GEN Genetic Engineering & Biotechnology News

Jan 1 2009 (Vol. 29, No. 1)

Feature Article

Linking Disease to Gene Variations

Predisposition Based on Genomics Is Slowly Being Ascertained

Vicki Glaser

Each human chromosome has dozens to hundreds of millions of base pairs of DNA, and variants in those base pairs can be associated with disease.

New technologies and strategies for genetic testing and genomic analysis were the focus of several presentations at the recent American Society of Human Genetics (ASHG) meeting in Philadelphia. Understanding the molecular basis of disease, discerning inheritance patterns of genetic disorders, and clarifying the implications of predisposing genetic factors are core goals of genome-wide association studies (GWAS).

As researchers and clinicians continue to unravel the mysteries of the human genome, they are looking to technology companies to provide next-generation sequencing and genome analysis tools to accelerate whole-genome and targeted DNA sequencing, SNP genotyping, and copy-number variation (CNV) analysis.

Aravinda Chakravarti, Ph.D., ASHG president and professor at the McKusik-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, defined the challenges that must be overcome before the inevitable future of personalized medicine can become reality. “This reality cannot become routine or useful unless we can predict the phenotype of each of our unique genomes within which the vast majority of variation is rare and unique.

“First, we need to understand the origin of variation. To gain an unbiased functional understanding we need to quantify the human mutation rate directly and how it is modulated. Second, progress in prediction will require not merely accumulating empirical facts but theoretical prediction of the functional content of any piece of DNA and the consequences of altering that sequence.”

To achieve this, the field needs to be able to assess many more genomes, to integrate structural variation with mutation/SNP data, and to evaluate variation in noncoding regions of the genome. The hope is that next-generation sequencing technology will allow for sequencing of more genomes with greater coverage and for targeted resequencing to enable detection of rare disease-associated alleles such as small insertions, deletions, or inversions.

Medical Sequencing

In a session entitled, “Using DNA Sequence to Detect Variation Related to Human Disease: The Promises and Challenges of Medical Sequencing,” Richard Wilson, Ph.D., professor at Washington University School of Medicine, explored new technologies for medical sequencing and mutation discovery in cancer. He described the disadvantages of traditional technologies: for example, array-based experiments only identify variation in exonic regions, and PCR-based resequencing is expensive.

Dr. Wilson’s group has used Illumina’s next-generation sequencing platform to perform whole genome sequence analysis and to look for disease-linked variability in acute myeloid leukemia. The researchers identified eight validated somatic mutations, all of which were heterozygous, and two somatic insertions/deletions. The group is pursuing similar studies in glioblastoma.

Stacey Gabriel, Ph.D., director of the Genetic Analysis Platform at the Broad Institute, gave conference attendees an

overview of the ongoing 1000 Genomes Project and described how GWAS are being used as an unbiased discovery tool for identifying rare mutations that may have value as novel drug targets. Dr. Gabriel's group developed a scalable process for exome sequencing that uses hybrid selection and capture of hybridized regions on beads to pull out the exonic regions of the genome.

The 1000 Genomes Project will catalog all SNPs and CNVs with an allelic frequency >1% and make them available in a public database. The Project comprises three pilot programs: 4x genomic coverage of 180 people; 20x coverage of three case/parent trios; and 1,000 genes studied in 1,000 different people. One of the main challenges at present is how to integrate the disparate data derived from multiple different technology platforms and the development of hybrid platforms that can perform SNP and CNV genotyping on the same sample.

Richard Gibbs, from Baylor College of Medicine, focused on directed sequencing approaches for exploring disease genomics and described his group's work using Roche's NimbleGen's "rebalanced" arrays to capture targeted regions of the genome and Roche 454's next-generation sequencing technology. The combination has allowed for greater than 94% coverage of the genome at 10x, according to Dr. Gibbs. He outlined the advantages of using high-density arrays of the exome instead of GWAS for medical sequencing studies: there is less data to manage; the experimental design has greater flexibility; the process is more readily scalable; it is easier to map variants; and the cost is substantially lower.

Richard Lifton from Yale University, commented on the dramatic decrease in the cost of DNA sequencing over the past decade: from $100 per 1,000 high-quality bases in 1998 to less than $0.01 at present, using Illumina's next-generation sequencing platform as an example.

Variation and Clinical Disease

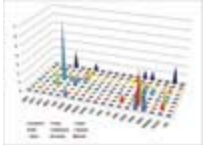
Leslie Biesecker, M.D., chief and senior investigator of the genetic disease research branch of the National Human Genome Research Institute (NHGRI), described a pilot project called "ClinSeq" to explore the translational genome space. The project's goal is to "understand the genetic architecture of disease" by interrogating large sections of the genome in 1,000 unrelated adults for whom clinical data is available.

Translational genomics aims to link genetic variability—whether common or rare variants—to specific disease phenotypes. In the pilot phase, researchers will sequence 400 candidate genes using PCR and capillary-based sequencing technology. As of November 2008, 531 patients had enrolled in the project. Dr. Biesecker reported on 1.6 million sequence reads that revealed about 10,000 variants, 82% nonexonic and 20% exonic.

Daniel Chasman, Ph.D., associate geneticist from Brigham and Women's Hospital, led off the session on "Cardiovascular Genetics and Blood Biomarkers," and presented the results of GWAS in nearly 17,000 women combined with nuclear magnetic resonance analysis of lipoprotein subfractions, which enables the differentiation of lipoprotein particles based on their size. Among the 39 candidate loci found to be associated with cardiovascular disease, 22 mapped to known genes and 12 were novel loci; eight of the novel loci could be linked to lipid metabolism and appeared to affect low density lipoprotein and very LDL concentrations.

A meta-analysis presented by Sekar Kathiresan, M.D., an affiliate of the Broad Institute, found common DNA sequence variants associated with polygenic dislipidemia at 30 loci. The study encompassed 19,840 genomes and about 2.4 million variants. Nearly all of the SNPs associated with abnormal blood lipid concentrations were present in noncoding regions of the genomes.

Therapeutic RNA Strategies



The distribution of mutations across tumor types using the OncoCarta™ Panel v1.0: The genes are listed against the tumor tissue type.

Gideon Dreyfuss, Ph.D., professor from the University of Pennsylvania School of Medicine, led a session devoted to "Innovative RNA Biology: A New Paradigm for the Treatment of Genetic Disorders," in which presenters described therapeutic strategies intended to disrupt the processing of targeted RNA species into proteins. All RNAs exist in cells as ribonucleotide-protein complexes, explained Dr. Dreyfuss, and the many proteins involved in RNA processing provide a wealth of potential therapeutic targets.

Allan Jacobson, Ph.D., professor at the University of Massachusetts Medical School, highlighted efforts to develop drugs that target post-transcriptional regulatory events and promote translation of RNA into protein. In the presence of a mutation that produces a nonsense codon, translation will terminate prematurely. Dr. Jacobson explained that the processes of normal versus premature translation termination differ mechanistically, and that premature termination is at least 100-fold less efficient. The

ability to exploit nonsense suppression as a therapeutic tactic would yield a broadly applicable method—independent of the mutation or affected gene—for treating diseases caused by a nonsense mutations.

Dr. Jacobson described a drug candidate in development by PTC Therapeutics, of which he is a co-founder. PTC124 induces dose-dependent read-through of premature stop codons, but not of normal termination codons. The drug promotes suppression of all three nonsense codons.

Preclinical studies in mouse models of cystic fibrosis and Duchenne's muscular dystrophy have demonstrated the ability of PTC124 to restore the missing disease-related proteins, which are properly localized and appear to be active based on the results of functional assays. The company has completed Phase IIa trials in patients with these disorders and has shown the drug to be well tolerated. In a Phase IIb trial in Duchenne's muscular dystrophy patients will receive high-dose or low-dose drug or placebo for 48 weeks. Completion of the trial is projected for early 2010.

Adrian Krainer, Ph.D., professor at Cold Spring Harbor Laboratory, described an antisense strategy being developed in collaboration with Isis Pharmaceuticals and Genzyme to correct a splicing error that causes spinal muscular atrophy, an autosomal recessive disorder that affects about 1 in 6,000 live births.

A single nucleotide difference results in exon skipping and the production of an abnormal RNA transcript. The antisense compound in development is intended to promote inclusion of the skipped exon to yield a full-length mRNA. Early-stage studies in a mouse model of spinal muscular atrophy have shown increased levels of normal mRNA and associated protein correction when the antisense drug was introduced into the spinal cord.

The use of an RNA interference (RNAi) strategy to silence the messenger RNA that codes for an aberrant protein causing Huntington's disease (HD) was the topic of a presentation by Beverly Davidson, Ph.D., professor at the Carver College of Medicine, University of Iowa. HD is a dominant heritable disorder that results from a polyglutamine insertion in a gene on chromosome 4. The aim of Dr. Davidson's group is to use an RNAi to knock out the mutant HD protein. The researchers inserted a short hairpin RNA (shRNA) sequence and a reporter protein (green fluorescent protein) into a recombinant adeno-associated virus, which acts as a delivery vehicle, inserting its genetic payload into the host cell nucleus.

The shRNA is exported from the nucleus into the cytoplasm, finds the target—the mutant HD mRNA transcript—and degrades it before it can be translated into protein. In a transgenic mouse model of HD, introduction of the recombinant virus resulted in reduced levels of mutant HD protein in the brain. Furthermore, Dr. Davidson reported that a 60% reduction in mutant HD protein yielded notable symptomatic improvement in the treated mice.

Probing Structural Variation

Don Conrad, graduate student at the University of Chicago, described a CNV discovery project under way aimed at developing a comprehensive map of common CNVs at 500 base pairs resolution. The group is using array comparative genomic hybridization (aCGH) technology for whole-genome analysis in European and African populations. Conrad noted that the 5´ ends of genes—where transcription initiates—tend to be CNV hotspots.

Adam Shlien, graduate student, and colleagues from the Hospital for Sick Children in Toronto, are exploring the link between CNV and susceptibility to Li-Fraumeni syndrome (LFS), an autosomal dominant, inherited predisposition to childhood cancer. Patients with LFS carry a variety of different mutations in the p53 oncogene.

Carriers of LFS-associated p53 mutations have about a fourfold increased number of CNVs compared to control subjects. Interestingly, p53 mutant carriers that do not have cancer were shown to have about as many CNVs as carriers with cancer. Shlien proposed that CNVs may contribute to tumorigenesis in LFS by acting as hotspots for additional changes in the tumor genome.

Representing the International Schizophrenia Consortium, Jennifer Stone, Ph.D., told session participants that individuals with schizophrenia have a greater burden of CNVs. Cases have a significantly greater average number of CNVs than controls in the genomes of patients analyzed by the consortium. Dr. Stone also noted emerging evidence that

CNVs may play a role in autism and mental retardation as well.

Alexandre Reymond, Ph.D., assistant professor at the University of Lausanne, presented several conclusions drawn from his group's study of CNV in the mouse genome using the NimbleGen mouse whole genome CGH array to catalog CNVs in 13 inbred strains and 21 wild-caught mice. They also used an Affymetrix gene-expression array to profile the transcriptome of six different tissues.

Reymond described significantly greater variation of gene expression for genes located in a region of CNV in all tissues, but only weak correlation of gene-expression levels and the number of gene copies present. In general, genes in CNV regions were expressed at lower levels than non-CNV genes, and CNV genes were expressed in fewer tissues than non-CNV genes. Furthermore, CNVs appeared to affect the expression of genes located megabases away on the chromosome.
