ScienceNews

NEWS LIFE

The new human pangenome could help unveil the biology of everyone

The pangenome includes the genetic instruction books of 47 people

Researchers have compiled genetic profiles of 47 people into a human pangenome. The work reveals levels of human genetic diversity never seen before.

HIROSHI WATANABE/STONE/GETTY IMAGES PLUS

By Tina Hesman Saey

23 HOURS AGO

More than 20 years after people got a peek at the first draft of the human genome, our genetic instruction book, researchers have unlocked the next level: the human pangenome.

In four studies published May 10 in *Nature*, researchers describe the achievement, how the pangenome was built and some of the new biology scientists are learning from it.

The more complete reference book, which includes almost all the DNA of 47 people, will allow researchers to explore types of variation that could never be examined before, such as large chunks of duplicated, lost or rearranged DNA. That work could possibly reveal more details about the genetic underpinnings of heart diseases, schizophrenia and various other diseases and disorders.

The pangenome adds 119 million DNA bases — the information-carrying units of DNA — not present in the existing human genome, called the reference genome. Much of that DNA is in never-before-explored parts of the genome containing multiple copies of genes that are duplicated from originals elsewhere in the DNA.

Those duplicated parts are changing faster than nonduplicated portions of the genome, says Evan Eichler, a human geneticist at the University of Washington in Seattle and one of the leaders of the Human Pangenome Reference Consortium. What's more, when Eichler and colleagues examined the types of variants that arise in these duplicated regions, they found "a very strong signal that the mutations that are occurring are fundamentally different from [mutations in] the rest of the genome," he says.

Some of these duplicated regions include ones implicated in humans' large brains relative to other species and other traits that set humans apart from other primates. Others have been implicated in certain traits or diseases.

Conversely, another study found that the very short arms of certain chromosomes, including chromosomes 13, 14 and 21, are becoming more like each other as they swap DNA. Those short arms are important because they contain genes for making ribosomal RNAs, which serve as the scaffolds for ribosomes, the machinery responsible for building every protein in the body.

But perhaps the biggest achievement of the pangenome project is that it is finally giving researchers a more complete look at the full spectrum of human genetic diversity.

How was the pangenome built?

The roughly two-decade-old human reference genome derives mostly from one man, but is a patchwork quilt of more than 60 people's DNA (*SN: 3/4/21*). It has been restitched and added to over the years but still has holes.

Last year, the first fully complete human genome was announced (*SN: 3/31/22*). That genome contains all of the DNA from tip to tip, or telomere to telomere, of each human chromosome. Except that genome wasn't from a

person. It came from a type of tumor known as a hydatidiform mole. These unusual tumors result when a human sperm fertilizes an empty egg and the father's chromosomes are duplicated.

The genetic information from such tumors represents "not even one individual. It's from one half of one individual," says human geneticist Timothy O'Connor of the University of Maryland School of Medicine in Baltimore who was not involved in either project.

The new pangenome draft is from actual people and contains almost complete DNA from 47 anonymous individuals from different parts of the world. That diversity is important "because it helps us to understand ourselves as a single human species, as a single human race," O'Connor says.

Past genetics research has been criticized for relying too heavily on DNA from people of European heritage. Studying just one population of people could mean missing genetic variants that have arisen in specific populations, O'Connor says. "Having a pangenome reference allows us to assess that population-specific variation in a much more detailed way. And hopefully, that will then lead to greater insight into the biology of everyone."

While the pangenome is a great first step to better represent all human genetic diversity, O'Connor says, "it still is missing key groups in the world. It's still underrepresenting Latin Americans and Indigenous Americans, and ... there's nobody included from Oceania.... There's still a lot more variation that needs to be added to the pangenome to really, truly be representative of everyone."

Added diversity is coming, human geneticist Karen Miga of the University of California, Santa Cruz said during a May 9 news conference. The consortium plans to complete a total of 350 genomes, including these 47, by mid-2024. The first phase of the project was aimed at developing the technology to build the pangenome.

Now, the consortium is in talks with Indigenous groups and scientists from around the world about "trying to develop a shared framework, so that it's not the U.S. trying to set the table. It's really providing a table and inviting other stakeholders who see the value in creating this type of reference resource to join us," said Miga, who helped lead the pangenome project.

How is the pangenome important for human health?

Having a more complete understanding of human genetic diversity could help researchers begin to unravel the genetic underpinnings of various diseases and disorders.

What's more, new DNA deciphering technologies have allowed pangenome researchers to examine types of genetic variants that have been difficult to study before.

In particular, duplicated regions of the genome were hard to study because researchers previously could read only short pieces of DNA. There was no way to tell where in the vast puzzle of the human genome those nearly identical pieces fit. Newer "long-read" DNA deciphering, or sequencing, technology makes it possible to read stretches of DNA many thousands of bases long (*SN: 2/22/21*).

Being able to assess where some people have extra DNA and others are missing DNA, called structural variants, adds a more nuanced view of human genetics, O'Connor says, revealing more of its complexity (*SN: 4/10/09*).

For instance, researchers used the pangenome map to trace how chromosomes fold up so that different parts are touching each other. Scientists could see some folds and chemical marks in structural variants that may affect how genes are turned on and off. That could affect traits or health. Eichler's group also mapped one version of a gene that has converted another copy into its own image. These gene conversions were surprisingly common with each person having, on average, more than 2,000 instances of them.

With this more nuanced and complex view of human genetics comes a promise for improved genetics-based medicine. But it may take a while before the pangenome makes a difference in medical clinics, Eichler says.

Researchers hope the pangenome will help them more easily diagnose the genetic changes that contribute to rare diseases and find treatments for common disorders, he says. Once that happens, clinicians may start incorporating data from the pangenome in their practices.

CITATIONS

W.-W. Liao *et al*. A draft human pangenome reference. *Nature*. Published online May 10, 2023. doi: 10.1038/s41586-023-05896-x.

M.R. Vollger *et al*. Increased mutation and gene conversion within human segmental duplications. *Nature*. Published online May 10, 2023. doi: 10.1038/s41586-023-05895-y.

A. Guarracino *et al*. Recombination between heterologous human acrocentric chromosomes. *Nature*. Published online May 10, 2023. doi: 10.1038/s41586-023-05976-y.

G. Hickey *et al*. Pangenome graph construction from genome alignments with Minigraph-Cactus. *Nature*. Published online May 10, 2023. doi: 10.1038/s41587-023-01793-w.