HapMap Catalogue of Human Genetic Variation Published

By Kevin Davies Bio-IT World

A landmark publication by an international consortium of more than 200 scientists has produced a comprehensive catalogue of human genetic variation.

The completion of the first phase of the so-called HapMap – a haplotype map of single nucleotide polymorphisms (SNPs) – was announced at a press conference in Salt Lake City, site of the American Society of Human Genetics convention. The data are published in this week's issue of *Nature*.

At the press conference, Secretary of Health and Human Services Mike Leavitt offered his warm congratulations to the international consortium of researchers. "This is a profound step forward, a triumph for collaborative science," Leavitt said.

The genomes of unrelated individuals vary on average at one DNA base per 1,000, or some 3 million locations across the human genome. In total, there are thought to be about 10 million single-base variations, or SNPs, in the human genome. Charting the identity and locations of these SNPs is important for many reasons, including the study of human evolution and migration, as evidenced by the Genographic Project.

But in medical circles, understanding genetic variation is vital for dissecting the genetic contributions to common, complex diseases that arise from uncertain combinations of genetic and environmental factors. Such variation is "the magnificent legacy of our evolutionary past," write Duke University geneticists David Goldstein and Gianpiero Cavalleri in a *Nature* commentary, "but it comes at a price. Along with making us different in benign and interesting ways, genetics also influences health."

"The HapMap is a powerful new tool for exploring the root causes of common diseases," said the Broad Institute's David Altshuler, a senior author of the study. "Such understanding is required for researchers to develop new and much-needed approaches to understand the still-elusive root causes of common diseases such as diabetes, bipolar disorder, cancer, and many others."

The $128-million International HapMap Consortium was established in October 2002, funded chiefly by the National Human Genome Research Institute and, with the Phase I announcement, has successfully met its initial three-year goal. The international group of academic researchers was joined by Perlegen Sciences. The project mapped 1 million SNPs (to ensure at least one SNP every 5,000 bases of DNA) in 269 individuals of different ethnic backgrounds – Han Chinese in Beijing, Yoribua in Nigeria, Japanese in Tokyo, and Utah residents with ancestors from north/western Europe.

Altshuler says the HapMap approach was "first conceptualized nearly a decade ago." The ultimate goal is to "systematically test genetic variants for frequency in populations with a clinical endpoint. If we can sample a subset of variants, the vast majority of information on common variants can be extracted." For that, there needed to be a catalogue of variants, including information on their location, frequency, and technology.

Tagging SNPs

While the costs of SNP genotyping have fallen dramatically in the past few years, it would still be a formidable challenge to type 1 million or more SNPs per subject for any research or clinical study. Thankfully, that is not necessary. Consecutive SNPs tend to fall in blocks, or haplotypes, like playing cards stuck together as the pack is shuffled. By mapping these haplotype blocks, researchers can select a much smaller number of

representative “tag SNPs” for each segment.

Japan’s leading genome researcher, Yusuke Nakamura of the University of Tokyo’s Human Genome Center, said, “The HapMap is a phenomenal tool that is making possible research that was impractical, if not unimaginable, only a few years ago. He added that it is “reducing the expense of searching for genome for hereditary factors in common disease by a factor of 10 to 20.”

While most of the impact of the HapMap will be in medical affairs, Altshuler stressed there is important relevance for the study of human evolution: “When mutations improve fitness, a fossil record is left in DNA. These data make it possible to search without preconceived notions for which genes might have been important [in human evolution].”

In addition to confirming the spread of known gene variants (beta globin, lactase), Altshuler said that “dozens of other genes have been identified for the first time as important in human evolution.”

Phase II
Much like the genome project itself, the HapMap is still undergoing improvement. The published Phase I version has more than 1 million mapped SNPs. However, the data for Phase II have already been generated, consisting of an additional 2.8 million SNPs genotyped by Perlegen Sciences, the offshoot of DNA chipmaker Affymetrix, and deposited into the public domain as of earlier this week.

“This phase was highly successful,” said Perlegen vice president Kelly Frazer. “Detailed analysis will occur over the next few months.” Frazer added: “Huge advances in genotyping technology have occurred, on the order of more than two magnitudes. These advances not only paved the way for HapMap Phase II, but will allow for an explosion in the number of follow-on studies.”

Calling the published Phase I map “a milestone for medical research,” Altshuler said: “This is the start of a new era. We now have the information and the technology.” But while the first phase of the HapMap undoubtedly marks another benchmark in the technological maturation of genomics, Duke University geneticists David Goldstein and Gianpiero Cavalleri say it is time to harness that power. They write in an accompanying commentary in *Nature*:

“The current state of genomic sciences may be considered as a sort of awkward adolescence. The power of the modern genomic toolkit is breathtaking. In a few years we have gone from knowing almost nothing that could be characterized as genomic… to having complete genome sequences for many organisms, and now a nearly complete catalogue of the common genetic differences among people. Technical prowess is not in itself a mark of maturity in science, however. The next phase for genomics research requires a greater focus on both biological understanding and clinical utility. It is time for genomicists to turn their attention from technology to application.”

The International HapMap Consortium. “A haplotype map of the human genome.” *Nature* **437**, 1299-1320 (27 October 2005) | doi: 10.1038/nature04226.