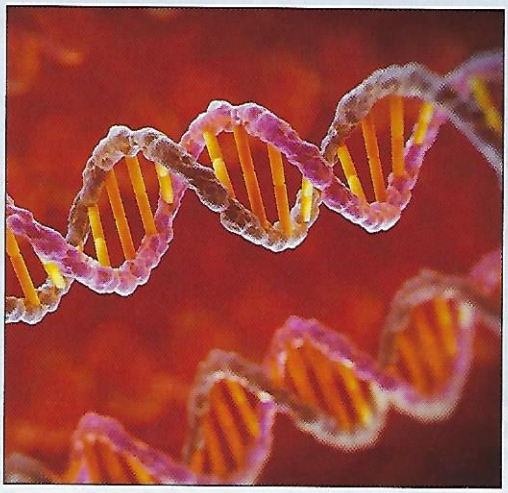
# The secrets of African DNA

*Africa's genetic data could spur a scientific revolution – but at the moment, it remains a largely untapped resource*

**Why is DNA research so important?**

DNA contains instructions for everything our cells do, from conception until death. Since 1953, when James Watson and Francis Crick discovered the double-helix structure of DNA – explaining how hereditary information is encoded – the central project of biology has been to try to understand how the shifting arrangement of four compounds in DNA determines the ways in which humans differ from each other, and from everything else alive. The Human Genome Project, completed in 2003, identified and mapped almost all the three billion or so base pairs – "letters" that spell out the genetic code – in our DNA. Since then, there have been great advances in our ability to sequence genetic data at scale, which promise to revolutionise how we understand our ancestry and history, how we identify physiological traits and, crucially, how we treat disease (see box).



*DNA: the operating manual for Homo sapiens*

**What makes African DNA special?**

The unique qualities of African DNA can be traced back to the origins of humanity. *Homo sapiens* evolved in Africa about 300,000 years ago, and began migrating to other parts of the world perhaps 115,000 years later. But only people from certain groups or regions left – those who left went through a population "bottleneck" that made them less genetically diverse, according to a 2008 Cornell University study published in the journal Nature. While the human race in Europe and beyond evolved from those early migrants, continental Africa retained the genomic diversity that was an early feature of the human race. As such, the continent is now the most genetically varied region on the planet: Nigeria alone is home to at least 250 distinct ethnic groups whose genetics have been conserved over thousands of years.

**How much do we know about African DNA?**

Not nearly enough. The Human Genome Project produced what is sometimes called the "book of life": the operating manual for *Homo sapiens*. But it was based on DNA from blood donations taken in Buffalo, New York: around 70% of the DNA sequenced came from one donor; most of the donors were of European descent. The "human reference genome" produced by the project – an approximation of most people's genetic make-up and a crucial tool for research – has been refined and is now more diverse, but is still mostly based on European DNA, since most of the big genome-mapping projects took place in North America and Europe. If data is the new oil, said Carlos Bustamante, professor of biomedical data science at Stanford University, then in terms of DNA, "we've only been looking in the North Sea".

**Why is that important?**

This lack of African genetic data "constitutes a significant obstacle to understanding how our bodies and diseases function", said Neil Munshi in the FT. Researchers look for small differences between the reference genome and the genomes of the individuals they are studying: even a single change in the base pairs, the "letters" within the DNA – known as a single-nucleotide polymorphism or SNP – can cause diseases, such as sickle-cell anaemia or cystic fibrosis. One recent study of people of African descent by Johns Hopkins found that 300 million base pairs of their DNA – about 10% of the estimated size of the entire human genome – simply did not feature in the human reference genome.

**Whom does that affect?**

First off, non-Europeans. Scientists miss mutations in these ethnic groups, meaning that much cutting-edge genetic research is worthless for them. "Until we understand the differences between genetics in Western populations and other regions of the world, we are blindly applying drugs, treatments and diagnostic insights from European populations onto African people," said Queen Mary University's Dr Deepti Gurdasani. But it also limits our understanding of disease in other populations, because our comprehension of the entire human genome is incomplete. "When the entire continent of Africa has been left out of genetic studies, it's everyone that loses out," said Dr Alicia Martin of Harvard Medical School. A study in Nature Genetics last year found that the lack of genetic diversity was "the most critical limitation to genetics in precision medicine".

**Can this be rectified?**

African populations offer "low-hanging fruit" for genetic research, the Nature Genetics study found – and researchers are now waking up to that untapped potential. In 2012, H3Africa (Human Heredity and Health in Africa) was launched to study genomics and improve health in African populations. The project brought together a consortium of more than 500 African scientists and is conducting research on the causes of health problems such as blindness, Alzheimer's, cancer and kidney disease. Meanwhile, Nigeria-based 54gene is seeking to create Africa's first private biobank. It hopes to have collected 200,000 DNA samples by this year; it will charge to give pharmaceutical companies access to African DNA for the development of drugs and medicines.

**Is this necessarily a good thing?**

There is a long history of "helicopter science", where researchers fly into less developed parts of the world, take samples, and leave; Africans (and African Americans) have sometimes been used as medical guinea pigs by unscrupulous scientists. Leading pharmaceutical companies are now keen to tap into Africa's genomic data; officials in the Trump administration have raised concerns that the Chinese state is trying to "steal" such data from Africans. The hope is that non-profit initiatives like H3Africa, and locally owned businesses like 54gene, will be able to ensure that Africans actually benefit this time. Dr Abasi Ene-Obong of 54gene insists that his company "will prioritise diseases that affect Africans disproportionately".

## The future of genetic medicine

Genetics has already transformed medicine, though arguably that process is only just beginning. Testing to diagnose inherited disorders – such as mutations in the BRCA breast cancer suppressor genes – has been around for decades. The human reference genome has helped researchers locate thousands of genes associated with particular illnesses, including hundreds that cause specific types of cancer. Recent advances in sequencing technology mean that an individual's genetic make-up can now be analysed relatively easily and cheaply, and rare diseases diagnosed.

A second major area is "bioinformatics": mining the human genome to help design drugs. This promises to give the process of developing drugs – traditionally long, difficult and vastly expensive – a much greater success rate. It also holds out the promise of personalised medicine. For instance, people who are given the antiretroviral drug Abacavir are tested to ensure they will not have a life-threatening response. A third major focus of research is gene therapy: the editing of defective genes to cure diseases. The area is still largely experimental, but it holds out great promise for treating cancer and single-gene defects such as cystic fibrosis and muscular dystrophy.