MedicalXpress

Finding variants in the human genome: HapMap 3 points the way forward for human genetics studies

September 1 2010

New findings show the value of genetic studies across human populations and the value of the latest DNA sequencing technologies to interrogate genetic variation. The results, from the latest phase of the international HapMap Project, are reported in *Nature*.

The researchers' extensive study of genetic variation in multiple populations will form a framework for future genetic studies of variation and disease: their findings highlight the need to examine various populations in order to tease out the widest collection of genetic variants, as well as the requirement to deploy sequencing technologies to find as many variants as possible.

The HapMap Project seeks to identify signposts on the human genome that will simplify the search for important genetic variants. In the latest phase - HapMap 3 - the researchers looked for variants across the genome in 1184 samples from 11 populations. They chose the large sample set and the wide range of populations to maximize the variation they could capture. The project includes both single-letter differences (single-letter polymorphisms, or SNPs) as well as large differences from the loss, gain or duplication of regions, called copy-number polymorphisms, or CNPs.

"Despite the remarkable achievements following from the Human Genome Project, our knowledge of human genetic variation remains

limited," says Professor Richard Gibbs, professor of molecular and human genetics at Baylor College of Medicine in Houston, Texas, and director of the BCM Human Genome Research Center. "Here we have studied more populations and were able to include CNPs in genomewide studies.

"These results tell us more about human genetic variation and about how to study variation successfully."

The results show that rarer variants are distributed more unevenly among populations. This might be expected - evolutionary theory implies that the common variants are generally the older ones, having had greater time to spread through a population - but also cautions that genetic studies should include a wide range of population groups to maximise discovery of more recent, population-specific variants.

"The closer we look at human genetic variation, the greater the granularity," explains Professor Manolis Dermitzakis, from the University of Geneva and one of the project coordinators, and formerly at the Wellcome Trust Sanger Institute. "An important task in genetics is to discriminate between the variants that are important for health and those that are part of the background.

"This new version of the HapMap will help us design ways to do that - to sort the wheat from the chaff."

In addition to the genotyping studies described above, HapMap 3 also sequenced ten segments of 100,000 bases from well-characterized regions of the human genome. Unlike discovery using DNA chips - as used in most studies to date - direct sequencing is not biased towards more common variants, but gives a direct estimate of the frequencies of variants.

The researchers found that most variants were relatively uncommon (found in less than one person in ten), but they also found a large number of rare variants (each found in less than one in 100 people) or 'private' variants (found in only one person). Almost eight of ten variants were new and almost four of ten of those seen in less than one in 100 people were found in only one population.

From the results, the researchers suggest that variants in some genes, including genes involved in the immune system, wound healing and sense of smell, are under selection in different populations. These genes can now be studied to learn about how these systems work and about disease resistance. These findings show the value of having large studies that include many populations and samples to achieve comprehensive understanding of human variation.

"Some have talked about how little has come from the Human Genome Project over the past ten years, but perhaps they forget how little we knew then," says Professor David Altshuler of Massachusetts General Hospital in Boston and the Broad Institute of Massachusetts Institute of Technology (MIT) and Harvard University in Cambridge, Mass. "It is amazing that we have gone from a genome less than 90 per cent completed to looking at genetic changes in one in 200 people or rarer. A few years ago, we had no idea of the extent of structural variation or how we might sample variants present at low frequency.

"The HapMap and other large-scale projects have transformed our understanding of the human genome and its relation to health and disease."

More information: The HapMap 3/ENCODE 3 data set is publicly available at www.hapmap.org

Article: The International HapMap 3 Consortium (2010) Integrating

common and rare genetic variation in diverse human populations. *Nature* doi: 10.1038/nature09298