

PERSPECTIVES & OPINIONS

Evan Eichler

CRUCIBLES OF
DYNAMISM

Scott Areman

Puzzling pockets of redundancy account for about 5 percent of the human genome. HHMI investigator Evan Eichler found a way to interpret what is happening in these areas of genetic repetition. The University of Washington researcher's discovery holds promise for our understanding of evolution and neurocognitive disease.

We're accustomed to thinking of evolution moving forward one base pair at a time: A single nucleotide is replaced in the ribbon of our DNA and a gene is gradually transformed from one function to another, or a disease emerges as the gene's function is perturbed by these small stepwise changes. But there is another, parallel story unfolding within our genes, which, until recently, no one was able to easily read. To all appearances, it's a more dramatic tale, with big events and rapid cataclysmic change.

If the single-nucleotide substitution view of genetic advances relates a drawing room drama of our evolution, the story my lab is uncovering is more like an action-adventure movie with multiple car crashes. We think it is in this part of the story where profound changes took place in our evolution. The regions we examine also appear to be a source of vulnerability to cognitive disease, where big genetic errors can lead to intellectual disability, autism, and schizophrenia, among other disorders.

Although it was long apparent that duplication reigned in parts of our genome, there was no way to accurately analyze these repeating elements. Our assays were based on hybridization, which relied on nucleotide probes to suss out complementary sequences, and such probes could not distinguish among identical regions. We could not count the multiple copies of genes accurately, and we could not relate genetic changes to function.

Using next-generation genetic sequencing and a new computational algorithm, my lab can now read about 70 percent of the chaotic chromosomal regions. The emerging picture is of regions that are crucibles of dynamism, where lots of change happens quickly, both adaptive and deleterious.

It appears that in the early common ancestor of humans and the great apes, certain genetic segments began hopping around the genome, inserting a duplicated version of themselves at new locations. This pattern is rarely seen in other mammals. While mice, cows, and dog genomes carry repeated genes, the repeats tend to follow one after the other. But our analysis of duplication in the genomes of macaque, orangutan, chimp, and human, published in *Nature* in 2005 and 2009, showed that in the early common ancestor, repeats may be more than a million bases apart.

When these hopscotching genes relocated—and we don't actually know how or why—they picked up whatever genetic code sat beside them, and then duplicated it and themselves somewhere else along the genome. We call these hopscotching genes "core duplicons" because they seem to be at the

heart of most regions of dynamic genetic change. From their new roost, the duplicons gathered additional flanking regions, and again continued their march along the chromosome, creating complex regions of duplications within duplications at many new locations.

This architecture can create problems. Any time you have identical sequences, it's easier to trick DNA's recombination machinery. These repeating segments and the regions that flank them are susceptible to unequal crossing-over events. That is, chromosome pairs, no longer perfectly aligned, swap genetic material unequally during meiosis, leading to big changes when sperm and egg form. Some of the most common genetic causes of epilepsy and autism are due to rearrangements in these duplicated genes.

When we looked more closely into these regions, we saw an incredible amount of variation between individuals and some intriguing genes that have expanded in primate evolution. Although we have been studying these regions for 15 years, the extent of genetic variation still surprises me.

Of course, the big challenge now is to sort out just what this means, to link this more complex variability to phenotype—that is, any observable traits. Early data suggest that the genes within the core duplicons are expressed preferentially in neurons, particularly at sites of new nerve growth. We know that the large-scale copy number variation can explain about 17 percent of the cognitive disability among children we study from University of Washington Autism Center and other clinics around the world.

One has to wonder why this kind of structural vulnerability hasn't been selected against over the generations. The persistence of this genomic architecture suggests that these regions also confer important evolutionary benefit. I believe there's a balancing act at play, especially considering the point in our evolutionary history when the burst of duplications occurred.

The circumstantial data argue for something fundamental that we're missing about evolution and disease—that this structural variability contributes to what distinguishes us from other primates. Is that the whole story of how we evolved? Absolutely not. But it is an important part that I think has been overlooked. Over the next couple of years, we'll either prove or disprove that hypothesis.

INTERVIEW BY JENNI LAIDMAN. *Evan Eichler is a professor of genome sciences at the University of Washington.*