

Embargo: 4:30 p.m. PT, Tuesday, Nov. 6, 2012

Surprising findings from NHLBI Exome Sequencing Project reported at American Society of Human Genetics 2012 meeting

A multi-institutional team of researchers has sequenced the DNA of 6,700 exomes, the portion of the genome that contains protein-coding genes, as part of the National Heart, Lung and Blood Institute (NHLBI)-funded Exome Sequencing Project, one of the largest medical sequencing studies ever undertaken.

Scientists participating in the project initially expected that individual rare variants would have a greater effect on over 80 heart, lung and blood related traits and diseases of high public health significance, said Suzanne M. Leal, Ph.D., professor and director, Center for Statistical Genetics in the Department of Molecular and Human Genetics of Baylor College of Medicine in Houston, TX.

The researchers found that many (1.1 million) of the 1.2 million coding variants that they identified in exome data from 4,420 European-Americans and 2,312 African-Americans occurred very infrequently in the population and often were only observed in a single individual, explained Dr. Leal, who presented the findings today at the American Society of Human Genetics 2012 meeting.

Dr. Leal added that most of the observed coding variants are population specific, occurring in either European or African Americans. “Of the identified variants, about 720,000 change the genetic code in a manner that could produce flawed proteins. Yet the role played by most of these variants in disease development has not been established,” she said.

The major goal of the project was to understand how variation in the exome affects heart, lung and blood related traits and diseases.

The study participants were selected from a sample of over 220,000 individuals who participated in another National Institute of Health (NIH) supported study that had collected extensive medical data on the participants. “Individuals were selected to have a disease endpoint of interest or an extreme trait value of public health importance,” said Dr. Leal.

By sequencing the exomes of 91 cystic fibrosis patients, Dr. Leal and her research colleagues discovered and replicated an association between variants in the *DCTN4* gene and when a patient first develops a *Pseudomonas aeruginosa* airway infection.*

The researchers were also able to replicate many known associations between individual DNA variants and traits, such as high blood levels of low-density lipoprotein, known as the ‘bad’ cholesterol, and C-reactive protein, which increases the body’s response to inflammation.

(more)

The majority of these findings are for variants that are common in the population, said Dr. Leal.

To detect associations with rare variants, analyses were performed by aggregating information from individual variants within a gene. This approach successfully detected an association with rare variants in the *APOC3* gene that lowers triglyceride levels, an unhealthy type of fat in the blood, said Dr. Leal.

“In order to detect associations with rare variants, due to their modest effects, very large samples sizes are required. In many cases the data from the Exome Sequencing Project gave us leads that had to be evaluated using more study subjects. One mechanism for doing this was by genotyping additional samples using the exome chip, which contains approximately 240,000 coding variants. The Exome Sequencing Project played a very important role in the development of the exome chip, by being the largest contributor of data,” she added.

According to the NHLBI, exome sequencing is an efficient way to search for rare variants associated with complex traits. In contrast to previous genome wide association studies (GWAS), which concentrated on common variants scattered throughout the genome, exome sequencing has the potential to accelerate the search for unambiguous genetic links to disease by focusing attention on the protein coding portion of the genome

In the journal *Science*^{**}, Dr. Leal and her colleagues wrote that GWAS have substantially improved knowledge about common genetic variation, but have been generally uninformative about the patterns of rare variation within the protein coding regions of the genome.

“This is a very new field for which new methodology had to be developed. We learned many lessons in the quality control and analysis of exome data, as well as what types of results one would expect to see when analyzing rare variants. Additionally, the Exome Sequencing Project has been extremely valuable in obtaining a better understanding of population genomics and the history of man,” Dr. Leal said.

The exome sequence variant data is publically available on the Exome Variant Server (<http://evs.gs.washington.edu/EVS>). Additionally phenotype and exome data can be requested from the NIH repository dbGaP (<http://www.ncbi.nlm.nih.gov/gap>).

The Exome Sequencing Project, sponsored by NHLBI, included researchers at the Broad Institute of MIT and Harvard University, Cambridge, MA; University of Washington, Seattle; Ohio State University Medical Center, Columbus; University of Virginia Health System, Charlottesville; Fred Hutchinson Cancer Research Center, Seattle, WA; Stanford University, Stanford, CA; University of Michigan, Ann Arbor; University of Vermont, Burlington; University of North Carolina, Chapel Hill; and Baylor College of Medicine in Houston, TX.

The researchers’ abstract is titled, “Lessons Learned from the NHLBI-Exome Sequencing Project.”

*“Exome sequencing of extreme phenotypes identifies *DCTN4* as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis,” *Nature Genetics* **44**, 886–889 (2012).

“Evolution and functional impact of rare coding variation from deep sequencing of human exomes.” *Science* **337, 64 (2012).

About ASHG

The American Society of Human Genetics is the primary professional membership organization for nearly 8,000 human genetics specialists worldwide. The ASHG Annual Meeting is the world's largest gathering of human genetics professionals and a forum for renowned experts in the field.

###