevalence of somatic mutations differed significantly among HCCs. Among 6,972 non-silent mutations (i.e., missense or nonsense mutations and frame shifts) identified in 13 HCC patients, we found that T:A→A:T transversion showed the highest percentage of non-silent mutations, followed by C:G→T:A transversion and C:G→A:T transversion. By comparison of somatic mutations identified in liver cancer tissues and matched cirrhotic tissues, somatic mutations were not accumulated in genes in HBV-associated cirrhotic liver tissues. Notably, two genes (TP53 and KRTAP4-3) were shown to be significantly mutated by MutSigCV analysis, suggesting that there were evidence for new mutations of these genes contributing to the development of liver cancer. Our study may provide an overview of patterns of somatic mutation in hepatitis B-virus associated HCC and its precancerous lesions (i.e., cirrhosis).

3375M
Highly sensitive, non-invasive detection of colorectal cancer mutations using single molecule, third generation sequencing. G. Russo1, A. Patrignani1, L. Poveda1, F. Hoehn2, B. Scholttka2, R. Schlabach1, A. Garvin2
1) Functional Genomics Center Zurich - ETHUZH, Zurich, Switzerland; 2) Droplet Diagnostics SAS, Mulhouse, France; 3) Department of Nutritional Toxicology, Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany.

Colorectal cancer (CRC) represents one of the most prevalent and lethal malignant neoplasms and every individual of age 50 and above should undergo regular CRC screening. Currently, the most effective procedure to detect adenomas, the precursors to CRC, is colonoscopy, which reduces CRC incidence by 80%. However, it is an invasive approach that is unpalatable for the patient, expensive, and poses some risk of complications such as colon perforation. A non-invasive screening approach with detection rates comparable to those of colonoscopy has not yet been established. The current study applies Pacific Biosciences third generation single molecule sequencing to the inspection of CRC-driving mutations. Our approach combines the screening power and the extremely high accuracy of circular consensus (CCS) third generation sequencing with the non-invasiveness of using stool DNA to detect CRC-associated mutations present at extremely low frequencies and establishes a foundation for a non-invasive, highly sensitive assay to screen the population for CRC and early stage adenomas. We performed a series of experiments using a pool of fifteen amplicons covering the genes most frequently mutated in CRC (APC, Beta Catelin, KRAS, BRAF, and TP53), ensuring a theoretical screening coverage of over 97% for both CRC and adenomas. The assay was able to detect mutations in DNA isolated from stool samples from patients diagnosed with CRC at frequencies below 0.5 % with no false positives. The mutations were then confirmed by exonic sequencing DNA isolated from the exonic tumor samples. Our assay should be sensitive enough to allow the early identification of adenomatous polyps using stool DNA as analyte. In conclusion, we have developed an assay to detect mutations in the genes associated with CRC and microsatellite instability using Pacific Biosciences RS Single Molecule Real Time Circular Consensus Sequencing (SMRT-CCS). With no systematic bias and a much higher raw base-calling quality (CCS) compared to other sequencing methods, the assay was able to detect mutations in stool DNA at frequencies below 0.5 % with no false positives. This level of sensitivity should be sufficient to allow the detection of most adenomatous polyps using stool DNA as analyte, a feature that would make our approach the first non-invasive assay with a sensitivity comparable to that of colonoscopy and a strong candidate for the non-invasive preventive CRC screening of the general population.
Integrating eQTLs from a range of normal human tissues with cancer genomics to help identify germline risk alleles in cancer driver genes.

A.V. Segre1, D.S. DeLuca1, T. Sullivan1, E. Gelfand1, L. Lappalainen2,3, M. Lawrence1, A. Kiezun1, D.G. MacArthur1,4, K. Ardlie1, G. Getz1,4, the GTEx consortium. 1) Broad Institute of Harvard and MIT, Cambridge, MA; 2) New York Genome Center, NY; 3) Department of Systems Biology, Columbia University, NY; 4) Massachusetts General Hospital, Boston, MA.

A comprehensive analysis of somatic point mutations in whole exome sequences from >4,700 human cancer and their matched normal tissue samples has increased the number of cancer driver genes to 260 for 21 different tumor types. The extent to which germline variants in the same cancer genes may predispose people to certain cancers is not yet known. Furthermore, the contribution of gene expression-altering variants to tumor development compared to changes in protein structure or function is not well understood. To address these questions and gain mechanistic insight into the potential role of germline variants in tumorgenesis, we are testing the hypothesis that increase in expression of oncogenes and/or decrease in expression of tumor suppressor genes in normal human tissues may increase risk of cancer development. To this end, we will annotate the cancer genes as potential oncogenes and/or tumor suppressor genes, based on large-scale experimental gene perturbation studies that measure the effect on cell proliferation, growth arrest or cell death. We will then examine which cancer genes are affected cis-eQTLs (expression quantitative trait loci) or allele specific expression (ASE), using the Genotype-Tissue Expression (GTEX) project with data in 40 normal human tissues and up to 450 unrelated individuals. Based on the pilot phase of GTEx with 178 individuals, 36% of the 260 somatic cancer genes, e.g. AKT1, CASP8 and CDKN2A have significant eQTLs and can be modeled as having a functional impact on the expression of genes, including skin, lung and nerve. Following characterization of cancer genes with eQTLs, we will test whether the germline genotypes of tumor samples in the pan-cancer cohort, such as melanoma, lung, cancer, and blood cancers, are enriched for eQTL or ASE alleles that increase oncogene expression or decrease tumor suppressor expression compared to a control panel. If so, this would suggest that the regulatory variants are cancer risk factors. Furthermore, we will integrate the results with somatic mutations from whole exomes and whole genome sequences in the pan-cancer cohort, employing heterozygosity (LOH) analyses, to evaluate how eQTLs acting on cancer genes may be contributing to tumor progression. This study may help propose new cancer germline risk variants and shed light on potential causal regulatory regions with significant eQTLs and ASE signals. In future phases of this study, we will test whether these findings can be extended to testing the regulatory role of somatic mutations during tumorgenesis.

Identification of mutations in oral cavity squamous cell carcinoma induced by betel quid chewing in Taiwan. Y. Yeh1, L. Wang1, J. Chen2,3, N. Limthong2, H. Chen4, D. Chen2,3, Y. Elshimali5, 1) La Sierra university, Riverside, CA; 2) Charles Drew Univ Med. Los Angeles, CA; 3) Dept of Path and Med David Geffen-School of Medicine at UCLA, Los Angeles, CA; 4) Pite Aid Pharmacy, Fontana, CA; 5) Rosemede Clinic, Rosemede, CA; 6) Materinal & Child Center, Taipei, Taiwan; 7) Dept of Path and Lab UCI Medical Center, Orange, CA.

Squamous cell carcinoma of the head and neck (HNSCC) is the sixth most common cancer worldwide, the most common site being the oral cavity. In Taiwan, squamous cell carcinoma of the oral cavity is the fourth most common type of cancer. Environmental carcinogens, including betel quid chewing and alcohol, are the major identifiable risk factors. The development of squamous cell carcinoma of the oral cavity arising from premalignant lesions, such as leukoplaikia, follows a multi-step process involving sequential activation and inactivation of oncogenes and tumor suppressor genes, resulting in a clonal population of cells. Chromosomal deletions of 3p and 9p21 are characterized as the first step of the tumorgenesis process. However, additional genomic abnormalities are well characterized. Such abnormalities can contribute to histological variations often observed in a tumor sample, such as invasiveness and, more specifically, the ability to penetrate normal tissue. Our findings thus suggest that tumor development compared to changes in protein structure or function is not yet known. Furthermore, the contribution of gene expression-altering variants to tumor development compared to changes in protein structure or function is not well understood. To address these questions and gain mechanistic insight into the potential role of germline variants in tumorgenesis, we are testing the hypothesis that increase in expression of oncogenes and/or decrease in expression of tumor suppressor genes in normal human tissues may increase risk of cancer development. To this end, we will annotate the cancer genes as potential oncogenes and/or tumor suppressor genes, based on large-scale experimental gene perturbation studies that measure the effect on cell proliferation, growth arrest or cell death. We will then examine which cancer genes are affected cis-eQTLs (expression quantitative trait loci) or allele specific expression (ASE), using the Genotype-Tissue Expression (GTEX) project with data in 40 normal human tissues and up to 450 unrelated individuals. Based on the pilot phase of GTEx with 178 individuals, 36% of the 260 somatic cancer genes, e.g. AKT1, CASP8 and CDKN2A have significant eQTLs and can be modeled as having a functional impact on the expression of genes, including skin, lung and nerve. Following characterization of cancer genes with eQTLs, we will test whether the germline genotypes of tumor samples in the pan-cancer cohort, such as melanoma, lung, cancer, and blood cancers, are enriched for eQTL or ASE alleles that increase oncogene expression or decrease tumor suppressor expression compared to a control panel. If so, this would suggest that the regulatory variants are cancer risk factors. Furthermore, we will integrate the results with somatic mutations from whole exomes and whole genome sequences in the pan-cancer cohort, employing heterozygosity (LOH) analyses, to evaluate how eQTLs acting on cancer genes may be contributing to tumor progression. This study may help propose new cancer germline risk variants and shed light on potential causal regulatory regions with significant eQTLs and ASE signals. In future phases of this study, we will test whether these findings can be extended to testing the regulatory role of somatic mutations during tumorgenesis.

Integrative analysis of regulatory aberrations in lung adenocarcinoma cell lines. A. Suzuki1, T. Kohn3, K. Tsuchihara1, Y. Suzuki1, 1) Department of Medical Genome Sciences, The University of Tokyo, Japan; 2) Division of TR, The Exploratory Oncology Research and Clinical Trial Center, National Cancer Center, Japan; 3) Division of Genome Biology, National Cancer Center Research Institute, Japan.

Intratumoral multi-omics analysis is a powerful approach to understand the biological relevance of the genomic, epigenomic and transcriptomic aberrations in cancer. In this study, we performed whole-genome sequencing, RNA-Seq, target-captured bisulfite sequencing and ChiP-Seq for eight different lung cancer cell lines. Using genome and whole exome sequencing data of clinical samples, the multi-omics analysis should be important to understand how variable genetic events driving cALL relapse, we hope to improve prognosis for refractory cALL and ultimately provide targeted personalized care for children diagnosed with ALL.

Changes in the fitness of somatic SNVs (rise or fall of allele frequencies) were assessed and clustering analysis was used to illustrate subclonal complexity and clonal evolution within the individual ALL tumours. Strikingly, our data suggest the presence of a fitter ancestral subclonal population, harbouring specific relapse-driving genes, that is present at diagnosis and that arises from 2-3% in an invader tumour cell and is resistant to therapy and was maintained at relapse. We queried putative relapse drivers against public pharmacogenomics databases (DGIdb, PharmGKB) and identified known drug-gene/pathway interactions which could lead to new avenues for improved targeted treatment of refractory cALL. By dissecting the clonal evolution of individual tumors and identifying the underlying genetic events driving cALL relapse, we hope to improve prognosis for refractory cALL and ultimately provide targeted personalized care for children diagnosed with ALL.
Molecular profiling in diagnosis and determining prognosis of "early" myelodysplastic syndrome. M. Thangavelu, E. Wei, M. Albitar, E. Mandell, B. Turkoglu, A. Ozer, C. Akyenti Boylu, M.C. Yakiciyes, A. Ozer, 1) Acibadem Genetic Diagnostic Center, ISTANBUL, Turkey; 2) Department of Medical Biology and Genetics, Marmara University, Istanbul, TURKEY; 3) Department of Medical Biology, Cumhuriyet University, Sivas, TURKEY; 4) Department of Dermatology, Acibadem Kadikoy Hospital, Istanbul, TURKEY; 5) Biochemistry and Molecular Biology, Department of Acibadem University, Istanbul, TURKEY.

Background: Myeloproliferative neoplasms (MPNs) are clonal myeloid cancers characterized by overproduction of mature blood cells. BCR/ABL1 and JAK2V617F are the most frequent mutations in MPNs. Other JAK2-exon12 and MPL-exon10 mutations were less common in suspected MPNs (2% and 3%-7%, respectively). Several other genes such as LNK, CBL, ASXL1, DNMT3A and TET2 are still being investigated.

Methods: 152 non-mutated BCR/ABL1 and JAK2 V617F patients were selected among a thousand patients who were referred from different centers to Acibadem Diagnostic Center, from May 2007 to February 2014. Genomic DNA and RNA samples were isolated from bone marrow and peripheral blood samples. RNA samples were used for detection of JAK2-exons 11-17 and TET2 mutations. DNA samples were used to detect the frequent mutations in other reported genes: CALR-exon9, CBL-exon5-8, DNMT3A-exon23, ASXL1-exon13 and LNK-whole exons by PCR and DNA sequencing. Results: We identified 11/152 JAK2 mutations (7.2%). We identified 10/152 MPL-exon10 mutations (3.9%). Although, all MPL mutations were reported previously, 6/11 JAK2 mutations were novel. We found 17/152 CALR mutations (11.2%); while 367f5'-46 (8/17) and K385f5'-47 (5/17) were the most common variants, the remaining 3 out of 4 variants were novel. Also, we identified 3 novel mutations all in the coding sequence of LNK (2%). In CBL, 2 patients had the same splice site variant in intron 8 (1.3%). We identified 10/152 ASXL1 mutations (6.6%) 2 of which were novel. None of the patients had the same mutation. Conclusion: To the best of our knowledge, this is the first mutation profile study associated with BCR/ABL1 and JAK2V617F negative MPNs in Turkish population. Our study enabled to improve the screening strategy of MPNs and provided genetic diagnosis of 49 of the selected 152 BCR/ABL1 and JAK2V617F negative MPNs. We observed new mutations as well as previously reported ones. Mutation profiles will contribute to the development of therapeutic drugs (such as tyrosine kinase inhibitors-TKis) and diagnostic strategies. Functional studies should be carried out in order to reveal the mechanism underlying these unique variants.

Whole genome sequencing of high-risk families to identify new mutational mechanisms of breast cancer predisposition. T. Walsh, S. Gulzumer, C. Casadei, M.K. Lee, J. Mandell, M.C. King, Division of Medical Genetics, University of Washington, Seattle, WA.

Multiple genes in addition to BRCA1 and BRCA2 are known to harbor mutations that significantly increase risk of breast cancer. The increasing use of comprehensive cancer gene panel allows women to be tested for all known breast cancer susceptibility genes in a single assay. For families severely affected with breast cancer, but without mutations in any known breast cancer gene, exome sequencing has been applied to identify new candidate genes, which could have a modifying effect. In our experience, approximately 50% of families with four or more relatives with breast cancer remain unresolved after comprehensive panel and exome sequencing. We hypothesize that in many of these families, breast cancer is due to individually rare alleles of moderate-to-severe effect located in the non-coding regions of the genome. A subset of these are likely influencing their mutational effect via alterations on expression or regulation of known breast cancer genes. We selected pairs of cousins with young-onset breast cancer from each of 20 very severely affected kindreds. Whole genome sequencing was carried out on DNA of the 40 individuals. We are filtering variants shared by the members of a pair with the following criteria: (i) rarity; (ii) within 5 MB of a known breast cancer gene; (iii) conservation of the site harboring the variant; and (iv) annotation of potential regulatory function. In parallel, we are performing targeted RNAseq of all known breast cancer genes from lymphoblast derived RNA of the same individuals (see abstract by Casadei S et al.). RNAseq data may reveal genes with asymmetrically expressed alleles in a family, thereby focusing regions of interest in the whole genome data from a tested family. Candidate variants will be tested for correlation with breast cancer in the host kindred. Our long-term goal is to integrate newly identified mutational mechanisms into clinical practice.
3383S
DICER1 mutations occurring in childhood anaplastic sarcoma of kidney. M. Wu1, H. Druker2, P. Thome3, J. Traubick4, R. Grant5, S. Albrecht6, E. Weber6, A. Charles1, 2, J.R. Pries7, R.R. Fabian8, N. Watanabe9, G. Vujanic10, W.D. Foulkes11, 12 1) Medical Genetics, Lady Davis Institute, McGill University, Montreal, Quebec, Canada; 2) Cancer Genetics Program, The Hospital for Sick Children and Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; 3) Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada; 4) Department of Diagnostic Imaging, Hospital for Sick Children and University of Toronto, Toronto, ON, Canada; 5) Department of Pathology, Montreal Children’s Hospital, McGill University, Montreal, QC, Canada; 6) Research Institute of the McGill University Health Centre, Montreal, QC, Canada; 7) School of Paediatrics and Child Health (SPACH), University of Western Australia (M561) Crawley, Western Australia; 8) Department of Paediatrics, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada; 9) Minneapolis, MN, USA; 10) Departments of Oncology and Experimental Medicine, McGill University, Montreal, QC, Canada; 11) Department of Pathology, Ni哄on University School of Medicine, Tokyo, Japan; 12) Institute of Cancer & Genetics, Cardinal University School of Medicine, Cardiff, UK; 13) Program in Cancer Genetics, Departments of Oncology and Human Genetics, McGill University, Montr哄al, QC, Canada.

Anaplastic sarcoma of the kidney (ASK) is an extremely rare renal neoplasm with fewer than 30 cases reported. We report two cases of ASK with DICER1. The first patient (Case 1) carried a germ line DICER1 mutation (c.2062C>T) which is predicted to cause a nonsense mutation (p.R688X) and her ASK bore a somatic in trans missense DICER1 mutation (c.5425G>A) which is predicted to cause an amino acid substitution at a conserved residue in the RNase IIIb domain. The second patient (Case 2) presented to the clinic with ASK with a pleuropulmonary blastoma (PPB) which bore a different somatic mutation in DICER1 (c.5125G>A; predicted p.D1709N). The ASK in Case 2 also possessed a somatic DICER1 RNAse IIIb domain mutation (c.5125G>A; predicted p.D1709N) and a predicted germline DICER1 mutation status of Case 2 is pending. Thus, the somatic mutations in both ASKs are in a DICER1 catalytic domain (RNase IIIb) and are predicted to affect the production of the “5p” class of microRNA by DICER1 protein. A previous report identified several cases of DICER1 renal sarcoma, two of which had germline and/or somatic DICER1 mutations. Taken together, we hypothesize that these tumors and ASKs currently described are the same entity. Our findings, therefore, build upon previous observations and suggest that DICER1 mutations may be an important cause of these rare tumors.

3384M
Single-cell mutation detection with multiple PCR-based targeted enrichment sequencing. Z. Wu, R. Vijay-Sankar, S. Sanders, W. M. Vujanic, W. H. Lader. Biological Research Content, QIAGEN Sciences, LLC, Frederick, MD, USA.

Mutations are the driving molecular causes of various biological processes such as development and cancer. Recent findings of genomic heterogeneity among ostensibly homogeneous cell populations such as cancer cells demand genomic characterization of mutations at the individual cell level to better understand the underlying biology. Additionally, single-cell technologies make feasible the characterization of rare cell types such as circulating tumor cells and in-vitro fertilized embryos. Due to its high sensitivity, next-generation sequencing (NGS) represents the ideal technology to analyze a collection of mutations in single cells. The challenge, however, is the limited amounts of DNA, which need to be amplified prior to NGS. To overcome this challenge, whole genome amplification (WGA), coupled with multiple PCR-based targeted enrichment, was tested for mutation detection in single cells isolated from two colon cancer cell lines, Lovo and HT29. A NGS panel targeting genes relevant to colon cancer biology was used to enrich DNA from WGA single-cell DNA (WSC), DNA from bulk cells (BC), or bulk-cell DNA which underwent WGA (WBC), followed by mutation detection with NGS. Forty (40) ng DNA was used across 4 PCR pools in a multiplex PCR set up to enrich for 38 genes known to harbor colon cancer-relevant mutations. The amplicons went through a standard library construction multiplex PCR set up to enrich for 38 genes known to harbor colon cancer-relevant mutations. The amplicons went through a standard library construction, followed by sequencing on a MiSeq®. Previously reported mutations such as KRAS (G13D) in LoVo and BRAF (V600E) in HT29 cells were successfully detected with this enrichment method on WSC as well as on BC and WBC samples. The surprising level of detection was further confirmed by the fact that we detected an extremely elevated prevalence of somatic mutations with a heterozygous nonsense mutation in the mutS homolog 2 (MSH2) gene (leading to MSH2 haploinsufficiency) in this patient as a model, we performed whole exome sequencing, transcriptome sequencing, and proteome profiling on liver cancer tissues and its matched liver cirrhotic tissues (i.e., adjacent ‘normal’ tissues). By whole-exome sequencing, we identified 20,382 genomic somatic mutations (single nucleotide variants, SNVs), and 4,998 were classified as non-silent mutations with a frequency over 1% in the 1,823 somatic mutations (36%) in genes were not expressed or covered, 1,892 (37.8%) where only the wild-type allele were expressed, 1,199 (24.0%) where both alleles were expressed, and 94 (1.7%) where only the mutant allele were expressed. In the current cancer level of proteome, we identified 118 mutations whose mutated peptides are identified in the liver cancer tissues, and their un-mutated peptides were identified in the liver cirrhotic tissues. These results suggested that only a small fraction of genomic variations were detected by targeted enrichment mutational analyses. This observation may provide important implications for conducting cancer-genome sequencing studies.

3385T
Patient-oriented functional genomics analysis of p53 mutations in cancer. O. Zill1, T. Shamu2, S. Fields2, 3, B.S. Taylor1. 1) Epidemiology & Biostatistics, University of California, San Francisco, San Francisco, CA; 2) Departments of Genome Sciences and Medicine, University of Washington, Seattle, WA; 3) Howard Hughes Medical Institute, Seattle, WA.

Over the last few years, thousands of cancer genomes have been sequenced and hundreds of recurrent cancer mutations have been identified. A defining challenge for cancer genetics is to determine which mutations in a patient’s tumor genome are biologically or clinically relevant and how two or more driver mutations co-occurring in that tumor might interact. We are using multi-dimensional computational analysis of cancer genomics data combined with human somatic-cell genetics experiments to address these questions. We have focused initially on the most frequently mutated cancer gene, TP53, because of the wealth of experimental data available for many mutations and to provide sufficient sample numbers to power association tests across many tumor types. Using a compendium of somatic mutations from the sequenced exomes and genomes of over 10,000 human tumors, we assessed the co-occurrence of codon-specific TP53 mutations with those in 188 significantly mutated genes in cancer. Co-mutation patterns fell into several distinct categories that were partly explained by tumor-type-specific TP53 mutational patterns. Interestingly, certain missense TP53 mutations preferentially co-occurred with mutations in specific oncogenes or tumor suppressors. To determine the functional co-occurrence of these specific-SMG-enriched TP53 mutations, we are introducing them into human cell lines and computing them against each other in FACS-based transcriptional reporter assays. For these mutant competition assays, we constructed transcriptional reporters using the promoters of p53 targets (e.g. MYC or Wnt1) and SMG promoters (e.g. OXPHOS or TP53). Enrichment of specific SMGs was assessed by competition assays. We are testing whether ID1-TP53 co-mutation contributes to specific, cancer-relevant phenotypes by introducing the mutations into human cell lines. The mutation combinations we identified are being evaluated for associations with alterations in co-occurred expression of SMGs. This work should help delineate the functional contributions of specific combinations of driver mutations to inform personalized cancer treatments.

3386S
Expression of genomic somatic mutations at the levels of transcriptome and proteome in a patient of hepatocellular carcinoma with MSH2 haploinsufficiency. K. Ding1, S. Wu1, W. Ying2, W. Yin1, S. Tong1, Q. Pan1, X. Li2, D. Zhao1, H. Xu1, X. Qian3, H. Ren1. 1) Key Laboratory of Molecular Pathology of Hepato-Biliary-Pancreatic Diseases, Ministry of Educati, Chongqing, China; 2) State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing, P.R. China; 3) Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China; 4) Department of Pathology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China.

Over the last few years, thousands of cancer genomes have been sequenced and hundreds of recurrent cancer mutations have been identified. A defining challenge for cancer genetics is to determine which mutations in a patient’s tumor genome are biologically or clinically relevant and how two or more driver mutations co-occurring in that tumor might interact. We are using multi-dimensional computational analysis of cancer genomics data combined with human somatic-cell genetics experiments to address these questions. We have focused initially on the most frequently mutated cancer gene, TP53, because of the wealth of experimental data available for many mutations and to provide sufficient sample numbers to power association tests across many tumor types. Using a compendium of somatic mutations from the sequenced exomes and genomes of over 10,000 human tumors, we assessed the co-occurrence of codon-specific TP53 mutations with those in 188 significantly mutated genes in cancer. Co-mutation patterns fell into several distinct categories that were partly explained by tumor-type-specific TP53 mutational patterns. Interestingly, certain missense TP53 mutations preferentially co-occurred with mutations in specific oncogenes or tumor suppressors. To determine the functional co-occurrence of these specific-SMG-enriched TP53 mutations, we are introducing them into human cell lines and computing them against each other in FACS-based transcriptional reporter assays. For these mutant competition assays, we constructed transcriptional reporters using the promoters of p53 targets (e.g. MYC or Wnt1) and SMG promoters (e.g. OXPHOS or TP53). Enrichment of specific SMGs was assessed by competition assays. We are testing whether ID1-TP53 co-mutation contributes to specific, cancer-relevant phenotypes by introducing the mutations into human cell lines. The mutation combinations we identified are being evaluated for associations with alterations in co-occurred expression of SMGs. This work should help delineate the functional contributions of specific combinations of driver mutations to inform personalized cancer treatments.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
A.K. Snow1, T.M.F. Tuohy1, N.R. Sargent1, L.J. Smith1, R.W. Burt2,3, D.W. Neklason1,2,4 1) High Risk Clinical Research, Huntsman Cancer Institute, Salt Lake City, UT 84112; 2) Department of Medicine, University of Utah, Salt Lake City, UT 84112; 3) Division of Gastroenterology, University of Utah, Salt Lake City, UT 84112; 4) Division of Genetic Epidemiology, University of Utah, Salt Lake City, UT 84112.

Familial adenomatous polyposis (FAP [MIM 175100]) is a colorectal cancer predisposition syndrome caused by mutations in the adenomatous polyposis coli (APC [MIM 611731]) gene. University of Utah has a large Hereditary Gastrointestinal Cancer Registry with over 600 FAP patients from 244 kindreds. A clear deleterious mutation is known in 122 of these kindreds, but in the remaining kindreds, APC genetic testing has either failed to detect a mutation (16%) or has not been pursued (40%). We have applied a simple and cost-effective approach to screen for families who may harbor identical pathogenic mutations due to a common founder. Members of our registry with FAP have been genotyped using a set of 4 short tandem repeat markers (STR) with high heterozygosity across the chr 5 APC locus. Following the inclusion of probes specific for APC promoter 1B, multiplex ligation-dependent probe amplification (MLPA) identified a proband from our registry with a deletion of promoter 1B. Six additional families shared the APC haplotype and the deletion, which is distinct from previously reported promoter 1B deletions. The clinical phenotype of 17 mutation carriers is classic colonic polyposis with colectomy at an average age of 24. The majority report having a large number of duodenal and gastric polyposis, and there were single reports of a desmoids and a hepatoblastoma. Measurements of allele-specific expression of APC mRNA using TaqMan assay specific to a common polymorphism, rs4987562, on cDNA from 3 tissue sources confirmed that the relative expression was reduced in the allele containing the promoter 1B deletion. The relative reduction in expression of the mutant allele was more pronounced in mRNA extracted from duodenal tissue (88%) and blood (98%) than in mRNA extracted from EBV transformed lymphoblastoid cell line. Previous work has shown down regulation of the alternative APC promoter 1A through methylation. This methylation, along with deletion of promoter 1B, is thought to be the underlying mechanism for the imbalance in mRNA expression. The impact of these promoter 1B deletions as a cause of FAP. Alterations in promoters and other regulatory elements should be considered critical targets of future genetic research as a cause of syndromes. This study also highlights the utility of using STR analysis to identify patients lacking a genetic diagnosis by linking them to a genetic haplotype of a known deleterious mutation.

Y. Wu, Z. Zhang, D. Zhou, Y. Liu, H. Huang. 1) The International Peace Maternity & Child Health Hospital of China welfare Institute, Shanghai, Jiao Tong University, Shanghai, 200030, China; 2) Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Graduate School of the Chinese Academy of Sciences, Shanghai, 200031, P. R. China; 3) School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China; 4) Bio-X Institutes, Key Laboratory for the Genetics of Neurodevelopmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai, 200030, P. R. China; 5) Institutes of Biomedical Sciences, Fudan University, Shanghai, 200032, P. R. China.

MRKH syndrome is characterized by Mullerian duct aplasia (uterus plus upper two-thirds of the vagina) with female phenotype and a 46, XX karyotype. MRKH is infrequent thus uterine leiomyoma of the rudimentary uterus in MRKH is an extremely rare case. We performed whole-genome sequencing (WGS) for uterine leiomyoma, rudimentary uterus and peripheral blood from a patient with MRKH syndrome, and conducted analyses for chromosomal structure variations (SVs), copy number variations (CNVs) as well as point mutations. We further performed array comparative genomic hybridization (CGH) for rudimentary uterus and peripheral blood to validate CNVs from the WGS findings. Our results confirmed 8q23.1 germline deletion in MRKH patient and we also observed several somatic CNVs in leiomyoma including deletions in 1q24.2, 4q34.3 and 19q13.33. Interestingly, the number of CNVs in rudimentary uterus showed more abundant than that in leiomyoma, suggesting that leiomyoma might be initiated from undifferentiated mesenchymal cells or immature smooth muscle cells instead of mature smooth muscle cells in rudimentary uterus. We also detected two translocation events t(16;21) and t(11;X) in rudimentary uterus and three translocation events t(1:7), t(2:10) and t(10:16) in uterine leiomyomas. Additionally, many somatic point mutations were newly identified in leiomyoma with probably damaging function involved in other type of cancers or tumors. Our study initially characterized rudimentary uterus and uterine leiomyoma from whole genome level, which may provide insight into development of rudimentary uterus in MRKH syndrome and pathogenesis of leiomyoma in further studies.
3391M
Comparison of NGS solutions for rapid and cost-effective analysis of degraded FFPE and cancer biobanked specimens with limited quantity.
A. Brooks1, M. Moreau1, L. Liu2, E. Fang3, G. Zhang2, L. Emrick2, M. Li2,3, 1) Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) Dan Duncan Cancer Center, Baylor College of Medicine, Houston, TX.

Recent studies have shown that a group of overgrowth syndromes, such as congenital lipoatrophic overgrowth with vascular, epidermal, and skeletal anomalies (CLOVES), Proteus syndromes, and two overlapping disorders, megalencephaly-polycystic kidney disease (MPKCD) and megalencephaly-capillary malformation (MCAP), are caused by somatic activating mutations in genes involved in the PI3K-akt signaling pathway. Due to the low-abundance nature of these mutations, routine Sanger sequencing often yields negative results. We have developed a next generation sequencing (NGS) panel that targets all known mutations in multiple genes involved in the PI3K-akt pathway for this group of overgrowth syndromes. Eight patients including two prenatal cases and six postnatal cases suspected to harbor mutations were tested using the NGS sequencing panel. A somatic mutation in the PIK3CA gene was identified in all 8 cases including one G542K mutation, two H1047L mutations, two H1047R mutations, two G914R mutations and one E110del mutation with the mutant allele frequencies ranging from 4.65% to 38.36%. These PIK3CA mutations were only present in the affected tissues in the majority of the cases demonstrating causal role of the mutations in the development of the diseases and the importance of obtaining affected tissue for testing. Phenotype-genotype correlation analysis showed that mutations at amino acid residue 1047 are often associated with CLOVES syndrome or isolated malformations of extremities or fatty mass; while the G914R mutation is preferentially related to syndromes with overgrowth of brain, such as MPKCD/MCAP. In addition, in vitro cell culture showed significant enrichment of the cells harboring the putative actionable mutation that can be used to further analyze these syndromes with low sensitivity failed to detect mutated clone, causing a false negative. Therefore, selection of a more sensitive molecular method would be helpful for accurate diagnosis. Investigation into more cases with JAK2V617F positive CEL should lead to a deeper understanding of the mechanism involved in leukemogenesis.

3392S
Small clone of JAK2V617F positive chronic eosinophilic leukemia detected by realtime PCR. N. Yu1, Y.J. Kim2, S.J. Park3, Y.R. Kim4, K.A. Lee1, J.W. Kim4, 1) Department of Laboratory Medicine, Yonsei University Wonju College of Medicine; 2) Samkwang medical laboratories; 3) Department of Internal Medicine, Yonsei University College of Medicine; 4) Department of Laboratory Medicine, Yonsei University Wonju College of Medicine.

Chronic eosinophilic leukemia (CEL) is defined as clonal proliferation of eosinophil precursors resulting in persistently increased numbers of eosinophils. Evidence of clonality is crucial in making a diagnosis of CEL. Herein we report a case of CEL with JAK2V617F positive by allele specific PCR and realtime PCR but negative by pyrosequencing and droplet sequencing. A 33-year-old Korean female patient was admitted to our hospital in March 2012 for evaluation of persistent abdominal pain, diarrhea, vomiting. The patient had no previous history of hematologic malignancies or cytotoxic therapy. Peripheral blood smear showed eosinophilia. Bone marrow aspira- tion and biopsy revealed normocellular marrow with increased number of eosinophils and its precursors (19.7% of all nucleated cells). Endoscopic gastric biopsy revealed eosinophilic infiltration. Conventional karyotyping showed 46,XX in all metaphases. In fluorescence in situ hybridization and gene rearrangement studies, no BCR-ABL, FGFR1 and PDGFR rearrange- ments were detected. Allele specific PCR for JAK2V617F mutation (Seeplex JAK2 Genotyping kit, Seegene, Seoul, Korea) revealed faint positive band. DNA sequencing for exon 14 of JAK2 gene (ABI3500 genetic analyzer, Applied Biosystems, CA, USA) did not reveal any point mutation that promp- ted further investigation. The pyrosequencing (JAK2 pyro kit, Qiagen, CA, USA) showed negative results while realtime PCR using Real-Q JAK2V617F detection kit (BioSewoom, Seoul, Korea) was confirmed positive for the mutation. To date, only 6 cases of JAK2V617F has been reported in the literature. However, JAK2V617F positive CEL cases always demonstrated negative results for other gene rearrangement; therefore, indicating that JAK2V617F could be an early event in the leukemogenesis. Although prog- nosis are not well defined due to paucity of CEL cases with JAK2V617F mutation, in our case, steroid alone led to a good response and outcome. In this case, the small mutated clone was detected using most sensitive methods such as allele specific PCR and realtime PCR which could detect JAK2V617F mutation with low sensitivity failed to detect mutated clone, causing a false negative. Therefore, selection of a more sensitive molecular method would be helpful for accurate diagnosis. Investigation into more cases with JAK2V617F positive CEL should lead to a deeper understanding of the mechanism involved in leukemogenesis.

3393T
Molecular characterization of over growth syndromes using NGS reveals potential phenotype-genotype correlation. F. Chang1, L. Liu2, E. Fang3, G. Zhang2, L. Emrick2, M. Li2,3, 1) Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX; 2) Dan Duncan Cancer Center, Baylor College of Medicine, Houston, TX.

In order to unlock genetic sequence information hidden in millions of tissue biopsies and cancer biobanks around the world, scientists require a robust Next-Generation Sequencing (NGS) platform that offers consistent, cost-effective, high quality and high coverage results. We examined the utility of two of the most common and prevalent next generation sequencing (NGS) criteria including input material needs, sequencing coverage specificity and sensitivity, workflow ease and robustness, scalability, and all-in total sample cost. We evaluated the combination of the RainDance ThunderBolts™ Cancer Panel and Illumina MiSeq System with the Ion Torrent AmpliSeq™ Cancer Hotspots v2 panel and Proton™. We used the RainDance ThunderBolts™ Cancer Panel to enrich 10ng of both FFPE and fresh frozen (FF) samples of colorectal cancer tumor and adjacent normal tissue. The ampli- con-based ThunderBolts Cancer Panel uses single molecule PCR technol- ogy to target 50 important cancer genes including tumor suppressors, hot- pots and drug resistance markers and features a DirectSert™ workflow method that integrates sequencing adapters and attaches sample indexes directly, without additional library preparation steps. The enriched samples were sequenced on an Illumina MiSeq System. The results were compared to those using 20ng of the samples with the Ion Torrent AmpliSeq Cancer Hotspots v2 panel and Proton. Sequence validation was performed using RainDrop digital PCR (dPCR) for positive targets derived from both sequenc- ing panels. The sequencing performance was validated using both variant calling and mutational enrichment studies. We also found significant enrichment of the cells harboring the putative actionable mutation that can be used to further analyze these syndromes with low sensitivity failed to detect mutated clone, causing a false negative. Therefore, selection of a more sensitive molecular method would be helpful for accurate diagnosis. Investigation into more cases with JAK2V617F positive CEL should lead to a deeper understanding of the mechanism involved in leukemogenesis.

3393M
Approaches to Integrating Germline and Tumor Genomic Data in Cancer Research. H.L. Spencer Feigelson1, A.B. Goddard2, C. Hollombe2, S.R. Tingle2, S.A. Nelson1, E.M. Gillanders2, 1) Epidemiology and Genomics Research Program, Division of Cancer Control and Popula- tion Sciences, National Cancer Institute, Rockville, MD; 2) Institute for Health Research, Kaiser Permanente Colorado, Denver, CO; 3) Center for Health Research, Kaiser Permanente Northwest, Portland, OR.

Cancer is characterized by a diversity of genetic and epigenetic alterations occurring in both the germline and somatic (tumor) genomes. Hundreds of germline variants associated with cancer risk have been identified, and similarly large amounts of data have been generated identifying mutations in the tumor genome that play important roles in tumorigenesis. Increasingly, these two genomes are being explored jointly to better understand how cancer risk alleles contribute to carcinogenesis and whether they influence development of specific tumor types or mutation profiles. To understand how data from germline risk studies and tumor genome profiling is being integrated, we reviewed 160 articles describing research that incorporated data from both genomes, published between 2009 and 2012, and summa- rized the current state of the field. We identified three types of research questions being addressed using these data: (1) use of tumor data to deter- mine the putative function of germline risk variants; (2) identification and analysis of relationships between host genetic background and particular tumor mutations or types; and (3) use of tumor molecular profiling data to reduce genetic heterogeneity or refine phenotypes for germline association studies. We also found several descriptive studies comparing germline and tumor variation in a gene, related genes, or gene families. We identified a large number of potential studies that can be used to test the hypothesis that alleles can be used for prenatal diagnosis of these syndromes. Our experience demonstrates that NGS technology is highly sensitive for the detection of low-level mosaic mutations and can be used for the diagnosis of these overgrowth syndromes in both prenatal and postna- tal settings.
NF2-associated VS arise from multiple independent tumor initiation events. These data suggest, for the first time, that the majority of NF2-associated multi-lobulated VS are not single tumors, but rather clusters of multiple smaller tumors, each arising from their own unique architecture by Illumina HumanOmniExpress SNP-arrays in 24 tumor samples using digital PCR. This multiplex assay targets seven KRAS point mutations prevalent at greater than 1%, resulting in an effective profiling tool for 98% of KRAS mutant colorectal cancers (Faulkner et al. 2010). No pre-amplification step is required. This KRAS screening assay was used to quantify KRAS mutational load in a panel of FFPE specimens from advanced metastatic colorectal cancer patients. KRAS mutations present at <1% fractional abundance were detected in multiple samples. This sensitive and inexpensive method reduces the risk of contamination and can be easily implemented in molecular diagnostic laboratories for rapid, routine screening of cancer patients.


Targeted therapies in many cancers have allowed unprecedented progress in the treatment of disease. However, routine implementation of genomic testing is limited due to: 1) limited amounts of sample (pg-ng range) per biological specimen, 2) diagnostic turnaround time and workflow, 3) cost, and 4) difficulties in detection of mutational loads below 5%. KRAS is mutated in approximately 40% of colorectal cancers. The majority of these KRAS mutations are activating mutations in codons 12, 13, and 61, and are predictive of a negative response to mEGFR therapy. To optimize therapy strategies for personalized care, it is critical to rapidly screen patient samples for the presence of multiple KRAS mutations. We have developed a multiplexing strategy to screen clinically-actionable KRAS mutations in cancer samples using digital PCR. This multiplex assay targets seven KRAS point mutations prevalent at greater than 1%, resulting in an effective profiling tool for 98% of KRAS mutant colorectal cancers (Faulkner et al. 2010). No pre-amplification step is required. This KRAS screening assay was used to quantify KRAS mutational load in a panel of FFPE specimens from advanced metastatic colorectal cancer patients. KRAS mutations present at <1% fractional abundance were detected in multiple samples. This sensitive and inexpensive method reduces the risk of contamination and can be easily implemented in molecular diagnostic laboratories for rapid, routine screening of cancer patients.

Evidence of multiple independent NF2 somatic inactivation and tumor initiation events in neurofibromatosis type 2-association vestibular schwannomas. A. Pemov1, R. Dewan2, J. Kim3, K. Morgan2, R. Vasquez2, P. Chittiboina2, X. Wang2, A. Ray-Chaudhury2, S. Chandrasekharapattu2, J. Butman1, A. Asthagiri2, D.R. Stewart2, J. Kim2, 1) NIH/NIC, Bethesda, MD; 2) NIH/NINDS, Bethesda, MD; 3) NIH/NIOQCT, Bethesda, MD; 4) NIH/NHGRI, Bethesda, MD; 5) NIH/NCI, Bethesda, MD; 6) University of Virginia, Charlottesville, VA.

Background and hypothesis. Neurofibromatosis type 2 (NF2) is a tumor predisposition syndrome that results from mutation of the NF2 tumor suppressor gene (chromosome 22q12). The hallmark of NF2 is the presence of bilateral vestibular schwannomas (VS). It is generally accepted that somatic inactivation of the second copy of NF2 is the initiating event for neoplastic transformation in NF2-associated tumors including VS. Though NF2-associated and sporadic VS share identical histopathologic findings and underlying cytogenetic alterations, NF2-associated VS are less responsive to radiosurgery, are associated with worse surgical outcomes and often appear as a cluster of grapes. We hypothesized that the distinct multi-lobulated morphology of NF2-associated VS is determined by the genetics of the tumors and that the “grapes” in the cluster are in fact individual tumors that arose independently with distinct somatic mutations in the NF2 gene and overall genomic architecture.

Study design and methods. To test this hypothesis, we analyzed the mutation status of NF2 by Sanger sequencing and genomic architecture by Illumina HumanOmniExpress SNP-arrays in 24 tumor samples procured from five VS, or 4.8 specimen per tumor on average, in four NF2 patients. Results. First, we identified NF2 germline mutations in each of four NF2 patients. Two patients carried deleterious point mutations in NF2 and the other two presented with large chromosomal deletions affecting the gene. Second, we identified 18 distinct NF2 somatic mutations (13 point mutations and five chromosome 22 LOH events) in five tumors, or 3.6 average independent somatic NF2 hits per tumor and ranging from 1 to 7 per tumor. Finally, SNP-array analysis revealed that a deletion or mitotic recombination affecting the entire chromosome 22 or its q-arm was the most frequent large-scale chromosomal alteration in the tumors (5/24, or 21% per cent). The rest of the genome in all but one tumor sample resembled that of normal diploid cells. Conclusions. These findings support our hypothesis that NF2-associated multi-lobulated VS are not single tumors, but rather are clusters of multiple smaller tumors, each arising from their own unique somatic hit event. These data suggest, for the first time, that the majority of NF2-associated VS arise from multiple independent tumor initiation events. These findings have important clinical implications and will be critical when establishing clinical trials endpoints.

Multiplex detection of KRAS mutations in colorectal cancer FFPE samples using droplet digital PCR. S. Cooper, W. Yang, D. Sheltan, J. Berman, B. Zhang, S. Tzonev, E. Hefner, J. Regan. Digital Biology Center, Bio-Rad Laboratories, Inc., Pleasanton, CA. DNA in normal mammalian somatic cells primarily exists as long chromosomes or in mitochondria. However, we previously reported the presence of thousands of short (mostly 200-400 bps) extra chromosomal circular DNAs (microDNAs), single- or double-stranded, arising mainly from the GCRich regions in the genome, and enriched in genic regions in embryonic mouse brain, heart and liver tissues (Shibata, Y. et al. Science 336, 82-86, 2012). More than 95% of microDNAs could be mapped uniquely to a genome. Chromosomal loci that are enriched sources of microDNA in the adult brain are somatically mosaic for microdeletions that appear to arise from the excision of microRNAs. MicroRNAs often have micro homology of 2-15 bases at the start and after the end of the corresponding chromosomal sequence suggesting that a micro homology mediated mechanism could be the mechanism of microDNA biogenesis. In the current research we studied the characteristics of microDNAs in chicken cell lines deleted in various DNA repair genes, including DNA ligase IV and Ku70 involved in non-homologous end joining (NHEJ); BRCA1, BRCA2, NBS1, and Rad54 involved in homologous recombination (HR); CtIP required for microhomology-mediated alternative end joining (MMEJ); and MSH3 involved in DNA mismatch repair (MMR). The MSH3-deleted cells had a striking change in their microDNA population, suggesting that the mismatch-repair pathway, and by extension replicative polymerase slippage is involved in the generation of a fraction of the microDNAs. To determine tissue specificity of microDNA generation, we examined a panel of mouse tissues (brain, heart, lung, liver, spleen, kidney, muscle, sperm, testis and thymus) and human prostate (C4-2, LnCap and PC-3) and ovarian (ES2 and OVCAR-8) cancer cell lines. MicroDNAs were observed in all tissue types, confirming that microDNAs are universally present even in germ cells. We discovered that there are hot spots of microDNA generation distributed throughout the genome that are common between tissues and correlate with areas of high gene density and high GC content. However, hierarchical clustering on the basis of microDNA co-ordinates classified the prostate and ovarian cancer cell lines as two separate clusters, suggesting that at least some microDNAs are tissue-specific and therefore their sites of origin are affected by tissue-specific gene expression patterns or epigenetic marks.

MicroDNA (Extra Chromosomal Circular DNA) in Mammalian Tissues and Cancer Cell Lines. P. Kumar, L.W Dillon, Y. Shibata, A. Dutta. DEPT. OF BIOCHEMISTRY & MOLECULAR GENETICS, SCHOOL OF MEDICINE, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA.

DNA in normal mammalian somatic cells primarily exists as long chromosomes or in mitochondria. However, we previously reported the presence of thousands of short (mostly 200-400 bps) extra chromosomal circular DNAs (microDNAs), single- or double-stranded, arising mainly from the GC-rich regions in the genome, and enriched in genic regions in embryonic mouse brain, heart and liver tissues (Shibata, Y. et al. Science 336, 82-86, 2012). More than 95% of microDNAs could be mapped uniquely to a genome. Chromosomal loci that are enriched sources of microDNA in the adult brain are somatically mosaic for microdeletions that appear to arise from the excision of microRNAs. MicroRNAs often have micro homology of 2-15 bases at the start and after the end of the corresponding chromosomal sequence suggesting that a micro homology mediated mechanism could be the mechanism of microDNA biogenesis. In the current research we studied the characteristics of microDNAs in chicken cell lines deleted in various DNA repair genes, including DNA ligase IV and Ku70 involved in non-homologous end joining (NHEJ); BRCA1, BRCA2, NBS1, and Rad54 involved in homologous recombination (HR); CtIP required for microhomology-mediated alternative end joining (MMEJ); and MSH3 involved in DNA mismatch repair (MMR). The MSH3-deleted cells had a striking change in their microDNA population, suggesting that the mismatch-repair pathway, and by extension replicative polymerase slippage is involved in the generation of a fraction of the microDNAs. To determine tissue specificity of microDNA generation, we examined a panel of mouse tissues (brain, heart, lung, liver, spleen, kidney, muscle, sperm, testis and thymus) and human prostate (C4-2, LnCap and PC-3) and ovarian (ES2 and OVCAR-8) cancer cell lines. MicroDNAs were observed in all tissue types, confirming that microDNAs are universally present even in germ cells. We discovered that there are hot spots of microDNA generation distributed throughout the genome that are common between tissues and correlate with areas of high gene density and high GC content. However, hierarchical clustering on the basis of microDNA co-ordinates classified the prostate and ovarian cancer cell lines as two separate clusters, suggesting that at least some microDNAs are tissue-specific and therefore their sites of origin are affected by tissue-specific gene expression patterns or epigenetic marks.
3397T
Heterozygous mutations in PALB2 predispose to breast cancer by causing DNA replication and damage response defects. R. Winqvist1,2, M. Boså1,2, J. Nikkia1,2, A.C. Papyls1, K. Borgmann2, Y. Huo2, K. Rapapko2, N. Laurila1,2, P. Nieminen1, B. Xia2, H. Pykkä2,1, H. Pospiech6,7, 1) Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, University of Oulu, Oulu, Finland; 2) Northern Finland Laboratory Centre NordLab, Oulu, Finland; 3) Laboratory of Radiobiology and Experimental Radiation Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; 4) Department of Radiation Oncology, Rutgers Cancer Institute of New Jersey and Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA; 5) Medical Informatics and Statistics Research Group, University of Oulu, Oulu, Finland; 6) Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland; 7) Research group Biochemistry, Leibniz Institute for Age Research—Fritz Lipmann Institute, Jena, Germany; 8) Present address: Gene Function Team, Breakthrough Breast Cancer Research Centre, ICR, London, UK.

Breast cancer is the most common cancer among females. Although the majority of the breast cancer cases appear to be sporadic, it has been estimated that as much as 5-10% is significantly contributed by powerful hereditary genomic risk factors. Besides mutations in BRCA1 and BRCA2, also defects in the PALB2 gene occur worldwide and are important in hereditary predisposition to breast cancer. PALB2 interacts with BRCA1 and BRCA2 to regulate homologous recombination and to mediate cellular DNA damage response. We have previously shown that in the Finnish population a relatively common heterozygous mutation in PALB2 increases the risk of malignancy in carrier individuals about sixfold. About 1% of the breast cancer patients in Finland carry this mutation. Unfortunately, however, the mecha-
nistic details for how heterozygous mutations in cancer predisposing genes such as PALB2 trigger disease development are currently largely obscure. By analysing lymphoblastoid cell lines from heterozygous female PALB2 mutation carriers, we have recently been able to demonstrate that haploin-
sufficiency for PALB2 causes aberrant DNA replication/damage response. Mutation carrier cells show increased origin firing and shorter distance between consecutive replication forks. Carrier cell lines also show compo-
nent-independent defects in checkpoint and radiosensitivity, which result in gradual accumulation of various genomic lesions. Elevated chromosome instability was also observed in primary blood lymphocytes of heterozygous PALB2 mutation carriers, indicating that the described mechanisms of genome destabilization operate also at the organ-
ism level. Thus the functional loss of one copy of the PALB2 gene causes instability in the genome, even though the other allele is still intact and functioning. These exciting findings provide a new mechanism for early stages of breast cancer development that may also apply to other heterozy-
gous homologous recombination signaling pathway gene mutations involved in hereditary cancer predisposition.

3398S
Investigation of de novo mutation rates in families with DNA Polymer-
ase ε and θ exonuclease domain mutations. S.E.W. Briggs, O. Vern, C. Palles, G. McVean, I.P. Tomlinson, The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, Oxfordshire, United Kingdom. DNA polymerases ε and θ have an essential role in eukaryotic DNA replication, proofreading newly synthesised DNA and identifying and exist-
ing mispaired nucleotides. Mutations in the exonuclease domains (EDMs) of DNA polymerases POLE (MIM #174762) and POLD1 (MIM #174761) have been identified as causal mutations in a novel cancer predisposition syndrome, polymerase proofreading associated polyposis (PPAP [MIM #666675]; [MIM #606691]). PPAP displays high penetrance and a variable pheno-
type including colorectal, endometrial and ovarian tumours. Sporadic colon and endometrial tumours with somatic POLE EDM mutations are defined by a hypermutated phenotype with a distinct mutational spectrum of increased transversion mutations.

To investigate the effect of germline mutations on de novo mutation rates we performed whole genome sequencing of two parent-offspring quintets. One quintet carries POLE L424V (c.1270G>C, NM_006231), and one POLD1 S478N (c.1433G>A, NM_002681); in each family the mother and two of three offspring carry the mutation.

The whole genomes of each family member were sequenced to a read depth of 7.9x. We applied alignment based calling to each family without incorporating relatedness, detected 4.9 million variants with a transition:
transversion ratio of 2.16. To these calls we applied a modified Lander-
Green algorithm to construct a transmission scaffold across chromosomes for each family. Probabilistic models were used to classify sites inconsistent with transmission, followed by filtering for false-positive errors. Candidate mutations were then validated using independent targeted sequencing to estimate the false positive rate, and direct paired-end phasing was con-
ducted.

We sought to test whether there is an elevated incidence of de novo mutations in the carriers. Preliminary results suggest that in both pedigrees the transition:transversion ratio is decreased (1.2 in POLE; 1.5 in POLD1), with an abundance of indels (23% in POLE; 34% in POLD1) suggesting an impact of germline mutation in line with the biological expectation. Paternal bias of de novo mutation origin has been reported in humans and chimpan-
zees. If DNA polymerase mutations do impact germline mutation, we may observe an increase in the female contribution given that the mothers are the carriers in both families.

3399M
NBN gene expression and cytogenetic changes in irradiated cells with NBN gene mutations. D. Januszkiewicz-Lewandowska1,2,3, B. Swiatek-Koscielna1, J. Rembowska1, A. Dzikiewicz-Krawczyk1, M. Żawada1, J. Nowak1, 1) Institute of Human Genetics, Polish Academy of Sciences, Poz-
nan, Poland; 2) University of Medical Sciences, Poznan, Poland; 3) Depart-
ment of Medical Diagnosis, Poznan, Poland.

NBN gene product is part of the MRE11/RAD50/NBN complex, which plays essential role in genomic stability. Homozygous mutations in the NBN gene c.657-661del causes Niemjen breakage syndrome in the course of which there is a high susceptibility to cancer. Heterozygous NBN gene mutations are considered to be a minor cancer risk factor. The study tries to answer the following questions: Does mutations of NBN gene influence the DNA synthesis and expression of NBN gene in irradiated immortalized cells? What is the effect of NBN gene mutation and irradiation on the chromosome stability in lymphoblastic cell line with germinal NBN gene mutations? Radio-
resistant DNA synthesis measured by 3H thymidine incorporation assay, NBN gene expression and chromosomal aberrations were assessed in cul-
tured cells before and after irradiation with doses 1-8 Gy. Our results indi-
cated that only cells with homo- and heterozygous c.657-661del mutation showed higher 3H thymidine incorporation after irradiation, particularly in the 48-hour culture and after higher doses of irradiation. These results suggest that those cells showed significantly lower rates of DNA synthesis inhibition. Radiosensitivity measured as a number of chromosomal aberrations per one metaphase was significantly higher in cells with all NBN gene mutations studied as compared to the control cell line and was clearly marked for homozygous and heterozygous c.657-661del cell lines. Cells with homozygous and heterozygous c.657-661del and p.R215W mutations possessed significantly higher relative NBN gene expression when compared to the control. After irradiation the relative expression of NBN was significa-
cantly higher only in homozygous c.657-661del and heterozygous p.R215W. Our observation may indicate on a compensatory increase in the expression of p70-NBN mutant in cells with homozygous c.657-661del and heterozy-
gous c.657-661del and p.R215W mutations. Similarly, irradiation may in an alternative way in these cell lines stimulate significantly higher increase of NBN gene expression. Financial support - National Science Centre 2011/ 01/B/NZ5/04322.
A new control mechanism for repair of DNA in human cells: MDC1 and ATR regulate DNA Double-Strand Break (DSB) resection independently of ATM. P S Bradshaw 1,2, M. Komosa 1, M S Meyn 1,2,3,4

1 Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada; 2 Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; 3 Division of Clinical and Metabolic Genetics, Department of Paediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada; 4 The Department of Paediatrics, University of Toronto, Toronto, Ontario, Canada.

To maintain genome stability, human cells have developed strategies to detect, repair or induce cell death/apoptosis in response to DNA damage. DNA DSBs frequently arise from exposure to exogenous and endogenous sources and are one of the most lethal forms of DNA damage. Detection of DNA DSBs results in the activation of the PIK-related kinase ATM and its phosphorylation mediated signaling pathway. ATM is mutated in the autosomal recessive disorder Ataxia-Telangiectasia (A-T) in which defective DSB repair leads to genome instability, radiation sensitivity, cerebellar ataxia, immunological defects, radiation sensitivity and high cancer risk. Here we report that ATM is rapidly recruited to DNA breaks in the nuclei of human cells, where it forms distinctive spatially well-defined, non-diffusible, foci. Using interactive deconvolution microscopy, we find that ATM belongs to a novel class of DSB repair proteins that rarely co-localize with Rad51 and RPA (proteins directly involved Homologous Recombination Repair - HRR), resected DSBs, or proteins that make up the mega-base region of γ-H2AX containing chromatin that surround DSBs. Human cells lacking ATM display reduced numbers of RPA and Rad51 foci following the generation of radiation-induced DSBs. RPA foci formation is dependent on CtIP a protein target for ATM phosphorylation and implicated in DNA resection and the generation of single stranded DNA (ssDNA). Our data indicate that ATM promotes the initiation of resection and subsequent HRR, through the CtIP dependent generation of ssDNA. In contrast, MDC1, a downstream member of the ATM-dependent signaling pathway, acts independently of ATM to restrain the initiation of DSB resection and prevent the formation of ssDNA. Interestingly, we find that MDC1 restrains DNA resection by inhibiting the ATR dependent activation of CtIP. Like ATM, ATR is a PIK-related kinase involved in DNA damage signaling. Mutations in ATR can cause Seckel Syndrome (SCLK1), an autosomal recessive disorder characterized by growth retardation, dwarfism, microcephaly, mental retardation and facial dysmorphology. Significantly, CtIP mutations can also cause Seckel Syndrome (SCLK1). Our work identifies that, in order to prevent DSB resection that functions independently of the classical ATM-mediated DNA damage signaling pathway. Further, our demonstration of MDC1/ATR control of CtIP mediated resection establishes a molecular link between SCLK1 and SCLK2 syndromes.
1. Medicine, Brigham and Women's Hospital, Boston, MA; 2) Center for Computational Biology, Dana Farber Cancer Institute, Boston, MA; 3) Pathology, Beth Israel Deaconess Medical Center, Boston, MA; 4) Division of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA; 5) Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA.

Genome-wide association studies (GWAS) of breast cancer have identified 71 single nucleotide polymorphisms (SNPs), majority of which are located in intergenic or intronic regions. To identify regulatory variants we conducted expression quantitative trait loci (eQTLs) analysis using invasive postmenopausal breast cancer cases in the Nurses' Health Study (NHI) diagnosed from 1990-2004 with GWAS data. RNA extracted from formalin fixed paraffin embedded (FFPE) breast tumor (tumor) and normal adjacent (normal) breast tissue was profiled using the Affymetrix Human Transcriptome Array (HTA 3.0v1). Multivariate linear regression was conducted separately for 376 tumor and 284 normal samples using 71 loci and 26,004 expression probes. We also developed the functional QTLs (fQTLs) method to gain pathway-level insight and evaluate the hypothesis that SNPs are associated with a specific pathway. We used the Molecular Functions (MFs) of the Gene Ontology and identified 2 SNPs associated with 2 MFs in normal tissue, 1 SNP associated with 1 MF in normal tissue, and 9 trans-acting SNPs in estrogen-receptor positive (ER+) tumors and 13 trans-acting SNPs associated ER- tumors, including 3 overlapping loci across normal adjacent, ER+ tumors and ER-negative tumors (false discovery rate (FDR)<10%). Although not statistically significant after FDR correction, we identified one putatively non-specifically associated ENR- tumors and no significant association in ER+ tumors (FDR<10%). Using bioinformatic tools we identified trans-acting loci that alter transcription-binding motifs and were associated with transcripts in chromatin enhancer regions or non-coding RNA. The overlapping loci are associated with multiple transcript isoforms in breast tissue and may be key regulatory SNPs in breast tissue. Further functional work is needed to elucidate the mechanism of these associations. Our results provide functional insights on the underlying biology of breast cancer loci in the specimen type that is most impactful in translation to clinical practice.

Breast cancer eQTLs from the Nurses' Health Study. A. Hazra,1,5, A. Quiroz-Zarate,2 B. Harshfield,3 R. Hu,4 N. Knoblauch,4 A.H. Beck,4 S.E. Hankinson1, V. Carey,4 R.M. Tamimi1,12, D.J. Hunter,1 J. Quackenbush1
1) Medicine, Brigham and Women's Hospital, Boston, MA; 2) Center for Computational Biology, Dana Farber Cancer Institute, Boston, MA; 3) Pathology, Beth Israel Deaconess Medical Center, Boston, MA; 4) Division of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA; 5) Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA.

Somatic mutations in the epidermal growth factor receptor (EGFR) gene are frequent drivers for non-small cell lung cancer (NSCLC). It has been suggested that EGFR mutation-positive (EGFRmut+) NSCLC might be a unique orphan disease with genetic susceptibility. EGFR mutations are associated with lung adenocarcinoma histology, female gender and non-smoking history. NSCLC with these features has been consistently associated with an intronic polymorphism rs2736100 in the telomerase reverse transcriptase (TERT) gene. We therefore hypothesized that rs2736100 may be a risk factor for EGFRmut+ NSCLC. To test this, we conducted a genetic association study in Chinese NSCLC patients (n=714) and healthy controls (n=2,520). Among the cases, 42% (n=303) were EGFRmut+. We further tested the association between the EGFR mutational status and mean telocyte leptomere length (TLT). The potential function of rs2736100 in lung epithelial cells was also explored. We found that while the rs2736100-C allele was significantly enriched in EGFRmut+ NSCLC patients compared to EGFRwt+ NSCLC patients (OR=1.24, 95%CI=1.1-1.39, p=4.10^-4), it was more strongly associated with EGFRmut+ NSCLC (OR=1.52, 95%CI=1.28-1.80, p=1.6×10^-4) compared to the EGFRmut- NSCLC (OR=1.07, 95%CI=0.92-1.24, p=0.4). There was also a significant difference in TERT expression between EGFRmut+ and EGFRmut- populations (OR=1.42, 95%CI=1.15-1.76, p=1.1×10^-5). The results remained significant after controlling for age, gender, smoking status and history (corrected p<0.039). Further analyses demonstrated that while NSCLC patients with EGFR mutational status have significantly longer TLT compared to healthy controls (p<10^-3), the EGFRmut+ patients have even longer TLT compared to EGFRmut- patients (p=0.008, corrected p=0.043). Meanwhile, rs2736100 was significantly associated with TERT mRNA expression in both normal and tumor lung tissues after controlling all covariates (corrected p<0.047). We also found that the rs2736100 DNA sequence has an allele-specific affinity to nuclear proteins extracted from lung epithelial cells. The polymorphism was also associated with an altered enhancer activity in vitro. Our study for the first time linked a TERT polymorphism to EGFRmut+ NSCLC. The data further revealed insight into the role of TERT and its polymorphisms in the carcinogenesis of EGFR mutation-driving lung cancer.

Cervical cancer is the most common cause of cancer mortality for women living in poverty, causing over 28,000 deaths annually in Latin America and 266,000 worldwide. We identified somatically mutated PIK3CA genes using exome and ultra-deep targeted sequencing of invasive cervical tumors from Guatemala, Venezuela, and Mexico. Analysis of 636 HPV positive cervical cancer revealed activation of PIK3CA genes in up to 34% of tumors. The novel PIK3CA mutations including del617/617 (ABD domain); PIK3T (link between ABD and RBD domains); T229I and M232L (RBD); V952G and E1034Q (Kinase domain) and mutations, D350N, D454Y (C2) and E726K (Kinase domain) lacking function study, were identified in the study. To future study the effect of these mutations on PIK3CA catalytic activity, we expressed wild type and mutant constructs in vitro and measured levels of AKT in both normal and tumor tissues after controlling all covariates. The result demonstrates that novel PIK3CA mutations led to significantly and slightly increased levels of phospho-AKT (Thr308) and phospho-AKT (Ser473), respectively. Furthermore PIK3CA and HPV E6 and E7 gene expression is significantly differentially expressed in HPV positive tumors. This study demonstrates that PIK3CA/PIK3AKT pathway activation contributes to the development of HPV positive cervical cancer in Latin America.
Carp Is a Potential Tumor Suppressor in Gastric Carcinoma. F. Lu, J. Xue, Y. Hu, L. Gan, H. Yang, Y. Wei. 1) Center for Molecular Genetics, Sichuan Provincial People's Hospital, chengdu, Sichuan, China; 2) State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, P.R. China.

Background: The caspase-associated recruitment domain-containing protein (CARD) is expressed in almost all tissues. Recently, the tumor-suppressive function of CARD was discovered and attracting increased attention. This study aimed to investigate the role of CARD in the carcinogenesis of human gastric carcinoma. Methods/Result: Compared with normal gastric tissue, the downregulation of CARD expression was observed in gastric carcinoma tissue by cDNA array and tissue microarray assay. In vitro, the gastric carcinoma cell line (BGC-823) was stably transfected with pcDNA3.1-B-CARD or plus CARD siRNA, and we used MTT, flow cytometry, cell migration on type I collagen, cell-matrix adhesion assay and western blot analysis to investigate the potential anti-tumor effects of CARD. Our data showed that overexpressing CARD suppressed the malignancy of gastric carcinoma BGC-823 cell line, including significant increases in apoptosis, as well as obvious decreases in cell proliferation, migration, adhesion ability, and tumor growth. The tumor-suppressive effects of CARD were almost restored by siRNA-directed CARD silence. In addition, overexpression of CARD induced G1 arrest, decreased the expressions of cyclin E and CDK2, and increased the expressions of p27, p53 and p21. In vivo, the tumor-suppressive effect of CARD was also verified. A single-nucleotide polymorphism (SNP) genotype of CARD (rs2297882) was located in the Kozak sequence of the CARD gene. The reporter gene assay showed that rs2297882 TT caused an obvious downregulation of activity of CARD gene promoter in gastric carcinoma cells. Furthermore, the association between rs2297882 and human gastric carcinoma susceptibility was analyzed in 352 cases and 889 controls. It displayed that the TT genotype of rs2297882 in the CARD gene was associated with an increased risk of gastric carcinoma. Conclusions: CARD is a potential tumor suppressor of gastric carcinoma, and the rs2297882 C>T phenotype of CARD may serve as a predictor of gastric carcinoma.

Association of Platelet Derived Growth Factor-B (PDGF-B) and Human Epidermal Growth Factor Receptor -2 (HER-2/neu) Single Nucleotide Polymorphisms (SNP}s) with Gallbladder Cancer (GBC). K. Mishra, V.K. Kapoor, A. Behari, S. Khatri, S. Agrawal. 1) Surgical Gastroenterology, Sanjay Gandhi Post-Graduate Institute of Medical Sciences (SGPGIMS), Lucknow, Uttar Pradesh, India; 2) Integral University, Kursi Road, Lucknow, Uttar Pradesh, India.

Purpose: Gall bladder cancer (GBC), a highly malignant gastrointestinal tumor, is very common in north India. Single nucleotide polymorphisms (SNPs) in PDGF-B and HER-2/neu-single nucleotide polymorphisms (SNPs) in PDGF-B and HER-2/neu were found to play an important role in human tumorigenesis. PDGF-B and HER-2/neu overexpression has been found in many cancers. We studied PDGF-B and HER-2/neu single nucleotide polymorphisms (SNPs) in GBC and gastric associated benign diseases viz. chronic cholecystitis (CC) and xanthogranulomatous cholecystitis (XG). Methods: DNA was extracted from blood in patients with GBC (n=195), CC (n=140), XGC (n=47) and normal controls (n=300). PDGF-B polymorphisms were investigated using ARMS PCR for +286A/G and +1135C/A, and for Her-2/Neu Ile656Val by PCR-RFLP method. Results: +286A/G polymorphism homozygous GC genotype, and +286G allele was found to be risk associated for GBC (OR=5.25; P=0.001) and GC genotype (OR=5.25; P=0.001). Recesive model (GG vs. AA+GA) of +286A/G polymorphism was risk associated (OR=4.78; P=0.001) whereas dominant model (AA vs. GG+GA) was risk protective (OR=0.56; P=0.003) with GBC. +1135A/C polymorphism CC genotype and +1135C allele was risk associated (OR=3.19, P=0.0001 and OR=1.81, P=0.0001) with GBC. In recessive model (CC vs. AA+AC) was risk associated (OR=2.75, P=0.0003) whereas dominant model (AA vs. CC+AC) showed protective association (OR=0.56, P=0.0024) with GBC. In Her-2 Ile656Val polymorphism, dominant model and Val allele were risk associated. XGC genotype frequencies of GBC were compared with CC and XGC. Homozygous GC genotype, recessive model and +286G allele of +286A/G genotype were risk associated but dominant model of this genotype was risk protective for GBC as compared with CC. In case of GBC vs XGC, GG genotype, recessive model and +286G allele were risk associated for GBC. In +1135A/C polymorphism, homozygous CC genotype, recessive model and +1135C allele were risk associated whereas dominant model was risk protective when GBC was compared with CC. In HER-2 Ile656Val polymorphism, Val allele was risk associated when GBC was compared with CC. In haplotype analysis of the SNPs, ACAC, GAlle (OR=1.48), GVal (OR=1.70) and GAAlle (OR=2.00) were risk associated with GBC. Conclusion: PDGF-B +286A/G and 1135A/C are risk susceptibility markers for GBC and need further evaluation.
3410S
1) Institute of Human Genetics, Medical University Graz, Graz, Austria; 2) Wellcome Trust Centre for Human Genetics, University of Oxford, UK; 3) Department of Zoology, University of Oxford, UK.
Germline mutations in the exonuclease domain of genes encoding the catalytic subunits Polymerase δ and ε, i.e. POLD1 and POLE, predispose to "polymerase proofreading associated polyposis" (PPAP) resulting in multiple colorectal adenomas and carcinoma with high penetrance and dominant inheritance. Moreover, somatic mutations in the ED of POLE have been frequently found in sporadic CRCs and endometrial carcinoma. Tumors, with both inherited and somatic EDMs, were microsatellite stable and showed a mutator phenotype with a dramatic increase of base substitutions indicating impaired proofreading. To assess the functional consequences of POLE germline and somatic mutations, we characterized exonuclease domain mutant alleles in S.pombe. We therefore generated constructs encoding the equivalent changes in the fission yeast protein and determined the effect of this change on reversion of the ade6-485 allele, 5-fluoroorotic acid and canavanine resistance. The somatic mutations including P286R, S297F, and S459F showed a dramatically increased mutation rate (up to more than 100-fold) compared to the wild type strain. V411L was only slightly increased (4-7 fold). Also the germline variant L424V showed increased mutation rates (2-11-fold) compared the wild type strain. As expected whole genome sequencing of POLE mutated strains and strains with a deficient MMR background revealed an elevated rate of base substitutions. This data indicate that EDMs in POLE indeed lead to proofreading deficiency and can induce replication errors during synthesis of oncogenes and tumor suppressors resulting in tumor formation.

3411M
From GWAS to therapy: Fatty acid synthase in uterine leiomyomata. Z. Ordulu1, M. Hayden1, S. Eggert2, M. Shinohara3, C. Serhan4, B.J. Quade2, C.C. Morton1,2,1. 1) Obstetrics, Gynecology and Reproductive Biology, Brigham and Women’s Hospital, Boston, MA; 2) Pathology, Brigham and Women’s Hospital, Boston, MA; 3) Anesthesiology, Perioperative and Pain Medicine, Brigham and Women’s Hospital, Boston, MA.
Uterine leiomyomata (UTL) pose a major public health problem given their high prevalence (>70%) in women of reproductive age and their indication for >200,000 hysterectomies in the U.S. annually. Genome-wide association studies for fibroid predisposition performed in white women identified a candidate SNP (rs4247357) under a linkage peak in 17q25.3 that spans >200,000 hysterectomies in the U.S. annually. Genome-wide association studies, University of Tabriz, Tabriz, Iran; 3) Department of Medical Genetics, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran; 4) Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran; 5) Cell and Molecular Biotechnology Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran; 6) Department of Biology, Faculty of Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.
Objective: Uterine leiomyoma is one of the most lethal form of human malignancies, mostly due to its late diagnosis. Goelastan province in Northeastern Iran, as a part of Caspian littorals, has been identified as one of the highest risk areas in the world. This high incidence, highlighting the necessity of identifying novel and potent diagnostic and prognostic biomarkers for early detection of the disease. Alterations in the expression of miRNAs have been widely reported in numerous diseases including almost all types of cancers. Act as oncogenes (oncomiRs) or tumor suppressors, miRNAs are playing prominent roles in cancer-related processes such as proliferation, apoptosis, metastasis and angiogenesis. Motivated by our recent data on downregulation of miR-338-3p expression in esophageal squamous cell carcinoma, we aimed to investigate its potential causative role in tumorigenesis of an Esophageal Carcinoma cell line, KYSE-30. Materials and Methods: The human KYSE-30 cell line was obtained from Pasteur Institute in Tehran, Iran. After seeding the cells, they were transfected with either pEGFP-C1 vector containing the miR-338-3p precursor sequence or the mock pEGFP-C1 vector with Lipofectamin 2000 transfection reagent. 48 h after transfection, the cells were harvested for RNA isolations and flow cytometry analysis. Experiments were repeated at least twice and cell cycle alterations were analyzed using FlowJo software. Results: In contrast to the cells transfected with a mock vector, as a negative control, the KYSE-30 cells overexpressing miR-338-3p showed a remarkable increase (15.5 times) in miR-338-3p expression. miR-338-3p caused a dramatic alteration in cell cycle distribution of transfected cells, including around two folds increase in the number of cells distributed in the sub-G1 phase of the cell cycle. Conclusion: The data suggests that miR-338-3p may act as an anti-oncogene in squamous cell carcinomas. A similar observation has been reported in other cancer cells such as pancreatic interepithelial neoplasia, gastric cancer and colorectal carcinoma. A functional causative role of miR-338-3p in ESCC is also supported by bioinformatic analyses using miRwalk and DIANALAB. These softwares predicted MMP2, MMP9, SMO, n-Ras, c-Myc and Cyclin D as potential targets of this miRNA. These findings and further studies may suggest miR-338-3p as a good biomarker for ESCC diagnosis and treatment target.

3412T
Evaluation of miR-338-3p role in the progression of Esophageal Squamous Cell Carcinoma. H. Mollaei1,2, M. Shafiee1,2, S.A. Aleyasin3, S.J. Mowia4, M. Moghaddam-Malin4,5, †The two first authors (M. Shafiee and H. Mollaei) contributed equally to this work. 1) Golestan Research Center of Gastroenterology and Hepatology (GRCGH), Golestan University of Medical Sciences, Gorgan, Iran; 2) Department of Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran; 3) Department of Medical Genetics, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran; 4) Department of Molecular Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran; 5) Cell and Molecular Biotechnology Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran.
Invasive esophageal squamous cell carcinoma (ESCC) is the most lethal form of esophageal malignancy, mostly due to its late diagnosis. Goelastan province in Northeastern Iran, as a part of Caspian littorals, has been identified as one of the highest risk areas in the world. This high incidence, highlighting the necessity of identifying novel and potent diagnostic and prognostic biomarkers for early detection of the disease. Alterations in the expression of miRNAs have been widely reported in numerous diseases including almost all types of cancers. Since ESCC is a well-known multifactorial disease, miRNAs are suggested as a potential diagnostic and therapeutic target for ESCC. miR-338-3p is a well-known oncogene with antiapoptotic, angiogenic, and metastatic properties. Several studies have reported the association of miR-338-3p with ESCC progression and chemotherapy. Moreover, miR-338-3p has been shown to have a role in regulating the progression of ESCC by targeting cyclin D and down-regulating RAS/MAP kinase pathway. In the present study, we aimed to investigate the role of miR-338-3p in ESCC progression and chemoresistance by using cell lines. Materials and Methods: We transfected human KYSE-30 cell line with miR-338-3p precursor and control (mock) vector. The transfected cell lines were treated with 5-fluorouracil (5-FU) for 48 h, followed by analysis of cell viability and cell cycle alteration using the Cell Counting Kit-8 (CCK-8) and Flow cytometry, respectively. Further, the role of miR-338-3p in chemotherapy resistance was assessed by luciferase reporter assay. Results: The data showed that overexpression of miR-338-3p led to a significant decrease in cell viability and cell cycle arrest in the sub-G1 phase of the cell cycle. Conclusion: Our results suggest that miR-338-3p may act as an anti-oncogene and chemosensitizer in ESCC. Further studies are recommended to investigate the mechanism of the anti-oncogenic property of miR-338-3p and its potential as a therapeutic target for ESCC.
Assessment of the clinical relevance of variants of uncertain significance in BRCA2 by functional and computational approaches. L. Giuliani,1 N.M. Lindow,1 V.S. Pankratz,1 D.L. Masica,1 R. Karchin1, F.J. Couch1, F.J. Couch1, 1) Lab Med and Path, Mayo Clinic, Rochester, MN; 2) Health Sciences Research, Mayo Clinic Arizona, Scottsdale, AZ; 3) Institute for Computational Medicine, Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD.

Genetic testing of individuals with a family history of breast and/or ovarian cancer has led to the identification of many unique BRCA2 missense variants of uncertain significance (VUS). Current methods for classification of BRCA2 mutations depend heavily on the availability of information on segregation of VUS with breast and ovarian cancer within families, and on predicted probabilities of pathogenicity based on evolutionary sequence conservation. The failure to assess the clinical relevance of VUS, owing to insufficient genetic information, may deprive patients of the benefits of risk assessment and enhanced risk management available to those found to carry pathogenic mutations in BRCA2. In order to classify greater numbers of VUS that may or may not have available family data for segregation analysis, we have developed a homology directed DNA repair (HDR) assay for evaluation of the pathogenicity of BRCA2 VUS. This assay measures the influence of VUS in the DNA binding domain of BRCA2 on the HDR activity of the protein and has high sensitivity and specificity for established pathogenic missense mutations. Here we report on the assessment of 106 VUS located inside the BRCA2 DNA Binding Domain (DBD) (amino acid 2459 to 3190) using this assay. Whereas 47 variants had a 99% probability of being non-pathogenic, and were classified as non-pathogenic, 41 variants had a 99% probability of being pathogenic, and were classified as pathogenic by the HDR assay. The results from the assay were subsequently compared to predictions from the Align-GVGD sequence based prediction model (r=0.49). The categories of Align-GVGD associated with high probabilities of pathogenicity (p=0.66 to 0.81) were significantly different from the results of the HDR assay (p=0.05). In contrast, the Phenotype Optimized Sequence Ensemble (POSE) sequence model, trained on functional properties, displayed a moderate to strong correlation with the HDR assay (r=0.69). Overall, the results suggest that the prior probability of pathogenicity assigned to a number of VUS by Align-GVGD and POSE were significantly lower than their corresponding HDR probabilities, suggesting that these methods may underestimate the probability of pathogenicity of some VUS. This study suggests that HDR is a useful method for selecting BRCA2 VUS for further evaluation by functional or family-based methods.

The role of the cilia protein Arl13b in activated-Smoothed medulloblastoma oncogenesis. S.N. Bay1,2, R.C. Castellino,2, T. Caspary,2 1) Genetics and Molecular Biology Program; 2) Dept of Human Genetics; 3) Dept of Pediatrics, Emory University, Atlanta, GA.

The proper balance of Shh signaling is critical to both development and adult homeostasis. Too little Shh signaling causes birth defects such as holoprosencephaly, while too much Shh signaling results in tumors, as occurs repeatedly in Gorlin syndrome patients. Indeed, overactive Shh signaling in a basal cell carcinoma and medulloblastoma - a tumor of the cerebellum that is the most common pediatric malignancy in the central nervous system. Many available treatments are inadequate; surgery, chemotherapy, and radiation are invasive and can cause long-term cognitive defects in patients, and current molecular therapies are prone to failure due to tumor resistance. The drawbacks of current treatments mean that new approaches are needed. Recent research has shown that Shh-derived medulloblastomas are “addicted” to cilia since cilia are required for proper Shh signaling. We study a ciliary GTPase called Arl13b that uniquely regulates Shh signaling, both upstream of Smoothened (Smoo) and between Smoo and Gli activation. Loss of Arl13b results in ligand-independent constitutive low-level pathway activation but prevents maximal signaling. Due to this unique relationship to Shh signaling, we predict that the loss of Arl13b will reduce the high levels of pathway output caused by constitutive activation of Smoo. To determine whether loss of Arl13b can prevent or delay tumor formation, we deleted Arl13b in a mouse model of medulloblastoma. We also use mouse embryonic fibroblasts to define how the loss of Arl13b affects Smoo localization and Shh pathway output in an activated-Smo signaling context. Together, these experiments will define whether Arl13b is a viable target for molecular therapies for the treatment of medulloblastoma and will reveal mechanistic details of Arl13b regulation of Shh signaling and a disease-relevant activated Shh signaling context.

Modeling cancer in zebrafish embryos. L. Francescato, N. Katsanis, Center for Human Disease Modeling, Duke University, Durham, NC.

Zebrafish has emerged as an invaluable model organism to study cancer. The capacity for rapid gestation time, the lack of an adaptive immune system during early development, and zebrafish embryos provide an in vivo model that is cost and time efficient, and can be used as a combinatorial immunotherapy for cancer. Model organism attractive to research experimentation, such as external embryonic development, transparency, high fecundity rate, and short generation time. We have successfully transplanted different types of cancer cells into the zebrafish embryo. The feasibility of this model prompted us to ask whether we could take primary cancer cells from patients and test them for tumorigenicity. We have successfully observed proliferation and migration of colon cancer cells in two patients, further confirming the potential utility of zebrafish in studying cancer. Using this model of cancer, we have undertaken a two-pronged approach that harnesses the power of this system. First, we have compared the exomes of distally attached human cells to non-migrating cell populations as a surrogate to identify candidate metastatic drivers. In parallel, we have initiated a small molecule screen to identify potential drugs that can inhibit either proliferation or migration of cancer cells. Zebrafish cancer xenotransplants provide an in vivo model that is cost and time efficient, and can be a first line of evidence for anticancer therapies.

Lifestyle issues of BRCA mutation carriers that may affect and health outcomes. A. Caceros, R. McMclaren, O. Ivanov, K. Wiercinski, C. Buffington, Florida Hospital Celebration Health, Celebration, FL.

The failure to assess the clinical relevance of VUS, owing to insufficient genetic information, may deprive patients of the benefits of risk assessment and enhanced risk management available to those found to carry pathogenic mutations in BRCA2. In order to classify greater numbers of VUS that may or may not have available family data for segregation analysis, we have developed a homology directed DNA repair (HDR) assay for evaluation of the pathogenicity of BRCA2 VUS. This assay measures the influence of VUS in the DNA binding domain of BRCA2 on the HDR activity of the protein and has high sensitivity and specificity for established pathogenic missense mutations. Here we report on the assessment of 106 VUS located inside the BRCA2 DNA Binding Domain (DBD) (amino acid 2459 to 3190) using this assay. Whereas 47 variants had a 99% probability of being non-pathogenic, and were classified as non-pathogenic, 41 variants had a 99% probability of being pathogenic, and were classified as pathogenic by the HDR assay. The results from the assay were subsequently compared to predictions from the Align-GVGD sequence based prediction model (r=0.49). The categories of Align-GVGD associated with high probabilities of pathogenicity (p=0.66 to 0.81) were significantly different from the results of the HDR assay (p=0.05). In contrast, the Phenotype Optimized Sequence Ensemble (POSE) sequence model, trained on functional properties, displayed a moderate to strong correlation with the HDR assay (r=0.69). Overall, the results suggest that the prior probability of pathogenicity assigned to a number of VUS by Align-GVGD and POSE were significantly lower than their corresponding HDR probabilities, suggesting that these methods may underestimate the probability of pathogenicity of some VUS. This study suggests that HDR is a useful method for selecting BRCA2 VUS for further evaluation by functional or family-based methods.
3418T
Accurate and inexpensive sequencing of BRCA1 and BRCA2: Application to a large US-wide study of breast cancer in Latinas. M. Dean1, J.F. Boland2, M. Yeager2, K.M. Im3, J. Mitchell2, D. Roberson2, K. Jones2, J. Sawitzke3, S. Bass3, X. Zhang4, H. Lee5, R. Eggebeen6, V. Robles2, C. Hollis5, C. Barajas5, E. Rath6, C. Arentz7, J.A. Figueroa8, D.D. Nguyen9, Z. Nathelh7, 1) Lab Experimental Immunology, NCI-FCRDC, Frederick, MD; 2) Cancer Genetics Research Laboratory, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD; 3) Basic Science Program, Leidos Biomedical Research, Inc., Frederick, MD; 4) Nueva Vida, Richmond, Richmond, Virginia; 5) Latino Community Development Agency, Oklahoma City, OK; 6) Texas Tech University Health Sciences Center, El Paso, TX; 7) Texas Tech University Health Sciences Center, Lubbock, TX.

Germline mutations in the BRCA1 and BRCA2 genes are the most common cause of inherited breast cancer and are found in 5-10% of all cases. Detection of BRCA mutation carriers can lead to therapeutic interventions such as mastectomy, oophorectomy, hormonal prevention therapy and improved screening; as well as targeted therapies such as PARP-inhibition. African Americans and Hispanics are 3-4 times less likely to receive BRCA screening, despite having approximately the same mutation frequency as non-Jewish European Americans. Furthermore, both of these minority groups have a higher mortality for breast cancer. To begin addressing this health disparity, we initiated a nationwide trial of BRCA testing of Latinas with breast cancer. Patients were recruited through community organizations, clinics, public events, and through the mail and internet. Subjects completed the consent process and questionnaire over the phone, and/or email contact, and provided a saliva sample by mail. DNA from 117 subjects was extracted from saliva and used to sequence the entire BRCA1 and BRCA2 coding regions and splice sites using a community designed panel at a total supply cost of $80/subject. A newly developed, more accurate polymerase (HQ) allowed for the detection of mutations and elimination of many false positive results. Subjects ranged in age from 23-81 years (mean of 51 years), 6% had bilateral disease, 57% were ER+PR+, 23% HER2+; and 17% had triple-negative disease. A total of 6 different predicted deleterious mutations, one newly described and the rest rare, were identified along with 4 variants of unknown effect. An alignment of all available primate BRCA1 and BRCA2 sequences indicates an oncogenetic strategy on a larger scale could lead to improved cancer care of minority and under-served populations.

3419S
How choriocarcinoma DNA identification can interfere in treatment decision? A report of two unexpected cases. G.JF. Gattas1, P. Exmar2, VD. Cantagalli1, FT. Gonçalves1, FN. Aguilar1, REO. Ramos2, MDPE. Diz2, 1) Dept of Legal Medicine, Bioethics and Occupational Health, Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil. Choriocarcinoma is a rare, highly malignant tumor, mainly with gestational origin, and rarely germ cell origin. The clinical presentation and histological characteristics of these tumors are identical but the origin differentiation is critical for treatment decision. DNA polymorphisms (STRs) used for forensic purposes can be useful to identify paternal cells in paraffin wax embedded choriocarcinoma tumor cells, compared to patient's healthy cells, to characterize gestational origin. We here describe two patients that were evaluated after histologic diagnosis of choriocarcinoma. The first patient was a 21 years old woman with previous abortion history and clinical presentation extremely suggestive of gestational disease. After no favorable treatment results the tumor DNA analysis, using MiniFiler Kit (Applied Biosystems), was realized and revealed identical profile when compared to the patient's blood sample. Thus, the treatment was changed for BEP (Bleomycin, Ethoposide and Cisplatin) that is considered the first line treatment for an ovarian germ cell tumor. The patient achieved pathologic complete response, with no disease evidence in 2 year follow-up. The second case was a 54 years old woman with initial presentation of metastasis within the pulmonary artery and absence of other disease site with biopsy confirmation as choriocarcinoma. The main hypothesis was gestational origin and she was firstly treated with BEP with complete response and no serum marker response. DNA STRs showed the presence of additional genetic material, beyond the maternal one for all 8 STRs evaluated, including Y chromosome. The results indicated paternal contribution (confirmed with the husband DNA) and the treatment was changed to dacarbanomycin, and patient achieved complete radiologic and of serum marker response. These two cases suggest that DNA STRs could be used as a necessary tool to confirm the gestational origin of choriochoriocarcinomas even when the clinical case is not indicative of it. Financial Support: LIN-40 (HC-FMUSP), FAPESP (09/54868-6).

3420M

Background – 15% of ovarian cancer (OC) patients carry a BRCA mutation, yet historically are under-referred to Cancer Genetics (CG) services. NICE (2013) recommended that BRCA testing be offered to individuals with ≥10% chance of mutation detection. This and development of targeted agents e.g PARP inhibitors make it essential that OC patients can access genetic testing within an appropriate time frame. We developed an ‘oncogenetic model’ to facilitate testing in routine oncology care. Methods Patients with non-mucinous OC ≤55 years at Royal Marsden Hospital were offered BRCA gene testing at their oncology appointment by cancer team members who had completed 30 min online training. The test results were interpreted by genetics and returned to the patient by the cancer team. All mutation carriers attended a Genetics appointment for detailed discussions and to organise testing for relatives. Any patient could contact genetics at any time, if required. Patients and clinicians were sent a questionnaire to assess their experience. Results 119 OC patients were tested in 6 months. Patients were tested during initial treatment (n=32), relapse (n=45) or follow-up (n=42). 20 (17%) carried a mutation: 8 BRCA1, 12 BRCA2. 12 had no family history, 8 met current genetic referral criteria. BRCA results changed management in 9/20 (45%) carriers. No patient requested additional Genetics input before testing. Questionnaires were sent to 71 patients; 63/57 responses were received. All were pleased to have had testing, and to have accessed testing within Oncology. All understood the results potentially had implications for themselves and their family. 25 clinicians undertook training. All welcomed the opportunity to offer testing and felt confident of the process and results. Conclusion: The Oncogenetic testing pathway provides flexible, patient-centred, equitable, high-throughput gene testing with considerable time and cost savings compared to traditional pathways. It is now the standard BRCA testing pathway for sequence variants. 120 OC patients at Royal Marsden and roll-out to other NHS centres is planned.

3421T
Inhibition of STAT3 and RelA expression levels Bortezomib treated K-562 leukemic cells and induction of apoptosis. N. Selvi Günel1, B. Taşcanlı Kaymak1, S. Kıpçak1, A. Dalmizırak1, B. Kosova1, G. Saydam2, 1) Department of Medical Biology, Ege University, Izmir, Turkey; 2) Department of Hematology, Ege University, Izmir, Turkey.

Signal transducer and activator of transcription (STAT) proteins; especially STAT3, are crucial for signalling pathways leading to discover the underlying cancer development mechanism. Activated STAT3 increases leukemic cell proliferation and survival; also suppresses anti-tumour immunity and apoptosis by promoting pro-oncogenic pathways, including nuclear factor-xB (NF-xB). RelA is one of the gene products of NF-xB that gives rise to increase in STAT3 activation. Bortezomib (BOR) is a first class of proteasome inhibitor, and used in of multiple myeloma and mantle cell lymphoma treatment. In this study, we aimed to identify the cytotoxic and apoptotic effects of BOR upon chronic myelogenous leukemia (CML) cell model K562, with determining STAT3 and RelA expressions both at mRNA and protein levels. Cell proliferation was assessed by XTT assay in order to determine cytotoxicity of BOR upon leukemic cells. While mRNA expression levels of STAT3 and RelA were analyzed by qRT-PCR; protein expressions were detected via western-blot method. IC50 was calculated as 17.7 ±m for 72th hour. Number of apoptotic cells were increased by 37% and 38% for 72th-96th hours (p=0.0041, p=0.0038). While STAT3 mRNA expression was significantly decreased by 66% [2.86 fold; p=0.0043] at 72th hour. Rel A was downregulated by 45.8% and 83.6% [1.83 fold; p= 0.003], [6.1 fold; p=0.0022] for 72th and 96th hours. As for protein results, both Rel A and STAT3 protein expressions were highly inhibited in a time dependent manner.
3422S


Introduction: Genomic profiling of cell-free DNA (cfDNA) through a simple blood test is non-invasive, sequential and can potentially be more predictive than tumor profiling. Due to high concordance of tumor and cfDNA, this blood test can be considered when biopsies are not an option due to the risk, cost, or time. Methods: We used single-molecule digital sequencing technology, Guardant360 for high-fidelity and comprehensive profiling of 54 actionable genes in cfDNA of more than 500 metastatic cancer patients. This technology allows concurrent profiling of somatic mutations, gene amplifications and fusions. The longitudinal blood samples from patients have been processed to assess residual disease post surgical resection or targeted treatments. Results: The overall detection rate of somatic alterations in cfDNA approached 90% for all indications (breast, lung, colorectal, melanoma and prostate). When blood samples were concurrent with biopsies, the concordance of alterations was 93%. Residual disease in 30 colorectal and melanoma patients post multi-surgical resection has been studied (2-7 draws). The alterations were positively correlated with clinical status of the patients. In 70 colorectal patients, plasma samples post targeted treatment were processed. A combination of previously reported and novel resistance-related alterations were found in cfDNA from the post targeted treatment samples. Conclusions: Comprehensive sequencing of patient’s cancer in real-time through simple blood test can empower oncologists in making more informed treatment decisions, especially when tissue biopsy is not an option.

3423M


Introduction. Constitutive mismatch repair deficiency syndrome (CMMR-D) is a recently described childhood cancer predisposition syndrome involving biallelic mutation of MMR genes (MLH1, MSH2, MSH6 and PMS2). More than 140 cases have been previously reported but only as case reports. Methods. We performed a retrospective review of all 31 cases of CMMR-D from 23 families diagnosed in French genetics laboratories in order to characterize clinical characteristics, treatment and outcomes, and biological diagnosis data of an unselected series of patients. Results. Overall, 67 tumors were diagnosed in these 31 patients, 17 (25%) hematologic malignancies, 22 (33%) brain tumors, 25 (37%) Lynch syndrome-associated malignancies, and 13 (5%) other tumors. Median age of onset of first tumor was 6.98 years [1.23-33.53], 23 (74%) patients had NF1-unrelated CALMs or hypopigmented macules and 4 (13%) had brain malformative features. Colorectal adenomas were found in all 16 patients who have had colonoscopy with synchronous colorectal cancer or advanced adenoma in all of them. Overall, 18 patients died, 7 (39%) due to the primary tumor. Median survival after diagnosis of the primary tumor was 23 months [0.26-213.2]. Among the patients who survived after their first malignancy, 20 developed a second malignancy. No obvious excess of toxicity to treatment was reported. A familial history of LS-associated cancer was found in only 15% of cases, consanguinity in 43% of cases. PMS2 mutations (18 patients) were more frequent than mutations of MLH1 (4 pts), MSH2 (3 pts) and MSH6 (6 pts). Conclusion. CMMR-D is a severe condition associated with multiple malignancies in childhood. Its rarity warrants international collaboration to define diagnosis criteria and guidelines for surveillance and prevention in order to decrease tumor-related mortality.
Germline TP53 mutation analysis in HER2-positive breast cancer patients from Southern Brazil. M. Fitarelli-Kiehl1, 2, A. Machaj1, A. Grudenz1, P. Ashton-Prolla1,2,3,3. 1) Experimental Research Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; 2) Department of Pathology, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; 3) Department of Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. TP53 mutations were associated with Li-Fraumeni Syndrome (LFS), an autosomal dominant disorder characterized by predisposition to multiple early-onset cancers including breast cancer (BC), the most prevalent tumors among LFS women. Recent reports show that TP53 mutation carriers have a significantly higher likelihood of developing a BC overexpressing HER2 (63-83%) when compared to mutation-negative BC cases (16-25%). Germline TP53 mutations are estimated to occur in 1 to 8.6% of patients with early onset BC unselected for family history of cancer, and HER2-positive status supposedly increases the odds of having a germline TP53 mutation by nearly 7-fold. In this work, we determined the prevalence of germline TP53 mutations in a cohort of women with HER2 overexpressing BC, diagnosed before age 60, with a positive family history of cancer. We retrospectively reviewed the immunohistochemistry data of BC from patients diagnosed at Hospital de Clínicas de Porto Alegre between years 2007 and 2012, and 106 women who fulfilled inclusion criteria were recruited and periferic blood was collected. Germline TP53 sequence variants were identified in 4 cases by next generation sequencing (NGS), and additionally sequencing of the entire coding sequence and flanking intronic regions. Among 106 patients analyzed, 2 (1.9%, 95% CI 0.5-6.6%) were found to carry a TP53 mutation with known pathogenic effect, and other 5 patients (4.7%, 95% CI 2.0-10.5%) showed sequence variants with unknown effect on p53 function (all variants with MAF<0.01). One patient with pathogenic mutation met classic LFS criteria. Although we found a low prevalence of germline TP53 pathogenic mutations in HER2 overexpressing BC patients, TP53 testing should be considered in these cases even in the absence of family history of cancer.

Further Defining the Polyposis Phenotype Associated with PTEN Mutations. L. Panos1, E. Weltmer2, H. LaDuca3, R. McFarland1, E. Chao1,2,3. 1) Ambry Genetics, Aliso Viejo, CA; 2) University of California, Irvine, School of Medicine, Irvine, CA. PTEN hamartoma tumor syndrome (PHTS) is associated with an increased risk for colon polyposis, particularly hamartomatous polyps. With the advent of multi-gene panel testing, PTEN mutations have been identified in patients with other poly histologies, potentially widening the spectrum of gastrointestinal disease burden in these patients. We sought to define the poly spectrum amongst PTEN mutation carriers identified by multi-gene panels. In a review of over 14,000 results of five multi-gene panels (BRCApulus, BreastNext, ColoNext, OvaNext, and CancerNext) that include mutation carriers identified by multi-gene panels, we found that 47.8% (10/21) of patients with PTEN positive cases presented with hamartomatous polyposis and 14% (3/21) presented with non-malignant breast tissue. Blood PTEN/pAKT protein levels and CC score were significant predictors of germline PTEN mutation status. 43 female adults had germline PTEN mutation; pathology materials were available for 25. Most breast cancers had ductal histology (80%), and were ER/PR+ (64%), AR+ (75%) and HER2+ (14%). Apocrine features, atypical apocrine adenosis, and atypical ductal hyperplasia were found more often in background breast tissue of PHTS patients compared to published frequencies. Germline PTEN mutation carriers would help further delineate the molecular profile in these PHTS patients were comparable to general population patterns and not consistent with a molecular apocrine profile. Distinctive PTEN-associated features were noted in surrounding non-malignant breast tissue. Blood PTEN/pAKT protein levels and CC score were significant predictors of germline PTEN mutation, and if replicated, could be applied in a clinical setting to rapidly identify breast cancer patients who may benefit from genetics referral.
3428S
Mutation and uncertain variant findings in ethnic minority patients undergoing multi-gene panel testing for cancer risk assessment at a safety-net hospital. C. Dunlap, V. Flower4, E.C. Chao1,2, R. Brown1,3, J.R. Morris1,2, R. La Duch1,2, E.C. Chao1,2, R. Gerdes1,2, M. El-Bahrawy2, R. Brown1,3, J.R. Morris1,2, E.C. Chao1,2, E.C. Chao1,2, E.C. Chao1,2, E.C. Chao1,2, E.C. Chao1,2

3430T
DNA methylation profiling to assess pathogenicity of BRCA1 unclassified variants in breast cancer. K. Flower1, N.S. Shenker1, M. El-Bahrawy2, D.E. Goldberg3, M.T. Parsons4, A.B. Spugnini4, J.R. Morris1,2, J.M. Flanagan1, KConFab Investigators, AFFECT study group. 1) Epigenetics Unit, Department of Surgery and Cancer, Imperial College London, London, UK; 2) Department of Histopathology, Hammersmith Hospital, Imperial College London, UK; 3) Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA; 4) QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia; 5) Genome Stability Unit, School of Cancer Sciences, University of Birmingham, UK; 6) Section of Molecular Pathology, Institute of Pathology, Lyon, France.

3429M
Majority of PTEN mutations identified on multi-gene panel tests are in non-classic patients: Expanding clinical phenotype or incomplete diagnostic criteria? E.C. Weltemer1, L. Palacios2, R. Joseph3, S. Zelcer1, R. Lanucci4, V. Vichi3, E.C. Chao1,2,3

3431S
Overexpression of MicroRNA-200c predicts poor outcome in patients with PR-negative breast cancer. K. Luostan1, M. Tuowania1, Y. Soini1, V. Katakas1,2, V.-M. Kosma1, A. Mannermaa1, J.R. Morris3,4, M. El-Bahrawy2, R. Brown1,3, J.R. Morris1,2, E.C. Chao1,2, E.C. Chao1,2, E.C. Chao1,2, E.C. Chao1,2

---

**Posters: Cancer Genetics**
Primary myelofibrosis (PMF) is a myeloproliferative condition characterized by the clonal proliferation of the hematopoietic precursors and progressive development of bone marrow fibrosis. This stromal alteration is an important clinical issue and specific prognostic markers are not yet available.

JAK2, MPL and CALR genes are frequently mutated in PMF and are mutually exclusive. Mutations of TET2, ASXL1, DNMT3A and IDH1/2 are present in about 20% of PMF cases. It is intriguing that these loci are involved in the epigenetic regulation of cells and are associated with epigenetic alterations in cancer, suggesting that epigenetic defects could also be present in PMF. In 65 bone marrow biopsies from 56 PMF patients, stratified by clinical classification and JAK2/CALR and IDH mutations, we explored the methylation pattern of genes encoding for cytokines involved in the stromal reaction: PDGF, TGFβ and FGF2. We also evaluated the methylation profile of the LINE-1. The methylation was analyzed using the pyrosequencing approach. In agreement with the European consensus of bone marrow fibrosis grading, the cases were categorized at diagnosis as cellular phase (MF-0 n=12) and fibrotic phases (MF-1 n=30; MF-2 n=18 and MF-3 n=5). The follow-up ranged from 6 to 287 months (average 76 months). As controls, we included a small number of fibrotic cases from patients who underwent surgery for iliac bone prosthesis. We found that PDGFB, FGF2 and LINE-1, but not TGFβ, methylation levels were heterogeneous and dynamic in the different phases of PMF compared to controls. The distribution of PDGFB, FGF2 and LINE-1 methylation levels in PMF patients and controls in MF-0 showed that for both genes the methylation values in MF-0 are significantly higher compared to controls (PDGFB, FGF2: p < 0.0005) and PMF cases showed that for both genes the methylation values in MF-0 are significantly higher compared to controls (PDGFB, FGF2: p < 0.0005) and MF-1 and MF-2 cases showed a subgroup of cases with hypomethylation. PDGFB hypomethylation (<15%) was correlated with a favorable PMF prognosis (p=0.03, p=0.01 and p=0.02 for fibrosis, the International Prognostic Scoring System and the Dynamic International Prognostic Scoring System, respectively).

In addition, low PDGFB methylation levels were mainly present in cases from young patients with normal karyotype and LINE-1 hypermethylation (<15%) was correlated with a favorable PMF prognosis. Patients with low methylation of PDGFB and FGF2 in MF-0 had significantly better disease-free and overall survival compared with patients with high methylation levels (p < 0.0026 and p < 0.0017, respectively). Patients with low methylation of PDGFB and FGF2 in MF-0 showed significantly better disease-free survival than patients with high methylation levels (p = 0.0007 and p = 0.0008).

In conclusion, our results suggest that PDGFB and FGF2 methylation levels may be useful for the identification of patients with a better prognosis, which may contribute to the management of PMF.

3433T

MiR-145 regulates stem cell characteristics of human lung non-small cell carcinoma Hep-2 cells. M. Ozlen1,2, O.F. Karatay1,2, I. Sueri,2, B. Yuceetkiz,2, M. Yilmaz,2, H. Canisiz2, M. Ittmann1, M. Solak1,3

1 Department of Pathology & Immunology Baylor College of Medicine, Houston, TX, 77030, USA; 2 Department of Medical Genetics, Istanbul University Cerrahpasa Medical School, Istanbul, Turkey; 3 Department of Pathophysiology and Biostatistics of Neurosurgery, Central Military Hospital and 1st Faculty of Medicine, Charles University, Prague, Czech Republic.

The Cancer Stem Cells (CSCs) are tumorigenic cells promoting initiation, progression and spread of the tumor. Accumulating evidences has proven the existence of CSCs in a variety of tumors including lung, brain, breast, prostate, colon, head and neck cancers. These are and progressively developing new strategies in cancer research and clinical practice. The main strategy in CSCs is the regulation of stem cell markers such as SOX2, OCT4, and CD133. The main strategy in CSCs is the regulation of stem cell markers such as SOX2, OCT4, and CD133.

The Cancer Stem Cells (CSCs) are tumorigenic cells promoting initiation, progression and spread of the tumor. Accumulating evidences has proven the existence of CSCs in a variety of tumors including lung, brain, breast, prostate, colon, head and neck cancers. These are and progressively developing new strategies in cancer research and clinical practice. The main strategy in CSCs is the regulation of stem cell markers such as SOX2, OCT4, and CD133.

The Cancer Stem Cells (CSCs) are tumorigenic cells promoting initiation, progression and spread of the tumor. Accumulating evidences has proven the existence of CSCs in a variety of tumors including lung, brain, breast, prostate, colon, head and neck cancers. These are and progressively developing new strategies in cancer research and clinical practice. The main strategy in CSCs is the regulation of stem cell markers such as SOX2, OCT4, and CD133.
3436T
Differential DNA Methylation Patterns in Hereditary Non-polyposis Colorectal Cancer with or without Germline MLH1/MSH2 Mutation. C.H. Chen1,2, S.S. Jiang1, L.L. Hsieh2, R. Tang1, I.S. Chang1,2,3, C.A. Hsiung2,3, H.J. Tsai4,5. 1) National Institute of Cancer Research, National Health Research Institutes, Zhunan, Taiwan; 2) Taiwan Bioinformatics Core, National Health Research Institutes, Zhunan, Taiwan; 3) Department of Public Health, Chang Gung University, Taoyuan County, Taiwan; 4) Colorectal Section, Chang Gung Memorial Hospital, Gueishan, Taoyuan County, Taiwan; 5) Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Taiwan.

Introduction: In spite of several DNA methylation studies reported in sporadic colorectal cancer, the aberrant global methylation patterns in hereditary non-polyposis colorectal cancer (HNPPC) remain unclear. It has been suggested that the global methylation patterns may vary among subtypes of HNPPC patients. Methods: A total of 40 HNPPC patients with tissue samples were included in this study. Global DNA methylation patterns were measured using Illumina Infinium HumanMethylation27 BeadChip in 40 HNPPC patients for both tumor and adjacent normal tissues. Paired t-test was used to identify CpG sites differentially methylated between tumor and adjacent normal tissues among HNPPC patients, compared with those with mismatch repair deficiency (MMR) and among those without such deficiency (MMN). Student’s t-test was applied to identify the differential methylated CpG sites between HNPPC patients with MMN and with MMR. Pathway analysis was applied to explore the global DNA methylation patterns between HNPPC patients with MMN and with MMR. Results: At false discovery rate 0.01 (q-values<0.01), 65.6% and 46.6% of CpG sites are hypomethylated in tumor tissues, compared with adjacent normal tissues, among MMN and MMR, respectively. Our agnostic approach also confirmed the previous findings and implicated the gene approach reporting that hypomethylation in RUNX3, MLH1, NEUROG1, SOCS1 and CACNA1G in HNPPC patients, compared with MMR patients. Conclusions: The findings provided suggestive evidence that global hypomethylation in tumor tissues are more prominent in HNPPC patients with MMN (Familial Colorectal Cancer Type X) than those with MMR (Lynch syndrome). The underlying pathogenesis between HNPPC patients with MMN and with MMR may be differentially regulated through methylation. Further investigation will be needed to gain more understanding in regulatory mechanisms of DNA methylation on HNPPC development and progression.

3437S

The study of the immune system in different cancers is of interest for the role it may play in the onset and development of the disease. The role in immunotherapy might play in the treatment of different cancer types. The epigenetic regulation of immunologically relevant genes in different cancer types may provide important information for the role they may play in cancer etiology. In order to develop an immunity panel for the determination of a DNA methylation pattern in cancer, we started with a candidate gene approach, in contrast to genome-wide methylation screening. Genes from different functional groups were selected as candidates for methylation analysis. The immunology panel includes interleukins and their receptors (i.e. IL1B, IL6, IL8, IL27RA), TNFs and their receptors (i.e. TNFA, TNFSF11, TNFRSF25), Interferon-gamma, chemokines and their receptors (i.e. CXCL10, CXCR1), and CD families (i.e. CD5, CD21), cell adhesion molecules (i.e. BCAM, ICAM1) and others such as the insulin gene and its receptor (INS, INSR) and PDCD1. A total of 50 genes with over 100 amplimers covering over 500 CpG sites were selected and assays for these regions were developed and validated using PCR/Pyrosequencing, individually. The development of a 5m-Seq™ Immunology NGS Panel using Ion Torrent PGM™ system requires that the PCR amplicons be grouped into different pools based on the amplicon size, amplicon GC contents, and PCR conditions. DNA methylation controls (0%, 5%, 10%, 25%, 50%, 75%, and 100%) were sequenced on Ion 318 Chip to validate each target region. The correlation between the calculated methylation levels and the detected methylation levels were examined. Additionally, 10 pairs of ovarian cancer DNA samples (tumor vs adjacent normal) were further tested using both an Ion Torrent PGM™ system and Pyrosequencing. The comparison of the results obtained by Pyrosequencing and Next-Gen Sequencing using PGM shows that PCR pools with common amplicons gave linear correlation. PCR pools with extreme conditions, i.e. AT-rich (less than 20% GC contents) or GC-rich (greater than 70% GC contents) required further optimization. In summary, the use of a 5m-Seq™ NGS Panel can generate equivalent or better results compared to Pyrosequencing with the potential of using less DNA at lower cost. These results support future use of a custom 5m-Seq™ panel for efficient and sensitive epigenetic detection and analysis of cancer samples.

3438M
HES1 gene expression in patients with Medullary Thyroid Cancer is independent of its promoter methylation. M.G. Cardoso1, M.M.L. Közy2, S.C. Filho3, J.H. Lee3, D.P. Ferreira4, C.P. Camacho2, R. Delceilo2, J.M. Cerutti5, R.M.B. Maciel2, M.G. Jasiulionis6, M.R. Dias-da-Silva7. 1) Biochemistry, Universidade Federal de São Paulo, São Paulo, Brazil; 2) Medicine, Universidade Federal de São Paulo, São Paulo, Brazil; 3) Pathology, Universidade Federal de São Paulo, São Paulo, Brazil; 4) Pharmacology, Universidade Federal de São Paulo, São Paulo, Brazil; 5) Genetics, Universidade Federal de São Paulo, São Paulo, Brazil.

Abnormal methylation seems to play an important role in differentiated thyroid carcinomas, in which several of the tumor suppressor genes are epigenetically silenced. Little is known about the methylation profiles of genes that are associated with medullary thyroid carcinoma (MTC [MIM 155240]). Studies have reported that aberrant Notch signaling in MTC results in the downregulation of the HES1 [MIM 138605] gene, increase in expression of the tumor markers chromogranin A and calcitonin. Therefore, we aimed to determine if the downregulation of HES1 is caused by hypermethylation at its regulatory region. From 128 patients followed for MTC, we obtained formalin-fixed, paraffin-embedded (FFPE) tumor tissues from 29 patients, and peripheral blood samples from 18 patients with persistent disease. We also studied a human MTC cell line (TT). For HES1 analysis in tumor DNA we used bisulfite sequencing and methylation-specific PCR. HES1 protein expression was evaluated by immunohistochemistry, and mRNA levels were evaluated by RT-qPCR of peripheral blood. We observed increased expression of HES1 in peripheral blood samples of patients with persistent MTC compared with healthy individuals (p<0.007). Likewise, immunohistochemical analysis revealed moderate staining for HES1 in MTC tumors, whereas HES1 protein expression was undetectable. Gene expression analysis showed that HES1 was silenced, but no evidence of methylation at this promoter was found in sporadic or familial MTC tissues, the TT cell line or nontumorigenic tissues. The finding of increased expression of HES1 observed in persistent MTC is not mediated by methylation of DNA in its promoter region. This expression of HES1 suggests that the Notch signaling pathway may be involved in the cellular response to the constitutively active MAPK signaling in MTC tumorigenesis. Funding: FAPESP.

3439T
Genetic characterization of near-haploid and low hypodiploid acute lymphoblastic leukemia. S. Safavi, B. Johansson, K. Paulsson. Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Sweden.

Near-haploid (23-29 chromosomes) and low hypodiploid (Hol; 30-39 chromosomes) acute lymphoblastic leukemia (ALL) are two rare subtypes associated with a dismal prognosis. The aim of the present study was to investigate the underlying genetic mechanisms in these cases, in order to detect genetic aberrations and imbalances specific for near-haploid/Hol. We performed methylation array analysis, whole transcriptome sequencing and whole exome sequencing on 9 near-haploid and Hol cases. Results from methylation array and gene expression analyses show that near-haploidy forms a group with an expression profile separate from Hol, emphasizing the divergent genetic nature of these two subtypes. Methylation array analysis detected 100 genes that showed significantly different methylation in near-haploid/Hol group compared with the control group (P < 0.05, false discovery rate (FDR) < 0.1). Forty-five genes showing a significantly different expression between the near-haploid/Hol group and the control group (P < 0.05, FDR < 0.2) were detected by gene expression analysis. PARD2B, involved in cell polarization, was one of four genes found to be upregulated in both subtypes and downregulated in the control group. By genetically characterizing near-haploid and Hol ALL we can potentially identify therapeutic targets that may be used in a clinical setting for an improved prognosis.
Heparanase is an endo-glucuronidase, which cleaves heparan sulfate proteoglycans in the extracellular matrix, thereby releasing growth factors. The expression of heparanase is often upregulated in various pathological conditions, including cancer, immune disorders, and inflammation. The heparanase promoter contains several binding sites for transcription factors, which can be activated by various stimuli. The regulation of heparanase expression is highly complex and involves both transcriptional and post-transcriptional mechanisms.

miR-22 expression level was significantly reduced in cancerous tissues compared to normal tissues. The knockdown of miR-22 using specific inhibitors led to an enhanced proliferation, migration, and invasion of cancer cells, indicating its role as a potential tumor suppressor.

miR-22 expression was also found to be lower in patients with GCT compared to healthy controls. The inhibition of miR-22 expression in GCT cells resulted in an increased proliferation and migration, further supporting its tumor suppressive role.

miR-22 expression is associated with the clinical outcome of patients with cancer. Studies have shown that higher miR-22 expression is associated with a better prognosis, while lower expression is linked to worse outcomes. Additionally, miR-22 has been identified as a potential biomarker for cancer diagnosis and prognosis.

miR-22 is involved in various cellular processes, including cell proliferation, apoptosis, migration, and invasion. It regulates the expression of several genes involved in these processes, such as Bcl-2 and AKT/mTOR. The inhibition of miR-22 expression leads to an increased expression of pro-apoptotic genes and a reduced expression of anti-apoptotic genes, resulting in an increased apoptosis rate.

miR-22 is also involved in the regulation of autophagy and autophagic cell death. It targets the mTOR pathway, which plays a crucial role in autophagy. The inhibition of miR-22 expression leads to an increased autophagy and autophagic cell death, which may be important for cancer treatment.

miR-22 is a promising therapeutic target for cancer treatment. The development of miR-22 inhibitors or activators could potentially improve the efficacy of current cancer therapies. Additionally, miR-22 expression could be used as a biomarker for cancer diagnosis and prognosis.
3444M
Resveratrol up-regulates tumor suppressor mir-31 expression via inhibiting histone deacetylase 1 gene expression in chronic myeloid leukemia. C. Biray Avcı1,2, Z. Mutlu1, C. Calışkan1, B. Gök1, S. Yilmaz Suslu1, F. Sahin2, G. Saydam3, C. Gunduz1. 1) Ege University Medical School Department of Medical Biology, Izmir, Turkey; 2) Ege University Medical School Department of Hematology, Izmir, Turkey.

Chronic myeloid leukemia (CML) is characterized by a reciprocal translocation between chromosomes 9 and 22. The Philadelphia chromosome is available as a result of this translocation and causes BCR-ABL1 fusion gene which produces actively tyrosine kinase. Resveratrol (RES) is a natural phytoalexin found in grapes and induces apoptosis, erythroid differentiation and autophagy in leukemic cells. MicroRNAs are small (~22 nucleotides), single strand, non-coding RNA molecules that regulate post-transcriptional gene expression. miRNAs cause modulation of oncogenic or tumor suppressive pathways in different cancer types. In this study we aimed to determine cytotoxic effect of RES in K562 human CML cell line and to evaluate the expressions of miRNAs that are related with leukemogenesis after the treatment with RES. Also we analyzed target genes of miRNAs which show notable expression levels. K562 cells were treated with 100 μM (IC50 dose) RES during 72 hours and cytotoxicity was evaluated by using WST-1 assay. The RT-qPCR is used for mRNA and gene expression analysis. miRNAs and gene expression levels were evaluated by using miScript miRNA PCR Array and RT2 Profiler PCR Array, respectively. Results showed that RES up-regulated tumor suppressor mir-31 level 3.60 fold and significantly down-regulated HDAC1 gene expression (p=0.003), according to the control cells. Our findings showed that Resveratrol acts as a HDAC inhibitor targeting HDAC1 gene expression level. HDACs play a key role in the regulation of genes in cell progressions such as tumorigenesis and cell proliferation. Downregulation of HDAC1 provides post-translational modification for expression of tumor suppressor genes and leads to cell cycle arrest and increases apoptosis. These results provide that Resveratrol could be a therapeutic candidate as a HDAC inhibitor for chronic myeloid leukemia treatment.

3444T
Association of Ile655Val polymorphism of the HER2 gene with Neutropenia in breast cancer patients treated with trastuzumab chemotherapy. D.I. Carrillo-Moreno1,2, O. Soto3, L. Gómez Flores1, G.M. Zúñiga1, A. Ramos1, R. Ramirez1, O. Soto1, I. Gutierrez1, 2, G.M. Zúñiga2, M.P. Gallegos1. 1) Laboratorio de Genética Molecular, División de Medicina Molecular, CiBQ, IMSS, Guadalajara, Jal., Mexico; 2) Doctorado en Genética Humana, Centro universitario de Ciencias de la Salud, Universidad de Guadalajara.; 3) División de Genética, CIBO, IMSS; 4) Doctorado en Farmacología, Centro universitario de Ciencias de la Salud, Universidad de Guadalajara;

Background: The influence of this deletion on TKI responses in Chinese individuals, we collected 141 newly diagnosed patients with chronic phase CML whose first-line therapy was a standard dose of IM and 200 healthy individuals. Patients were classified as resistant or sensitive to IM according to the European Leukemia Net (ELN) criteria. The 2,903bp deletion of BIM was detected by PCR. Results: 1) 20 patients were heterozygous for the deletion and the other 121 did not harbor the deletion. 2) Of the patients with the deletion, 15 (75%) were sensitive to IM and only 5 (25%) were resistant. For patients without the deletion, 86 (71%) were sensitive to IM and 35 (29%) were resistant. 3) 41 of 200 healthy individuals carried the deletion. One was homozygous and the other 40 were heterozygous. The carrier frequency was 20.5%, which was much higher than previously reported (12.3%) and indicated a different genetic background. 4) Logistic regression analysis shows the overall odds ratio for resistance among patients with the deletion compared to those without it was 1.221 (P=0.719, 95% CI 0.412-3.616).

Conclusion: The BIM deletion polymorphism cannot account for intrinsic TKI resistance of Chinese patients with CML.

3447M
Association of CD44 expression before, during and after treatment in patients with head and neck cancer in comparison with healthy controls. K. Chukka1,2, Z. Vishnuvardhan2, S. Dasari3, U. Radhakrishna4. 1) Department of Biotechnology, Acharya Nagarejuna University, Nagarjunna Nagar, Guntur, India; 2) Department of Botany & Microbiology, Acharya Nagarjuna University, Nagarjunna Nagar, Guntur, India; 3) Department of Biotechnology, Dravidian University, Kuppam, India; 4) Green cross Pathology and Molecular Biology laboratory, Paldi, Ahmedabad, India.

Background: The influence of Ile655Val polymorphism in the HER2 gene involved in trastuzumab metabolism has been studied in breast cancer. HER2 (erbB-2, neu) is a proto-oncogene which encodes a transmembrane protein with tyrosine kinase activity but with no identified physiological ligand. The HER2 gene is amplified in 30% of invasive breast cancers and correlated with reduced patient survival. Our aim was to evaluate the association of Ile655Val polymorphisms in the HER2 gene with toxicity effects in breast cancer patients treated with trastuzumab chemotherapy. Methods: DNA genomic samples from 175 patients (UMAE gynecology and obstetrician Hospital, CMNO, IMSS), that received trastuzumab chemotherapy; were selected 141 newly diagnosed patients with chronic phase CML whose first-line therapy was a standard dose of IM and 200 healthy individuals. Patients were classified as resistant or sensitive to IM according to the European Leukemia Net (ELN) criteria. The 2,903bp deletion of BIM was detected by PCR. Results: 1) 20 patients were heterozygous for the deletion and the other 121 did not harbor the deletion. 2) Of the patients with the deletion, 15 (75%) were sensitive to IM and only 5 (25%) were resistant. For patients without the deletion, 86 (71%) were sensitive to IM and 35 (29%) were resistant. 3) 41 of 200 healthy individuals carried the deletion. One was homozygous and the other 40 were heterozygous. The carrier frequency was 20.5%, which was much higher than previously reported (12.3%) and indicated a different genetic background. 4) Logistic regression analysis shows the overall odds ratio for resistance among patients with the deletion compared to those without it was 1.221 (P=0.719, 95% CI 0.412-3.616).

Conclusion: The BIM deletion polymorphism cannot account for intrinsic TKI resistance of Chinese patients with CML.
3448T

PCA3 prostate cancer biomarker long non-coding transcription unit: Transcriptional interference of overlapping genes. R. Clarke1,2, M. Lavin1, R. Gardiner2, C. Chow2, R. Stirling Richards2, Z. Fang1. 1) Ingham Institute of Western Sydney School of Medicine, Liverpool, NSW, Australia; 2) University of Queensland Centre for Clinical Research, Brisbane, Qld, Australia; 3) Department of Urology, Royal Brisbane & Women’s Hospital.

PCA3 is arguably the most specific diagnostic biomarker for prostate cancer with great potential to improve clinical screening of men at risk. Given the over-riding need for prognostic and therapeutic biomarkers for prostate cancer the question remains as to whether PCA3 has a functional role in the disease and whether it could serve as a target or facilitator of therapy. The PCA3 gene appears to be a long non-coding transcription unit which is transcribed at high levels in prostate cancer. The PCA3 gene currently emerged in primates to give rise to a processed transcript with weak conservation and no detectable PCA3 protein product suggesting that PCA3 may have no direct functional role in prostate cancer. However, PCA3 is nested within another gene, BMCC1, and the transcription of PCA3 is discordantly regulated with BMCC1 in prostate cancer cells and after androgen treatment. BMCC1 is involved in Rho signalling and vesicular trafficking and may play roles in cellular transformation and metastasis. We apply novel technology to the BMCC1/PCA3 overlapping gene complex that suggests that the nested status of PCA3 within BMCC1 modulates transcription in a discordant fashion. These results suggest that PCA3 transcription may be a surrogate biomarker for BMCC1 and useful not only for the early detection of prostate cancer but to better understand prostate cancer initiation and development. If functional, interference of PCA3 transcription may represent a possible therapeutic application for early prostate cancer. Alternative novel PCA3 transcriptional products specific for prostate cancer could be used to selectively trigger a lethal exogenous target vector.

3449S

Creation of an open data sharing exchange to optimize BRCA clinical variant assessment. N. Conti1, C. Strom2. 1) Quest Diagnostics, Madison, NJ; 2) Quest Diagnostics Nichols Institute, San Juan Capistrano, CA.

Following the Supreme Court of the United States (SCOTUS) decision invalidating certain BRCA gene patents, several clinical laboratories in the US began to offer BRCA sequence testing. As with any sequence based test, some patients will have sequence variants that require clinical assessment. Clinical assessment encompasses literature review, population frequency determination, database searches, segregation and coinheritance analyses, computer modeling, and functional studies. Databases aid in this process by summarizing non-published data (such as segregation analyses) and previous interpretive conclusions, and serving as a gateway to published literature. Careful review of databases and confirmation, in general, are required. The Universal Mutation Database (UMD) collects data from 16 contributing laboratories throughout France. The UMD is actively curated. As of May 2014, UMD contained 6902 BRCA1 (MIM 113705) variants and 8062 BRCA2 (MIM 600185) variants. UMD is comparatively richer in BRCA content than US databases, because European labs have been doing BRCA testing for more than a decade. Until the recent SCOTUS decision, BRCA testing in the US had been limited to a single laboratory that no longer contributes variants to public databases. Our objective was to optimize patient care through facilitated access to UMD and to create a mechanism to assure contributions of new variant information to the database. Quest Diagnostics obtained a commercial use license for UMD with the stipulation that any research laboratory would continue to have free access to the database. Any commercial laboratory willing to share their data and provide user group fees through a user group will also have access to UMD. The fees will allow UMD to expand access with both hardware and software upgrades, develop processes for incorporation of new variants, and increase the number of curators for the expected increase in uploading of new variants. The user group agreement also allows participating laboratories to request functional studies for families with potentially clinically relevant variants. The planned transparency of the processes will permit independent assessment of the level of evidence for any variant in the UMD database and provide the scientific and medical communities with a valuable resource to optimize research and patient care.

3450M


We are investigating whether the variable response to chemotherapy in breast cancer (BC) can be explained by the diversity of somatic mutations among patients. Our laboratory proposed that there is a minimal genome required for BC cell survival. A stable set of 5,804 genes was derived by finding the minimum number of genes required to sustain tumor cell growth (70% or greater viability after three months) in 104 BC cell lines (20 BC with genes normal in expression levels in tumours (Mol. Oncol. 6: 347-359). A subset of these genes is targets of well-established therapies (e.g. paclitaxel, gemcitabine). We hypothesize that the effectiveness of these or other chemotherapy treatments may rely on the integrity of stable genes. We analyzed integrated genomic data for BC tumours from The Cancer Genome Atlas to re-define and/or narrow the minimal BC genome. Consistent with the design of our previous study, we identified genes stable in CN and gene expression, but now also exclude genes with deleterious somatic point/indel mutations (in ≥ 90% tumours). An analysis of 318 tumours revealed 6,994 stable genes, of which 2,364 genes were identified in the previous stable gene set. There are 3,856 unstable genes with deleterious somatic point/indel mutations, of which 32 genes exhibit ≥ 10 mutations in different tumours. Removing somatic point/indel mutations reduces the stable gene set by 122 genes. This is because unstable genes were often mutated through different mechanisms: 3,353 genes with somatic mutations also showed either normal CN or gene expression. Dysregulated pathways present in the minimal BC genome were enriched n = 37, mRNA processing and transport (n = 41), translation (n = 16), signaling (n = 19), and the formation of protein complexes (n = 19). Growth inhibition studies (GI50 values) of 27 BC cell lines show that chemotherapy sensitivity is not consistent across all cell lines (14 up-regulated in 2013, 14-R110). Paclitaxel and gemcitabine metabolic pathways include 33 (9 stable) and 14 (stable) genes, respectively. We related GI50 of BC cell lines, to CN, gene expression, and somatic mutation data for these genes. Paclitaxel sensitivity variations correlated with MUTP, BCL2, and CAC2 expression levels, and somatic mutations in CYP2C8, FG2, and TWIST1. BCL2 and MAPT are direct targets, and the remaining genes have previously been associated with paclitaxel resistance. Growth inhibition by gemcitabine was inversely correlated with RRM2 and RRM2B expression, which interact with known drug targets.

3451T

Whole exome sequencing approach in sib pairs identifies oligogenic germline mutations predisposing to early lung adenocarcinoma in non-smokers. E. Frullanti1, M.A. Mancarelli1, F. Cetta2, M. Baldassarri1, F. Mari1, S. Furrini3, P. Piu3, T.A. Dragani3, F. Ariani1, A. Renieri1, L. Azienda Ospedaliera Universitaria Senese, Siena, Siena, Italy; 2) IRCCS Multimedica, Milan, Italy; 3) Medical Genetics, University of Siena, Siena, Italy; 4) Department of Medical Biotechnology, University of Siena, Siena, Italy; 5) Department of Medicine, Surgery & Neuroscience, University of Siena, Siena, Italy; 6) Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.

Polygenic models are commonly assumed for the predisposition to common cancers. Despite the great bulk of research to identify genetic susceptibility genes in lung cancer by genome-wide association studies, only three loci have been identified (Lung Cancer Exempt from Mini Review 2013, 6: 347-359). A polygenic model is commonly assumed for the predisposition to common cancers. Despite the great bulk of research to identify genetic susceptibility genes in lung cancer by genome-wide association studies, only three loci have been identified. Recurrent somatic mutations in the GSAP gene were found in 15q25, 5p15.33, and 6p21, respectively. It is noteworthy that findings have not been replicated consistently in subsequent studies. In addition to confer a very low risk, they have been associated with lung cancer in smokers, but not in non-smokers. The polygenic nature of common cancers has frequently been suggested, but its biological basis still remains elusive. We tested the hypothesis that genetic susceptibility may rely on a restricted number of genes, disrupted by germ-line mutations. A combination between an advanced technical tool, i.e. the exome sequencing, and a new patient selection strategy was used. The strategy relies on the selection among 954 lung adenocarcinoma patients those with early onset disease (mean age 43) in absence of cigarette smoking, and having a first degree relative with lung cancer (BC) can be explained by the diversity of somatic mutations among patients. Our laboratory proposed that there is a minimal genome required for BC cell survival. A stable set of 5,804 genes was derived by finding the minimum number of genes required to sustain tumor cell growth (70% or greater viability after three months) in 104 BC cell lines (20 BC with genes normal in expression levels in tumours (Mol. Oncol. 6: 347-359). A subset of these genes is targets of well-established therapies (e.g. paclitaxel, gemcitabine). We hypothesize that the effectiveness of these or other chemotherapy treatments may rely on the integrity of stable genes. We analyzed integrated genomic data for BC tumours from The Cancer Genome Atlas to re-define and/or narrow the minimal BC genome. Consistent with the design of our previous study, we identified genes stable in CN and gene expression, but now also exclude genes with deleterious somatic point/indel mutations (in ≥ 90% tumours). An analysis of 318 tumours revealed 6,994 stable genes, of which 2,364 genes were identified in the previous stable gene set. There are 3,856 unstable genes with deleterious somatic point/indel mutations, of which 32 genes exhibit ≥ 10 mutations in different tumours. Removing somatic point/indel mutations reduces the stable gene set by 122 genes. This is because unstable genes were often mutated through different mechanisms: 3,353 genes with somatic mutations also showed either normal CN or gene expression. Dysregulated pathways present in the minimal BC genome were enriched n = 37, mRNA processing and transport (n = 41), translation (n = 16), signaling (n = 19), and the formation of protein complexes (n = 19). Growth inhibition studies (GI50 values) of 27 BC cell lines show that chemotherapy sensitivity is not consistent across all cell lines (14 up-regulated in 2013, 14-R110). Paclitaxel and gemcitabine metabolic pathways include 33 (9 stable) and 14 (stable) genes, respectively. We related GI50 of BC cell lines, to CN, gene expression, and somatic mutation data for these genes. Paclitaxel sensitivity variations correlated with MUTP, BCL2, and CAC2 expression levels, and somatic mutations in CYP2C8, FG2, and TWIST1. BCL2 and MAPT are direct targets, and the remaining genes have previously been associated with paclitaxel resistance. Growth inhibition by gemcitabine was inversely correlated with RRM2 and RRM2B expression, which interact with known drug targets.
Epigallocatechin-3-gallate induces apoptosis and autophagy via up-regulation of TNF and GABARAPL2 gene expression in chronic myeloid leukemia cells. B. Goker1, C. Caliskan2, Z. Mutlu2, B. Erbaykent2, Tepedelen2, M. Korkmaz5, G. Saydam4, C. Gunduz1, B. Biray Avci1. 1) Department of Medical Biology, School of Medicine, Ege University, Izmir, Turkey; 2) Department of Molecular Biology and Genetics, Faculty of Science And Letters, Avrasya University, Trabzon, Turkey; 3) Department of Medical Biology, School of Medicine, Celal Bayar University, Manisa; 4) Department of Hematology, School of Medicine, Ege University, Izmir, Turkey.

Chronic myeloid leukemia (CML) is a clonal hematopoietic stem cell disease. Philadelphia chromosome (Ph) is the result of a reciprocal (9;22) translocation is responsible for developing leukemic phenotype of CML. Autophagy related genes with the treatment of EGCG. The cytotoxic effect of EGCG on K562 CML cell line treated with EGCG and untreated cells as control group. Reverse transcription of autophagy related genes with EGCG was determined as 50 µM. IC50 (24th hours) dose of EGCG was evaluated with ApoDIRECT and gene expressions were shown by RT-qPCR. EGCG has significant therapeutic effects on different human tumor types is a major flavonoid of green tea. Autoapagy known as programmed cell death is one of the main response of cell to stimulus. Raise of autophagic activity in cancers is novel therapeutic target with using different active substance. In this study we aimed to determine cytotoxic and autophagic changes in K562 CML cell line treated with EGCG compared to untreated control group, and detect the expression changes of autophagy related genes with the treatment of EGCG. The cytotoxic effect of EGCG on K562 cells was determined in time and dose dependent manner by using WST-1 analysis. Total RNA was isolated from K562 cells treated with EGCG and untreated cells as control group. Reverse transcription procedure was performed for cDNA synthesis. Apoptotic effect of EGCG was evaluated with ApoDIRECT and gene expressions were shown by RT-qPCR. IC50 (24th hours) dose of EGCG was determined as 50µM. EGCG induced apoptosis 10.9 fold compared to control cells. EGCG up-regulated TNF and GABARAPL2 genes 13.1775 and 9.5137 fold according to the control cells, respectively. Increase of TNF, most significant cytokine is placed in apoptotic and autophagic pathways, and GABARAPL2, responsible for autophagy induction, expression and autophagic cell death in several human cancer cells. Our current findings manifested that EGCG treatment increases cancer cell death via up-regulation of genes are part of both apoptotic and autophagic signaling pathways. Therefore EGCG suggests a nontoxic agent for chronic myeloid leukemia treatment.

Hypomorphic CYP2C9 ‘*2 and ‘*3 alleles associate with improved non-small-cell lung cancer (NSCLC) prognosis. L.N. Gordon1, A. Pozzi2, J.H. Capdevila3, S.M. Williams4,1. 1) Dartmouth College, Department of Genetics Hanover, NH 03755; 2) Vanderbilt University, Department of Medicine Nashville, TN 37232, USA; 3) Vanderbilt University, Department of Nephrology Nashville, TN 37232, USA; 4) Dartmouth College, Institute for Quantitative Biomedical Sciences Hanover, NH 03755.

Lung cancer is the leading cause of cancer death in the United States, despite the higher incidences of other cancers. Given this disproportionate mortality, more effective treatments are needed. In mice knockout of the Cyp2c44 epoxigenase is associated with reduced angiogenesis and tumor growth. Interestingly, expression of the Cyp2c44 gene can be downregulated by treatment with PPARα ligands with a corresponding decrease in angiogenesis tumor growth, and distant metastasis. The human homologue of Cyp2c44, CYP2C9 epoxigenase, and its variants CYP2C9*2 and CYP2C9*3 have been linked to breast and colon cancer. However, whether CYP2C9 influences survival in non-small-cell lung cancer (NSCLC) patients and if it is linked to PPARα is not known, but could constitute an excellent drug target. We genotyped CYP2C9 alleles ‘*2 and ‘*3 along with tag SNPs in CYP2C9 in a sample of 398 Caucasian NSCLC patients. Possession of either CYP2C9 ‘*2 or ‘*3 in female patients associated with increased survival (HR = 0.383, p-value = 0.035) adjusted for cancer staging, chemotherapy and tumor resection. However, not only is this genetic association limited to females, but these alleles may interact with gender because the interaction between the hypomorphic alleles and gender borders the significance threshold (p-value = 0.100). Together, these results suggest not only that genetic variation in CYP2C9 influences NSCLC survival, but also that this gene with PPARα ligands, such as clinically used fibrates, may be a viable clinical treatment.

Evaluation of miR-27a, miR-181a, and miR-570 Genetic Variants with Gallbladder Cancer Susceptibility and Prognosis in North Indian Population. A. Gupta, B. Mittal, A. Yadav, A. Sharma, K.L. Sharma, N. Rastogi, S. Agrawal, A. Kumar. Sanjai Gandhi Post Graduate Institute of Medical Sciences, Lucknow, Uttar Pradesh, India.

Purpose of study: miRNAs are small endogenously expressed short non-coding RNA molecules. They appear to be critical mediators of cell biology as their aberrant expression is well characterized in cancer progression. Many miRNAs genetic variants have been found to play important role in many cancers but their contribution in gallbladder carcinoma has not been fully explored. So, in present study, we investigated the role of cancer-specific miR-27a, miR-181a, and miR-570 genetic variants with gallbladder cancer (GBC) susceptibility, therapeutic response and toxicities to chemoradiotherapies. Methods: The present study was carried out in 515 GBC patients and 200 healthy controls in North Indian population. Among them, 126 patients receiving adjuvant or neoadjuvant therapies as per requirement were followed up for treatment outcome. Treatment response was recorded according to Response Evaluation Criteria in Solid Tumors (RECIST) criteria. Hematological and gastrointestinal toxicities profile were recorded as per CTCAE version 3.0. 2006. Genotypes were determined by TaqMan probes. Statistical analysis was done by SPSS ver. 16. Results: Logistic regression analysis showed no significant association of miR-27a, miR-181a, and miR-570 polymorphism with GBC survival (p > 0.05). On stratifying data on the basis of gall stone status, the [AG+GG] genotypes of miR-27a rs895819 (A>G) were significantly associated with increased risk of GBC in patients without stone (p-value=0.003; OR=0.06; OR=1.83; CI=1.23-2.72). The risk genic miR-27a, rs895819 (A>G) was modulated by tumor confusion (p-value=0.005, OR=0.01, OR=1.94; CI=1.22-3.08). Univariate analysis revealed that AG+GG genotype of miR-27a and CT+TT genotype of miR-181a were significantly associated with nonresponsiveness to chemoradiotherapies (p-value=0.03 and 0.04 respectively) whereas miR-570 variants did not show any association. None of the studied polymorphism had any influence on haematological and gastrointestinal toxicities. Conclusion: We found significant association of miR-27a rs895819A>G with gallbladder cancer risk through palliative independent pathway and tumor confusion. Our results are also suggest that variant allele of miR-27a (rs895819 A>G) and miR-181a (rs12537 C>T) may be associated with poor therapeutic response in GBC patients. However study needs to be validated in independent cohorts.


Chronic myeloid leukemia (CML [MIM 608232]) is one of the most frequent hematopoietic malignancies in the elderly population. However, genetic factors associated with an increase risk to CML development are unknown. Recent studies have shown that SNPs affecting miRNA biogenesis or miRNA/miRNA interaction are important risk factors in the development of different types of cancer. Thus, we carried out a case-control study to test the association of SNPs located in the microRNA machinery genes AGO1 (rs7399917) and KRAS (rs1764370) with CML susceptibility. We genotyped 828 Mexican-mestizo individuals (497 healthy subjects and 331 CML cases) using TaqMan probes. We found a significant association between the minor homocytoge of the KRAS rs1764370 SNP (G/G) genotype and an increased risk for CML susceptibility (OR = 3.39, p = 0.015). After gender stratification, this association remains only in male individuals (OR = 3.41, p < 0.016). None of the other studied SNPs showed a significant association with CML, even after gender stratification. In addition, the minor allele (C) of GEMIN4 rs2740348 significantly associated with a high-risk Sokal score (OR = 3.15, p = 0.01254). To the best of our knowledge this is the first study to show a significant association of the KRAS rs16174370 SNP with CML susceptibility in the Mexican population. To further determine the participation of these SNPs in CML susceptibility it is necessary to replicate our findings in different populations.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3456M
The role of Stem Cell Markers in Prostate Cancer Recurrence. E. Guzel1,2,3, OF. Karatil1,2, MB. Duz1, M. Ittimann3,4, M. Solak4, M. Ozden4,1, Istanbul University, Cerrahpasa Medical School, Istanbul, Turkey; 2) Department of Electroneurophysiology Biruni University, Istanbul, Turkey; 3) Molecular Biology and Genetics Department, Erzurum Technical University, Erzurum, Turkey; 4) Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX, 77030, USA; 5) Baylor College of Medicine, Michael E. DeBakey VAMC, Houston, TX, 77030, USA; 6) Afyon Kocatepe University Medical School Department of Medical Genetics.
Prostate cancer (PCa) is one of the most common tumor types related to mortality in males in the developed countries. Studies have demonstrated that therapeutic tools mostly ineffective to give positive outcome especially for PCa. Cancer stem cells are composed of a small cell population, which are supposed to have roles in tumorogenesis, metastasis, and tumor recurrence after chemo-radiotherapy. The aim of this study is to explore the differential expressions of stem cell markers in recurrent PCa and non-recurrent PCa tumors as well as in adjacent normal prostate tissues. Here we compared the expression of important stemness regulators like Sox2, OCT4, KLF4, and ABCG2 genes in recurrent, non-recurrent PCa and adjacent normal tissue samples using quantitative real-time polymerase chain reaction. Our results demonstrated that SOX2 and OCT4 are strongly overexpressed in PCa samples. Recurrent PCa samples are markedly positive for stem cell markers Sox2, OCT4 and KLF4. Furthermore, non-recurrent PCa samples presented low levels of ABCG2, a multidrug resistance protein, compared to both normal and recurrent samples, which might be associated with chemoresistance. Enhanced expression of ABCG2 and stem cell markers in the recurrent PCa tissues postulates the suggestion that enrichment for cells with stem cell characteristics in these tissues might be critical for enhancement of chemoresistance and recurrence of cancer.

3457T
KEAP1 genetic polymorphisms associate with breast cancer risk and survival outcomes. J.M. Hartikainen1,2, M. Tengström1,2, R. Winqvist1,2, A. Jukkola-Vuorinen3,4, K. Pykäät5, V.-M. Kosma5,6, Y. Soini1,7, A. Mannnermaa1,7,1) School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, and Cancer Center of Eastern Finland, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Kuopio, Finland; 2) School of Medicine, Institute of Clinical Medicine, Oncology, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland; 3) Cancer Center, Kuopio University Hospital, P. O. Box 1777, FI-70211 Kuopio, Finland; 4) Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, P. O. Box 5000, FI-90014 University of Oulu, Oulu, Finland; 5) Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Nordlab, Oulu University Hospital, Oulu, Finland; 6) Department of Oncology, University of Oulu, Oulu University Hospital, P.O. Box 5000, FI-90014 University of Oulu, Finland; 7) Imaging Center, Clinical Pathology, Kuopio University Hospital, P. O. Box 1777, FI-70211 Kuopio, Finland.
KEAP1, a key protein involved in the regulation of the NRF2-KEAP1 pathway, is involved in the regulation of multiple cell processes, including survival, proliferation, and chemoresistance. The aim of this study was to investigate the role of KEAP1 gene polymorphisms in breast cancer survival. The strongest association was with the MICAL-L2 gene (Trend P = 0.016), and rs34197572 with overall survival (OS) (P = 0.016), and rs34197572 with overall survival (OS) (P = 0.045). The minor allele of rs34197572 associated with increased risk of breast cancer (OR = 1.021, 95% CI: 1.004–1.038, P = 0.015). The minor allele A of rs11085735 associated with lower KEAP1 protein expression (P = 0.040) and high extent nuclear NRF2 protein expression (P = 0.009). It also associated with worse survival in all invasive cases (P = 0.023). When treatment data were included rs11085735 associated with response to radiotherapy. The most significant associations were observed in cancers classified as non-small cell lung cancer (P = 0.009). The results suggest that KEAP1 gene polymorphisms may influence breast cancer outcome, and further studies are needed to confirm these findings.

3458S
Rare and Common Variants Contribute to Lung Cancer Survival in African Americans. C.C. Iverson1, W.S. Bush1, D.C. Crawford1, H.H. Dilks2, J. Long3, W.J. Blot3,4, E.L. Gregg5, M.C. Aldrich1,2, 1) Center for Human Genetics Research, Vanderbilt University Medical School, Nashville, TN; 2) VANGAMT, Vanderbilt University, Nashville, TN; 3) Division of Epidemiology, Vanderbilt University Medical School, Nashville, TN; 4) International Epidemiology Institute, Rockville, MD; 5) Department of Thoracic Surgery, Vanderbilt University Medical School, Nashville, TN; 6) Tennessee Valley Health System Veterans Affairs, Nashville, TN.
Lung cancer is the leading cause of cancer-related mortality in the U.S. Survival rates differ by race, with blacks experiencing poorer survival than whites, yet few studies have focused on blacks. Germline genetic variation may influence overall lung cancer survival. A total of 305 incident non-small cell lung cancer African American cases were identified from the prospective Southern Community Cohort Study through linkage with 12 state cancer registries. Vital status was determined by linking with the National Death Index or Social Security Administration. After 8.6 years of follow-up, 87% of lung cancer cases were deceased. We performed genotyping using the Illumina HumanExome BeadChip. After standard quality control, 301 individuals (60% male) and 274,438 variants remained for analysis. We identified variants previously associated with lung cancer survival from the NHGRI GWAS catalog on the ExomeChip array. For each of these SNPs, we ran a Cox proportional hazards model adjusted for age, sex, patient ancestry (estimated from ancestry informative markers), stage at diagnosis, and treatment. We found the C allele at rs1878022, in the chemokine receptor-like 1 (CMKLR1) gene, was associated with reduced mortality [hazard ratio (HR): 0.72, 95% confidence interval (CI): 0.54-0.97, p = 0.03]. The improved survival in contrast to a prior lung cancer survival GWAS of a similar sample size conducted in whites, suggesting ethnic-specific associations. We then sought to identify rare variants in protein coding regions associated with lung cancer survival. We identified variants with a MAF < 5% (n = 114,646) and mapped them to 9,725 genes. We used the sequence kernel association test (SKAT) and burden test, adjusting for age, sex, and African ancestry, to examine associations between rare variants and lung cancer survival. The strongest association was with the MICAL-L2 gene (Trend P = 0.024 and 0.025, respectively). rs9676881 and rs34197572 associated with OS also among radiation therapy-treated cases, providing further support to the involvement of the NRF2-KEAP1 pathway in breast cancer susceptibility and patient outcome.

3459M
Cytogenetic abnormalities of 50 AML patients by FISH detection and conventional karyotype analysis. E. Karaca, A. Aykut, B. Durmac, A. Durmac, I.M. Tekin, O. Cogulu, H. Akin. Ege University Faculty of Medicine, Department of Medical Genetics, Izmir, Turkey.
Cytogenetic analyses in acute myeloid leukaemia (AML) have come to light a great number of chromosome abnormalities. The cytogenetic analysis is an important laboratory tool for diagnosis, prognosis, in clinical decision-making and in follow-up for pediatric and adult patients with AML. Fluorescence in situ hybridization (FISH) is increasingly being used in the cytogenetic diagnosis of AML. In this study the detectable rate of cytogenetic abnormalities including t(8:21)(q22;q22), t(15:17)(q24;q21), (9;22)(q34;q11), inv(16)(p13q22) and MLL rearrangements detection rates were 4%, 2%, 5%, and 1%, respectively. In follow up request for AML patients. A more anomalies and using an algorithm to reach effective diagnose and follow up in AML patients.
3460T DNA methylation profiling reveals novel diagnostic biomarkers in renal cell carcinoma. B.N. Lasseigne1,2, T.C. Burwell1, M.A. Patil1, D.M. Absher1, J.D. Brooks2, R.M. Myers1. 1) HudsonAlpha Institute for Biotechnology, Huntsville, AL; 2) Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL; 3) Department of Urology, Stanford University, Stanford, CA.

Renal cell carcinoma (RCC) is the 10th most commonly diagnosed cancer in the United States. While it is usually lethal when metastatic, RCC is successfully treated with surgery when tumors are confined to the kidney and have low tumor volume. Because most early stage renal tumors do not result in symptoms, there is a strong need for biomarkers that can be used to detect the presence of the cancer as well as to monitor patients during and after therapy. We examined genome-wide DNA methylation alterations in renal cell carcinomas of diverse histologies and benign adjacent kidney tissues from 96 patients. We observed widespread methylation differences between tumors and benign adjacent tissues, particularly in immune, G-protein coupled receptor, and metabolism-related genes. Additionally, we identified a single panel of DNA methylation biomarkers that reliably distinguishes tumor from benign adjacent tissue in all of the most common kidney cancer histologic subtypes (area under receiver operating characteristic curve, 0.991), and a second panel does the same specifically for clear cell renal cell carcinoma tumors (area under receiver operating characteristic curve, 0.990). This set of biomarkers were validated independently with excellent performance characteristics in more than 1,000 tissues in The Cancer Genome Atlas clear cell, papillary, and chromophobe renal cell carcinoma datasets (area under receiver operating characteristic curves, >0.97). These DNA methylation profiles provide insights into the etiology of RCC and may lead to clinically applicable biomarkers for use in early detection of kidney cancer.

3461S Association of LEP rs7799039 (G-2548A) polymorphism with obesity in breast cancer patients. A. Méndez-Hernández1,2, JA. Espinosa1, F. Pérez2, MP. Gallegos2. 1) Laboratorio de Investigación, Facultad de Medicina, Universidad Juárez de El Estado, Durango, Mexico; 2) Laboratorio de Genética Molecular, División de Medicina Molecular, CIBO, IMSS, Guadalajara, Jalisco; 3) Facultad de Ciencias Quimicas, Universidad Juárez del Estado de Durango, Gómez Palacio, Dgo.

Background: LEP rs7799039 polymorphism (G-2548A) has been correlated with changes in leptin levels, the degree of obesity and is involved in the evolution of breast cancer. The LEP-2548AA genotype was associated with increased leptin concentrations, increased expression of leptin mRNA in adipose tissue, and more high body mass index, thereby increasing the risk for overweight and obesity and evolution breast cancer. Our goal is to evaluate the association of polymorphism rs7799039 LEP (G-2548A) in obese patients with breast cancer. Methods: DNA genomic of 245 samples (UMAE Hospital Gynec-Obstetrician, CMNO, IMSS) were included in the study. The LEP rs7799039 (G-2548A) genotyping was determinate by Taq-Man® probe and allelic discriminated was determinate by StepOne Software v.2., of Applied Biosystems. The association was determinate by odds ratio (OR). Results: The genotype LEP-2548AA was associated with overweight and obesity [OR = 2.1 (95% CI 0.96 - 4.71) , p = 0.05 ] in patients with breast cancer, while the LEP-2548GG genotype was found as a protective factor to overweight and obesity [OR = 0.47 (95% CI 0.20 - 0.82) , p = 0.011]. Conclusion: LEP rs7799039 polymorphism (G-2548A) may be an indicator of risk for overweight and obesity and explain the possible mechanism of obesity in the progression of breast cancer.

3462M Changes in Colorectal Cancer Genomes under Anti-EGFR Therapy Identified by Whole-Genome Plasma DNA Sequencing. S. Mohan1, E. Heitzer2, P. Utz3, I. Lauer1, S. Lax1, M. Auer1, M. Pichler1, A. Gerger1, F. Eisenr1, G. Hoefler1, T. Baumhofer1, J.B. Geigl1, M.R Speicher1. 1) Institute of Human Genetics, Medical University of Graz, Graz, Austria; 2) Department of Pathology, General Hospital West, Graz, Austria; 3) Division of Oncology, Medical University of Graz, Graz, Austria.

Monoclonal antibodies targeting the Epidermal Growth Factor Receptor (EGFR), such as cetuximab and panitumumab, have evolved to important therapeutic options in metastatic colorectal cancer (CRC). However, almost all patients with clinical response to anti-EGFR therapies show disease progression within a few months and little is known about mechanism and timing of resistance evolution. Here we analyzed plasma DNA from 10 patients treated with anti-EGFR therapy by whole genome sequencing (plasma-Seq) and ultra-sensitive deep sequencing of genes associated with resistance to anti-EGFR treatment such as KRAS, BRAF, PIK3CA, and EGFR. Surprisingly we observed that the development of resistance to anti-EGFR therapies was associated with acquired gains of KRAS in 4 patients (40%), which occurred either as novel focal amplifications (n=3) or as high level polysomy of 12p (n=1). In addition, we observed focal amplifications of other genes recently shown to be involved in acquired resistance to anti-EGFR therapies, such as MET (n=2) and ERBB2 (n=1). Overrepresentation of the EGFR gene was associated with a good initial anti-EGFR efficacy. Overall, we identified predictive biomarkers associated with anti-EGFR efficacy in 7 patients (70%), which correlated well with treatment response. In contrast, ultra-sensitive deep sequencing of KRAS, BRAF, PIK3CA, and EGFR did not reveal the occurrence of novel, acquired mutations. Thus, plasma-Seq enables the identification of novel mutant clones and may therefore facilitate early adjustments of therapies that may delay or prevent disease progression.

3463T Zoledronic acid treatment up-regulates miR-15a via targeting anti-apoptotic BCL2 gene expression in chronic myeloid leukemia. Z. Mutlu1, C. Calskan1, B. Goker1, C. Kayabasi1, T. Terecan2, G. Saydam2, C. Gunduz1, C. Binay Avci1. 1) Ege University Medical School Department of Medical Biology, Izmir, Turkey; 2) Ege University Medical School Department of Hematology, Izmir, Turkey.

Chronic myeloid leukemia (CML) is associated with a characteristic chromosomal translocation called the Philadelphia chromosome (Ph). This genetic alteration for the diagnosis of CML causes the formation of the BCR-ABL1 fusion gene which produces actively tyrosine kinase. Apoptosis is a form of programmed cell death and autophagy is a lysosomal degradation pathway essential for homeostasis that contributes to cell death. Zoledronic acid (ZA), a nitrogen-containing bisphosphonate, is a potential inhibitor of the monoclonal antibodies targeting the Epidermal Growth Factor Receptor (EGFR). As such, this study aimed to evaluate the cytotoxic, autophagic and apoptotic effects of ZA and to examine expression levels of miRNAs treatment with ZA in K562 CML cells. In this study, we evaluated the cytotoxic, autophagic and apoptotic effects of ZA and to examine expression levels of miRNAs treatment with ZA in K562 CML cells. Also we investigate the roles of miRNA target genes and their clinical relevance in CML. Because, apoptomirs, such a miR-15a were down-regulated and further experiments are required for the usage of ZA in CML.
3464S
Towards a national implementation of DNA-based personalized cancer treatment in the Netherlands. I.J. Nijman on behalf of the Netherlands Center for Personalized Cancer Treatment. Center for Personalized Cancer Treatment.

In the past 30 years, cancer treatment has improved enormously, partly due to the development and implementation of targeted therapies. These therapies are designed to tackle specific characteristics of a cancer cell, such as growth factor receptors. Unfortunately, the presence or absence of such characteristics only paints a partial picture of the tumor’s responsiveness. Instead, treatment outcome can be better predicted by a combination of genetic mutations. As each tumor has its own genetic characteristics, the appropriate treatment decision differs per patient and is therefore called “personalized treatment.” This approach can improve treatment outcome for patients with metastatic cancer, decrease the number of patients unnecessarily exposed to toxic agents, and reduce costs. Being able to classify patient populations into groups with different likelihoods to benefit from a particular treatment is key in this process. The Center for Personalized Cancer Treatment brings this concept to the patient today by forging collaborations between all academic cancer centers in the Netherlands. The CPCT focusses on two DNA analysis tracks. First, rapid screening for actionable mutations and amplifications to select the appropriate standard targeted treatment or to allocate patient into investigator and Pharma-driven Phase I trials. At the moment over 1200 patients have been analyzed through this route. Secondly, cohorts of patients with valuable response to targeted treatments are sequenced and analyzed in a system biology perspective to identify predictive biomarkers. Detailed follow up data for over 700 patients is now being collected in an electronic clinical response form (eCRF) and integration with clinical pathology and pharmacological is ongoing. Combined with public datasets, this integrated dataset provides a huge and growing national resource that can be used for biomarker- or target evaluation, systems biology analyses, as well as for improving care for individual patients.

3465M
Integration of microarray meta-analysis with RNASeq and genomewide genetic data to identify variants associated with endometrial cancer histological subtypes. T.A. Burton, A.B. Spurdle, Genetica and Computational Biology Dept, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia.

Non-endometrioid endometrial cancer comprises approximately 20% of new endometrial cancer diagnoses, but is responsible for an estimated 50% of deaths from this disease. Microarrays have been successfully used to measure mRNA expression in biological samples to identify differentially expressed genes. However, it is recognized that results from individual microarray studies are often not reproducible. Public availability of gene expression microarray data has facilitated research assessing the increased reliability of combined datasets. Furthermore, the availability of genomewide scale “omics” data on the same sample sets, such as that provided by The Cancer Genome Atlas (TCGA), allows for a powerful, integrative approach to identify genes and genetic variants important in cancer initiation and/or progression. We have used patient derived microarray data from multiple datasets to identify differences between endometroid (EEC) and non-endometrioid (NEEC) histological subtypes. Biologically consistent genes were then validated these differences using an independent endometrial cancer set with RNAseq data from the TCGA. A literature review and repository search conducted in April 2014 identified 15 endometrial cancer microarray studies. Raw microarray data was accessed from publication supplementary data, the NCBI Gene Expression Omnibus (GEO) and the TCGA data portal, or otherwise requested by directly contacting authors. Overall raw microarray data was accessed for 9 of 15 studies, with a maximum of 79 EEC and 12 NEEC cases in one study. Following QC and meta-analysis with the MetaOms package in R, 1,910 genes from 6 studies were identified by Fisher’s test as displaying significant differential expression between the two subtypes (FDR < 5%). The expression of these genes was then analyzed using normalized RNASeq data for 322 EEC and 97 NEEC non-overlapping samples from the TCGA. For the 1,186 genes captured by the RNASeq analysis, moderated t-test identified differences in expression for 912 genes at P < 0.05 (77% of genes analysed), of which 261 were at P < 1×10^{-5}. For the 1,186 genes captured by the RNASeq analysis, moderated t-test identified differences in expression for 912 genes at P < 0.05 (77% of genes analysed), of which 261 were at P < 1×10^{-5}. These findings demonstrate a strong relationship between microarray meta-analysis and RNASeq analysis of the TCGA dataset. Genes that were differentially expressed expressed in aggressive endometrial cancer histological subtypes. Future analysis will investigate genetic variation associated with differential gene expression in the TCGA dataset, to prioritize genetic variants for downstream prognostic studies.

3466T
The type II transmembrane serine proteases hepsin and TMPRSS3 are associated with breast cancer survival. M. Pelkonen1, 2, R. K. Luostarinen1, 2, V. Katja2, 3, Y. Soimik1, 2, V. M. Kosma2, 3, A. Mannermaa1, 2, 3, 4, 1) Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland; 2) Biocenter Kuopio and Cancer Center of Eastern Finland, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland; 3) Imaging Center, Clinical Pathology, Kuopio University Hospital, P.O. Box 1777, FI-70211 Kuopio, Finland; 4) Institute of Clinical Medicine, Oncology, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland; 5) Cancer Center, Kuopio University Hospital, P.O. Box 1777, FI-70211 Kuopio, Finland.

Background: Hepsin, (also called TMPRSS1) and TMPRSS3 are type II transmembrane serine proteases (TTSPs) that are involved in cancer progression. TTSPs can remodel extracellular matrix (ECM) and, when dysregulated, promote tumor progression and metastasis by inducing defects in basement membrane and ECM molecules. This study investigated whether the gene and protein expression levels of these TTSPs were associated with breast cancer characteristics or survival. Methods: Immunohistochemical staining was used to evaluate hepsin levels in 372 breast cancer samples and TMPRSS3 levels in 373 samples. TMPRSS1 mRNA expression was determined in 125 invasive and 16 benign breast tumor samples, and TMPRSS3 mRNA expression was determined in 162 invasive and 24 benign breast tumor samples. The gene and protein expression levels were analyzed for associations with breast cancer-specific survival and clinicopathological parameters. Results: Low TMPRSS1 and TMPRSS3 mRNA expression levels were independent prognostic factors for poor breast cancer survival during the 20-year follow-up (log rank P-values < 0.0001). Grade III tumors were associated with lower survival compared to high levels (log rank P-values 0.013 and 0.023, respectively). TTSPs were associated with low mRNA and protein expression levels. Conclusions: The results suggest that the TTSPs hepsin and TMPRSS3 may have similar biological functions in the molecular pathology of breast cancer and that they represent a potential new therapeutic target.

3467S
Genetic variants in the gene ARID5B associated with susceptibility to childhood acute lymphoblastic leukemia. A. Reyes-León1, C. Salas-Labadia1, J. García-Cruz1, M. Zapata-Tarrés1, M. Ramírez-Martínez1, R. Paredes-Aguilera1, R. Rivera-Luna1, S. Jiménez-Morales1, L. Orozco-Orozco1, A. Reyes-León1, A. M. de la Fuente, 1, 2, 3, P. Pérez-Vera1, 2, 3, P. García-Márquez1, 2, 3, A. M. de la Fuente, 1, 2, 3, P. Pérez-Vera1, 2, 3, A. M. de la Fuente, 1, 2, 3, 4, 1) Instituto Nacional de Pediatría, Mexico City, 2) Instituto Nacional de Medicina Genómica, Mexico City, Mexico; 3) National Cancer Institute, Frederick, Maryland, USA.

In Mexico the acute lymphoblastic leukemia (ALL) is the most common cancer in children, according to the National Popular Medical Insurance Program, in our country the incidence is 60 cases per million per year. Some individual genetic variations might contribute to the increased incidence, and simultaneously influence susceptibility to the disease development. We performed a genome-wide association study to identify polymorphisms (SNPs) of ARID5B gene that could be associated with susceptibility to childhood ALL in Latin American population; however these observations are not representative of the real situation in Mexico. The purpose of this study was to determine the presence of 2 SNPs of ARID5B in ALL children for establishing their possible association with susceptibility to develop the disease. The study population included 112 children with ALL and 114 controls. Informed consent was obtained before taking saliva samples. The DNA was extracted, and genotyping analysis was performed using TaqMan probes for the SNPs rs10821936 and rs10994982 of ARID5B. The allele and genotype frequencies of ALL cases and controls were calculated. Odds Ratio (OR) with 95% confidence intervals [95% CI] were performed using Fisher’s exact test. P-values less than 0.05 were considered statistically significant. For the SNP rs10821936, 112 ALL cases and 114 controls were analyzed and for SNP rs10994982 only 52 ALL cases and 99 controls were studied. The frequency of the risk allele C of the SNP rs10821936 was significantly higher (p=0.0001) in cases (89.3%) than in controls (70.2%); the risk allele A of the SNP rs10994982 was significantly higher (p=0.003) in controls (78.8%) than in cases (65.2%). The frequency of the risk allele C and in A patients was higher than reported in other populations. The frequency of these alleles in the control group is similar to that observed in healthy indigenous population from Mexico and Guatemala. These results suggest that the genetic component of our patients could be influencing the susceptibility to develop ALL and could also explain its high incidence in Mexico. This work was supported by Fonundos del Presupuesto Federal para la Investigación 2012 (085/2012).

Copyright © 2014 The American Society of Human Genetics. All rights reserved.

Prostate cancer (PCA) is clinically heterogeneous, making it difficult to accurately predict which patients will have an indolent vs. aggressive disease course. Therefore markers that can help distinguish aggressive tumors in high-risk patients are urgently needed. In 2012 Cuzick et al. reported a cell cycle progression (CCP) score (based on expression of 31 genes in tumor tissue) as a predictor of PCA death in a conservatively managed patient cohort. To further evaluate the CCP score we tested its performance in a population-based cohort of 383 men diagnosed with clinically localized PCa in 1993-1996 or 2002-2005 who underwent radical prostatectomy (RP). Recurrence information was obtained from participant surveys, medical records, and the SEER cancer registry. Over a follow-up period of 12 years, 278 men (73%) had no evidence of recurrence and 105 (27%) had recurrence events, including 27 (7%) patients who developed metastasis or died of PCA (lethal PCa). Tumor RNA was isolated from FFPE blocks of RP samples and gene expression data were obtained using the Illumina Human Genome DASL HT Assay. CCP score was based on average expression levels of 30 of the 31 progression genes (c18024 was excluded due to low quality). Based on Cox proportional hazards models, the hazard ratio (HR) associated with a change from the 25th to 75th percentile of the CCP score (range 7.97-9.96) was 1.25 (95% CI 0.96-1.62; P = 0.10) for overall recurrence, and for the subset of patients with lethal PCa it was 2.01 (95% CI 1.26-3.21; P = 0.004), adjusted for clinical variables (ie, age at diagnosis, Gleason score (range 7.97-9.96) was 1.25 (95% CI 0.96-1.62; P = 0.10) for overall recurrence, and for the subset of patients with lethal PCa it was 2.01 (95% CI 1.26-3.21; P = 0.004), adjusted for clinical variables (ie, age at diagnosis, Gleason score, tumor stage and pathologic stage). Additional progression markers K67 and to the model did not change the results. CCP score was only weakly correlated (by Pearson correlation) with Gleason score (0.13 and P0.16). Secondary analyses showed similar associations when patients were stratified by Gleason scores. This hierarchy of progression markers is off. Several reports have been identified the CTAs expression in various human tumors, such as, ovarian, endo-metrial, cervical, esophageal, and breast cancers. CTAs due to their limited expression pattern, are as promising targets for cancer diagnosis and immuno-therapy. Method: We study the expression of AKAP4, RGS22, SPAG9 and NY-ESO-1 genes in colorectal cancer. A. Tavakoli Koudesh, F. Mahjoubi, B. Mahjoubi, R. Mirzaee.

Background: Colonic rectal cancer (CRC) is the most common gastrointestinal cancer and the second leading cause of death in women after breast cancer and the third leading cause of death in men after lung and prostate carcinomas in the world. Because of the slow progression of CRC and the ability of treatment in primary levels (before lymph-vascular invasion), there is an urgent need to identify non-invasive tumor bio-markers (biological markers) that can prognostic the CRC in primary levels and subsequently decrease the prevalence and mortality of it. Cancer-Testis Antigens (CTAs) are a group of tumor-associated proteins which typically are expressed in normal reproductive cells of men, but their expression in normal somatic cells is off. Several reports have been identified the CTAs expression in various human tumors, such as, ovarian, endo-metrial, cervical, esophageal, and breast cancers. CTAs due to their limited expression pattern, are as promising targets for cancer diagnosis and immuno-therapy. Method: We study the expression of AKAP4, RGS22, SPAG9 and CTAG1B genes from the CTAs family in both tumor and normal tissues of 80 Iranian CRC patients by RT-PCR with the aim of comparing the genes expression and finding a biomarker for early detection and anticipated progress to CRC. According to studies, AKAP4, RGS22, and NY-ESO-1 genes expression in colorectal cancer tissues has not been investigated so far. The SPAG9 gene expression by RT-PCR was first investigated in Iranian patients. Result: Results are pending.
Gene polymorphisms as risk factors for Cervical Cancer in a South Indian Population. P. Upendram1, S. Poornima1, G. Apoorva1, K. Jayanthi1, V. Kiran Kumar1, A. Shah2, Q. Hasan1.1) Department of Genetics & Molecular Medicine, Kaminieni Hospitals, L.B.Nagar, Hyderabad, Andhra Pradesh, India; 2) Department of Oncology, Kaminieni Hospitals, L.B.Nagar, Hyderabad, Andhra Pradesh, India.

Carcinoma of the uterine cervix is the most frequent gynecological malignancy affecting women in developing countries despite being a potentially preventable disease. Globally, it is the second most frequent cancer affecting women with a steady rise in incidence among younger women. It is evident from current literature that specific gene polymorphisms in hormone receptor genes and the cell cycle regulating genes along with ineffective host immune response may enhance the risk for gynecological malignancies. The aim of this case control study from South India was to evaluate four gene polymorphisms (i) T/C SNP of estrogen receptor (ER) alpha gene recognized by Pvull enzyme (rs2234693), (ii) 306bp Alu insertion in the progesterone receptor (PGR) gene (rs1042838), (iii) G870A SNP of CCND1 gene (rs9344) and (iv) C850T SNP of TNF alpha gene (rs909253) in cervical cancer patients. DNA was isolated from a total of 200 women comprising 100 cases with cervical cancer and 100 age-matched healthy controls using the salting out technique. Polymorphisms were evaluated by Polymerase Chain Reaction followed by restriction enzyme digestion / gel electrophoresis. Our results showed a statistically significant association of the C allele of ER gene (OR= 5.8072; 95% CI=3.7671 to 8.9521; p<0.001) with cervical cancer. While, the T2 allele of PR gene did not show any significant association with cervical cancer (OR=1.3700; 95% CI=0.7193 to 2.6093; p= 0.33). The allele CCND1 gene was associated with cervical cancer (OR= 1.917; 95% CI=1.286 to 2.857; p<0.0014) and the T allele of TNF alpha gene was not significantly associated with this cancer (OR= 1.413; 95% CI=0.95 to 2.102; p=0.087), however the heterozygous CT genotype was protective. In Multifactor Dimensionality Reduction (MDR) analysis, the ER and CCND1 gene polymorphisms showed an interaction with cervical cancer (p=0.0001). Our results suggest that the ER Pvull polymorphism and CCND1 G870A polymorphism can be used as biomarkers in identifying women with an increased risk of getting cervical cancer. A larger study in different ethnic groups is warranted for establishing them as biomarkers for cervical cancer.

Association of common Cancer stem cells (CSCs) genes variants with gallbladder cancer susceptibility and prognosis in North Indian population. Ashok Kumar1, A. Shah1, V. Upendram1, L. Mittal1, A. Gupta1, K. Shah1, A. Kiran Kumar1, A. Kumar2.1) Medical genetics, Sanjay Gandhi Post Graduate Institute of medical Sciences, Lucknow, Uttar pradesh, India; 2) Dr. Balraj Mittal, Medical genetics, Sanjay Gandhi Post Graduate Institute of medical Sciences, Lucknow, Uttar pradesh, India; 3) Annapurna Gupta, Department of Medical Genetics, Medical genetics, Sanjay Gandhi Post Graduate Institute of medical Sciences, Lucknow, Uttar pradesh, India; 4) Kiran lata Sharma, Medical genetics, Sanjay Gandhi Post Graduate Institute of medical Sciences, Lucknow, Uttar pradesh, India; 5) Dr. Vijay Kumar, King George's Medical University, Lucknow, Uttar pradesh; 6) Dr. Ashok Kumar, Medical genetics, Sanjay Gandhi Post Graduate Institute of medical Sciences, Lucknow, Uttar pradesh, India.

Cancer of the uterine cervix is the most frequent gynecological malignancy affecting women in developing countries despite being a potentially preventable disease. Globally, it is the second most frequent cancer affecting women with a steady rise in incidence among younger women. It is evident from current literature that specific gene polymorphisms in hormone receptor genes and the cell cycle regulating genes along with ineffective host immune response may enhance the risk for gynecological malignancies. The aim of this case control study from South India was to evaluate four gene polymorphisms (i) T/C SNP of estrogen receptor (ER) alpha gene recognized by Pvull enzyme (rs2234693), (ii) 306bp Alu insertion in the progesterone receptor (PGR) gene (rs1042838), (iii) G870A SNP of CCND1 gene (rs9344) and (iv) C850T SNP of TNF alpha gene (rs909253) in cervical cancer patients. DNA was isolated from a total of 200 women comprising 100 cases with cervical cancer and 100 age-matched healthy controls using the salting out technique. Polymorphisms were evaluated by Polymerase Chain Reaction followed by restriction enzyme digestion / gel electrophoresis. Our results showed a statistically significant association of the C allele of ER gene (OR= 5.8072; 95% CI=3.7671 to 8.9521; p<0.001) with cervical cancer. While, the T2 allele of PR gene did not show any significant association with cervical cancer (OR=1.3700; 95% CI=0.7193 to 2.6093; p= 0.33). The allele CCND1 gene was associated with cervical cancer (OR= 1.917; 95% CI=1.286 to 2.857; p<0.0014) and the T allele of TNF alpha gene was not significantly associated with this cancer (OR= 1.413; 95% CI=0.95 to 2.102; p=0.087), however the heterozygous CT genotype was protective. In Multifactor Dimensionality Reduction (MDR) analysis, the ER and CCND1 gene polymorphisms showed an interaction with cervical cancer (p=0.0001). Our results suggest that the ER Pvull polymorphism and CCND1 G870A polymorphism can be used as biomarkers in identifying women with an increased risk of getting cervical cancer. A larger study in different ethnic groups is warranted for establishing them as biomarkers for cervical cancer.

Association of GST polymorphism with susceptibility to Leukemia and differential chemotherapy response. S. Caplash1, S. Kaur1, R. Arora2.1) Department of Human Genetics, Punjabi University, Patiala, India; 2) Oswal Cancer Hospital, Ludhiana.

Glutathione S-transferases (GST) isoenzymes play a significant role in phase II biotransformation and detoxification of many xenobiotics including environmental carcinogens, pollutants and drugs. Both the genetic polymorphisms and expression pattern of GST genes may have a major impact on cancer susceptibility, inter-individual variability in the prognosis, drug effects and toxicity. Of these, GSTT1 and GSTM1 isoenzymes are highly polymorphic with homozygous deletion of either or both genes resulting in absence of enzyme activity. We carried out a case-control study involving 150 Leukemia patients and 194 normal healthy controls from Punjab (North India). Multiplex PCR was carried out to determine GSTM1 and GSTT1 polymorphism. The frequency of individuals carrying GSTM1 and GSTT1 null genotypes was higher among Leukemia patients (42 % and 23 %) as compared to the control group (41 % and 19 %), although the difference found was not statistically significant (p<0.05). Heterogeneity in patients’s response to chemotherapy is consistently observed across populations so follow up of patients is in process to analyze the association of GST genotypes with differential chemotherapy drug response.
3474M

BCR-JAK2 Translocation with Complexes Chromosomal Karyotype, A. Oztürk Kaymak1, C. Sonmez2, S. Ozomurcu3k, G. Guntas2, C. Civan Bozdağ2, F. Altuntas1, 1) Medical Genetics, Dr. A.Y. demetevler Oncology Research and Education Hospital, ANKARA, Turkey; 2) Biochemistry, Dr. A.Y. demetevler Oncology Research and Education Hospital, ANKARA, Turkey; 3) Hematology, Dr. A.Y. demetevler Oncology Research and Education Hospital, ANKARA, Turkey.

sequencing is routinely used for diagnosis in chronic myeloproliferative neoplasms(2). Polycythemia vera, primary myelofibrosis and essential thrombocythemia have been occurred because of V617F, the common gain of function mutation. Although the high ratio of JAK2 point mutations, JAK2 rearrangements are rare and have been associated with ALL and CML. Today only about 22 cases are reported with JAK2 rearrangement. Previous cases with JAK2 rearrangements are diagnosed with acute myeloid leukemia and lymphoblastic leukemia, atypical chronic leukemia, myelofibrosis, myelodyplastic syndrome (4, 5). The partners of chromosome 9 are various such as chromosome 2, 3, 4, 5, 6, 8, 12 and 22 at JAK2 rearrangements. The fusion genes are ETV6, PC1M, P1N2, P1F2, PAX5, SSBP2 and BCR- BCR-JAK2 fusion has leukemogenic role in myeloproliferative disorders. Our patient also had chromosome 9 deletions that is one of the most common structural rearrangements in MDS. The last anomaly of our patient was 8p11 deletion that is also associated with myeloproliferative neoplasms. Here we present BCR- JAK2 translocation with complexes chromosomal karyotype. With our case, the total number of cases with BCR-JAK2 fusion in leukemia become five. But our case is the first one with complexes chromosomal karyotype.

3474T

Evaluation of rapid whole-body magnetic resonance as screening strategy for early cancer detection in L-Fraumeni syndrome patients, M. Achatz1, 1), D. Paixao3, M.D. Guimaraes1, A. Nogueira1, R. Chojnacki1, 1) Department of Oncogenetics, A.C.Camargo Cancer Center, Sao Paulo, Brazil; 2) International Research Center, A.C. Camargo Cancer Center; 3) Department of Imaging, A.C. Camargo Cancer Center.

Li-Fraumeni Syndrome (LFS) is a rare autosomal dominant syndrome that predisposes to a high-risk for developing multiple early onset cancers that includes breast, sarcomas, brain and adrenocortical carcinoma. It is related to germline mutations in TP53 gene. Screening strategies for early diagnosis in carriers constitutes a major challenge due to the wide tumor spectrum. Rapid whole body MRI (RWB-MRI) has been proposed as a screening strategy according to the Toronto protocol and its effectiveness needs to be accessed in different populations. In Brazil, there is a high prevalence of TP53 mutation carriers due to the occurrence of a founder mutation, p.R337H TP53, detected in 0.3% of South and Southeastern Brazilian population. This mutation occurs in the oligomerization domain and has a lower penetrance than regular DNA binding domain mutations in the TP53 gene. Moreover, due to genetic modifiers, tumors occur at a later age than in the regular LFS carriers. This was probably one of the reasons why the mutation spread to such proportions in Brazil. This constitutes a public health issue since its occurrence may be present in a large number of people in the most populated area in the country. The aim of this study is to evaluate the efficacy of RWB-MRI for early cancer detection in p.R337H TP53 mutation carriers and compare it to classical LFS mutation carriers.RWB-MRI was performed in 33 TP53 germline mutation carriers, including 29 cases with p.R337H TP53, detected in 0.3% of South and Southeastern Brazilian populations. In Brazil, there is a high prevalence of TP53 mutation carriers due to the occurrence of a founder mutation, p.R337H TP53, detected in 0.3% of South and Southeastern Brazilian populations. In Brazil, there is a high prevalence of TP53 mutation carriers due to the occurrence of a founder mutation, p.R337H TP53, detected in 0.3% of South and Southeastern Brazilian populations. This mutation occurs in the oligomerization domain and has a lower penetrance than regular DNA binding domain mutations in the TP53 gene. Moreover, due to genetic modifiers, tumors occur at a later age than in the regular LFS carriers. This was probably one of the reasons why the mutation spread to such proportions in Brazil. This constitutes a public health issue since its occurrence may be present in a large number of people in the most populated area in the country. The aim of this study is to evaluate the efficacy of RWB-MRI for early cancer detection in p.R337H TP53 mutation carriers and compare it to classical LFS mutation carriers.RWB-MRI was performed in 33 TP53 germline mutation carriers, including 29 cases with p.R337H TP3 gene. One malignant lesion was detected in an eighteen-years-old p.R337H female carrier. RWB-MRI evidenced bilateral renal cortical alterations. Images were further confirmed with abdominal MRI which showed a solid lesion with enhancement in right kidney and a lesion with benign aspects in the left kidney. Lesion in the right kidney was surgically removed and anatomic-pathological findings indicated a papillary renal cell carcinoma. Incidental findings were detected in two out of the 33 patients. Further imaging techniques were done to detail diagnosis and no unnecessary biopsies were performed. No lesions were detected among the classical LFS carriers. The same cohort will further undergo WB-MRI in a 12- and 14 month interval. Preliminary results from this study demonstrated the effectiveness of RWB-MRI in detecting one case of a malignant lesion among 29 p.R337H LFS patients. It indicates that this might be an important imaging tool for early detection in LFS p.R337H TP53 mutation carriers.

3476S

Spectrum of mutations in BRCA1 and BRCA2 genes in Hereditary Breast/Ovarian Cancer families from Algeria: current knowledge and implications in genetic counseling and testing, F. Cherabti1, R. Bakour2, C. Mehenni3, H. Gaceb1, I. Derouiche1, K. Akli2, K. Bougaida1, 1) Unit of Genetics, LMCB, Faculty of Biological Sciences,USTHB, Algiers, Algeria; 2) Anti Cancer Center, Bila, Algeria.

Breast and ovarian cancer is one of the major challenges in the clinical management of patients with prostate cancer is to gauge the risk of developing progressive disease. We therefore used genetic markers to understand the dynamics of genomic alterations during disease progression. Using a novel, advanced multicolor fluorescence in situ hybridization (FISH) approach, we enumerated copy number of six genes previously identified by array CGH to be involved in aggressive prostate cancer - TBL1XR1, CTNNBP2, MYC, PTEN, MEN1, and PDGFB - in six non-recurrent and seven recurrent radical prostatectomy cases. An ERG break-apart probe to detect TMPRSS2-ERG fusions was included. Subsequent hybridization with biotin-labeled DNA of cells resulted in signal counts for all probes in each individual cell analyzed. We neither observed differences in the degree of chromosomal and genomic instability, i.e., intratumor heterogeneity, nor in the percentage of cells with TMPRSS2-ERG fusion between tumor samples with or without ERG. However, tumors from patients that progressed had more chromosomal gains and losses, and showed a higher degree of selection for a predominant clonal pattern. The loss of PTEN was the most frequent aberration in progression (10/11 cases, 91%), followed by high-level 20q gain (7/11 cases, 64%). The 30% high gain of MYC was observed in one progressor, which was the only lesion with an ERG gain, but no TMPRSS2-ERG fusion. According to our results, a probe set consisting of PTEN, MYC and TBL1XR1 would detect progressors with 86% sensitivity and 100% specificity. This will be evaluated further in larger studies.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3478T
Sequence variation in known cancer susceptibility genes identified in high-risk breast cancer cases from the French GENESIS study. F. Lesueur1,2, E. Girard3,4, F. Damialet3,3, M.G. Dondon1, L. Jarhoux5, S. Bon Marchais1, M. Gauthier-Villars5,3, B. Buecher4, C. Lasset4, P. Berthet4, C. Delattet5, J.P. Fricker5, M. Longy5, O. Caron5, P. Pujol5,6, A. Chevrier7, P. Gesta8,9, E. Mouret-Fourme10,11, C. Maugard12, N. Servant12, V. Meyer12,13, A. Bolańez-Ruige14, E. Barillot5,15, J.F. Delattet5,16, S. Mazoyer5, N. Andrieu5, O. Sinilnikova5,16, D. Stoppa-Lyonnet17, CEGE platform, GENESIS investigators. 1) Inserm U900, Institut Curie, Mines ParisTech, Paris, France; 2) Genetics of Breast Cancer group, Cancer Research Center of Lyon, INSERM U1052, CNRS UMR5206, Université de Lyon, Centre Léon Bérard, Lyon, France; 3) Service de Génétique, Institut Curie, Paris, France; 4) Unit of Genetic Epidemiology and Prevention, Centre Léon Bérard, Lyon, France; 5) Unit of gynecological pathology, centre François Baclesse, Caen, France; 6) Centre René Gauducheau, Nantes, France; 7) Unit of Oncology, Centre Paul Strauss, Strasbourg, France; 8) Laboratory of Molecular Genetics, Institut Bergonié, Bordeaux, France; 9) Department of Medical Oncology, Institut Gustave Roussy, Villejuif, France; 10) Unité d’oncogénétique, Hôpital Arnaud de Villeneuve, Montpellier, France; 11) Centre Henri Becquerel, Rouen, France; 12) Centre Georges Renon, Niort, France; 13) Department of Public Health, Institut Curie, France; 14) Laboratoire de diagnostic génétique, Nouvel Hôpital civil, Strasbourg, France; 15) Centre National de Génotypage, Evry, France; 16) Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon, Centre Léon Bérard, Lyon, France. In France, a pathogenic BRCA1 or BRCA2 mutation is found in at most 16% of tested index cases of families predisposed to breast and ovarian cancer. Whole exome sequencing (WES) of BC probands identified (PALB2, BRIP1, RAD51C, RAD51D, XRCC2, etc.), and with the introduction of massive-parallel sequencing (MPS) in clinical laboratories some of these genes are already being screened in patients referred to the clinics as part of a diagnosis. However, the statistical significance of these genes in clinical practice because their associated cancer risks have not been precisely estimated yet due to the too small number of families segregating the mutations identified so far. Our goal was to assess the frequencies of protein-truncating mutations (TM) and missense (MS) identified in the newly identified BC susceptibility genes in a sample of high-risk women participating to the GENESIS study, in order to estimate the number of families that would need to be tested to get a precise estimate of cancer risks in mutation carriers. GENESIS is a unique French national existing resource for familial BC research where >1700 index cases affected with BC and negative for BRCA1/2, >800 affected sisters and >1,500 unrelated matched healthy women have been enrolled. Blood samples, clinical and epidemiological data have been collected for each study participant through the national network of cancer genetics clinics. High-throughput genotyping and MPS projects have been initiated in order to decipher the genetic landscape of BC. Here we examined rare variants (with reported frequency <0.5% in the public sequencing databases) in known BC susceptibility genes, mismatch repair genes and Fanconi Anemia genes identified through exome sequencing in a subset of 100 GENESIS high-risk cases. We found 2 protein-truncating mutations (TM) and 3 likely deleterious missense substitutions (MS) in ATM, 1 TM and 4 MS in CHEK2, and 1 TM and 5 MS in genes of the MRN complex (MRE11A, RAD50 and NBN) occurring in unrelated families. We will next conduct full mutation screening of the gene panel in GENESIS cases and controls to estimate the frequency and the distribution of the different classes of variants in the two groups. National family studies will be set up to assess penetrance and to study cancer spectrum associated with the pathogenic variants. This should help clarifying relevance of mutation screening of these genes to familial cancer clinics and counselors.

3479S
Genome-wide association study (GWAS) for transaminase elevations in pazopanib-treated patients. X. Wang1, Z. Xue1, C. Carpenter2, P. Harten1, M. King3, S. Sinnett3, T. Johnson4, J. Shen4, S. Spraggs4, M.R. Nelson1, M. Chen5, K. Deen5, I. Mitic5, L. Pandite5, N. Kaplowitz2, C-F. Xu6, 1) GlaxoSmithKline, Research Triangle Park, NC, USA; 2) GlaxoSmithKline, Collegeville, PA, USA; 3) Klinikum Essen-Mitte, Essen Germany; 4) GlaxoSmithKline, Stevenage, UK; 5) School of Medicine, University of Southern California, Los Angeles, CA, USA. Background: Pazopanib (Votrient™, GlaxoSmithKline), an oral angiogenesis inhibitor, has been approved for the treatment of advanced renal cell carcinoma and soft tissue sarcoma. Transaminase elevations have been commonly observed in pazopanib-treated patients in clinical studies. We conducted a GWAS using data from 8 phase II and III pazopanib clinical studies to evaluate association between genetic variants and treatment-emergent aspartate aminotransferase (AST) elevation. Methods: 1271 pazopanib-treated (monotherapy) patients with cancer were included in the analysis. Thirty million genotyped or imputed variants were tested for association with on-treatment peak ALT and time to first ALT ≥3x upper limit of normal (ULN) or ≥5xULN. Further analysis with sequence-based 4-digit HLA alleles from 5 genes was also performed. Results: One genetic variant, near NNT, rs80228453 (minor allele frequency 7.5%) was associated with peak ALT (P=2x10-8) at a genome-wide significance level for common variants (P≤5x10-8), with the variant genotype associated with increased ALT. Among patients with ALT<3xULN, 18% carried the NNT variant genotype (vs 6% in patients with ALT≥3xULN). NNT encodes nicotinamide nucleotide transhydrogenase, which is essential for mitochondrial defense against oxidative stress and redox detoxification. Additional variants associated with ALT ≥5xULN were identified, some of which are known to be germline variants in regions of known implications for drug-induced liver injury, including the MHC region. Subsequent HLA analysis identified an association between HLA-B*57:01 and maximum ALT (P≤1x10-4) as well as time to 5xULN (P≤1x10-4). A statistical significant association was also observed for the number of HLA alleles tested (P=5x10-4). HLA-B*57:01 carriers were seen in 5% of all patients, 10% of patients with ALT≥3xULN and 12% with ALT≥5xULN. The incidence rates of ALT≥3xULN and ALT≥5xULN were 36% and 27%, respectively, in HLA-B*57:01 carriers, and 23% and 9%, respectively, in HLA-B*57:01 non-carriers, with an odds ratio (95% confidence interval, P value) of 3.2 (1.6-6.6, 9x10^{-4}) and 4.6 (2.1-9.9, 1x10^{-4}), respectively. Conclusions: We identified statistically significant associations at NNT and HLA-B for ALT elevation in pazopanib-treated patients. Complicated, these findings would provide novel insight into independent mechanisms of pazopanib liver toxicity and may influence clinical management.

3480M
Correlation among MDR1, MRP and hTERT expression level and clinical response in colon cancer patients. s. sha’bani1, f. mahjoubi1, s. samanian1, b. mahjoubi2, r. mirzaei3, a. tavakoli4. 1) national institute of genetigic engineering and biotechnology, tehran, Select a Country; 2) iran university of medical science. Introduction: Colon cancer is one of the common cancers in the world. Despite current advances in the treatments of cancer, the clinical result is far away from expectation yet. Drug resistance is still a major obstacle in treatment of cancer. In this study, we attempted to investigate the possible correlation among MRP1, MRP and hTERT expression level and multidrug resistance in colon cancer patients. Materials and Methods: Tumor and adjacent normal tissues from 35 colorectal cancer patients were assessed for the mRNA expression level of MDR1, MRP and hTERT by Real Time RT-PCR. Results: A statistically significant increase in MDR1 and hTERT expression level was observed in tumoral tissues in comparison with normal tissues. However MRP expression level was not significantly increased in tumoral tissues. Furthermore, no correlation was seen among MDR1, MRP and hTERT expression level. Conclusion: MDR1 and hTERT have no direct correlation, but mRNA expression of these two genes in addition to other factors indirectly helps to tumorgenesis and cancer progression.
Association of IL-1β gene polymorphism with HCC related to viral causes. H.A. Abdalla1, A.A. Badawy1, R. Morin1, K. Farag1. 1) Medical Biochemistry, Faculty of Medicine, Mansoura University, Mansoura, Egypt, 2) Oncology Center, Mansoura University, Egypt

Hepatocellular carcinoma (HCC) is the 5th most common malignancy in the world. In Egypt, the incidence of HCC is doubled over the last decade. Recently, numerous functional gene polymorphisms among pro-inflammatory and pro-fibrogenetic factors and their inhibitors have been thought to influence carcinogenesis including Interleukine-1 (IL-1β), which is considered one of the potent cytokines that has effects on cell survival and proliferation. The aim of the current study was to investigate the association of IL-1β gene polymorphism with HCC patients on top of viral versus non viral cases. The study included 134 subjects; 89 patients suffering from HCC; 61 patients with HCC on top of viral hepatitis (37 HCV and 24 HBV) and 28 patients with HCC without viral hepatitis, and 45 healthy individuals with no liver affection. IL-1β-31 gene polymorphism was studied by PCR amplification followed by RFLP analysis. Quantitative determination of serum IL-1β level was performed by ELISA. In patients with HCC on top of chronic HCV infection, IL-1β-TT genotype was found to be more frequent than other genotypes; whereas in patients with HCC on top of chronic HBV infection, IL-1β-CT genotype was more frequent than other genotypes. IL-1β is significantly elevated in HCC patient on top of both types of virus hepatitis when compared with HCC due to non-viral causes (p<0.05). It is shown that IL-1β was highly significant elevated in patients with TT genotype of IL-1β gene rather than other IL-1β genotypes (p=0.004). In conclusion, People with T homozygous of IL-1β-31 gene polymorphism in combination with serum level of IL-1β could be considered as candidate molecular biomarkers for development of HCC on top chronic HCV or HBV infection. Further studies are recommended to explain and confirm these conclusions.

Expression regulation of ABC transporters (ABC) genes in leukemia cells during granulocytoid maturation processes on imatinib therapy. L.S.R.A Pedroza1, C.A.F. Nunes1,2, J.A.R. Lemos1,2, 1) Federal University of Pará State (UFPA) Belem, Brazil; 2) HEMOPA and Fundação Centro de Hemoterapia e Hematologia do Pará (HEMOPA), Brazil

Despite the results of imatinib (TKI) therapy to control patients Chronic Myeloide Leukemia (CML) progress, prolong overall survival of 88% of new diagnosed patients, 20% will fail. Indeed, TKI are not curative options, cause it can not eliminated all cancer stem cells that differentiated to mature hematopoietic blood cells, and disease can relapse. The search for mRNA related to prognostic predicts, and membrane transports responsible for imatinib efflux from the stem cells became a way to overcome resistance. Multiresistant protein - MRP and Multidrug resistant - MDR, associated to drug resistance in a tumor/drug specific way and ABCG2/B1 responsible for the imatinib extrusion are members of ABC superfamily. To gain more insight about the ABCs expression on leukemic cells on imatinib therapy we assessed the transcription of C24D3+ cells into a pool of 4 CML patients in sustained major molecular response - MMR on imatinib therapy and understanding the ABCs possible function during imatinib therapy and understanding the ABCs possible function in hematopoietic stem cells and mature blood cells. We aim to validate these findings on peripheral blood and in cito. Understand the pathways involved in ABCs function and know patients ABC expression profile considering pharmacogenetic interaction are crucial to test new therapy protocol in vitro or to clinical trial and management.

Inherited NK Cell Defective Mutations in Chinese Lymphoma Patients with HHV infection. Y. Zhang1, H. Liu1, F. Wang2, W. Teng3, X. Chen2, Z. Zhu2, C. Teng3, Y. Lin3, J. Yang2, D. Lu1. 1) Clinical Laboratory Division, Ludaopei Hematology & Oncology Center, Beijing, China; 2) Department of Hematology, Peking University First Hospital, Beijing

Background: It has been reported that some patients with lymphoma may harbor mutations of PRF1, UNC13D and STXBP2, which lead to NK cell deficiency and are causative genes of familial hemophagocytic lymphohistiocytosis (FHL). Data of the association between genetic defects and lymphoma is limited for Chinese patients. Methods and Cases: This study included 45 unrelated Chinese patients with either Hodgkin or non-Hodgkin lymphomas (26 males, 19 females; age range 3 to 60 years) and with active HHV1-8 infection. PCR and Sanger sequencing of all coding exons and flanking sequences of UNC13D, PRF1, STXBP2 and STX11 were performed. Active HHV1-8 infections were tested by PCR method. Results: Mutations were observed in 12/45 patients (26.7%). 9/12 (75%) patients carried a total of 8 different UNC13D mutations, including 6 with monoallelic mutations, 1 with homozgyous mutation, and 2 with compound heterozygous mutations. 2 patients had PRF1 mutations, one with monoallelic mutation and the other with biallelic mutations. 1 patient was detected to carry STX11 monoallelic missense mutation. No patient carried STXBP2 mutation. All mutations in these patients were confirmed to be germline derived by pedigree analysis or sequencing of non-lysosomal tissue of the same patient. Conclusions: Germline UNC13D mutations are frequent in Chinese lymphoma patients with HHV infection. Involvement is less frequent for PRF1 and STX11. Inherited NK cell defective mutations might be the innate predisposing factor of active HHV infection and together lead to the lymphomagenesis.

Whole genome exome sequencing to identify novel candidates for hereditary predisposition to UM uveal melanoma. M.H. Abdell-Rahman1,2, D. Hedges1, D. Pilarski1, G. Boru1, J.B. Massengill1, K. Rai1, F.H. Davidorf1, C.M. Cebulla1. 1) Dept Ophthalmology, The Ohio State University, Columbus, OH; 2) Division of Human Genetics, Department of Internal Medicine, The Ohio State University, Columbus, OH; 3) Pathology Department, Menoufiya University, Shebin Elkom, Egypt.

Background: About 12% of uveal melanoma (UM) patients have features suggestive of hereditary cancer predisposition. Germline mutation in BAP1 explains only a small subset of these patients suggesting the existence of other candidate genes. Methods: Whole exome sequencing (WES) was carried out on six BAP1 mutation-negative UM patients from 5 different families. Three patients had family history of UM and two patients had personal history of UM and personal or family history of cancers associated with the BAP1 tumor predisposition syndrome (TPDS). In one family, the index case and a paternal 1st cousin once removed with UM were sequenced. Results: We focused our primary analysis on identifying unreported, deleterious or potentially deleterious (stop loss, stop gain, nonsynonymous, splice variance, frame shift) germline variants in any of the 522 COSMIC “cancer census genes”. We identified deleterious or possibly deleterious mutations in MAML2, TPR, ARID1A, PBRM1, PRDM1 and MUTFYH. These variants were subsequently confirmed by direct sequencing. Segregation analysis ruled out contribution of MAML2, ARID1A, PBRM1, PRDM1 and MUTFYH. Also, sequencing of the tumor tissue from the patient with germline TPR (translocated promoter region) mutation showed significant loss of the mutant allele in tumor suggesting that it is likely not contributing to the pathogenesis of the disease. No pathogenic variant was detected in any of the cancer predisposition gene in more than one patient. However, we identified several potential candidate genes that have not been previously implicated in hereditary cancer predisposition. Conclusions: This pilot whole exome sequencing (WES) study suggests that hereditary cancer predisposition in these patients is not caused by coding mutation in any of the 522 COSMIC cancer census genes. It also suggests locus heterogeneity across families, causative genetic alterations in non-coding regions and/or both as mechanisms for hereditary predisposition to UM. We have identified several additional potential candidate genes that we are further exploring. Our study provides important preliminary data to further identify potential candidate genes for hereditary predisposition to UM.
Two novel germline BAP1 mutations in two unrelated families with features of the BAP1 Tumor Predisposition Syndrome. C.M. Celagina, R. Pilar, K. Rie, J.B. Massengill, G. Boru, F.H. Davidson, M.H. Abdel-Rahman 1,2. 1) Ophthalmology and Visual Science, The Ohio State University, Columbus, OH; 2) Division of Human Genetics, Department of Internal Medicine and Comprehensive Cancer Center, The Ohio State University, Columbus, OH; 3) Department of Pathology, Menoufia University, Egypt. Objective: The BAP1 Tumor Predisposition Syndrome (TPDS) is a recently recognized hereditary cancer predisposition syndrome. Four main cancers, uveal melanoma (UM), cutaneous melanoma, mesothelioma, and renal carcinoma, are associated with this syndrome. Here we report the clinical phenotype of two new independent families with novel germline BAP1 mutations and features suggestive of the BAP1 TPDS. Methods: Germline BAP1 mutations were evaluated by direct sequencing of all BAP1 exons and adjacent intronic regions. Two new families were evaluated, one presenting with cutaneous melanoma and mesothelioma in the proband and a family history of UM and several other cancers. In the second family, the proband presented with peritoneal mesothelioma and a strong family history of pleural mesothelioma as well as UM. Results: In the first family we identified a nonsynonymous mutation in the ubiquitin carboxyl-terminal hydrolase domain of BAP1 (c. 539T>C, p. Leu180Pro). The mutation was present in a first degree relative with UM (obligate carrier) and in an unaffected second degree relative of the proband. However, the mutation was not identified in other first and second degree relatives in the family, including those with cutaneous melanoma and liposarcoma. In the second family a 3 bp intronic deletion was identified (g.2270_2272delCA, c. 256-4, 256-2del). This mutation was predicted to be a splice site pathogenic variant. We also tested the same gene region in unaffected relatives and found no evidence of the same mutation. Physical examination of mutation positive individuals showed coarser skin, and broader hands and feet as compared to mutation negative family members. In addition, the mutation positive members lost most of their body hair and the PDGFRA gene was also tested in these individuals to have an important role in skeletal development in addition to other developmental pathways and we hypothesize that this gain of function mutation is the cause of abovementioned phenotypic features. To our knowledge, this is a first report of two families with a comprehensive genetic, morphological, pathological and clinical characterization of Familial TPDS Syndrome. Such information may not only help in the clinical recognition of affected individuals, but also in follow-up investigations for IFPs and possibly related GISTs.

Predisposition to Burkitt Lymphoma in Williams-Beuren Syndrome. D. Guenat,1,2 J. Soulier,3 C. Rizzari,4 H. Fryssira,5 C. Lundin,6 C. Borg,7 PS. Rohrich. 1) Laboratory of Cellular and Molecular Biology, University Hospital of Besançon, Besançon, France; 2) UMR1098 Inserm/EFS-BFC/UFC, LabEx LipSTIC, Besançon, France; 3) Saint-Louis Hospital APHP and Hematology University Institute (IUH), University Paris-Diderot, Paris, France; 4) Department of Pediatrics, San Gerardo Hospital, University of Milano-Bicocca, Monza, Italy; 5) Department of Medical Genetics, Medical School, University of Athens, Greece; 6) Department of Clinical Genetics, University and Regional Laboratories, Skåne University Hospital, Lund University, Lund, Sweden; 7) Department of Medical Oncology, University Hospital of Besançon, Besançon, France; 8) Department of Hematology, University Hospital of Nice, Nice, France. Chromosomal disorders are frequently associated with predisposition to cancer as shown by the increased incidence of leukemia in Down syndrome, Williams-Beuren Syndrome (WBS [MIM 194050]), and WBS fragilis syndrome (WBFs). WBS is caused by a hemizygous microdeletion on chromosome 7q11.23 spanning 1.5 Mb and encompassing 28 genes and 2 mRNA loci. WBS is not currently considered as a risk factor for cancer. However, a thorough review of the literature of development of primary cancer in WBS patients revealed that 4 primary cancers and 4 primary tumors were described in the WBS patients and, strikingly, 66% of them were Burkitt Lymphomas (BL). We report here 2 novel cases of BL in children with WBS. DNA isolated from lymphoma cells and normal tissue of the 2 novel cases (one from France and one from Italy) was analyzed. In addition, tumor and normal DNA of 2 other children were investigated: a Greek child with WBS who presented a BL whose case was reported in 2004 and a Swedish child without WBS but with a sporadic lymphoma that has shown a 7q11.23 deletion in a previous study. CGH-array and Next Generation Sequencing BAP1 were used to characterize the size of the WBS deletion and to seek for the presence of a loss of heterozygosity (LOH) in the 7q11.23 locus. The constitutional microdeletion observed in the normal DNA of the 3 WBS patients corresponded to the typical WBS critical region of length 1.5 Mb and in accordance with the location of LOH in the WBFs DNA. Interestingly, a somatic deletion observed in the sporadic BL was mono-allelic and similar to the one observed in WBS patients. No other rearrangement was observed in this tumor. Thus, BL may arise in the presence of either constitutional or somatic microdeletions of the 7q11.23 region, and in 2 of 3 cases in the WBS critical region. A number of genes mapping to this region are involved in DNA repair pathways (BAZ1B, RFC2, GTF2I family genes) or B-cell proliferation and differentiation (LAT2, FZD9). The haplosufficiency of one or more members of this gene family may be the substrate for tumor suppressor gene mutation and lymphoma initiation. Given these observations and the difficult management of epidemiological data concerning association of rare diseases such as WBS and BL, we suggest that the deletion of the WBS critical region is involved in the development of Burkitt Lymphomas, regardless of the WBS critical region lacking of Burkitt Lymphoma occurring at an early age. Functional studies are needed to identify the precise molecular pathways that drive the B-cell lymphomagenesis in WBS patients.
3489M
Genome-wide DNA methylation patterns and genetic ancestry in sporadic breast cancer patients from a Latino population. M. Cappella1, L. Brignon1, N. Artagaveitia2, O. Stefansson3, M. Esteve1,2, B. Berton1, M. Berdasco1. 1) Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay; 2) Departamento Básico de Medicina, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay; 3) Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute, L’Hospitalet de Llobregat, Barcelona, Catalonia, Spain; 4) Department of Physiological Sciences II, School of Medicine, University of Barcelona, Barcelona, Catalonia, Spain; 5) Instituto Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain.

Genetic variants alone are not enough to explain a complex disease like cancer. Alterations in DNA methylation patterns have been associated with different types of tumors. Although methylation profiles are often tissue-specific, recent data indicate that epigenetic changes in peripheral blood leukocytes are promising risk markers for solid tumors.

In order to detect risk markers for sporadic breast cancer in the Uruguayan population, we integrated genetic and epigenetic information of individuals. We determined the level of global leukocyte DNA methylation (gIDNAmet) in cancer patients by relative quantification of SmC by HPLC-MS; and then, we focused on site-specific studies using HumanMethylation450 BeadChip.

A global DNA hypomethylation in breast cancer patients was detected when it was compared with healthy controls. This result suggests its potential use as a risk marker. Since the Uruguayan population is admixed, we studied the correlation between gIDNAmet and individual genetic ancestry (determined by analyzing ancestry informative markers). We found a negative correlation between African ancestry and gIDNAmet in cancer patients. These results suggest that the ancestral admixture genome structure could be modeling DNA methylation patterns and its effect and underscore the importance of considering genetic ancestry as a modifying variable in epigenetic association studies in admixed populations like Latino ones.

We detected 77 differentially methylated CpG sites in blood of breast cancer patients including known cancer genes involved in disease-specific pathways and also novel candidates with possible implication in breast tumorigenesis. This set was characterized and validated in an independent sample of breast cancer patients who were being treated at the Hospital Lipschutz Nenard. These CpG sites are able to differentiate leukocytes of sporadic breast cancer patients from controls. It also distinguishes healthy tissue from breast tumor, even in cases of hereditary breast cancer.

We identified potential DNA methylation at global and site-specific level in leukocytes of patients with sporadic breast cancer, which suggests the existence of a systematic variation in DNA methylation associated with susceptibility to tumor development.

3489T

The aim of this study was to develop a massively parallel sequencing workflow for routine diagnosis for hereditary breast and ovarian cancer syndrome. The aim was to reduce the delay of results with a highly specific and sensitive method. A workflow was designed using a multiplex amplification approach (BRCA MASTR Dx) with 5 multiplex PCR reactions covering the complete coding regions of BRCA1 and BRCA2, followed by synthesis sequencing on Miseq. Bioinformatic analysis of the results was made through the use of adapted settings in SeqNext software. A training set of 56 DNA samples containing polymorphisms, unclassified variants and pathogenic mutations previously identified by others technologies, was used. This group consisted of 448 variants controls polymorphisms, 23 deleterious mutations (2 missense, 2 nonsense, one splice mutation and 18 frameshift mutations) including 4 located in homopolymers. The workflow developed on the Miseq permitted the identification of all variants (n=491), including those located in or close to homopolymers. The routine diagnosis was established for series of 48 patients per run to achieve a high depth reading (averaged per amplicon 4000X). Eighteen series of 48 patients (n=864) were analyzed by this approach. A series of 48 patients from DNA extraction to results is performed by a technician in 10 days, included medical validation. The workflow meets the sensitivity and specificity requirements for genetic diagnosis of breast and ovarian cancer and improves on the cost-effectiveness of current approaches. It could be respond to situations in which the expectation of the mutational status is an emergency such as the practice of contralateral prophylactic surgery or the development of targeted therapies (Anti PARP).

3490S
New genome-wide technologies and low volume, archival, formalin-fixed paraffin embedded material: are the two compatible? L.M. FitzGer- ald1, E.M. Wong2, J.E. Joo3, J. Pedersen1, J. Mills1, CH. Jung2, J. Chung2, G. Severi3, M.C. Southey4, G.G. Giles5. 1) Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria, 3004, Australia; 2) Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Victoria 3010, Australia; 3) TissuPath Specialist Pathology, Victoria 3149, Australia; 4) Victorian Life Sciences Computation Initiative, Carlton, Melbourne, Victoria 3010, Australia; 5) Human Genetics Foundation, Torino, 10126, Italy.

Purpose: Archival formalin-fixed, paraffin-embedded (FFPE) tumour material is a widely available, valuable resource for genomic research but one that has been under-utilised due to low yields of highly degraded nucleic acids. This limitation has prevented assaying such samples on early generation, whole-genome arrays, which required high quality and quantities of starting material. However, with improved whole transcriptome amplification systems and the evolution of next-generation array technologies, it is becoming more feasible to use low quantities of FFPE material. In a pilot study, we investigated whether low volume, FFPE diagnostic samples of prostate cancers (PCa) could be successfully applied to the Infinium HumanMethylation450 BeadChip (HM450K) array and the Agilent SurePrint G3 Human Gene Expression v2 Microarray. Here we present the results of both the individual and combined analyses of these assays.

Methods: Fourteen FFPE diagnostic PCA samples, consisting of six transrectal ultrasonography biopsies and six transurethral resection samples were selected from the Cancer Council Victoria Prostate Cancer Program. DNA and RNA were extracted from macro-dissected prostate tumour tissue. DNA samples were whole genome amplified and assayed on the Infinium HM450K array. RNA samples were amplified in duplicate using two different whole-transcriptome amplification kits and assayed on the Agilent SurePrint array. Quality control measurements and analyses were performed using publicly available and in-house analysis tools. DNA and RNA were successfully amplified for 11 samples. These samples were applied to the Infinium HM450K and the Agilent SurePrint arrays, and quality control assessment and analyses are currently underway. Conclusion: One of the greatest challenges in PCa research is distinguishing at diagnosis those men who are at risk of progressing to life-threatening disease from those whose disease will remain indolent. Diagnostic tumour samples are a vital resource for biomarker discovery but have been under-utilised due to the nature of the material. This study suggests that low volume FFPE material can be successfully interrogated using next-generation, genome-wide arrays, thereby opening up a previously under-utilised resource for epidemiological studies.

Copyright © 2014 The American Society of Human Genetics. All rights reserved.
3492M

Background: Retinoblastoma (RB) is a neoplasm of the retina that occurs in infants and young children. Sporadic bilateral retinoblastoma (RB), results from a de novo germline mutation in the RB1 gene. This de novo mutation occurs in the father’s gamete before the child’s conception in about 94% of cases. However, paternal age is not significantly associated with risk of having a child with de novo germline mutation in RB. We have previously shown that paternal diet, occupation, and medical exposures to mutagens before the child’s conception affect this risk. Methods: We investigated the mutation load of the paternal genome and compared it to that of the mother’s and unrelated control adults using whole genome or exome sequencing from 12 parent-child trios with sporadic bilateral RB and four controls. Whole genome sequencing library was prepared using Illumina TrueSeq DNA library protocol. Exome capture was performed using NimbleGen V3.0 oligonucleotide libraries. Following sequencing on Illumina HiSeq, variant calling was done with the GATK Unified Genotyper and filtered using the GATK VariantFilter and SelectVariants tools. Exome sequencing reads were filtered against the target regions to remove off target reads. Remaining variants were annotated with Annovar before further filtering via Python script and in Excel. Results and Discussion: The results indicate that the distribution of mutations including non-synonymous, synonymous, frameshift and non-frameshift mutations was significantly different between the paternal and maternal genomes (p< 0.001). The paternal genomes included increased number of non-synonymous damaging changes. Analysis of the mutated genes indicated novel mutations in genes in nucleotide excision repair and metabolism pathways. These results provide a strong association with our epidemiologic data on risk enhancing paternal exposures that includes medi-cal radiation, diet and occupation. In conclusion, this study validates our hypothesis that paternal exposure and mutation load increase the risk of having a child with sporadic RB. We acknowledge that the data set is very small; however considering that RB is a rare disease, it is important to emphasize that we have the largest number of RB parent-child trios currently available in any institution, and we are in the process of enlarging the data set.

3493T
Differential Gene Expression In Key Oncolytic Pathways Observed Between Caucasian-American and African-American Women with Triple-Negative Breast Cancer. J.E. Getz1, L.L. Baumbach-Beaumont2, C. Gomez1, M.E. Ahearn2, M. Jord2, C.R. Legendre2, W. Tembe1, S. Nasser2, V. Yellapantula1, M.D. Pegram3, J.D. Carpent2. 1) Integrated Cancer Genomics Division, Translational Genomics Research Institute, Phoenix, AZ; 2) Miller School of Medicine, University of Miami Medical School, Miami, FL; 3) Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA.

In the U.S., breast cancer (BC) incidences among African American women (AA) and Caucasian women (CA) are similar, however, AA have a significantly higher mortality rate (20%). AA often presents with tumors of higher grade, later stage, and are more likely to be triple-negative breast cancer (TNBC). Although multiple factors contribute to these cancer disparities, it is essential that we identify the molecular characteristics and underlying biological differences between CA and AA TNBC. In this study, we performed gene expression array analysis on archived formalin fixed paraffin embedded (FFPE) blocks from a multi-ethnic U.S. cohort of node-negative (NO) TNBC samples. Total RNA was isolated from 10 μm scrolls from each FFPE block. Following cDNA synthesis, each sample was hybridized to a breast-enriched gene expression array (Affymetrix, BC DSA Research Tool). Expression analysis was conducted using GeneSpring 12.1 analytical software and after contemporary data analyses, the final study cohort consisted of 10-AA, 13-CA. PCA analysis revealed that the samples clustered well with respect to ethnicity and unsupervised cluster analysis, based on ethnicity and genes, was performed. The resulting dendogram segregated into distinct subgroups based on ethnicity, revealing a pattern of differential gene expression between the subcohorts. A list of differentially expressed genes (DEG) were selected using ANOVA analysis (fold change > 3.0, p value <0.05) followed by the Benjamin/Hochberg method for multiple-testing correction. Finally, the lists of DEG were uploaded into GeneGo MetaCore to identify pathway level differences. These analyses revealed differentially expressed genes pathways enriched for cytoskeletal remodeling, cell adhesion and EMT pathways. In particular, significantly deregulated genes associated with the Wnt/beta-catenin pathway were observed in the AA cohort as compared to the CA, suggesting that this pathway may contribute to the more aggressive phenotype in AA women diagnosed with TNBC. In summary, our results indicate gene expression differences within several key oncolytic pathways across these ethnic groups and these results are currently being validated through several parallel approaches. These studies have important implications for further understanding BC and TNBC health disparities, as well as future tailored approaches to prediction, prevention and therapeutic advances.

3494S
Shared genetic background between chronic gastroesophageal reflux and Barrett’s esophagus and esophageal adenocarcinoma, consistent with a causal relationship. P. Ghahrekhani1, J. Tung1, T.L. Vaughn2, D.C. Whittemore1, N. MacGregor3, Barrett’s and Esophageal Adenocarcinoma Consortium. 1) QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia; 2) 23andMe, Mountain View, CA, USA; 3) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

Esophageal adenocarcinoma (EA) is a cancer with rising incidence and high mortality rate. Every year, 0.12-0.5% of patients with Barrett’s esophagus (BE), a precancerous metaplastic change, develop EA. Chronic gastro-esophageal reflux (GERD) is a risk factor for both BE and EA. However, it is not clear whether GERD actually causes those diseases. In a previous study, we found genetic overlap between BE and EA but did not detect significant evidence of common genetic background between GERD and BE/EA. The previous study was limited by small sample size, which may have resulted in false negative findings. Here we aimed to further investigate the polygenic overlap between GERD and BE/EA using larger datasets, and to use the Mendelian randomization approach to investigate the causality of GERD for those diseases. The top genome-wide GERD-associated SNPs from a study of 23andMe in 8,743 GERD cases and 43,932 controls were used to calculate polygenic risk scores for GERD in our study population (2,051 BE cases, 1,239 EA cases, and 2,088 controls of European descent from the Barrett’s and Esophageal Adenocarcinoma Consortium (BEACON)). We showed the 23andMe risk scores significantly (P<5e-4) predicted BEACON reflux status, providing the strongest evidence to date for a role of genetic variation in reflux predisposition. Further, the scores were used as an instrumental variable in a Mendelian randomization approach using logistic structural equation model estimator to investigate the causality of GERD for BE and EA. Unlike our earlier smaller study of GERD, we found a significant polygenic overlap between GERD and BE/EA. Our preliminary Mendelian randomization analysis provides estimates for the role of GERD on BE and EA risk which are consistent with causality. Findings on causality are potentially clinically important as they support the use of GERD interventions to control cancer risk.
The Kaiser Permanente Biobank: A multi-region, multi-ethnic resource linking specimens and electronic medical records for broad research in an integrated health care delivery system. KA Goddard1, N. Car- 
rente, Denver, CO; 5) Center for Health Research Southeast, Kaiser Perma-
rente Georgia, Atlanta, GA; 6) Center for Health Research, Hawai‘i, Kaiser Perma-
rente Hawaii, Honolulu, HI; 7) Mid-Atlantic Permanente Research Institute, Mid-Atlantic Permanente Medical Group, Rockville, MD; 8) Kaiser Perma-
rente So Southern California, Department of Research and Evaluation, Pasaden, CA.

Kaiser Permanente (KP) is an integrated healthcare delivery system spanning seven geographic regions in the US and including 9.3 million members. The diverse membership reflects the US population, where about half of the members are non-Hispanic white. Membership includes between 4% and 29% of the total population in the coverage areas, and broadly repre-
sents populations in overall demographic and socioeconomic status measures. In the last decade, KP established several regional biobanks that now include specimens and phenotypic data for a total of 210,000 members. We describe the design and features of a new initiative to consoli-
date these efforts into a single entity, called the KP Biobank, and to expand the cohort to a total of 500,000 participants. The KP Biobank will include a general cohort (410,000 participants), designed to represent the diversity of KP members, and two specialized cohorts: a cancer cohort with 30,000, participants, and a pregnancy cohort with 60,000 participants. Specimens will be stored centrally and linked to phenotypic information from the medical record. Clinical and electronic data systems capture comprehensive and longitudinal records on every medical encounter, including diagnoses, procedures, pharmacy data, tumor registry, pathology records, imaging, and laboratory results. Participants have about 23 years of membership on average (range 0 - 71 years), with comprehensive electronic records for the last 15 years.

We also collect patient reported information on demographics and behavioral and environmental exposures via survey. The KP Biobank is designed for both discovery and translational research. Our setting of research units embedded within a healthcare delivery system supports clinical integration of new discoveries and research across the translational continuum. The KP Biobank Translational Research Center will specifically support this function. The consolidated governance of the KP Biobank will streamline and facilitate collaboration and partnerships through a single Access Review Committee and application procedure. Any researcher, including external scientists, may apply for use of specimens. Participants consent to broad future uses of specimens and data, and allow for broad data sharing, including through dbGaP. This novel resource will facilitate future genomic research, particularly on research questions addressing diverse populations, longitudinal measures to assess changes over time, and long-term outcomes.


Lymphoid cancers are collectively the sixth most common cancer in the USA and impose a large health and financial burden on individuals and health care systems. Lymphoid cancers are a biologically diverse group of neoplasms, yet familial aggregation of various lymphoproliferative disorders has been observed, suggesting shared genetic risk factors. Specifically, non-Hodgkin lymphoma, Hodgkin lymphoma, multiple myeloma and chronic lymphocytic leukemia aggregation may be specific to gender, age, ethnicity, infectious agents and familial relationship. Our collection of over 120 families with more than one lymphoid cancer diagnosis will aid in understanding the inter-relationship of lymphoid cancers between and within lymphoid cancer families. The relationships between different lymphoid cancers observed in these families was investigated by categorizing affected individuals by type of lymphoid cancer, age of onset, affected generation and rarity of each type of lymphoid cancer. We have observed that within these families, some types of lymphoid cancers co-occur more frequently than expected based on their incidence in the general population. This co-aggregation of specific types of lymphoid cancers may be due to genetic and/or other risk factors that are shared within families. In addition, we observe anticipation in the set of families; the mean age of diagnosis among generations was significantly different for non-Hodgkin lymphoma, Hodgkin lymphoma, chronic lympho-
ctic leukemia and all lymphoid cancers collectively.
3498M
The prostate cancer risk mutation G84E in HOXB13 is associated with the subtype of ETS fusion negative adenocarcinoma with early age of diagnosis. M. Luedeke², A. Rinckleb¹, D. Stanford², L. FitzGerald, J. Schlieueker³, T. Wahltors, R. Eeles⁴, ⁷, Z. Kote-Jarai, S. Weikert, H. Krause, K. Herkommer, J. Hoegel, C. Maier, and. the ICPGC. 1) Department of Urology, University Hospital of Ulm, Ulm, Germany; 2) Fred Hutchinson Cancer Research Center, Division of Public Health Science, Seattle, Washington, USA; 3) Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA; 4) Institute of Biomedical Technology, University of Tampere and FinLab Laboratories, Tampere, Finland; 5) Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland; 6) The Institute of Cancer Research, Sutton, UK; 7) Royal Marsden National Health Service Foundation Trust, London and Sutton, UK; 8) Department of Urology, University Hospital Charité, Berlin, Germany; 9) Department of Urology, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany; 10) Institute of Human Genetics, University Hospital of Ulm, Ulm, Germany; 11) International Consortium for Prostate Cancer Genetics.

HOXB13 was discovered as the first prostate cancer (PrCa) specific high-risk susceptibility gene. The most prevalent HOXB13 germline mutation in PrCa patients of European descent is HOXB13 G84E, which likely originated in Northern Europe. Previous molecular examination of a set of G84E driven tumors suggested a distinct somatic phenotype, where oncogenic ETS gene fusions appear at unusually low frequencies as compared to the general prevalence of ETS fusions in PrCa (22 % vs approx. 50%). Hypothesizing that HOXB13 could predispose to ETS fusion negative PrCa we have analyzed 942 cases from three European ancestry populations for the coincidence of HOXB13 G84E and the most common ETS fusion, TMPRSS2-ERG (T2E), in corresponding tumor samples.

While the prevalence of T2E fusions was similar among study sites (range: 57.2% - 60.1%), the frequency of G84E genotypes differed markedly between US (1.5%), German (3.6%) and Finnish samples (9.3%). Despite the expected frequency gradient among study populations, all subsamples showed a strong enrichment of G84E mutation carriers among T2E fusion negative cases as compared to fusion positive cases (center adjusted OR = 4.96; 95% CI = 2.30 - 11.9; p = 0.0001). Consistent with the previous study, the crude frequency of the T2E fusion in HOXB13 G84E carriers was 23.5% (range 16.7% - 28.5%). Examination of disease characteristics highlighted age at diagnosis, with fusion positive cases being diagnosed 1.85 years earlier (95% CI = 1.05 - 2.74 years) (p = 0.0001). Age at diagnosis in G84E carriers did not differ significantly from non-carriers (p = 0.13). However, within the subtype of fusion negative carcinoma, which is usually associated with later ages at diagnosis, carriers of G84E were diagnosed on average 3.45 (0.61 - 6.3) years earlier (p = 0.018). No associations were seen for tumor stage, tumor grade or diagnostic PSA levels.

In conclusion, this study has demonstrated a tumor type specific association for HOXB13 G84E mutation carriers having a higher frequency of T2E fusion negative PrCa. While the T2E fusion negative subtype is known to be associated with later ages of diagnosis, HOXB13 driven tumors within this subtype may represent an early onset subgroup.

3499T

Allogenic bone marrow transplant (allo-BMT) recipients and patients with a hematologic malignancy present unique challenges when their blood/buccal samples are submitted for genetic testing. While blood cells from allo-BMT recipients are usually entirely derived from the transplant donor, buccal samples from such patients can be chimeric, containing cells derived from both the recipient and the bone marrow donor. Therefore, molecular genetic testing performed on DNA derived from blood/buccal samples of allo-BMT patients may reflect the genetic status of the transplant donor and not the recipient. Blood/buccal samples from individuals with a hematologic malignancy (HM) are also not recommended for genetic testing. This is because genetic analysis may detect somatic, non-heritable mutations caused by the hematologic malignancy, which are present only in the patient’s blood. Genetic testing on blood/buccal samples submitted by patients with hematologic malignancies may therefore reflect the genetic status of the tumor rather than the germline genetic status of the patient. Here we present data from three representative cases that illustrate the complexity of genetic testing for allo-BMT recipients and patients with a hematologic malignancy. In one case, we tested a buccal sample from an allo-BMT patient that was sufficiently chimeric such that the donor DNA was detected on the genetic test results. In the second case, we illustrate the potential to report a false negative result on a blood sample from an allo-BMT patient. In the final case, we demonstrate the detection of a mosaic large rearrangement in a patient with a hematologic malignancy. This large rearrangement was subsequently determined to be a somatic mutation present only in the patient’s blood cells. We have therefore determined that cultured cells derived from skin fibroblasts are the most suitable type for allo-BMT patients, which is in accordance with the published 2012 NCCN guidelines. Fibroblasts are also recommended for HM patients where there is a high likelihood of leukemic cells circulating at a quantity that is detectable by our genetic analyses. We have collaborated with ARUP to assist in the procurement of fibroblast cells. It is imperative that germline testing be performed on these patients prior to testing, as they may impact the interpretation of test results and subsequent clinical management of the patient and their family members.

3500S
Familial lung cancer: A genetic epidemiology study. D. Mandal¹, A. Bencaz⁴, J. Chambless⁵, J.E. Bailey-Wilson⁶. 1) Department of Genetics, LSU Health Sciences Center, New Orleans, LA; 2) National Human Genome Research Institute, National Institutes of Health, Baltimore, MD.

The association between lung cancer (LC) and smoking is well known. However, only 15% of smokers are diagnosed with LC. Inherited genetic factors have a major role in lung cancer etiology. About 10% of LC cases (22,000 cases per year in the U.S.) have at least one first-degree relative affected with LC, and 25% of cases at least one first- or second-degree relative affected, indicating that family history is a significant risk factor. In the last few years histology specific genetic alterations in lung cancer have been reported in epidemiologic studies. In our previous studies, we have reported adenocarcinoma (38%) being the most common histologic subtype, followed by squamous cell (25%) and small cell (15%) in familial cancer cases. The objective of the present study is to describe the rapid ascertainment of familial lung cancer cases and to determine the co-segregation of histologic subtypes transmitted in familial lung cancer families. Eligible subjects (N=203) with two or more relatives affected with primary lung cancer were recruited from 24 parishes across southern Louisiana. Diagnosis of primary LC was confirmed by medical records, and histologic subtype (N=173) was abstracted from pathology reports. About 20 of these families were developed as multigenerational families and biological specimens were collected on affected lung cancer cases and unaffected family members. Annual follow up on these families confirm a change of health status and new diagnosis in the families. Histological data on these families will be analyzed to co-segregation of specific histologic subtype and lung cancer affection status will be determined in this unique collection of families. Methods for rapid ascertainment of lung cancer cases maintaining HIPAA compliance and communication about dbGaP policies with general public in the development of bio-repository will be discussed in detail.
3501M Identification of a novel founder MSH2 c.705delA mutation causing colon cancer in a Druze population. M. Melas1, C.C. Studenmund1, K. McDonnell1, L. Raskin1, P.S. Boonstra1, M. Zawistowski2, B. Mukherjee3, F. Lejbkowicz4, H.S. Rennert4, G. Rennert4, S.B. Gruber1. 1) USC Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA; 2) USC Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, Tennessee; 3) Department of Biostatistics, University of Michigan, Ann Arbor, MI; 4) CHS National Israeli Cancer Control Center, Haifa, Israel.

Lynch Syndrome is an autosomal dominant cancer genetic syndrome resulting from mutations in the DNA mismatch repair genes MSH1, MSH2, MSH6, PMS2 and EPCAM. Individuals carrying mutations in these genes characteristically develop cancer of the colon, rectum, endometrium and other cancers at an early age. A number of founder mutations in Lynch causing genes have previously been described as important in identifying individuals at risk within specific populations. The Druze are a unique religious and ethnic population that originated approximately 11,000 years ago in the Arabian Peninsula. The population subsequently migrated throughout the Middle East, primarily to modern Syria, Lebanon, Israel and Jordan, between 1,000 and 500 years ago. The Druze are culturally discouraged from marrying outside their communities and endogamy rates are relatively high. In the present study we report a novel MSH2 mutation, c.705delA, in a group of Druze individuals who developed Lynch Syndrome. To establish whether this variant is a founder mutation we performed haplotype analysis using SNP genotyping data from 12 Druze carrying MSH2 c.705delA mutation (Infinium HumanCore Beadchip) and 47 Druze individuals from the Human Genome Diversity Project (HGDP) with chromosome 2 phasing. Concomitantly, recombination length was determined by carrying out a short tandem repeat genetic analysis of the same Druze carriers together with 9 Druze non-carriers and 8 Christian Arab non-carriers employing six highly polymorphic microsatellite markers near the MSH2 gene: D2S367, D2S370, D2S378, D2S391, D2S1248 and D2S2238. Haplotype phasing using Shapetl2 software showed a 13 Mb haplotype common to all MSH2 c.705delA mutation carriers. This haplotype was not found in non-carriers and also not in the 47 Druze individuals from the HGDP. Attempts to determine the age of the mutation were confounded by significant consanguinity and complicated population structure of the Druze, which highlights the difficulty of estimating the age of a mutation that depends on the assumption of independent ascertainment. In summary, the present study describes a novel founder MSH2 c.705delA mutation among the Druze. This establishes this deletion as a founder mutation and illustrates the challenges of statistical genetic analyses in highly inbred populations. Additionally, the results of this study may contribute to improved cancer genetic screening and cancer care in the Druze.

3502T Investigating the Genetic Basis of Breast Cancer Disparities Using Whole Genome Sequencing and Parallel Computing. J.J. Pitt1,2, TY. Yoshimatsu1, Y. Zheng3, AJ. Grundstad4, J. Tuteja5, A. Odetunde6, G. Khramtsova7, TO. Ogundiran5, CP. Babalola4, OA. Ojengbede6, CO. Olopade7, D. Huo6, KP. White1,9, Ol. Olopade1,9. 1) Committee on Genetics, Genomics, and Systems Biology, The University of Chicago, Chicago, IL, USA; 2) Institute for Genomics and Systems Biology, The University of Chicago, Chicago, IL, USA; 3) Center for Clinical Cancer Genetics and Global Health, Department of Medicine, The University of Chicago, Chicago, IL, USA; 4) Institute for Advanced Medical Research and Training, College of Medicine & Faculty of Pharmacy, University of Ibadan, Ibadan, Oyo, Nigeria; 5) Department of Surgery, University of Ibadan, Ibadan, Oyo, Nigeria; 6) Centre for Population & Reproductive Health, College of Medicine, University of Ibadan, Ibadan, Oyo, Nigeria; 7) Section of Pulmonary and Critical Care, Department of Medicine, The University of Chicago, Chicago, IL, USA; 8) Department of Health Studies, The University of Chicago, Chicago, IL, USA; 9) Department of Human Genetics, The University of Chicago, Chicago, IL, USA.

Breast cancer incidence and mortality rates vary widely based on geography and ethnicity. Compared to Caucasians, individuals with African ancestry are more likely to die from breast cancer, which is in part attributable to higher prevalence of the triple negative subtype. It is hypothesized that these disparities manifest through population-specific inherited and somatic genetic variation. Breast cancer specimens and germline DNA have been collected from the University College Hospital Ibadan in southwest Nigeria. To date, 17 tumor-normal pairs and 50 germline breast cancer samples have undergone whole genome sequencing on the Illumina platform, with neoplastic and non-neoplastic sequenced to average depths of 90x and 30x, respectively. In addition, 25 asthmatics and 75 healthy Nigerians from the surrounding community have been sequenced as controls. To handle the computational burden inherent to large-scale sequencing analyses, we have developed SwiftSeq, a modular, highly-parallel workflow for fast, efficient, and robust processing of DNA sequencing data. Using Genome Analysis Toolkit’s best practices, SwiftSeq is able to completely align, process, genotype, and annotate a 30x genome in ~36-40 hours. By scaling with compute resources, our framework can analyze hundreds of genomes in days, rather than weeks. Currently, 19 germline case and 21 control genomes (mean depth ~27x) have been pushed through SwiftSeq. Each harbor, on average, 3,921,264 SNPs; 621,141 indels; and 306,882 variants. While early onset and/or familial cases lacking deleterious BRCA1/2 mutations were explicitly selected for sequencing, damaging mutations were found in other known cancer susceptibility genes such as BRIP1, FAHNC, and SMARCE1. This underlies a potentially expanded role for these genes in African breast cancer pathogenicity. With continued sequencing, analysis, and comparison to Caucasian tumor-normal genomes from The Cancer Genome Atlas, we will elucidate the unique characteristics of African breast cancer genomes.

Background: Genetic testing for germline mutations in BRCA1/2 has been available for many years, but until recently technical complexity and cost has limited testing access to high-risk cases. Further, many epithelial ovarian cancer (EOC) patients who would be eligible for testing are not referred for a genetic assessment. The advent of targeted therapies for BRCA1/2 associated tumours, and the need for cost-effective new service delivery models in oncology, mandates a population-based approach to direct genetic testing in all women recently diagnosed with EOC.

Methods: To date we have recruited 76 participants out of a total of 139 eligible women. The mean age at diagnosis was 65.5 years (range 29-85), mean time since diagnosis to recruitment = 142 days (range 15-355 days). BRCA1/2 gene testing has been completed on 65 samples, resulting in detection of pathogenic mutations in 6 women and variants of uncertain clinical significance (VUS) in 6 women. Of the women found to have a pathogenic mutation, 3 had no significant previous family history of cancer to their knowledge. We will present our preliminary psychosocial findings and economic analysis, and highlight the challenges and success of our model and implications for genetic testing in oncology settings. Conclusion: The initial GTEOC study recruitment rate is 55%. Pathogenic BRCA1/2 mutations have been identified in 9.2% of women and VUS in 9.2%. This method of population-based genetic testing appears to be acceptable to patients and is less resource-intensive than standard practice where all patients have a full assessment by the genetics team prior to testing.

The InSiGHT Variant Interpretation Committee (InSiGHT VIC) undertook a collaborative effort to develop, test and apply a standardised classification scheme to germline variants in the Lynch syndrome-associated genes MLH1, MSH2, MSH6 and PMS2 (accounting for ~2% of all colorectal and endometrial cancers). Unpublished data submission was encouraged to assist in variant classification and was recognised through microtribution. The scheme was refined by multidisciplinary expert committee review of clinical and functional data available for variants, applied to 2,360 sequence alterations, and disseminated online. Assessment using validated criteria altered classifications for 66% of 12,006 database entries. Clinical recommendations based on transparent evaluation are now possible for 1,370 variants that were not obviously protein truncating from nomenclature. All data were merged into one publicly available database, which will continue to be curated according to the classification scheme. This large-scale endeavour has important implications for the clinical management of suspected Lynch syndrome families; at the same time, it provides an important model of international multidisciplinary collaboration for DNA variant interpretation.

3503S

Insecticide Exposure Induces Leukemia-Associated Gene Aberrations. M.P. Navarrete Meneses1, M. Betancourt2, E. Bonilla3, M. Altimanaro4, C. Salas5, A. Reyes5, M. Sanabrais5, P. Pérez Vera5. 1) Instituto Nacional de Pediatría, Mexico City; 2) Universidad Autónoma Metropolitana-Iztapalapa, Mexico City; 3) FES-Zaragoza, UNAM, Mexico City.

Background: Leukemia is the most common childhood cancer. It is characterised by the presence of cells with chromosomal abnormalities which are proposed to arise in utero. It is associated with prenatal exposure to pesticides. Epidemiological studies have demonstrated that environmental exposures to these chemicals could be implicated in leukemogenesis, however there is limited biological evidence showing that pesticides can induce leukaemia-associated gene aberrations. Malathion and permethrin are two widely used insecticides; nevertheless their genotoxic and carcinogenic potentials are controversial. The aim of this study was to detect numerical and structural rearrangements in leukemia-related genes on human peripheral blood lymphocytes exposed in vitro to insecticides. Methods: Mononuclear cells from two healthy volunteers were cultured in triplicates for 72h and exposed to 200μM of permethrin or malathion for the last 24h. MLL, ETV6 and RUNX1 genes were analyzed with fluorescence in situ hybridization. Besides, numerical aberrations were assessed using centromeric probes for chromosomes 12, 18 and 14/21. Structural and numerical analyses were performed separately. 1000 nuclei were scored. Groups were compared with U-Mann Whitney test. Results: Permethrin exposure increased the frequency of cells with MLL structural aberrations and copy number deviations significantly, as well as the diversity and complexity of damage. It also induced aneuploidy significantly. An increased frequency of cells with abnormalities in ETV6 and RUNX1 was detected, however it was not significant. On the other hand, a significant increase in the frequency of cells with structural damage in MLL gene was observed with malathion exposure. This effect was not observed on ETV6 and RUNX1 genes. Neither chromosome gains nor losses were significantly induced by malathion exposure. Discussion: Permethrin exposure can affect leukemia-related genes in structure and number, whereas malathion does not. The results show that this pyrethroid insecticide has clastogenic and aneuploidal potential. Similar patterns of alterations have been reported in cells treated with leukemogenic agents such as etoposide. Malathion and permethrin are common insecticides which can be detected in the food for human consumption. Further study is necessary to confirm these effects.

3504M

A systematic approach to clinical classification of DNA sequence variants in mismatch repair genes: the InSiGHT initiative. B.A. Thompson1,2, A.B. Spurridge1, J. Pizzera1, M.S. Greeland1, P. Moller1, F. Macrae3, M. Cenci1,2, M. de Medici1,2, InSiGHT Variant Interpretation Committee. 1) Genetics & Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia; 2) School of Medicine, University of Queensland, Brisbane, Australia; 3) Department of Colorectal Medicine and Genetics, Royal Melbourne Hospital, Melbourne, Victoria, Australia; 4) Vermont Cancer Center, University of Vermont College of Medicine, Burlington, VT, USA; 5) Research Group on Inherited Cancer, Department of Medical Genetics, Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway; 6) Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Italy; 7) Fioren Foundation for Pharmacogenomics, Sesto Fiorentino, Italy.

Clinical classification of sequence variants identified in hereditary disease genes directly affects management of patients and their relatives. The International Society for Gastrointestinal Hereditary Tumours Variant Interpretation Committee (InSiGHT VIC) undertook a collaborative effort to develop, test and apply a standardised classification scheme to germline variants in the Lynch syndrome-associated genes MLH1, MSH2, MSH6 and PMS2. Unpublished data submission was encouraged to assist in variant classification and was recognised through microtribution. The scheme was refined by multidisciplinary expert committee review of clinical and functional data available for variants, applied to 2,360 sequence alterations, and disseminated online. Assessment using validated criteria altered classifications for 66% of 12,006 database entries. Clinical recommendations based on transparent evaluation are now possible for 1,370 variants that were not obviously protein truncating from nomenclature. All data were merged into one publicly available database, which will continue to be curated according to the classification scheme. This large-scale endeavour has important implications for the clinical management of suspected Lynch syndrome families; at the same time, it provides an important model of international multidisciplinary collaboration for DNA variant interpretation.
such as the LOVD and ClinVar are essential for genes with important clinical
and data standardization is increased. Collaborations between resources
which results in a higher-quality level of data over merely providing weblinks,
provides an example of facilitated data sharing between different databases,
significance, 41% as Likely benign, and 10% as Benign. Conclusions This
are asserted as Pathogenic, 3% as Likely pathogenic, 32% as Uncertain
a handful of variants described in OMIM. 27% of the submissions in PALB2
other PALB2 ClinVar submissions are unpublished variants from clinical
submission makes up 77% of the 342 PALB2 submissions to ClinVar. The
database currently contains 242 listed variants in 671 individuals. The LOVD
disorders or independent variant/disorder relationships. Results The LOVD
whether a list of multiple phenotypes indicated individuals with all listed
Genetics (ACMG) standard terms; and variants were aggregated based on
the pathogenicity values were mapped to the American College of Medical
with ClinVar. In the ClinVar submission process, variant data were validated;
the LOVD and ClinVar data models e.g. determining whether a list of multiple phenotypes indicated individuals with all listed
orders or independent variant/disorder relationships. Results The LOVD
database currently contains 242 listed variants in 671 individuals. The LOVD
submission makes up 77% of the 342 PALB2 submissions to ClinVar. The
other PALB2 ClinVar submissions are unpublished variants from clinical
laboratories, structural variants that include PALB2 with other genes, and
a handful of variants described in OMIM. 27% of the submissions in PALB2
are asserted as Pathogenic, 3% as Likely pathogenic, 32% as Uncertain
significance, 41% as Likely benign, and 10% as Benign. Conclusions This
This example of facilitated data sharing between different databases,
which results in a high-quality level of data over merely providing weblinks,
because the variants have been assessed independently by different teams
and data standardization is increased. Collaborations between resources
such as the LOVD and ClinVar are essential for genes with important clinical
consequences such as PALB2.

Combined contribution of intermediate-risk gene rare variants and
modest-risk SNP genotypes to early-onset breast cancer. E.L. Young¹,
B.J. Feng², A.W. Stark³, T.C. Francy⁴, A.M. Paquette¹, J.S. Rosenthal¹,
N. Forey⁵, G. Durand⁶, S. McKay-Chopin³, M. Hashibe¹, J. Gertz¹, F. Le
Calvez-Kelm⁷, F. Lesueur², D.E. Goldgar², S.V. Tavtigian¹, Breast Cancer
Family Registries. 1) Department of Oncological Sciences, Huntsman Can-
cancer Institute, University of Utah School of Medicine, Salt Lake City, USA;
2) Department of Dermatology, Huntsman Cancer Institute, University of
Utah School of Medicine, Salt Lake City, USA; 3) Genetic Cancer Susceptibil-
ity group, International Agency for Research on Cancer, Lyon, France; 4)
Department of Family and Preventive Medicine, Huntsman Cancer Institute,
University of Utah School of Medicine, Salt Lake City, USA; 5) Genetic
Epidemiology of Cancer team, Inserm, U900, Institut Curie, Paris, France.
Since the introduction of genetic testing of BRCA1 and BRCA2 in women
with family histories of breast cancer and/or ovarian cancer, many additional
genes have been identified as breast cancer susceptibility genes, and many
of these genes are included in multigene panel testing. Building from our
published case-control mutation screening studies of ATM, CHEK2,
MRE11A, RAD50, NBN, RAD51, RINT1, and XRCC2, we evaluated the
contribution of these eight genes, plus mutation screening of BARD1 and
genotyping of 18 Breast Cancer Association Consortium (BCAC) confirmed
modest-risk SNPs, to early onset breast cancer among a series of 1300
cases and 1120 ethnically matched controls that did not carry pathogenic
variants in BRCA1, BRCA2, or PALB2. Rare missense substitutions were
graded on severity using Align-GVGD, Polyphen2, CADD, and MAPP. To
avoid gene-by-gene over-fitting, we applied uniform missense substitution
analysis models across all nine genes. Using attributable fraction, familial
relative risk, and receiver operator characteristic (ROC) analyses, we: (1)
compared the contribution of known and likely pathogenic missense substitu-
tions to the contribution of protein truncating variants across the nine gene
set, and (2) compared the contribution of the nine gene mutation screening
data to the 18 BCAC SNP genotyping data. Finally, setting an odds ratio
(OR) of 2.5 as a threshold for the point at which a genotype-defined risk
estimate is high enough to justify early screening (early mammography or
breast MRI), we estimated the fraction of cases and controls that reach the
OR ≥ 2.5 threshold on the basis of mutation screening data or SNP genotyping
data. From preliminary results across several analysis models, mutation
screening of the nine genes suggests that approximately 8.8% of cases
and 3.7% of controls carried one or more rare variants associated with an
OR ≥ 2.5. We plan on combining the intermediate-risk gene mutation screening data with the modest-risk
SNP genotyping data to paint a clearer picture of clinically actionable breast
cancer susceptibility explained by genes other than the established high-
risk susceptibility genes.
3508T
Germline Next Generation Full Gene Sequencing of MLH1, MSH2 and MSH6 Detects Pathogenic Mutations in Cases Previously Tested Negative for a Germline Mutation. R.P. Graham 1, M.S. Delpitoke 1, A. French 2, S. Gallinger 3, M. Cotterchio 3, R. Haile 4, G. Casey 5, M.A. Jenkins 6, J.L. Hopper 7, M. Woods 8, L. Le Marchand 9, J. Potter 10, P.A. Newcomb 11, D. Duggan 12, E.L. Goode 13, N. Lindor 14, S.N. Thibodeau 15. On behalf of the Colon Cancer Family Registry; 1) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Department of Surgery, Mount Sinai Hospital, Toronto, Ontario, Canada; 3) Population Studies and Surveillance, Cancer Care Ontario, Ontario, Canada; 4) Department of Medicine, Division of Oncology, Stanford Cancer Institute, Palo Alto, CA; 5) Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, CA; 6) Centre for Epidemiology & Biostatistics, Melbourne School of Population & Global Health, University of Melbourne, Melbourne, Victoria, Australia; 7) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Victoria, Australia; 8) Memorial University of Newfoundland, St. John’s, Newfoundland, Canada; 9) Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI; 10) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 11) Translational Genomics Research Institute (TGen), Phoenix, AZ; 12) Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ.

Lynch syndrome (LS) is due to germline mutations in mismatch repair (MMR) genes and tumors of affected patients show microsatellite instability. A subset of likely hereditary cases with mismatch repair deficient (dMMR) tumors has no identified germline mutations in clinical screening of coding regions of MMR genes (MLH1, MSH2, and MSH6). In 98 such cases, we examined both coding and non-coding regions of key MMR genes by next generation sequencing (NGS). Germline DNA was obtained from 98 patients with loss of MLH1 [n=71], MSH2 [n=18] and isolated MSH6 [n=9] identified in the Colon Cancer Family Registry and 90 unaffected spousal controls was sequenced using Agilent’s custom capture and an Illumina HiSeq2000. We targeted the coding and non-coding regions (intron, upstream and downstream) for MLH1, MSH2 and MSH6. Pseudogene homology impaired adequate PMS2 capture. Quality control filtering excluded variants based on read depth (<20), genotype quality (<30), and minor allele frequency (>5%). The minimum depth for variants was 10x. The variants were categorized as deleterious, suspected deleterious variants were predicted to alter splicing. IHC data correlated with variant interpretation for deleterious and suspected deleterious variants. Among 98 suspected LS cases analyzed, for which standard testing was negative, we identified 4 deleterious and 6 suspected deleterious alterations (10% rate of diagnosis). The unresolved cases may still be due to the non-coding variants discovered. Alternatively, mutations in other MMR genes, structural or epigenetic changes, or somatic alterations in these genes may explain these cases. This work was supported by grant U1CA167551 from the National Cancer Institute.

3509S
Identification of men with a genetic predisposition to prostate cancer: targeted screening of BRCA1/2 mutation carriers and controls. The IMPACT study of quality of life study. L. N. A. Bancroft 1, 2, 3, N. Aaronson 4, C. Mikropoulos 7, 8, 9, S. Saya 10, E. Page 11, J. Pope 12, E. Farrow 13, E. Castro 14, N. Gadea 15, C. Selikirk 1, S. Buys 16, J. Cook 17, K. Ong 18, R. Davidson 19, D. Eccles 20, M. Tischkowitz 21, L. Greenhalgh 22, J. Barwell 23, 17. Detects Pathogenic Mutations in Cases Previously Tested Negative for a Germline Mutation. R.P. Graham 1, M.S. Delpitoke 1, A. French 2, S. Gallinger 3, M. Cotterchio 3, R. Haile 4, G. Casey 5, M.A. Jenkins 6, J.L. Hopper 7, M. Woods 8, L. Le Marchand 9, J. Potter 10, P.A. Newcomb 11, D. Duggan 12, E.L. Goode 13, N. Lindor 14, S.N. Thibodeau 15. On behalf of the Colon Cancer Family Registry; 1) Department of Health Sciences Research, Mayo Clinic, Rochester, MN; 2) Department of Surgery, Mount Sinai Hospital, Toronto, Ontario, Canada; 3) Population Studies and Surveillance, Cancer Care Ontario, Ontario, Canada; 4) Department of Medicine, Division of Oncology, Stanford Cancer Institute, Palo Alto, CA; 5) Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, CA; 6) Centre for Epidemiology & Biostatistics, Melbourne School of Population & Global Health, University of Melbourne, Melbourne, Victoria, Australia; 7) Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Victoria, Australia; 8) Memorial University of Newfoundland, St. John’s, Newfoundland, Canada; 9) Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI; 10) Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA; 11) Translational Genomics Research Institute (TGen), Phoenix, AZ; 12) Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ.

Lynch syndrome (LS) is due to germline mutations in mismatch repair (MMR) genes and tumors of affected patients show microsatellite instability. A subset of likely hereditary cases with mismatch repair deficient (dMMR) tumors has no identified germline mutations in clinical screening of coding regions of MMR genes (MLH1, MSH2, and MSH6). In 98 such cases, we examined both coding and non-coding regions of key MMR genes by next generation sequencing (NGS). Germline DNA was obtained from 98 patients with loss of MLH1 [n=71], MSH2 [n=18] and isolated MSH6 [n=9] identified in the Colon Cancer Family Registry and 90 unaffected spousal controls was sequenced using Agilent’s custom capture and an Illumina HiSeq2000. We targeted the coding and non-coding regions (intron, upstream and downstream) for MLH1, MSH2 and MSH6. Pseudogene homology impaired adequate PMS2 capture. Quality control filtering excluded variants based on read depth (<20), genotype quality (<30), and minor allele frequency (>5%) in 100 bp windows. We evaluated potential deleterious variants by splicing prediction, published reports and their frequencies in suspected LS patients relative to controls. After filtering, there were 10 coding and 262 non-coding variants in suspected LS patients. Three of the coding variants (MLH1 p.L653R; MSH2 p.N586del; MSH6 p.F365X) were pathogenic and the remainder predicted to be benign. There were 69 non-coding variants in MLH1, 144 in MSH2 and 50 in MSH6. All MLH1 variants were detected once, 40 of which were novel and predicted cryptic splice sites. There was one known pathogenic MLH1 intrinsic variant, c.454-13A>G. Two novel MSH2 variants were predicted to form cryptic donor sites and 3 variants were present in >10% of cases compared to 1-3% of controls. A single private MSH6 variant was predicted to alter splicing. IHC data correlated with variant interpretation for deleterious and suspected deleterious variants. The remaining variants in MLH1, MSH2 and MSH6 were novel and predicted to be neutral. Among 98 suspected LS cases analyzed, for which standard testing was negative, we identified 4 deleterious and 6 suspected deleterious alterations (10% rate of diagnosis). The unresolved cases may still be due to the non-coding variants discovered. Alternatively, mutations in other MMR genes, structural or epigenetic changes, or somatic alterations in these genes may explain these cases. This work was supported by grant U1CA167551 from the National Cancer Institute.

Posters: Cancer Genetics

Copyright © 2014 The American Society of Human Genetics. All rights reserved.