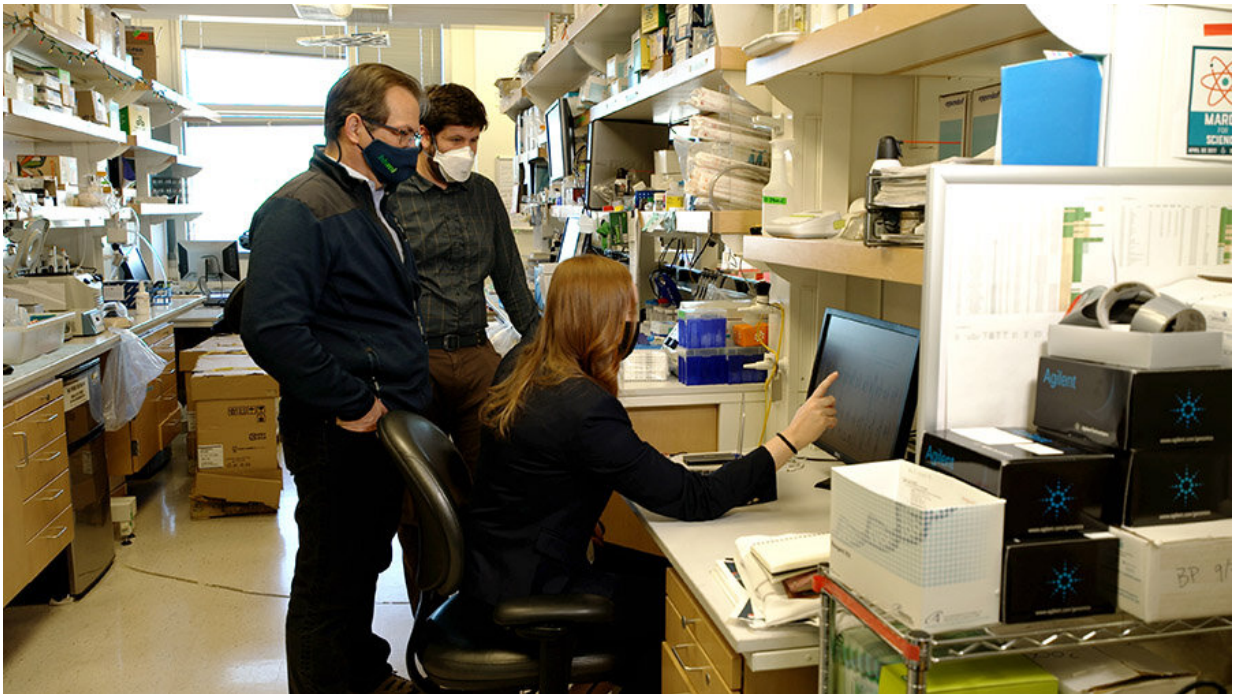


Some hard-to-crack genome areas carry genes that make us distinctly humans

March 31 2022



Inside Evan Eichler's genome sequencing lab in the Department of Genome Sciences at the University of Washington School of Medicine. Eichler is on the left, postdoctoral student Mitchell Vollger is standing, and postdoctoral student Glennis Logsdon is seated at the computer. Credit: Randy Carnell/UW Medicine

UW Medicine genome scientists were among the leading contributors to the publication of the first complete, gapless sequence of a human genome announced this week by the National Human Genome Research

Institute.

The lab of Evan Eichler, professor of genome sciences at the University of Washington in Seattle, was one of the major contributors to the main paper, "The complete sequence of the human genome," published in *Science* April 1. The achievement is the culmination of a large consortium, the Telomere-to-Telomere, or T2T, which aimed to obtain complete sequences of all 23 [human chromosomes](#), end-to-end.

Eichler's team and collaborators from other institutions also produced a companion paper offering the first comprehensive view of highly identical, large repeat regions, called segmental duplications, and their variation in [human genomes](#).

These areas of the human genome are critical to understanding [human evolution](#) and [genetic diversity](#), as well as resistance or susceptibility to many diseases. Of the 20,000 [genes](#) in the human genome, about 950 originate in segmental duplications.

However, segmental duplications were among the last regions of the human genome assembly to be fully sequenced, due to their complexity.

The desire to resolve these regions was part of the impetus for advancing sequencing technologies, such as the ability to read long stretches of DNA. These technologies, along with many laboratory tools, computational biology approaches and other essential research resources, were not available during the first drafting of the human genome more than two decades ago.

The team led by the Eichler lab reported their results and analysis in a companion *Science* paper published this week, titled "Segmental duplications and their variation in a complete human genome." The lead author on this paper is Mitchell R. Vollger, a postdoctoral fellow in

genome sciences at the UW School of Medicine. He applied skills in computer science, data visualization, and mathematics to analyze the new genomic repeats to further our understanding of human variation within segmental duplications. Working with Phil Dishuck, a graduate student in the Eichler lab, they showed that the completion of the human genome added about 180 "new" protein-coding [genes](#), almost all of which mapped to segmental duplications.

"As a kid, I saw the magazine covers for a complete human genome in 2001," Vollger recalled. "I remember thinking that was the coolest project, and how I was disappointed that I would never get to do something that cool. I've thought about that a lot during this project, that I got to contribute sequence to the human genome, and that excites me a lot, that I had the opportunity to do that."

Several intriguing findings emerged from the recent accomplishments in sequencing these regions.



Postdoctoral students in Evan Eichler's genome sciences lab at UW Medicine who worked on completion of human genome. Left is Mitchell Vollger, right is Glennis Logsdon. Credit: Randy Carnell/UW Medicine

In addition to the medical research implications of the completed assembly, it also is helping to answer: What is contained in our genomes that makes us distinctly human? Some of the genes that were gaps in the original genome are now thought to be critically important in helping to make a bigger brain in humans compared to other apes.

Eichler's lab also generated long-read assemblies from other nonhuman primate genomes and compared them to the new gapless human genome assembly. They systematically reconstructed the evolution of some biomedically relevant genes, as well as certain human-specific duplicated genes.

These human-specific segmental duplications are reservoirs for new genes that drive the formation of more neurons in developing brains and increase connectivity of synapses in the [frontal cortex](#)—the anatomical part of the brain where some of the higher-level thinking, reasoning, logic, and language functions that seem to be characteristically human take place.

In *TBC1D3*, a [gene family](#) related to the expansion of the human prefrontal cortex, analysis by Xavi Guitart, a graduate student in the Eichler lab, revealed that recurrent and independent expansions occurred at different points in primate evolution. The most recent was about 2 million to 2.6 million years ago, about when the genus *Homo* emerged. Surprisingly, the human *TBC1D3* gene family showed remarkable, large-scale structural variation in a subset of samples.

"Different humans carry radically different complements and arrangements of the *TBC1D3* gene family," the researchers explained in their paper and that was unexpected for a gene thought to be so important to brain function. The scientists also found diversity in the complex structure of the *LPA* gene, in which variability in part of this lipoprotein gene underlies the most significant genetic risk factor for cardiovascular disease from abnormal lipid levels in the blood.

The researchers also looked at *SMN* (a motor neuron gene) whose mutations are linked to certain neuromuscular disorders. Having better sequence resolution of the spinal muscular atrophy region—one of the most difficult regions to finish on chromosome 5—could be of practical advantage in both disease risk determinations and in treatment as the duplicate gene *SMN2* is a target for one of the most effective gene therapies.

Based on these and other findings, the scientists noted that the new reference genome "reveals unprecedented levels of human genetic variation in genes important for neurodevelopment and human diseases."

In addition to being a source of new knowledge about human biology, the recently completed human genome also is likely to answer some basic questions about cell biology. For example, the assembly will help to better understand the differences in the centromeres present in each of the human chromosomes. Problems in centromeres can cause difficulties during in cell division.

Studying the sequences of the centromeres could get to the root of medical conditions where cell division, and the allocation of genetic material between cells, goes awry. These include cancer as well as abnormalities that affect prenatal development, such as Down syndrome or Robertsonian translocations.

Glennis A. Logsdon, a postdoctoral fellow in genome sciences at the UW School of Medicine, has made several discoveries related to centromere sequencing.

"We had to develop new ways to target these regions," she explained. "We took advantage of new technology that had been on the horizon, such as ultra-long-read sequencing, in order to get across these regions. We also put effort into polishing the genome sequence to make sure that it was highly accurate."

Eichler commented on the training and experience early-career human genome researchers received during the T2T projects.

"I consider it a privilege to actually help build the next generation of scientists," he said. "It's so much fun to see them start as students, contribute to a big project, and then carry it to the next level."

Eichler was part of the original Human Genome Project back in 2001. He was fascinated by regions of the genome that were complex from the perspective that they were highly repetitive, but also encoded genes.

When the conclusion of the [human genome](#) sequence project was declared, a lot of those regions weren't done.

Eichler added that, since then, he has had an intense desire to finish them.

"I've always come back to that point that, to understand genetic variation comprehensively, we need to have a reference that's complete. Otherwise, we're missing pieces of the puzzle. 95% of the puzzle being solved is good enough for some people. But I guess for me, getting that last 5% was so important because I believe so much of what we don't understand about disease, or we don't understand about evolution is

disproportionately represented in that 5% of the of the genome that we didn't sequence first off."

This is not the end, he said. "Even though people would say, 'Well, we're done with finishing the genome.' We finished a [genome](#). There will be hundreds, probably thousands of genomes over the next few years. I think our view of how humans differ from each other is going to be transformed, and how more complex genetic variation is important not only for making us human, but also making us different."

More information: Mitchell R. Vollger et al, Segmental duplications and their variation in a complete human genome, *Science* (2022). [DOI: 10.1126/science.abj6965](https://doi.org/10.1126/science.abj6965).
www.science.org/doi/10.1126/science.abj6965

Provided by University of Washington School of Medicine

Citation: Some hard-to-crack genome areas carry genes that make us distinctly humans (2022, March 31) retrieved 13 February 2023 from <https://medicalxpress.com/news/2022-03-hard-to-crack-genome-areas-genes-distinctly.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.