

Researchers create largest ever human family tree

February 24 2022



Credit: CC0 Public Domain

Researchers from the University of Oxford's Big Data Institute have taken a major step towards mapping the entirety of genetic relationships among humans: a single genealogy that traces the ancestry of all of us. The study has been published today in *Science*.

The past two decades have seen extraordinary advancements in human genetic research, generating [genomic data](#) for hundreds of thousands of individuals, including from thousands of prehistoric people. This raises the exciting possibility of tracing the origins of human genetic diversity to produce a complete map of how individuals across the world are related to each other.

Until now, the main challenges to this vision were working out a way to combine [genome](#) sequences from many different databases and developing algorithms to handle data of this size. However, a new method published today by researchers from the University of Oxford's Big Data Institute can easily combine data from multiple sources and scale to accommodate millions of genome sequences.

Dr. Yan Wong, an evolutionary geneticist at the Big Data Institute, and one of the principal authors, explained: "We have basically built a huge family tree, a genealogy for all of humanity that models as exactly as we can the history that generated all the genetic variation we find in humans today. This genealogy allows us to see how every person's genetic sequence relates to every other, along all the points of the genome."

Since individual genomic regions are only inherited from one parent, either the mother or the father, the ancestry of each point on the genome can be thought of as a tree. The set of [trees](#), known as a "tree sequence" or "ancestral recombination graph," links genetic regions back through time to ancestors where the genetic variation first appeared.

Lead author Dr. Anthony Wilder Wohns, who undertook the research as part of his Ph.D. at the Big Data Institute and is now a postdoctoral researcher at the Broad Institute of MIT and Harvard, said: "Essentially, we are reconstructing the genomes of our ancestors and using them to form a vast network of relationships. We can then estimate when and where these ancestors lived. The power of our approach is that it makes

very few assumptions about the underlying data and can also include both modern and ancient DNA samples."

The study integrated data on modern and ancient human genomes from eight different databases and included a total of 3,609 individual genome sequences from 215 populations. The ancient genomes included samples found across the world with ages ranging from 1,000s to over 100,000 years. The algorithms predicted where common ancestors must be present in the evolutionary trees to explain the patterns of genetic variation. The resulting network contained almost 27 million ancestors.

After adding location data on these sample genomes, the authors used the network to estimate where the predicted common ancestors had lived. The results successfully recaptured key events in human evolutionary history, including the migration out of Africa.

Although the genealogical map is already an extremely rich resource, the research team plans to make it even more comprehensive by continuing to incorporate genetic data as it becomes available. Because tree sequences store data in a highly efficient way, the dataset could easily accommodate millions of additional genomes.

Dr. Wong said: "This study is laying the groundwork for the next generation of DNA sequencing. As the quality of genome sequences from modern and ancient DNA samples improves, the trees will become even more accurate and we will eventually be able to generate a single, unified map that explains the descent of all the human [genetic variation](#) we see today."

Dr. Wohns added: "While humans are the focus of this study, the method is valid for most living things; from orangutans to bacteria. It could be particularly beneficial in [medical genetics](#), in separating out true associations between genetic regions and diseases from spurious

connections arising from our shared ancestral history."

More information: Anthony Wilder Wohns et al, A unified genealogy of modern and ancient genomes, *Science* (2022). [DOI: 10.1126/science.abi8264](https://doi.org/10.1126/science.abi8264). www.science.org/doi/10.1126/science.abi8264

Provided by University of Oxford

Citation: Researchers create largest ever human family tree (2022, February 24) retrieved 27 April 2024 from <https://phys.org/news/2022-02-largest-human-family-tree.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.