GENOMICS

New Mapping Project Splits the Community

A new type of genomic map, known as the haplotype map, promises to speed the search for elusive genes involved in complex diseases. But some geneticists question whether it will work

It will upend the practice of medicine and save lives the world over, asserts Francis Collins, the director of the National Human Genome Research Institute (NHGRI). "This is the single most important genomic resource for understanding human disease, after the sequence," he says. "We needed it yesterday, as far as I am concerned." Indeed, his institute is leading the National Institutes of Health's (NIH's) $40 million downpayment on the project.

But to many biologists, it's an untested concept hardly worthy of the $110 million it will consume. "The whole thing is a big waste of taxpayer money," says Joseph Terwilliger, a statistical geneticist at Columbia University in New York City.

Welcome to the haplotype map, a new type of genome map that, depending on where you look, is eliciting exuberance or exasperation.

Proponents of the map, who include Collins, Eric Lander of the Whitehead Institute Center for Genome Research in Cambridge, Massachusetts, Panos Deloukas of the Sanger Institute in Hinxton, U.K.—and just about every big name from the Human Genome Project—say it's the best hope for tracking down the genes involved in common diseases, such as heart disease and cancer, that claim so many lives and have eluded most gene-hunting strategies. As an added perk, they say, it provides a tantalizing glimpse at human evolution and migrations.

On the other side, many researchers—mostly population geneticists—say the map's promise is inflated and it may fail to deliver, adding that its proponents are forging ahead with too little data on how best to proceed. "I think there was lots of good, pure, scientific motivation for wanting to do this haplotype map," says Nancy Cox, a human geneticist at the University of Chicago who's not involved in the project. Still, she adds, "people fell into a trap that we should have been smart enough to avoid: that was saying, 'If we have this, we'll get the genes for complex common disorders.' I think it's premature to know how much that will help."

The most virulent critics even contend that the HapMap, as it is known, is nothing more than a full-employment enterprise for the big sequencing labs that might find themselves out of a job once the human genome is completed. "Is the HapMap being designed to satisfy the need [to get] another big project going?" asks Kenneth Weiss, an anthropologist and geneticist at Pennsylvania State University, University Park. Terwilliger minces no words. "The haplotype map is just an excuse for Lander and others to keep funding coming for large factories set up for the sequencing of the genome"—an idea Lander dismisses as sour grapes.

Not everyone's smiling. A plan to study haplotypes in these populations is prompting angry words.

Despite these misgivings, the HapMap is well under way. In addition to NIH's $40 million, Canada recently kicked in CAN$9.5 million to fund one of its own researchers. Still some $60 million short, NHGRI will award the first grants this fall.

Unexpected architecture

The idea for the HapMap emerged from the gradual realization that the genome has a surprisingly structured architecture. Rather than being thrown together randomly, thousands of DNA bases—as well as patterns of single-base variations found among them—line up in roughly the same order in many different people. Like an interior decorator debating among four kitchen designs, a person's genome has just one of a few potential blocks of DNA to slap into a defined space on a chromosome. Each DNA block—or kitchen design—is a haplotype.

Mark Daly, a computational biologist at Whitehead, stumbled upon these blocks a couple of years ago as he was scouring chromosome 5 for susceptibility genes for Crohn's disease. Stretches of DNA from 129 families affected by the disease kept falling out in one of about four patterns, Daly wrote in the October 2001 *Nature Genetics*.

His Whitehead colleagues David Altshuler, Lander, and others wasted no time investigating. In a paper published last spring in *Nature*, they described common haplotypes in a group of northern Europeans and a Nigerian population called Yorubans, arguing that haplotypes varied somewhat between the two—a vestige of evolutionary history. A paper in *Science* (23 November 2001, p. 1719) by a group led by David Cox of Perlegen Sciences in Mountain View, California, found just a few haplotypes on chromosome 21.

Presented at a Cold Spring Harbor Laboratory meeting last spring, these findings initially generated skepticism. They began to win converts, however, and by summer, Collins made the HapMap—a map identifying haplotypes across the genome—a priority and was rustling up money and collaborators. In March, NIH began soliciting grants for the first of the map's two stages. The first will be to create haplotype maps of the genomes of three populations: those of northern and western European ancestry, Japanese and Chinese, and Yorubans. In the second stage, scientists will test whether the haplotypes they find in those very large populations also appear in about 10 others.

From SNPs to haplotypes

Haplotypes gained popularity as scientists realized that mining the genome was much tougher than expected. A few years ago, geneticists heralded single-nucleotide polymorphisms (SNPs) as the long-sought answer to finding genes involved in complex

CREDITS: (CLOCKWISE FROM LEFT) BOB ROWAN, PROGRESSIVE IMAGE/CORBIS; OWEN FRANKEN/CORBIS; PAUL A. SOUDERS/CORBIS

diseases. Unlike cystic fibrosis or Huntington's disease, say, which are caused by a single genetic defect, complex diseases seem to arise from an uncertain mix of multiple gene mutations and environmental factors. Because each gene likely has a subtle effect, its "signal" is exceedingly tricky to detect.

SNPs are simply places where the genomes of different people vary by one base; they occur at least every 300 bases. A genome blanketed with millions of these markers, the reasoning went, would enable researchers to compare which SNPs predominate in people with a certain disease. Those that do may point to a disease mutation on nearby DNA that is inherited along with the SNP.

Over the past few years, a public-private consortium has deposited some 2.4 million unique SNPs in a public database. Geneticists everywhere are using them. But parsing millions of SNPs to find a few implicated in disease can be challenging—and can cost a fortune. Genotyping, or identifying which bases a SNP can be, costs about 30 cents per SNP, making large disease gene studies unaffordable.

Instead of searching for individual SNPs, the HapMap focuses on patterns of a few SNPs that define each haplotype. If a specific haplotype is more common among those with a certain disease, the mutation linked to that disease should be on that same block of DNA. The HapMap is slated to have about 1/20 the number of SNPs in any given region of the genome that a full SNP map would, says Altshuler, vastly reducing the number that need to be analyzed and saving a truckload of money in the process.

If all goes as planned, researchers working with a finished HapMap could do what Whitehead's Daly is now doing. After finding that one haplotype in the region he was studying is about twice as common among those afflicted with Crohn's, he's now scouring that 200,000-base-long stretch of chromosome for a gene or genes that boost susceptibility to the disease.

Since a workshop last summer in Washington, D.C., roughly 40 scientists and ethicists from institutions across the United States, Europe, and Asia have been solidifying plans for the map. They have settled on the populations to study in the large-scale map: the two that the Whitehead group identified in last year's *Nature* paper, and Asians. The game plan, says Lisa Brooks, program director for the genetic variation program at NHGRI, is to pool the DNA of roughly 50 unrelated people in each population and then search for haplotypes across the entire genome.

Before NHGRI had worked out this plan, the Whitehead scientists launched a mini-haplotype map study to test the idea, analyzing haplotypes in European Caucasians, Japanese and Chinese, Yorubans, and African Americans. In a paper published online by *Science* this week (www.sciencexpress.org), they report that their HapMaps, which focused on 0.4% of the genome, showed that between six and eight SNPs can identify a haplotype and that a given stretch of DNA includes one of three to five haplotypes.

To save money and time, HapMap researchers will start with SNPs in the public database, selecting those that provide good coverage across the chromosomes. Then they'll whittle that set down to those found in at least 10% of their sample populations. If there are large gaps on the chromosomes where no SNPs appear, they'll seek additional ones for the HapMap—something Lander predicts will be necessary.

Building blocks. Persons B and D share a haplotype unlike the other three, characterized by three different SNPs.

But will it work?

Regardless of its final design, will the HapMap actually turn up susceptibility genes? Many say the answer just isn't known. "There's virtually no empirical evidence" that haplotypes will help in the search for genes behind common ailments, says Jonathan Pritchard, a population geneticist at the University of Chicago.

The success of the HapMap hinges on one critical, hotly debated assumption: that common mutations are behind most common diseases. In other words, many people susceptible to, say, colon cancer share specific mutations that increase their likelihood of developing the disease. Because the haplotypes being mapped will include only common SNPs, the disease mutations associated with these SNPs will, by definition, be equally common.

But another possibility, particularly espoused by population geneticists, is that common diseases arise from combinations of rarer mutations. And the HapMap is unlikely to reveal these. "There are innumerable diseases where nothing has been found yet, which in itself is an argument that that [common disease–common mutation] hypothesis has got to be rejected," says Kenneth Kidd, a population geneticist at Yale University in New Haven, Connecticut.

Lander concedes that the number of mutations that are common is unknown, but he believes "it's reasonable to guess there are many hundreds." And "even if the map is only useful for the common mutations, that'll be just fine," he says. Despite his confidence, however, some supporters "have this nagging fear" that there may be problems with the map, says Kwok.

Proponents also add that there's no way to know whether the $110 million gamble is worth it without, well, taking the gamble. David Valle, a pediatrician and human geneticist at Johns Hopkins University, argues that even an apparent flop confers critical information. The HapMap "will be the proof or disproof of common versus rare variants" theory, he says.

HapMappers will need about $60 million or so more to vindicate their theory. Collins is optimistic that the money will flow in. China, Britain, and Japan have all expressed interest, he says.

HapMap proponents are also passing the hat among the drug companies that supported the SNP initiative. "It's under consideration," says a GlaxoSmithKline spokesperson. But these companies may be reluctant to gamble: Many are now scrambling to preserve revenue as they lose patent protection on popular drugs. Furthermore, says Zenta Tsuchihashi, group leader in pharmacogenomics at Bristol-Myers Squibb in Princeton, New Jersey, "until someone really does it, it's a little bit hard to judge the value" of the HapMap.

–JENNIFER COUZIN

CREDIT: ADAPTED FROM PERLEGEN SCIENCES