

First phase of HapMap project already helping drug discovery

Genetic data helps better predict differences in disease and treatment response.

Mariam Andrawiss

Three years after its launch, Phase I of the International HapMap project (HapMap) is now complete (The International HapMap Consortium. *Nature* 437, 1299–1320; 2005). The information is already helping researchers to identify the root genetic cause of common diseases, and to predict differences in drug effectiveness and side effects between patients.

The HapMap provides a new dimension to the pharmacogenomics era. The project, a collaboration between academic researchers worldwide and Perlegen Sciences, Affymetrix and Illumina, identified one million of the estimated 10 million changes in individual bases of DNA in the human genome, called single-nucleotide polymorphisms (SNPs), in 269 individuals from different ethnic backgrounds.

SNPs often cluster together in the genome closely enough to be inherited as a unit. The arrangement of these SNPs within the unit is called a haplotype. As identifying one variation within these blocks acts as a tag to identify a whole series of mutations, the entire genome can now be scanned more easily to find all the factors involved in disease. “If you really want to go down to the genetics, you will probably go to a larger, more complex encyclopaedia,” says Allen Roses, senior Vice-President, Genetics Research at GlaxoSmithKline.

There have been some concerns as to how useful this haplotype data will actually turn out to be for drug discovery and development. Companies and academic laboratories that have taken advantage of the fact that the accumulating data have always been publicly available say that they are already reaping benefits.

For instance, the Iceland-based company deCODE Genetics recently discovered a haplotype around the gene encoding leukotriene A4 hydrolase that is associated with a significantly greater risk of heart attack among African Americans (Helgadóttir, A. *et al. Nature Genet.* Published online: 10 November 2005; doi:10.1038/ng1692).

“One of the critical instruments that we used to pull out these new findings are the HapMap data,” says Kári Stefánsson,

CEO and director of deCODE, and senior investigator of the study.

Mark Rieder, research assistant professor in Genome Sciences at the University of Washington, Seattle, says the HapMap data proved useful in his studies on differences in patient response to the anticoagulant drug warfarin.

“You need good analytical instruments, you need people who understand statistics and who are willing to deal with complex multiple testing.”

Warfarin is used worldwide for the long-term prevention of blood clots, but it is notoriously difficult to control the dose and many medications interfere with the effectiveness of the drug. Rieder and co-workers published a study earlier this year showing that variations in the gene encoding for the vitamin K epoxide reductase complex 1 can be used to group patients into low-, intermediate- and high-dose warfarin groups (Rieder, M. J. *et al. NEJM* 352, 2285–2293; 2005).

“We used the first release of the HapMap data, but at that stage it was probably not efficient in terms of

density of SNPs compared to what we already knew about for this particular gene,” says Rieder.

“But this is changing,” he adds. “I am sure the HapMap will allow us to go wider and we will try to pull out other genes involved if that is the case.” Data on an additional 2.5 million SNPs from Phase II of the HapMap project have already been generated, and were deposited into the public domain last month.

However, researchers benefiting from the HapMap data admit that ease of use could be a problem. “You need good analytical instruments, you need people who understand statistics and who are willing to deal with complex multiple testing,” says Stefánsson.

Pharmaceutical companies may have the facilities and the capabilities necessary to use the HapMap data efficiently, but at the moment most academic settings will struggle with the vast amount of information, the level of expertise and the amount of computer power needed.

“For the average researcher, at this point, it is actually difficult to pitch out the most important SNPs. There is still a large barrier to drill down into the data and really say ‘I want these SNPs for genotyping this study,’” says Rieder.



The HapMap project has so far identified several million SNPs in DNA from multiple individuals.