



# Pangenome Reference Provides 'Inclusive' View of the Human Genome

May 10, 2023 | [Aayushi Pratap](#)

NEW YORK – An international team of researchers has come up with an updated version of the human reference genome — a pangenome — that combines phased diploid assemblies from 47 individuals of diverse ancestries.

The new reference is more inclusive and provides a more complete image of the human genome than the original human DNA blueprint that was released in 2001, according to the investigators.

The initial human genome was mostly based on just one individual of mixed race, along with a few others who were primarily of European ancestry, meaning it did not represent the vast amount of genomic variation found across human populations.

"A single genome cannot represent the genetic diversity present within the human species, due to the presence of structural variants and alternative alleles, some of which were not present in the original reference genome," the authors wrote.

In a paper published in *Nature* on Wednesday, previously [made available as a preprint](#), the Human Pangenome Reference Consortium presented the first draft of the pangenome, which adds 119 million base pairs of euchromatin sequence and 1,115 gene duplications to the existing reference human genome, GRCh38.

The project used a graph assembly approach to represent the 47 individual genomes in the pangenome reference, detailed in a separate [publication](#) in *Nature Biotechnology*.

The implications of the pangenome are far reaching and range from better diagnostics to improved personalized medicine for people of multiple ancestries, according to the researchers.

"For example, the identification of genetic variations that are associated with human diseases will be both more sensitive and more specific, directly improving disease diagnosis and therapeutics," Ting Wang, a professor of medicine at Washington University School of Medicine in St. Louis and a coauthor of the paper, said in a statement. "The new reference also provides a foundation to investigate functional consequences of genetic variations."

Additionally, the new pangenome could help reduce the inequities in genomic analyses. "The new pangenome increasingly represents the diversity of the human population and will enable scientists and healthcare professionals to better understand genomic variants that influence health and diseases," Eric Green, director of the US National Human Genome Research Institute, said in a conference call. "This is crucial in advancing the field of genomics in an equitable way."

Technological advancements and falling costs of DNA sequencing greatly enabled the creation of the pangenome, the researchers noted. While the previous reference genome was sequenced using short reads only, the new reference genome also incorporated long-read sequencing technology that allowed the team to see structural variations more clearly.

Most of the individuals whose genomes were sequenced for the pangenome project were originally recruited as part of the 1,000 Genomes Project, an NIH program that aimed to improve the catalog of genomic variants in diverse populations.

Meanwhile, in [an accompanying paper](#), Evan Eichler, a professor in the department of genome sciences at the University of Washington School of Medicine, and colleagues presented a map of single-nucleotide variants within segmental duplications, characterizing millions of unmapped SNVs and finding a distinct mutational spectrum that differs from unique DNA. "We reason that these distinct mutational properties help to maintain an overall higher GC content of [segmental duplications] compared to that of unique DNA," they wrote.

In yet [another paper](#), researchers used data from the new pangenome to identify patterns of recombination between the short arms of heterologous acrocentric chromosomes. This provided the first observational evidence for a mechanism of DNA exchange between these chromosomes, senior author Erik Garrison, assistant professor in the department of genetics at the University of Tennessee Health Science Center, noted along with his colleagues.



The new pangenome draft, which is publicly available, is just a start, according to the researchers. Over the course of the next few years, they plan to expand the project to include data from a total of 350 diverse individuals. "This will give us a more comprehensive representation of all types of human variation," they wrote.

Challenges remain, though, when it comes to the adoption of the new reference. "[W]idespread adoption of the pangenome by scientists could take time, because new methods supporting pangenome analysis are continually being developed, and scientists will often require training to use them," Arya Massarat and Melissa Gymrek of the University of California San Diego cautioned in a [related News & Views](#) article in *Nature*.

Nevertheless, the effort will likely change human genetics research profoundly. "This is not the end of a project, but it is the beginning of a new era to much more meaningfully incorporate human diversity in biological, biomedical, and clinical sciences," Wang said. "The new reference will continue to grow, expand, and [be polished] to accurately depict the genetic blueprint of our species — this requires a global effort."

Filed Under

- Genetic Research
- Sequencing
- Sequencing
- Journal Study
- personalized medicine

- North America
- NIH
- UCSC
- NHGRI
- University of Washington
- Washington University
- genome assembly

- UCSD

[Privacy Policy](#). [Terms & Conditions](#). Copyright © 2023 GenomeWeb, a business unit of Crain Communications. All Rights Reserved.