March 25, 2021

Evolution & Behaviour

Extending the genomic record of human diversity

by Anders Bergström[1] | Postdoctoral Research Fellow

doi.org/10.25250/thescbr.brk500

[1]: Wellcome Sanger Institute, Cambridge, UK; The Francis Crick Institute, London, UK

This Break was edited by Ayala Sela, Scientific Editor - TheScienceBreaker

We studied the genetic variation among people from diverse groups across the globe, to learn about the complex evolutionary history that has shaped our genomes. Our study provided insights into the timescale of early human evolution and the genetic contributions of Neanderthals and Denisovans to present-day people.

Image credits: Pixabay

The genetic material of any two humans is 99.9% identical, but the small differences that do exist between our genomes provide a record of the complex evolutionary history we have undergone as a species. Over the past decade, scientists have sequenced a large number of genomes from populations across the globe, which have given us a detailed understanding of human genetic variation. These datasets are also of great use to medical geneticists, who use them to estimate the frequencies of genetic variants associated with disease in different populations.

However, currently available genome sequences are heavily skewed towards populations of European ancestry, and do not capture the full breadth of human diversity. The few large-scale studies that have extended their sampling more globally, such as the 1000 Genomes Project, have still been limited to relatively small numbers of large, metropolitan populations. If genomics research is not broadened to include a greater diversity of groups, we risk ending up with an incomplete understanding of human genetic variation. Furthermore, sometimes smaller, more isolated populations can provide vital clues towards the broader picture of how humans have diversified and dispersed across the world.

In our study, we sequenced the genomes of 929 people from 54 diverse human groups across five continents, including many populations of particular geographical, anthropological and linguistic interest. The dataset includes, among many others, hunter-gatherer groups from Central and Southern Africa, indigenous groups from the Americas and Siberia, highlander groups from Papua New Guinea, and linguistically unique groups from Europe and South Asia.

We discovered many genetic variants that had not been found by previous studies. While most of these novel variants are rare, some of them are common in at least one group. This demonstrates the value of anthropologically-informed sampling for characterizing human genetic diversity. Our study was not a medical genetics study, so we cannot say if any of these novel variants influence disease risks, but it is likely that at least some of them will be medically relevant. Our findings thus highlight how the effects of such variants could be overlooked, unless medical genetics studies also extend to more diverse groups.

The global patterns of genetic variation also testify to the different historical trajectories that unfolded in different parts of the world. We found that populations in the Americas, Oceania and parts of Africa each have tens of thousands of common genetic variants that are not found elsewhere, while populations in different regions of Eurasia have much fewer such variants. This likely reflects how Eurasian populations have been more genetically interconnected with each other, especially in the last 10,000 years - after the large-scale migrations that followed the advent of agriculture. By contrast, populations in the other continents have remained more isolated and thus retained more unique variation.

Neanderthals and Denisovans, two ancient human groups that roamed Eurasia until about 40-50,000 years ago, have also been important in shaping our diversity. As a result of prehistoric mixing, many people today carry small pieces of DNA inherited from these groups. Our data shows that Neanderthal DNA segments found in present-day people show low diversity. This is consistent with the idea that these Neanderthal segments come from one major episode of mixing, but that it must have involved several Neanderthal individuals. By contrast, the Denisovan segments found in modern human genomes display greater diversity and seem to derive from two separate mixing episodes. Mixing with different past human groups has thus introduced many new genetic variants into present-day populations, but so have entirely new mutations occurring in our more recent evolutionary history. Overall, we found that a larger number of genetic variants were created by new mutations than the number inherited through mixing with other human groups.

All humans can trace their ancestry back to Africa, but we only have a fuzzy understanding of how our early common ancestors within the continent branched out to give rise to the diversity of people we see today. By applying new sequencing technology, we could more accurately study the timescale of this early phase of human evolution. We found that the process of diversification unfolded in a highly gradual and complex fashion. It is thus not very meaningful to think about the human family tree as having started to branch out at some specific, single point in time. Instead, our results suggest that the bulk of this process occurred slowly during the last 250,000 years, but with substantial genetic contact between all populations during much of this period. However, a small fraction of present-day diversity also seems to trace back to groups that had branched off earlier than this, perhaps even as early as before 500,000 years ago. The details of this intricate process, including where within Africa it unfolded, remain largely unknown.

Our study represents an important step towards a more comprehensive understanding of our genetic variation and history, but the genomic map of human diversity still contains many blank spots. Sequencing the genomes of even more people from diverse groups, combined with studying ancient DNA from fossil remains, will hopefully resolve how our complex past shaped our genomes.