The Case of the Disappearing DNA Hotspots

COLD SPRING HARBOR, NEW YORK—New approaches to genome studies are showing that DNA is full of surprises, researchers learned here 12 to 16 May at the Biology of Genomes meeting.

DNAwise, chimps and humans are virtually identical. But two research teams have found, to their surprise, key differences among these close cousins in the locations of DNA recombination hotspots: places where matching chromosomes exchange DNA much more frequently than normal. Evidence is preliminary, but in at least some cases, chimps don't have hotspots in the same places as humans. And it's now unclear whether chimps have as many hotspots as people.

Kissing cousins. Chimps and humans have DNA differences after all.

More than just a curious phenomenon, hotspots contribute to evolution and perhaps to human disease. Recombination, in or out of hotspots, introduces variety into a genome and can eliminate bad genes—or, possibly, introduce deleterious mutations. For example, when chromosomes recombine, one sometimes loses a gene and the other gains an extra copy, a problem that can lead to diseases such as the blood disorder thalassemia. Researchers also want to understand hotspots to make best use of a multimillion dollar project, the HapMap, that is cataloging variation in the human genome as a tool for tracking down disease genes (*Science*, 30 April, p. 671).

Until recently, researchers were unsure whether hotspots—typically short regions less than 2000 bases long—are common. They assumed that recombination occurs at a fairly constant rate along much of the genome. To try to learn more about what makes a hotspot hot, Molly Przeworski, a population geneticist at Brown University in Providence, Rhode Island, compared them in humans and chimps.

She and her colleagues collected DNA samples from two dozen western chimps, a subspecies that lives in western Africa. They looked at the comparable 10,000-base stretch of DNA where a human hotspot is known to exist. After sequencing that piece of DNA in each chimp, they looked for high recombination rates there. This stretch of chimp chromosome turned out to be not so hot after all, she reported. She couldn't tell whether an equivalent chimp hotspot had been shifted to a different place along that stretch of DNA or whether it didn't exist at all. Either way, the result was puzzling: "You would predict that the human and chimp would be the same [in terms of hotspot location] because the sequence is the same," says Edward Rubin, director of the Department of Energy Joint Genome Institute in Walnut Creek, California.

After a more extensive comparative study, Harvard graduate student Wendy Winckler found similar results, she reported at the meeting. Independent of Przeworski, she and her colleagues looked at equivalent places in the chimp genome where hotspots turn up in human DNA. First Winckler, working with David Altshuler of Boston's Massachusetts General Hospital and others, sequenced presumed hotspots in 22 chimps and 24 humans to catalog single-base changes, or SNPs, in each. They then looked for those same SNPs in another 38 western chimps and 94 humans, half Africans and half Caucasians.

As expected, both human populations shared the same six hotspots. But in the chimp, all the hotspots were missing. Much more work needs to be done to determine whether chimps have fewer hotspots overall.

The implication is unclear, but at the very least, the work points to a previously unrecognized genomic trait that distinguishes the two primates and raises questions about the role hotspots play in evolution. It also drives home how much we have left to learn about genomes. "I find it amazing that both groups find that human recombination hotspots are totally gone when they check in chimps," says Svante Pääbo, a geneticist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. "Since there are no dramatic sequence differences between the species in these areas, this suggests to me that there must be some mechanism that we do not yet understand that is responsible" for creating hotspots.

"It's not sequence," agrees Arend Sidow of Stanford University. He wonders whether the particular arrangement of the proteins surrounding those DNA regions sets the stage for recombination. Pääbo suggests that perhaps hotspots are chemically marked in some way.

Whatever the cause, both Winckler and Przeworski are trying to determine the history of these and other hotspots. It could be that at one time, both species shared hotspots, but that some disappeared in the chimp. Or it could be that they arose separately in humans and chimps after the species started evolving their separate ways. Or there could be another explanation altogether.

Disposable DNA Puzzles Researchers

For a long time, any DNA that didn't make up genes was considered junk, even though it constituted the bulk of the human genome. Gradually, though, genome biologists have found gems among this non–protein-coding sequence, suggesting that "junk" was a serious misnomer. But new research suggests that vast tracts of this sequence may be disposable after all: Marcelo Nóbrega, a geneticist at Lawrence Berkeley National Laboratory (LBNL) in Berkeley, California, finds that mice can do just fine with millions of these bases deleted from their genomes.

About 2 years ago, Edward Rubin, director of the Department of Energy Joint Genome Institute in nearby Walnut Creek, and his colleagues discovered that some DNA sequence in human gene deserts—

CREDIT: YANN ARTHUS-BERTRAND/CORBIS

Useless? Even with conserved areas deleted, some genes work fine (as shown in blue) in transgenic mice.

long stretches of DNA between genes—was almost the same as the sequence in comparable mouse deserts (*Science*, 31 May 2002, p. 1601). This conservation across species that shared a common ancestor more than 80 million years ago seemed highly unlikely—unless those regions served a purpose. Since then, "we've had the assumption that all these regions are doing something," says Michael Zody, a computational biologist at the Broad Institute in Cambridge, Massachusetts; that something is probably gene regulation, he says.

Nóbrega and his colleagues found support for this idea when they compared desert regions conserved between fish and humans, quite distant relations. "Most sequences conserved between fish and humans are not only functional, but they help regulate genes," Nóbrega reported. But their mouse-human comparison told a different story. They analyzed 15 comparable desert sequences and found that only one was a regulatory region in both species.

Puzzled, Nóbrega, Rubin, and their colleagues decided to delete deserts from mouse genomes, hoping to learn what other function these regions might serve. His LBNL colleague, geneticist Yiwen Zhu, knocked out two regions, one about 2 million bases long and the other about 1 million, both of which were conserved in humans and mice but not fish. After inserting the altered genome into embryonic mouse stem cells, Zhu added the cells to mouse embryos and looked for abnormalities in their descendents. "There was no sign of any difference in survival" between the genetically altered and normal mice, Nóbrega reported at the meeting. "There's no sign of overt pathology."

Other researchers are dumbfounded. "To knock out 2 megabases and not have an effect—that's remarkable," says Jim Hudson, a geneticist at Open Biosystems in Huntsville, Alabama. "It can't be true," says a skeptical Arend Sidow of Stanford University. Both Hudson and Sidow wonder whether these noncoding regions have a function that just doesn't show up in the tests Nóbrega did. So the question remains open, says Rubin: "Is the genome like a trash novel, from which you can remove 100 pages and it doesn't matter, or is it like a Hemingway, where if you remove a page, the story line is lost?"

Surveys Reveal Vast Numbers of Genes

From the inside of the mouth to farmland to the surface of the ocean, microbes reign supreme. We know there are lots, but not how many; we also know they live remarkable lives, but we don't fully understand how they survive in often-adverse conditions. The enigmas remain because most of these organisms can't be grown in the lab. But now, their genes at least are within reach.

At the meeting, three research teams described successes in surveying various microbial communities by sequencing the organisms' DNA collected in environmental samples. And everywhere these researchers look, they are finding not only more kinds of microbes than expected but a staggering number of new genes. It's becoming clear, says David Relman, a molecular microbiologist at Stanford University, that with this new approach, called community or environmental genomics, "you can discover far more novel genes and novel biology" than was possible before.

Already, "we are amazed by the degree of gene discovery," adds J. Craig Venter, who runs The Center for the Advancement of Genomics in Rockville, Maryland. Venter is leading a team that is sampling the microbial diversity in oceans around the world.

Until recently, researchers have tended to sequence the genome of a particular organism—be it yeast, mouse, or mold—then analyze its genetic makeup. But now a few genome biologists are sequencing whatever DNA they can get from a test site and then piecing together these DNA sequences into genomes, or at least bits of genomes. Each set of combined sequence represents a different organism. "The approach is to understand the complex ecology [of microbes]," says geneticist Stefan Schreiber of Christian Albrecht University in Kiel, Germany. "The potential is enormous."

Relman has been looking at the ecology of the mouth. He's surveyed the DNA of microbes that live in the pockets between the teeth and the surrounding gum—the site of gum disease. Not much has been known about that community because so few of its members can be grown in the lab. But by focusing on DNA, Relman has found what could be a genetic signature of gum disease. One type of bacterium, called TM7, thrives in gums with mild disease. But when gums disintegrate, another group—methane-producing archaea—may take over.

Relman, along with Steven Gill and Karen Nelson of The Institute for Genomic Research in Rockville, has yet to look at all the genes of these oral microbes, but other work is showing just how fruitful gene searches can be. Edward Rubin, director of the Department of Energy Joint Genome Institute in Walnut Creek, California, has preliminary data on soil samples collected from a Minnesota farm. He and his colleagues have sequenced 100 million bases but can't seem to fit them together because "there are so many different kinds of organisms," he explains. But they can pick out as many as 150,000 previously undescribed genes hidden in that DNA, just from one site.

Venter is also finding incredible diversity. At first he focused on the Sargasso Sea, sequencing more than 1 billion bases and grouping them into genomes representing 1800 species, including 148 groups of unknown bacteria. The DNA includes 1.2 million new genes, he reported in April (*Science*, 2 April, p. 66).

Now, Venter's group has much more data, and off Long Island, New York, alone, only 950 putative organisms—1%—match those found in the Sargasso Sea. Eventually, the team expects to find as many as a billion new genes, he reported at the meeting. All these genes "will provide a different view of evolution," Venter points out, as they suggest that new genes or new microbes arise much more often than had been previously thought.

Hard to swallow. Even gums have a surprising diversity of microbes and, likely, genes.

–ELIZABETH PENNISI

CREDITS: (TOP TO BOTTOM) M. NÓBREGA *ET AL.*, *SCIENCE* **302**, 413 (2003); LESTER V. BERGMAN/CORBIS