WHAT'S HOT IN... BIOLOGY , September/October 2008

HapMap, The Next Generation: More SNPs, More Insights

by Jeremy Cherfas

PDF

Sequencing the human genome was an important and vital achievement in its own right, but it was possibly even more valuable for what it enabled: a better understanding of human differences. Among the most interesting of these are single nucleotide polymorphisms (SNPs), differences in a single letter of the DNA. Researchers have mapped millions of these SNPs. Very few of these are directly associated with disease, like the SNP that causes sickle cell anemia. The overwhelming majority have no known function, but that does not limit their usefulness.

Adjacent SNPs on the DNA tend to be inherited in blocks, and two individuals often have long stretches of SNPs in common. These larger blocks of DNA are known as haplotypes, and in 2002 an international consortium set out to create a map of SNPs and associated haplotypes. The International HapMap Consortium's first map, published in 2005, was an instant citation success (see *Science Watch*, 17 [5]: 8, September/October 2006). Now they're back with a sequel, and, unlike so many sequels, the Phase II HapMap at #9 does not disappoint.

HapMap I placed one SNP at roughly every 5,000 DNA letters. HapMap II sequenced an additional 2 million SNPs, increasing the map's resolution to one SNP per kilobase (kb). That has offered several insights, primarily into crossing-over, the phenomenon that creates haplotype blocks. During sexual reproduction, the

Biology Top Ten Papers

Rank	Papers	Cites Mar-Apr 08	Rank Jan-Feb 08
1	E. Bettelli, *et al.*, "Reciprocal developmental pathways for the generation of pathogenic effector T_H17 and regulatory T cells," *Nature*, 441(7090): 235-8, 11 May 2006. [Harvard Med. Sch., Boston, MA] *040YP	52	1
2	K. Takahashi, *et al.*, "Induction of pluripotent stem cells from adult human fibroblasts by defined factors," *Cell*, 131(5): 861-72, 30 November 2007. [Kyoto U., Japan; CREST, Kawaguchi, Japan; Gladstone Inst. Cardio. Dis., San Francisco, CA] *243MG	41	†
3	M. Wernig, *et al.*, "*In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state," *Nature*, 448(7151): 318-24, 19 July 2007. [5 U.S. institutions] *191GC	40	7
4	The ENCODE Project Consortium (E. Birney, *et al.*), "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project," *Nature*, 447(7146): 799-816, 14 June 2007. [80 institutions worldwide] *178FV	39	2

maternal and paternal chromosomes come together and cross over, recombining stretches of DNA. But the crossing points are not randomly distributed along the DNA. They are concentrated into hotspots, where crossing over is much more likely. In fact, hotspots account for some 60% of recombination. The stretches between hotspots are the basic building blocks of the haplotypes.

With a SNP every kb, it is possible to investigate the hotspots in detail. For example, the model of crossing over and haplotypes described above is very simple. Crossing over does not occur at every hotspot in each generation. So if two individuals recently share a common ancestor, the haplotype blocks will be much longer and may span many hotspots. The pedigrees of the individual people whose DNA is the basis of the HapMap are not known. Nevertheless, by looking closely at the pattern of SNPs, the International HapMap Consortium showed that between 10 and 30% of the pairs in each population share an ancestor within the past 10 to 100 generations. The regions that are identical by descent can extend over tens of megabases and encompass hundreds of SNPs.

With regard to hotspots, the IHC looked in detail at the location of hotspots relative to aspects of the gene sequence. Recombination is less likely in the actual coding sequence of a gene, but more likely just upstream of the start of gene transcription. The region just downstream of a transcribed gene is generally less likely to contain a hotspot.

5	P.R. Mangan, *et al.*, "Transforming growth factor-ß induces development of the T_H17 lineage," *Nature*, 441 (7090): 231-4, 11 May 2006. [U. Alabama, Birmingham; NIDCD, NIH, Bethesda, MD] *040YP	38	4
6	A. Barski, *et al.*, "High-resolution profiling of histone methylations in the human genome," *Cell*, 129(4): 823-37, 18 May 2007. [NHLBI, NIH, Bethesda, MD; U. Calif., Los Angeles] *172FA	38	†
7	D.F. Easton, *et al.*, "Genome-wide association study identifies novel breast cancer susceptibility loci," *Nature*, 447 (7148): 1087-93, 28 June 2007. [87 institutions worldwide] *183HT	37	†
8	K. Okita, T. Ichisaka, S. Yamanaka, "Generation of germline-competent induced pluripotent stem cells," *Nature*, 448(7151): 313-7, 19 July 2007. [Kyoto U., Japan; Japan Sci. Tech. Agency, Kawaguchi] *191GC	35	5
9	Intl. HapMap Consortium (K.A. Frazer, *et al.*), "A second generation human haplotype map of over 3.1 million SNPs," *Nature*, 449(7164): 851-61, 18 October 2007. [72 institutions worldwide] *221LY	31	†
10	Hara, *et al.*, "Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice," *Nature*, 441 (7095): 885-9, 15 June 2006. [10 Japanese institutions] *052SL	30	†

SOURCE: Thomson Reuter's Hot Papers Database. Read the Legend.

Genes associated with defense and immunity see the highest levels of recombination; crossing over is six times more likely for them than for genes associated with internal functions such as DNA repair. This makes sense, inasmuch as one of the evolutionary advantages ascribed to sexual reproduction is that recombination protects against the rapid evolution of parasites and pathogens by throwing up new defense and immunity combinations each generations. HapMap II provides many other insights into natural selection

The existence of the HapMap spawned an industry supplying the tools to map SNPs. HapMap II will improve the utility of those tools, but even the first version has enabled new kinds of insight. Before the HapMap, for example, researchers looking for the genetic basis of disease had to find a very strong genetic link with the disease and then home in on a candidate gene. In the 1990s this approach identified two breast cancer genes, BRCA1 and BRCA2. But even though there is a strong genetic basis to breast cancer, which is twice as common in first-degree relatives, variation in BRCA1 and BRCA2 and a couple of other genes accounts for less than 25% of familial risk. The idea developed that maybe there were many other breast cancer genes, each with a relatively small effect. Multiple genes with small effects are never going to be easy to find with ordinary gene sequencing. The HapMap and SNPs, however, make it more possible, and at #7 is a paper that does so.

Douglas Easton and a huge team scanned the entire genome looking for SNPs associated with breast cancer. In Phase 1 they mapped more than 200,000 SNPs in 400 cases and controls. About 5% of the SNPs were linked to the disease. These 12,000 SNPs were mapped in another 4,000 cases and controls; 30 SNPs were tightly linked. These were now sequenced in 20,000 cases and controls. Five novel genes emerged, one of which had previously been linked to breast cancer, and four of the five are genes that could plausibly result in breast cancer.

The new genes are not of much use for predictive screening. They do not explain much of the additional risk. Additional genes may make screening more useful, but for now the main impact—beyond showing the value of large HapMap studies to identify disease genes—is that it may indicate new therapeutic avenues to explore.

Dr. Jeremy Cherfas is Science Writer at Bioversity International in Rome, Italy.

Keywords: HapMap, haplotype map, International HapMap Consortium, single-nucleotide polymorphism, SNPs, gene sequencing, Douglas Easton, breast cancer susceptibility.

PDF

back to top

What's Hot In... : What's Hot In Biology Menu : HapMap, The Next Generation: More SNPs, More Insights

Scientific Home | About Scientific | Site Search | Site Map

Copyright Notices | Terms of Use | Privacy Statement