The New York Times https://www.nytimes.com/2023/05/10/science/pangenome-human-dna-genetics.html

Scientists Unveil a More Diverse Human Genome

The "pangenome," which collated genetic sequences from 47 people of diverse ethnic backgrounds, could greatly expand the reach of personalized medicine.

By Elie Dolgin
May 10, 2023

More than 20 years after scientists first released a draft sequence of the human genome, the book of life has been given a long-overdue rewrite.

A more accurate and inclusive edition of our genetic code was published on Wednesday, marking a major step toward a deeper understanding of human biology and personalized medicine for people from a wide range of racial and ethnic backgrounds.

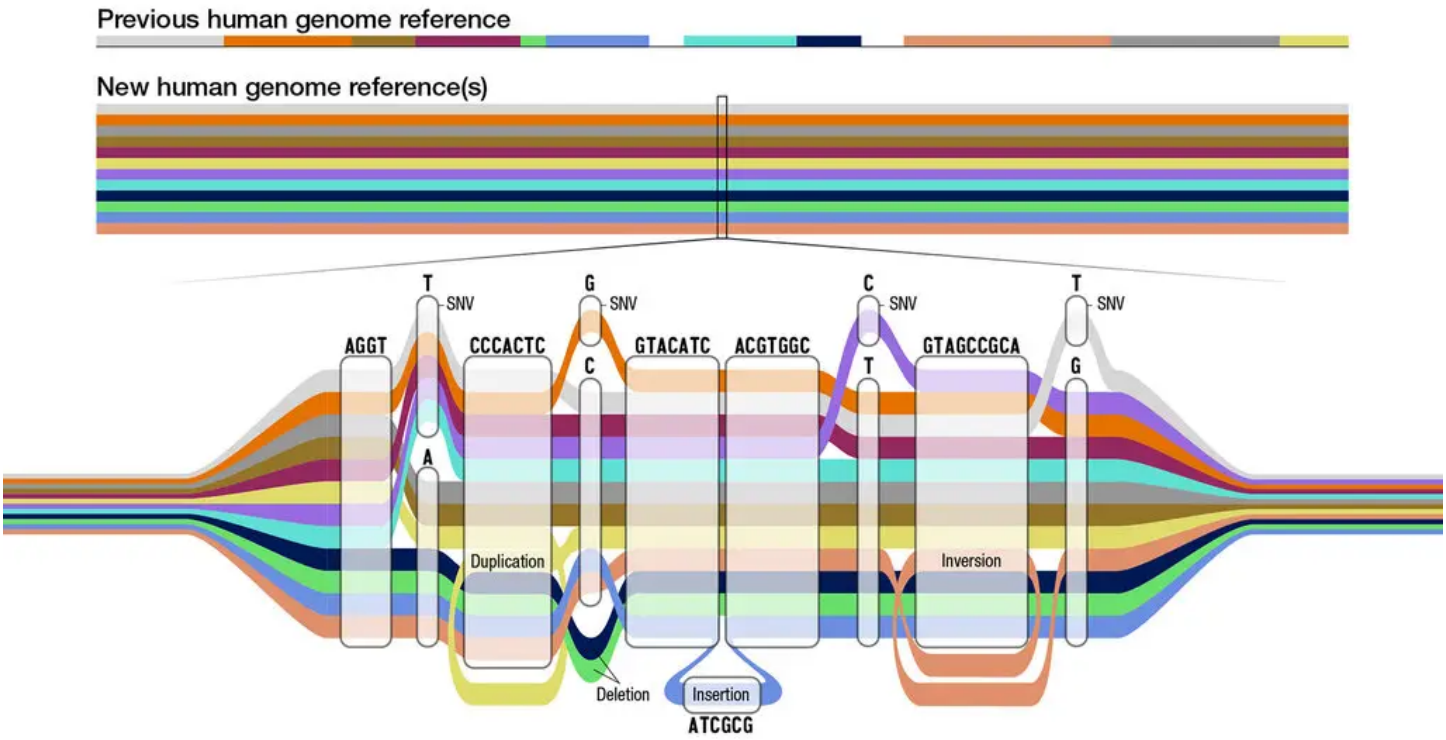
Unlike the previous reference — which was largely based on the DNA of one mixed-race man from Buffalo, with inputs from a few dozen other individuals, mostly of European descent — the new "pangenome" incorporates near-complete genetic sequences from 47 men and women of diverse origins, including African Americans, Caribbean Islanders, East Asians, West Africans and South Americans.

The revamped genome map represents a crucial tool for scientists and clinicians hoping to identify genetic variations associated with disease. It also promises to deliver treatments that can benefit all people, regardless of their race, ethnicity or ancestry, researchers said.

"It's been long needed — and they've done a very good job," said Ewan Birney, a geneticist and deputy director general of the European Molecular Biology Laboratory, who was not involved in the effort. "This will improve our fine-grained understanding of variation, and then that research will open new opportunities toward clinical applications."

Powered by the latest in DNA sequencing technology, the pangenome collates all 47 unique genomes into a single resource, providing the most detailed picture yet of the code that powers our cells. Gaps in the earlier reference are now filled, with nearly 120 million previously missing DNA letters added to the three-billion-letter-long code.

Gone is the idea of a totemic strand of DNA that extends six feet when uncoiled and stretched out in a straight line. Now, the rebooted reference resembles a corn maze, with alternative paths and side trails that allow scientists to explore a broader range of the genetic diversity found in people the world over.



The new pangenome reference shows many possible routes for a sequence to take, represented by the different colors: Yellow is a duplication variant; pink is an inversion variant; green and blue represent a deletion variant, and light blue is an insertion variant. Darryl Leja/NHGRI

Dr. Eric Green, director of the National Human Genome Research Institute, the government agency that funded the work, likens the pangenome to a new kind of bodywork manual for automotive repair shops. Whereas before, every mechanic only had the design specs for one kind of car, now there is a master plan that covers different makes and models.

“We’ve gone from having one really nice blueprint of the Chevy to now having blueprints of 47 representative cars from each of 47 different manufacturers,” he said.

Knowing what to do with this Kelley Blue Book of genomics will involve a steep learning curve. New analytical tools are needed. Coordinate systems must be redefined. Widespread adoption will take time.

“Making this easy to be used by the community is work to be done,” said Heidi Rehm, chief genomics officer at Massachusetts General Hospital in Boston, who was not involved in the project.

But in due course, experts said, the pangenome will revolutionize the field of genomic medicine.

“We’re going to have the benefit of actually understanding ourselves as a species much, much better,” said Evan Eichler, a genome scientist at the University of Washington. Dr. Eichler was among more than 100 scientists and bioethicists who described the new pangenome reference in the journal Nature.

The architects of the project are continuing to add more population groups, with the goal of including at least 350 high-quality genomes that encompass the bulk of global human diversity.

“We want to represent all the branches of the human tree,” said Ira Hall, a geneticist who leads the Yale Center for Genomic Health.

Some of the new genomes will come from New Yorkers who previously participated in a research program at the Mount Sinai Health System. If their preliminary DNA data seems to reflect certain underrepresented genetic backgrounds, those individuals will be invited to participate in the pangenome project.

Some gaps might never get plugged in the publicly available reference, though — by design.

Previous attempts to capture human genetic diversity often extracted sequence data from marginalized populations without regard for their needs and preferences. Informed by those ethical missteps, pangenome coordinators are now collaborating with Indigenous groups to develop formal policies around data ownership.

“We are still grappling with the issue of native and tribal sovereignty,” said Barbara Koenig, a bioethicist at the University of California, San Francisco, who was involved in the project.

In Australia, researchers are incorporating DNA sequences from various Aboriginal peoples into a similar depository that will be combined with the open-source pangenome, but then kept behind a firewall. According to Hardip Patel of Australia’s National Centre for Indigenous Genomics in Canberra, the scientists next plan to consult with community leaders about if or how to make the data accessible through request.

Some Indigenous advocates want to see the pangenome project go further. Keolu Fox, a genomics scientist at the University of California, San Diego, who is Native Hawaiian has suggested training the next generation of Indigenous scientists to have greater agency over the genomic data.

“It’s finally time that we decentralize power and control and redistribute that among the communities themselves,” Dr. Fox said.