Technology Review
PUBLISHED BY MIT

Wednesday, October 03, 2007

Deciphering Human Differences

Chunks of shuffled DNA in the human genome could underlie many diseases.

By Alexandra M. Goho

With the help of new high-speed DNA sequencing technology, scientists have uncovered extensive regions in the human genome where chunks of DNA have been deleted, copied, or completely rearranged. Mapping and characterizing these structural variants could be key to understanding human diversity and the origins of many diseases.

Until recently, most researchers believed that differences among individuals were mainly due to changes in single bases, or "DNA letters." Over the past couple of years, however, several studies have shown that rearrangements of large chunks of DNA (imagine shuffling around entire sentences, pages, or chapters in a book) could play a more important role. "These are pretty big changes," says Michael Snyder (http://www.yale.edu/snyder/) , a molecular biologist at Yale University. "When you flip around 8,000 bases of DNA, that's likely to have a dramatic impact."

Already, studies have linked structural variants in the human genome with several different diseases. For instance, extra copies of a particular gene have been associated with a reduced susceptibility to HIV infection, whereas the full or partial deletion of other genes has been linked to certain forms of heart disease.

In collaboration with 454 Life Sciences (http://www.454.com/) , based in Branford, CT, which has developed a novel method of rapidly sequencing DNA (see "Sequencing in a Flash (http://www.technologyreview.com/Biotech/18625/) "), Snyder and his colleagues analyzed the genomes of two individuals: one of African descent and one of European descent.

The researchers chopped each genome into millions of fragments, each 3,000 bases long, and then tagged the fragments' ends. They then sequenced the ends and, using specially designed computer algorithms, mapped the millions of end sequences back to a reference genome, derived from the Human Genome Project.

Most of the time, the end sequences for each fragment lined up just right along the reference genome. However, a certain number did not. "They were either too far apart, too close, or flipped relative to the reference genome," says Snyder. If a pair of end sequences were too far apart, this meant that an extra chunk of DNA was inserted into that region. If two sequences were closer than expected, a piece of DNA was likely deleted.

Reporting in the September 27 edition of *Science Express* (the online publication of *Science* magazine), the researchers found close to 1,300 structural variants between the two genomes--many more than other studies have found. "This is a major form of variation in people," says Snyder. "It's much more extensive than people realized."

Almost half of the structural variants were common between the two individuals, suggesting that these changes in the genome are ancient. Seventeen percent of the structural variants directly affected genes. Most of those affecting genes tended to be involved in the immune system and in the way that the body interacts with the environment, as opposed to genes involved in development.

For instance, the researchers found that two different genes implicated in smell had been altered by DNA shuffling. In one individual, the two genes lay side by side; in the other, they had fused together to form a single gene. "The presumption is that this reflects differences in the way people smell things, although we don't know that yet for certain," says Snyder.

To reconstruct how these alterations formed, Snyder and his colleagues went back and sequenced the ends of the chunks of DNA that were either inserted or deleted. The team did this for some 200 of the variants identified. "That was quite impressive," says Stephen Scherer, a geneticist at the Hospital for Sick Children, in Toronto.

Many of the structural variants may have been the result of sequences of DNA that spontaneously jump from one place to another in the genome, explains Snyder. Other changes may have occurred through the unequal swapping of DNA between pairs of chromosomes.

Evan Eichler (http://eichlerlab.gs.washington.edu/index.html) , a geneticist at the University of Washington, in Seattle, describes the work as a technological feat. "New DNA sequencing technologies are coming out fast and furious," he says. With 454's sequencing technology, he adds, the researchers obtained a large amount of information about the human genome in a relatively short period of time.

Snyder emphasizes that there are many more structural variants to be found, not only in their samples (the group's technique could not discern changes smaller than 3,000 bases), but also in the genomes of other individuals. Michael Egholm, 454's vice president of research and development, hopes to repeat the experiment on 100 individuals in a year, and then on 1,000. "Our technology is getting faster," he says. "We did this in a fraction of the time it took others to do this kind of analysis, and at higher resolution and at greater depths."

The ultimate goal, says Egholm, is to create a complete inventory of all the structural variants in the genome and then begin testing each one to see which are associated with disease.
