

# WHAT OUR DNA CAN TELL US ABOUT THE HISTORY OF HUMANS

Leo Speidel<sup>1\*</sup> and Clare Bycroft<sup>2\*</sup>

<sup>1</sup>Department of Statistics, University of Oxford, Oxford, United Kingdom

<sup>2</sup>Genomics plc, Oxford, United Kingdom

## YOUNG REVIEWERS:



ANNA-MARIE

AGE: 15



MARIANA

AGE: 15



ZARA

AGE: 14

## DNA

Molecule that stores the instructions for how our bodies work. It is passed on from parents to their children, and sometimes mutates, causing small differences in the DNA of present-day people. Its full name is deoxyribonucleic acid.

Almost every cell in our bodies contains DNA. DNA is a molecule that stores the instructions for how our bodies work and it is passed on from parents to their children. In this article, we show you how DNA can be used as a time machine, taking us back many thousands of years and revealing stories of our ancestors. For instance, we can find out about the ancient history of humans, and tell where and with whom our ancestors likely lived. DNA can also tell us about a country's recent history, uncovering stories of how ordinary people lived or moved about.

## INTRODUCTION

Our **DNA** is like a long diary of human history, passed down from generation to generation. This diary contains many fascinating stories of our ancestors, with each new generation adding its small contribution. Using mathematics, statistics, and computers, scientists can uncover these stories by making sense of small differences in our

## Figure 1

(A) We show two different chunks of DNA in red and blue and how they were passed through our ancestors. As they are passed from one generation to the next they sometimes mutate, so that there are small differences between the chunks inherited by present-day people. (B) DNA family trees are different for the red and blue chunks, which can be seen by tracing back who each chunk was inherited from in (A). Scientists can reconstruct these trees using the idea that chunks with fewer differences are more closely related than chunks with more differences.

## HOMO SAPIENS

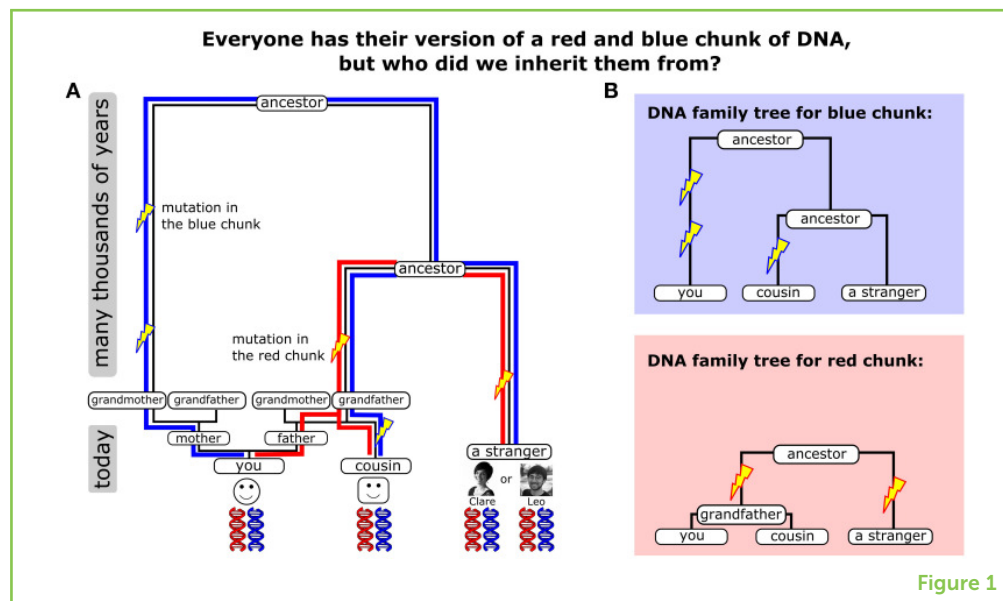
The name of the only remaining human species to which we all belong.

## MOLECULES

These are the chemical building blocks of our bodies and much of the living world. They are made up of multiple tiny particles called atoms that stick very tightly together.

## ANCESTRAL LINES

Ancestral lines describe which ancestors we inherited our DNA from (for instance, through mum—grandad—greatgrandma and so on). We inherit many chunks of DNA, each from a different ancestral line. Interestingly, only a few of your distant ancestors passed on



DNA. These can be stories about ancestors of **Homo sapiens** who lived hundreds of thousands of years ago, or they can be about our ancestors over the last few thousand years.

Let us start with where our DNA **Molecules** come from. We each inherited a unique combination of DNA from our parents. This is why we look different from our siblings, except for identical twins, whose DNA is the same. Each chunk of DNA can be traced to one of our four grandparents, and because they all have different ancestors themselves, each chunk of DNA tells a slightly different part of the story of the past.

Chunks of our DNA have been passed on through our ancestors over thousands of years. Let's imagine looking back in time to trace our **ancestral lines** for each chunk (red and blue lines in Figure 1A). When two people share an ancestor, their ancestral lines meet. For people who are closely related (like cousins), these ancestral lines can meet very recently, but for people who are not closely related, ancestral lines meet much further in the past. Ancestral lines form **DNA family trees**, showing how people's DNA are related to each other. Each chunk of DNA is inherited through different ancestors, so the DNA family tree can be different for different chunks (Figure 1B). The same two people can be closely related in some DNA family trees and distantly related in others. In fact, for some chunks, the closest DNA ancestor of you and your cousin can go back millions of years to the very origin of human-like apes.

Scientists have worked out ways to reconstruct DNA family trees by looking at the small differences in the DNA between people living today [1, 2]. Typically, the further back in time two people share an

their DNA to you, just by chance. This is because a parent can only pass on half of their DNA to each child, and so over many generations of halving, halving, and halving again, not all of their descendants will end up receiving a chunk of their DNA.

### DNA FAMILY TREE

They describe how and when the ancestral lines of different people meet back in time, and can tell us how we are related to each other over many thousands of years.

### NEANDERTHAL

A distant relative of *Homo sapiens*, who lived across Eurasia long before *Homo sapiens* did. When *Homo sapiens* started to populate Eurasia (migrating out of Africa), they mixed with Neanderthals, so that most people with non-African ancestry carry a small amount of Neanderthal DNA. Neanderthals went extinct around 40,000 years ago.

ancestor, the more genetic differences there will be between those two people.

In this article, we will tell you three different stories contained in these DNA family trees. The first story is about how we can uncover the ancient past, including when our ancestors settled in different parts of the world. The second story looks at our recent history and what we can learn from the DNA of people living in the same country today. The third story is about how groups of people have moved and merged, and how we are all really a mixture of many ancestral populations.

## UNCOVERING THE ANCIENT PAST USING DNA FAMILY TREES

Written records only go back a few thousand years. They cannot tell very ancient stories but, amazingly, our own DNA can. To understand how scientists read the stories from DNA, let us consider you and a friend. If you are both in a small room with a few other people, you will easily find your friend. If you are in a larger room full of hundreds of people, it might be much harder to find your friend. We can think about your ancestors using similar reasoning. At any given time in the past, if there were only a small number of humans alive or if your ancestors inhabited the same geographic region, then the chance that you share an ancestor with your friend at that time is high. On the other hand, if the number of humans living at that time was large or if your ancestors lived far away from each other, the chance of sharing an ancestor at that time will be small.

How often you and your friend shared an ancestor at different times in the DNA family trees can tell us whether your ancestors lived close to each other, how they moved about, and how many other humans occupied our world back then.

Today, almost 8 billion humans live on our planet, spread widely across all continents. Until relatively recently, humans were a much rarer species, and going back as far as 200,000 years ago, humans were mostly living on the African continent. DNA family trees indicate that a relatively small number of people migrated out of Africa and settled in other parts of the world. At that time, some distant relatives of ours, most famously **Neanderthals**, were already living in Europe and Asia. For a long time, we did not know how Neanderthal people were related to us. Only when scientists compared their DNA to ours, and used the idea of DNA family trees, did they discover that humans who migrated out of Africa mixed with Neanderthals [3].

## A NATION'S HISTORY WRITTEN IN DNA

Our DNA can also tell us about the much more recent past. If we concentrate on the most recent bits of our DNA family trees, we can learn about the history of our modern human ancestors—when, where, and with whom ordinary people lived or moved about. When people living within a localized region have children together over many generations, their descendants will share common ancestors more recently with each other than with people from elsewhere. This shows up as small differences in how DNA chunks are shared between people whose ancestors lived in different places.

### GENGHIS KHAN

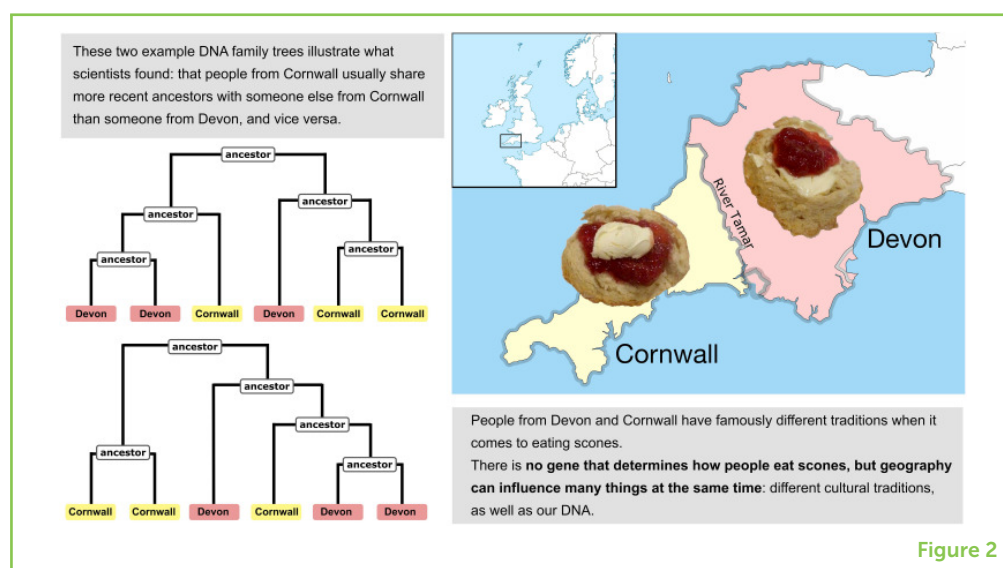
A military leader who conquered and ruled over a large area of the world known as the "Mongol Empire." The empire began around 1200 CE and lasted about 200 years. For some of the time it covered a vast area that reached roughly from parts of Eastern Europe to the eastern-most parts of China.

We have known for a while that such regional differences came about on a world-wide scale over many hundreds of generations (many thousands of years). For example, two Han Chinese people tend to have more similar DNA than a Han Chinese person and an Irish person, because overall, the ancestral lines of two Han Chinese people tend to merge more recently than with the Irish person in the DNA family trees. You might wonder if these regional differences happen on a much smaller scale, such as within countries. The DNA among people who share many recent ancestors tends to be extremely similar: for example, if the shared ancestor lived around the time of **Genghis Khan** (roughly 800 years ago), we expect fewer than one in a million DNA positions to be different!

When scientists looked at the DNA of people living in the British Isles today [4], they found DNA differences between groups of people who lived only short distances apart. They could see, for instance, the border between Cornwall and Devon (two regions in England) because it is possible to tell apart the DNA of people whose grandparents were born in each region. This means that the ancestors of people in Devon tended to have children with other people from Devon rather with

**Figure 2**

Scientists have found small differences in DNA even among people from very nearby regions, such as Devon and Cornwall in England.



**Figure 2**

### Figure 3

DNA can be used like a clock to figure out when two populations likely came together. For example, if a group of people from Iberia and a group of people from North Africa had children together, the DNA of their descendants would look like a mosaic of chunks from the two ancestral groups. Using DNA from lots of people today, scientists can figure out which chunks were likely inherited from the different ancestral groups, and the size of the chunks tells them when the population mixing likely first happened.

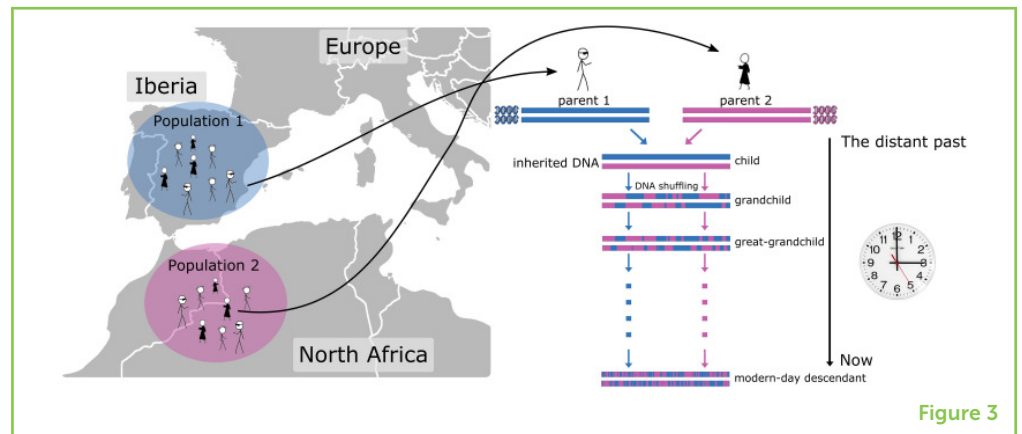


Figure 3

than the Cornish, just across the River Tamar (Figure 2). There are many other areas of relative isolation, such as North and South Wales, West-Yorkshire, and geographically isolated places, such as the Orkney Islands. Interestingly, these patterns reflect certain events in history, such as areas of different settlements of Britons and Saxons around 600 CE. We now know that small regional differences like these are not unique to Britain, as they have also been found in many other parts of the world, such as Spain, Ireland, Italy, Finland, and Japan, each with a unique story to tell.

## THE HUMAN MIGRATORY MELTING POT

By looking at DNA we know that geographical features, such as the sea or mountains, can be barriers to people meeting and having children together—but not always. There are many examples in human history of groups of people moving, sometimes across vast oceans and huge mountains, and having children with other people they meet. We can see this in the DNA of people living today.

One clear example of this is in Iberia (including today's Spain). We know from written history that there was a migration of people from North Africa into Iberia during the time of Muslim rule (around 700–1200 CE). When scientists compared the DNA of Spanish people today with that of people from elsewhere in Europe and in North Africa, they could see that the North African migrants most likely had children with the locals [5]. This is because, for some chunks of DNA, the ancestral lines of a Spanish person merge most recently with a modern-day person from North Africa (usually North Morocco), but other chunks merge first with a person from Europe (usually France).

The DNA of modern-day people can also tell us when the mixing of historical human populations is likely to have occurred. Over many generations, the chunks of DNA inherited from each of the original ancestors get shuffled up among the descendants into smaller and smaller chunks. We can use this shuffling like a genetic clock, to count how many generations have passed since the initial mixing [6] (Figure 3). In the case of Iberia, DNA revealed that most of the

mixing probably happened between 860 and 1120 CE, beginning a few generations after the initial migration from North Africa [5].

## CONCLUSION

Our DNA contains stories of the ancient past, long before any written records existed. These stories add to archeological finds by telling us things like how many people likely lived in the ancient past, and when different groups of humans likely came into contact with each other. DNA can also add another dimension to our history books, by telling us about the events that shaped the lives of ordinary people, not just those who happened to be in charge.

Learning about our shared history is interesting in itself, but studying small DNA differences (and their origins) between people is also crucial for understanding human health. Scientists are busy studying DNA to better understand the processes that keep us well or make us sick, and to discover new treatments to make us better [7]. By deciphering the genetic code of people living today—with the help of mathematics, statistics and computers—we can both uncover the story of our past and improve our health in the future.

## ACKNOWLEDGMENTS

LS and CB would like to thank Megan (and Simon) Myers for very helpful feedback. LS acknowledges the Wellcome Trust (200186/Z/15/Z) for funding.

## REFERENCES

1. Speidel, L., Forest, M., Shi, S., and Myers, S. R. 2019. A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet.* 51:1321–9. doi: 10.1038/s41588-019-0484-x
2. Kelleher, J., Wong, Y., Wohns, A. W., Fadil, C., Albers, P. K., and McVean, G. 2019. Inferring whole-genome histories in large population datasets. *Nat. Genet.* 51: 1330–8. doi: 10.1038/s41588-019-0483-y
3. Ottenburghs, J. 2019. Why do some humans have neanderthal DNA? *Front. Young Minds* 7:104. doi: 10.3389/frym.2019.00104
4. Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., et al. 2015. The fine-scale genetic structure of the British population. *Nature* 519:309–14. doi: 10.1038/nature14230
5. Bycroft, C., Fernandez-Rozadilla, C., Ruiz-Ponte, C., Quintela, I., Carracedo, Á., Donnelly, P., et al. 2019. Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula. *Nat. Commun.* 10:551. doi: 10.1038/s41467-018-08272-w
6. Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., et al. 2014. A genetic atlas of human admixture history. *Science* 343:747–51. doi: 10.1126/science.1243518

7. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562:203–9. doi: 10.1038/s41586-018-0579-z

**SUBMITTED:** 09 December 2019; **ACCEPTED:** 10 July 2020;

**PUBLISHED ONLINE:** 10 September 2020.

**EDITED BY:** Vishal Shah, West Chester University, United States

**CITATION:** Speidel L and Bycroft C (2020) What Our DNA Can Tell Us About the History of Humans. *Front. Young Minds* 8:106. doi: 10.3389/frym.2020.00106

**CONFLICT OF INTEREST:** CB is a scientist at Genomics plc. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**COPYRIGHT** © 2020 Speidel and Bycroft. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## YOUNG REVIEWERS



### ANNA-MARIE, AGE: 15

My favorite subject is biology, I like bionics. In future I would like to work on new materials, new substances. I am curious about the mysteries of the universe, there is so much to discover.



### MARIANA, AGE: 15

Hi! My name is Mariana and I am 15 years old. I like reading books, spending time with animals, and I love swimming. I consider every day brings an opportunity for me to learn something new and for being happy.



### ZARA, AGE: 14

My name is Zara and I am 14 years old. In my spare time, I enjoy reading, helping my community, and playing volleyball. I think it is important to work hard for achieving our dreams and for being a better person every day.

## AUTHORS

### LEO SPEIDEL

I studied mathematics at university in my home countries of Germany and Japan. Being of mixed ancestry, I always had an interest in where we come from and how we are all related. When I discovered that maths can be used to find out about our past,



I decided that this is what I wanted to do. Now I am a researcher at the University of Oxford, where I use maths and stats to study genetics. \*speidel@stats.ox.ac.uk

**CLARE BYCROFT**

I have always enjoyed puzzles and logic, so I studied mathematics at university in my home country, New Zealand. It turns out that maths can be really useful for studying DNA—nature’s instruction manual for life. At the University of Oxford (UK) I used DNA to research the population history of Spain. I also helped curate a big research database, the “UK Biobank,” that many scientists like myself are now using to discover which bits of our DNA are most important for human health and disease. \*bycroftc@gmail.com