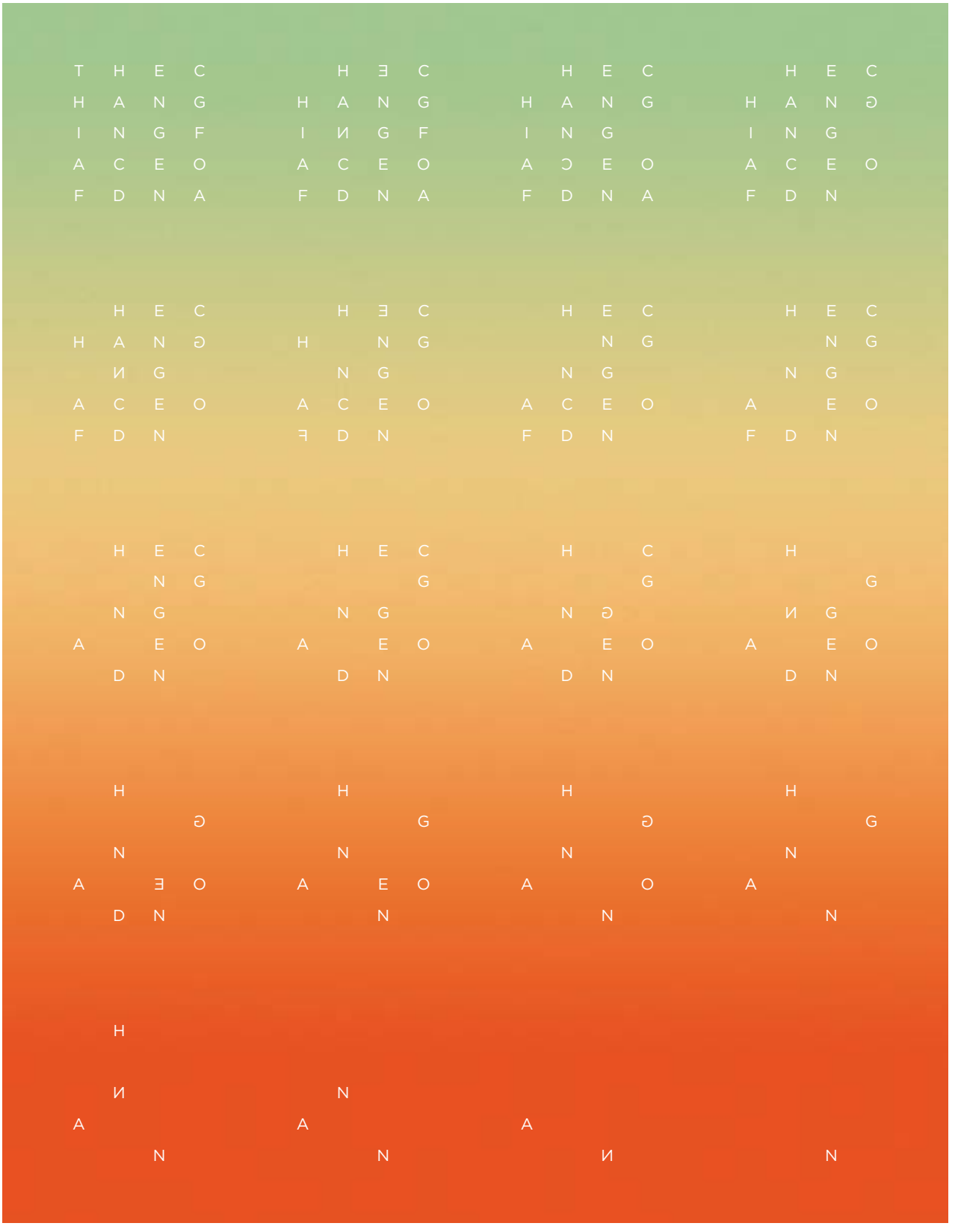

THE CHANG_NG FAƆE OF DNA

BY STƎVE OLSON

Each individual's genome is its own unique landscape. And variations much larger than slight misspellings in our DNA base pairs may explain our looks as well as our susceptibility to disease.

In the basement of the Foege Building on the campus of the University of Washington, the DNA of a woman known only as G248 lies in thousands of tiny wells inside a freezer cooled to –84°C. Five floors above, HHMI investigator Evan E. Eichler points to her DNA as the harbinger of a new way of thinking about human genetics.

"*The* human genome" is a misnomer, according to Eichler. G248 has big sections of DNA that other people don't have, and she's missing DNA that most people do have. "In the last few years, it's been shown that big changes in DNA—insertions and duplications and deletions and inversions—are extremely common in the population," Eichler says. "That's the first important point. The second important point is that these changes play a role in human disease—everything from HIV susceptibility to autism to mental retardation to epilepsy."

A few years ago, most human geneticists would have been very skeptical about such a statement. At that time, geneticists focused almost exclusively on spelling differences in the human genome—places where the chemical bases that make up DNA, represented by the letters A, T, C, and G, differ from one person to another. According to the thinking of the day, these individual changes in DNA codes largely accounted for differences in our genetic susceptibility to disease and in our physical appearance.

But in the first half of this decade, a handful of geneticists, working independently at laboratories scattered across the United States and Canada, began to notice something strange. As they looked more carefully at human DNA, they found that some people had multiple copies of big sections of DNA, hundreds or thousands of base pairs long. Sometimes these structural variants, as they came to be known, were in DNA regions that didn't seem to be doing anything. But sometimes they were in regions rich with genes, so that some people had more copies of particular genes than other people.

"We were finding a huge amount of copy number variation—that was the message," says another pioneer in the study of structural variation, Stephen W. Scherer, a former HHMI international research scholar who directs the Centre for Applied Genomics at the Hospital for Sick Children in Toronto, Canada.

The discovery has been a revelation for many geneticists. "A lot of the more complex disorders are not explained by coding variation, which is what people were looking for," says HHMI investigator Val C. Sheffield, who for years has suspected that structural variation might play a prominent role in the eye diseases he studies in his University of Iowa lab. "But until recently we haven't had the technologies to look at variation on a genome-wide scale."

The new picture that Eichler, Scherer, and a handful of other geneticists have been painting differs radically from the traditional view of our genome. Instead of the book of life, DNA is more like the scrapbook of life. Sentences, paragraphs, or entire chapters are copied and haphazardly inserted into various parts of our genome. In some people, the same page repeats over and over, while other people don't have that page at all. And geneticists have been tying this structural variation to an increasing number of diseases. "It's amazing," says Scherer. "At human genetics meetings, 30 to 40 percent of the talks have a direct focus on copy number variation."

"WE'RE FIND_NG NEW PARTS OF THE HUMAN GENOME, AND THAT'S A POWERFUL THING."

EVAN EICHLER

The discovery of structural variation was partly a consequence of better technologies and new data. But it was also a case, says Eichler, of "good luck favoring the prepared mind."

DNA CHURNING

"I KNEW THAT I WANTED TO DO GENETICS when I was in grade 9, and by grade 10 I knew that I wanted to do human genetics," says Eichler. He grew up on a farm in far northern Canada, where winter locks the landscape in icy splendor. His father grew wheat and canola in summer and taught French in a nearby town the rest of the year. His mother raised Angora rabbits, whose wool she would spin into yarn for sweaters. "My mother was one of those people who didn't like dyes, so she decided that she wanted a natural variation of colors," Eichler says. "She said to me, 'Can you figure out how to get these other colors, these creams and buffs and so on?' That's where I learned the basic genetic coat color system. I got a book, drew my first Punnett squares, and within about a year I was producing true lines of different colors. I knew at that point that this was probably the coolest field ever."

After receiving a baccalaureate from the University of Saskatchewan and working in a molecular virology laboratory in Munich for a year, Eichler enrolled in 1991 in the genetics program at Baylor University. Though he and his Canadian wife struggled with the climate and culture shock of living in Houston, it was the perfect place for Eichler scientifically. He began investigating the genetic disorder fragile X syndrome and "absolutely fell in love with research." His faculty adviser, David Nelson, was "a brilliant scientist and mentor who encouraged a lot of free thinking," Eichler says. "He didn't lord over me at all but let me hang myself with my own proposals."

Fragile X introduced Eichler to the instability of the genome. It occurs when mutations make a particular part of the X chromosome much longer than usual, inactivating a gene critical to development of the brain and other parts of the body. "The idea that an unstable region of the genome could increase the probability of a disease a hundredfold or a thousandfold—that's the idea I fell in love with," Eichler says. "I haven't strayed far from those roots."

While at Baylor, Eichler also began working on a study associated with the Human Genome Project, which was just then getting under way. He was attaching short DNA probes to portions of the X chromosome when he noticed that the probes also were binding to parts of chromosomes 2, 12, 16, and 22. "That was odd," Eichler recalls. It was as if portions of human DNA had been copied and scattered across the genome. "I began to think, 'How widespread is this?'"

In 1997 Eichler moved to Case Western Reserve University, where he continued investigating the genome's structure. During those years, duplications in the human genome were becoming a big problem for the Human Genome Project. When DNA is broken into pieces for sequencing, duplications make it hard to put the pieces back together, because one copy can be mistaken for another. Eichler and his coworkers took on the computer-intensive job of calculating the frequency of duplications from data being generated by both the public and the private sequencing efforts. Using PCs from CompUSA and fans from K-Mart to keep the computers cool, they found "there was a lot more duplication than anyone had thought," Eichler says.

A TRUE INDIVIDUAL

To be completely accurate, the Human Genome Project should have been called the "Half-of-a-Composite-Human Genome Project" because the DNA sequenced came from a single set of chromosomes drawn from several donors. Our cells contain two copies of every chromosome, with one copy coming from our mothers and one from our fathers. Therefore, the structural variation carried in our parents' chromosomes is passed down to us—making for a lot of variation in our own cells.

Recently, a team of researchers that included J. Craig Venter and Stephen Scherer compared the chromosome pairs in Venter's DNA. They found much more structural variation between his chromosomes than most geneticists expected. Forty-four percent of the genes on Venter's chromosomes differ from the corresponding genes on the other member of the chromosome pair. And three-fourths of the total variable DNA content between chromosome pairs arises from structural variants, with one-fourth coming from changes in individual DNA letters. —*S.O.*

Geneticists like Evan Eichler (top) and Stephen Scherer are tying structural variation seen in our genomes to an increasing number of diseases.

His team continued to study duplications after the release of the draft human genome in 2000, and they discovered that many were occurring in particular "bad neighborhoods" of the genome. There, multiple copies of DNA sequences made the genome susceptible to further rearrangements through a process known as nonallelic homologous recombination (see sidebar, page 31). The DNA in those regions seemed to be "churning," continually rearranging itself from one generation to the next. Eichler was sure those rearrangements had consequences for human evolution and health. But what were they?

THE IMPACT ON HEALTH

ABOUT THIS SAME TIME, AT TORONTO'S HOSPITAL FOR Sick Children, Scherer was equally puzzled by what he was seeing in the genome. He and his colleagues were searching chromosome 7 to uncover genes involved in disease. In the process, they were uncovering massive and unexpected differences in the chromosomes of different people. "Most geneticists thought that if you had a large genetic change, it would be associated with disease all the time," says Scherer. But he and his team were finding big differences that didn't seem to have an obvious effect on health—including million-base-pair insertions or deletions, "which was really unbelievable."

Many geneticists were skeptical. At that time, the technologies they were using were so new that the differences might have come from experimental design or malfunctioning equipment. "We were criticized a lot," Scherer says. "My first grant application [to study structural variation] was rejected, because people said it couldn't possibly be true."

But as analytic techniques improved, so did the evidence for substantial structural variation. Charles Lee of Brigham and Women's Hospital in Boston had found similar DNA differences, as had a group at Cold Spring Harbor Laboratory led by Michael H. Wigler. By about 2003, the case for widespread structural variation in the human genome was becoming unassailable.

Furthermore, evidence was accumulating that some of these variants influence health. The human genome regulates itself through a process still largely unknown. But variable numbers of a gene can produce a greater or smaller amount of a protein important to the body, and a duplicated section of DNA can disrupt the function of an important gene.

As Scherer and others investigated the genomes of people with genetic disorders, they found that structural variation often seemed a more likely contributor to the disorder than DNA spelling differences. Schizophrenia, Alzheimer's, Parkinson's, autism, kidney disease, and many other diseases were linked to structural variation. "We've been shocked to see how quickly the idea has been adopted and how many diseases are being associated with large structural variants," Scherer says.

In the past couple of years, Scherer has focused on structural variants in patients with birth defects and neurological disorders. For example, at the Hospital for Sick Children screens of children with unexplained genetic disorders have shown that some 20 percent have structural changes in their DNA that may contribute to their conditions. He also has been participating in studies to identify and characterize structural variation in the

Eichler: Ron Wurzer / AP, ©HHMI Scherer: Evan Dion

genome, including differences between chromosomes in the same cell (see sidebar, page 29). "It's incredible how many people are using these data, from commercial companies to clinical geneticists to everyone in between," he says.

WHY VARIATION?

AT THE UNIVERSITY OF WASHINGTON, WHERE HE MOVED in 2004, Eichler and his colleagues also have been delving into the link between structural variation and disease. In one particularly intriguing study, they examined the DNA of 290 British children with neurological disabilities. "We were looking for recurrent deletions in regions of the genome that are highly dynamic," says Andrew Sharp, the postdoctoral fellow in Eichler's lab who headed the project.

Of the 290 children, 16 had deletions or duplications that are "likely to be pathogenic," according to the group's September 2006 paper in *Nature Genetics*. Remarkably, four had very similar but not identical deletions on the long arm of chromosome 17. All four, though unrelated, had very similar features, including silvery hair, blue eyes, and a bulbous nose—"they could be brothers and sisters," says Sharp—but their shared characteristics hadn't been noticed before. And the region of their deletions included several genes implicated previously in neurological and behavioral conditions.

Building on that success, Eichler's group has begun examining the connection between structural variation and a range of more common diseases. "The million dollar question is: What is the genetic basis of diseases like diabetes, hypertension, and high cholesterol levels?" he says. "We know there is a genetic factor, but what is the role of single base pair changes versus structural changes?"

To answer that question, Eichler and a group of colleagues known as the Human Genome Structural Variation Working Group have decided to get a better fix on where the structural variation in our genome occurs. The freezer in the basement of Eichler's laboratory containing the DNA of G248 is one of 62 freezers scattered around the United States, each containing the DNA of a single individual. The working group will compare each donor's DNA with the reference sequence from the Human Genome Project, looking for locations where the DNA doesn't line up. Wherever they find a discrepancy, they'll sequence the DNA to identify the differences.

Understanding human disease is the main objective, but Eichler wants to know something else. Why did variable regions of our genome evolve, and what purpose do they serve?

Eichler's hypothesis is that structural variation is a way for our genomes to remain fluid and adaptable. As our ancestors encountered new environments and new circumstances, continual rearrangement of their DNA would have generated lots of evolutionary experiments. In fact, initial comparisons have shown that humans and other primates have much more structural variation than do other mammals. Eichler speculates that the unique abilities of primates—our elaborate social structures and communication abilities—may be related to the amount of structural variation in our genomes. "Maybe the cost of having these new abilities is the possibility of disease caused by genes that allow us to adapt to the right environments at the right time," he says.

The discovery of structural variation has shattered the image of the human genome as an inert and largely stable object. Instead, there are as many human genomes as there are humans, and each unique assemblage of DNA has its own strengths and weaknesses. "My wife and I had a baby just two months ago, and I joke with her that it's amazing that any of us ever comes out normal, knowing what we know now," Eichler says. "But I think the right answer is that none of us is normal. And that's an enlightening feeling, to realize that no one has the perfect genome." ■

SWAPPING SEGMENTS

Many structural variants in the human genome arise when the male and female sex cells (egg and sperm) are preparing chromosomes to pass on to the next generation. During this process, the two members of each chromosome pair line up next to each other and swap segments through a mechanism known as recombination. But sometimes the chromosomes misalign. Duplications in the genome cause the wrong parts of chromosomes to line up next to each other, so that when the chromosomes swap parts, genes are added to one chromosome and deleted from another. The result is a new structural variant—an evolutionary experiment ready to be tested against nature. —*S.O.*