

## BREAKTHROUGH OF THE YEAR

# Human Genetic Variation

**Equipped with faster, cheaper technologies for sequencing DNA and assessing variation in genomes on scales ranging from one to millions of bases, researchers are finding out how truly different we are from one another**

THE UNVEILING OF THE HUMAN GENOME ALMOST 7 YEARS AGO cast the first faint light on our complete genetic makeup. Since then, each new genome sequenced and each new individual studied has illuminated our genomic landscape in ever more detail. In 2007, researchers came to appreciate the extent to which our genomes differ from person to person and the implications of this variation for deciphering the genetics of complex diseases and personal traits.

Less than a year ago, the big news was triangulating variation between us and our primate cousins to get a better handle on genetic changes along the evolutionary tree that led to humans. Now, we have moved from asking what in our DNA makes us human to striving to know what in my DNA makes me me.

### BREAKTHROUGH ONLINE

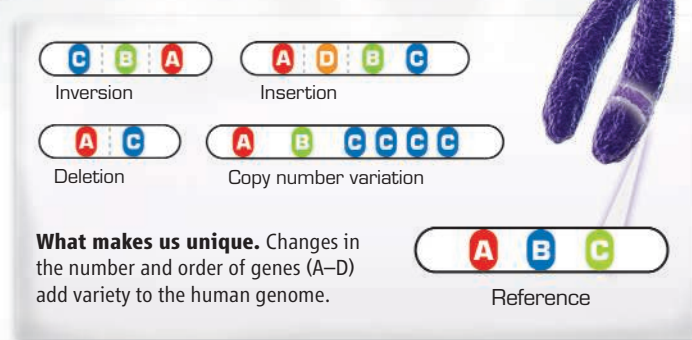
For an expanded version of this section, with references and links, see [www.sciencemag.org/sciext/btoy2007](http://www.sciencemag.org/sciext/btoy2007)

are showing that these changes are more common than expected and play important roles in how our genomes work—or don't work. By looking at variations in genes for hair and skin color and in the "speech" gene, we have also gained a better sense of how we are similar to and different from Neandertals.

Already, the genomes of several individuals have been sequenced, and rapid improvements in sequencing technologies are making the sequencing of "me" a real possibility. The potential to discover what contributes to red hair, freckles, pudginess, or a love of chocolate—let alone quantifying one's genetic risk for cancer, asthma, or diabetes—is both exhilarating and terrifying. It comes not only with great promise for improving health through personalized medicine and understanding our individuality but also with risks for discrimination and loss of privacy (see sidebar, p. 1843).

### Turning on the flood lamps

Even with most of the 3 billion DNA bases lined up in the right order, there was still much that researchers couldn't see in the newly sequenced human genome in 2001. Early comparative studies threw conserved regulatory regions, RNA genes, and other features into relief, bringing meaning to much of our genome, including the



**What makes us unique.** Changes in the number and order of genes (A–D) add variety to the human genome.

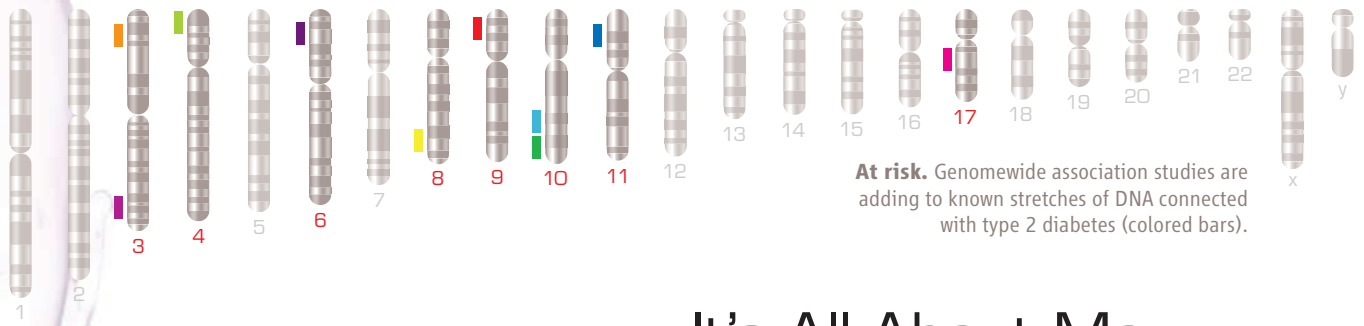
98% that lies outside protein-coding regions. These and other studies, including a pilot study called ENCODE, completed this year, drove home how complex the genome is.

There are an estimated 15 million places along our genomes where one base can differ from one person or population to the next. By mid-2007, more than 3 million such locations, known as single-nucleotide polymorphisms (SNPs), had been charted. Called the HapMap, this catalog has made the use of SNPs to track down genes involved in complex diseases—so-called genome-wide association studies—a reality. More than a dozen such studies were published this year.

Traditionally, geneticists have hunted down genes by tracking the inheritance of a genetic disease through large families or by searching for suspected problematic genes among patients. Genome-wide association studies go much further. They compare the distribution of SNPs—using arrays that can examine some 500,000 SNPs at a time—in hundreds or even thousands of people with and without a particular disease. By tallying which SNPs co-occur with symptoms, researchers can determine how much increased risk is associated with each SNP.

In the past, such links have been hard-won, and most have vanished on further study. This year, however, researchers linked variants of more than 50 genes to increased risk for a dozen diseases. Almost all the variants exert relatively small effects, in concert with many other genetic factors and environmental conditions, and in many cases the variant's real role has not yet been pinned down. But the sheer numbers of people studied have made even skeptics hopeful that some of these genetic risk factors will prove real and will help reveal underlying causes.

The Wellcome Trust, the U.K.'s largest biomedical charity, began to put its weight behind genome-wide association studies in 2005 and recruited 200 researchers to analyze the DNA of 17,000 people from



**At risk.** Genomewide association studies are adding to known stretches of DNA connected with type 2 diabetes (colored bars).

across the United Kingdom. The results are part of an avalanche of genetic information becoming available as more and more geneticists agree to share data and as funding agencies require such exchanges. In June, the consortium published a mammoth analysis of seven diseases, including rheumatoid arthritis, bipolar disorder, and coronary artery disease. It also found several gene variants that predispose individuals to type 1 diabetes and three new genes for Crohn's disease.

Several large studies have also pinpointed type 2 diabetes genes. One French study involving nonobese diabetics found that a version of a gene for a protein that transports zinc in the pancreas increased the risk of this disease. Three simultaneous reports involving more than 32,000 participants uncovered four new diabetes-associated gene variants, bringing to 10 the number of known non-Mendelian genetic risk factors for type 2 diabetes. These finds strongly point to pancreatic beta cells as the source of this increasingly common chronic disorder.

New gene associations now exist for heart disease, breast cancer, restless leg syndrome, atrial fibrillation, glaucoma, amyotrophic lateral sclerosis, multiple sclerosis, rheumatoid arthritis, colorectal cancer, ankylosing spondylitis, and autoimmune diseases. One study even identified two genes in which particular variants can slow the onset of AIDS, demonstrating the potential of this approach for understanding why people vary in their susceptibility to infectious diseases.

### Genomic hiccups

Genomes can differ in many other ways. Bits of DNA ranging from a few to many thousands, even millions, of bases can get lost, added, or turned around in an individual's genome. Such revisions can change the number of copies of a gene or piece of regulatory DNA or jam two genes together, changing the genes' products or shutting them down. This year marked a tipping point, as researchers became aware that these changes, which can alter a genome in just a few generations, affect more bases than SNPs.

In one study, geneticists discovered 3600 so-called copy number variants among 95 individuals studied. Quite a few overlapped genes, including some implicated in our individuality—blood type, smell, hearing, taste, and metabolism, for example. Individual genomes differed in size by as many as 9 million bases. This fall, another group performed an extensive analysis using a technique, called paired-end mapping, that can quickly uncover even smaller structural variations.

These differences matter. One survey concluded that in some populations almost 20% of differences in gene activity are due to copy-number variants; SNPs account for the rest. People with high-starch diets—such as in Japan—have extra copies of a gene for a starch-digesting protein compared with members of hunting-gathering societies. By scanning the genomes of autistic and healthy children and their parents for copy-number variation, other geneticists have found that newly appeared DNA alterations pose a risk for autism.

New technologies that are slashing the costs of sequencing and genome analyses will make possible the simultaneous genome-wide search for SNPs and other DNA alterations in individuals. Already, the unexpected variation within one individual's published genome has revealed that we have yet to fully comprehend the degree to which our DNA differs from one person to the next. Such structural and genetic variety is truly the spice of our individuality.

—ELIZABETH PENNISI

## It's All About Me

Along with the flood of discoveries in human genetics, 2007 saw the birth of a new industry: personal genomics. Depending on your budget, you can either buy a rough scan of your genome or have the whole thing sequenced. The companies say the information will help customers learn about themselves and improve their health. But researchers worry that these services open up a Pandora's box of ethical issues.

At \$300,000 to \$1 million per genome, sequencing all 3 billion base pairs is still too costly for all but a few. Although dozens more personal genomes will probably be sequenced in the coming year, most will be done by public and private research organizations—including the institute run by genome maverick J. Craig Venter, whose personal genome was one of three completed in 2007 in the United States and China. In a lower-budget effort, Harvard's George Church this month will deliver initial DNA sequences for the protein-coding sections (1% of the genome) to the first 10 volunteers for his Personal Genome Project. Meanwhile, a new company called Knome is offering full-genome sequencing to 20 customers willing to pay \$350,000.

A glimpse of one's genome is already within the reach of ordinary people, thanks to several companies. They include 23andMe, which has financing from Google and may let users link to others with shared traits; Navigenics, which will screen for about 20 medical conditions; and deCODE Genetics in Iceland, a pioneer in disease gene hunting. For \$1000 to \$2500, these companies will have consumers send in a saliva sample or cheek swab, then use "SNP chips" to scan their DNA for as many as 1 million markers. The companies will then match the results with the latest publications on traits, common diseases, and ancestry.

Although many customers may view this exercise as a way to learn fun facts about themselves—recreational genomics, some call it—bioethicists are wary. Most common disease markers identified so far raise risks only slightly, but they could cause needless worry. At the same time, some people may be terrified to learn they have a relatively high risk for an incurable disease such as Alzheimer's.

The rush toward personal genome sequences also sharpens long-held worries about discrimination. A bill to prevent insurers and employers from misusing genetic data is stalled in Congress. Complicating matters, your genetic information exposes your relatives' DNA, too.

The most profound implications of having one's genome analyzed may not be what it reveals now—which isn't much—but what it may show later on. Perhaps to sidestep such questions, some companies will limit which markers to disclose. Others, however, will hand customers their entire genetic identity, along with all the secrets it may hold.

—JOCELYN KAISER



**Pandora's box?** This cheek-swab kit could reveal your intimate secrets.

CREDITS (TOP TO BOTTOM): K. KRAUSE/SCIENCE (SOURCE: RICHA SAXENA AND DAVID ALTSHULER, BROAD INSTITUTE OF HARVARD AND MIT); COURTESY OF 23ANDME

Downloaded from www.sciencemag.org on January 2, 2008