Sense About Genetic Ancestry Testing

There are now many companies which offer to tell you about your ancestors from a DNA test. You send off a sample of your DNA and £100–£200 ($150–300), and in return you receive a report. The results of these tests may find a connection with a well-known historical figure. They might tell you whether you are descended from groups such as Vikings or Zulus, where your ancient relatives came from or when they migrated.

Adverts for these tests give the impression that your results are unique and that the tests will tell you about your specific personal history. But the very same history that you receive could equally be given to thousands of other people. Conversely, the results from your DNA tests could be matched with all sorts of different stories to the one you are given.

It is well known that horoscopes use vague statements which recipients think are more tailored than they really are (referred to as the 'Forer effect'). Genetic ancestry tests do a similar thing, and many exaggerate far beyond the available evidence about human origins. You cannot look at DNA and read it like a book or a map of a journey. For the most part these tests cannot tell you the things they claim to – they are little more than genetic astrology.

What can we know about your personal ancestors by looking at your DNA?

Not much. Genetic ancestry tests use some techniques that have been developed by researchers for studying differences in DNA across many groups of people. The things we know about genetic ancestry, almost without exception, are about the genetic history of whole populations.

Companies use techniques from this field and sell their findings to people who want to find out about their personal history. The techniques were not designed for this. The information they give is not unique to any individual. While there are other, more specific flaws with these testing services, that fundamental point alone means that the very concept of individual genetic ancestry tests is unsound.

What is genetic ancestry testing?

There are different types of test. All of them use a small sample of a person's DNA, usually taken from a mouthwash or a cheek swab, and compare sections of it to the DNA of others for whom we have information about ethnicity and geographic location. Different tests look at different parts of an individual's DNA:

- Y chromosome DNA (this is only found in men and is inherited along the male line).
- mitochondrial DNA, 'mtDNA' (this is found in men and women, inherited along the female line).
- autosomal DNA (this is 98% of your DNA and can come from any ancestor).

Each of us has just one ancestral lineage for mtDNA, and each man just one for Y chromosome DNA. These are inherited as a unit – so in ancestry terms they are passed on like a single gene. On the other hand autosomal DNA is made up of thousands of sections of DNA, each with its own history.

What do we mean when we talk about ancestors?

The DNA ancestry tests appeal to our interest in our family trees. However, our DNA is not the story of our family tree. It is a mosaic of genetic sequences that have been inherited via many different ancestors. With every generation you (nearly) double your number of ancestors because every individual has two parents – going back just 10 generations (200-300 years) you are likely to have around a thousand ancestors. We don't have to look back very far in time before we each have more ancestors than we have sections of DNA, and this means we have ancestors from whom we have inherited no DNA.

When genetics researchers talk about common ancestry between people they usually mean that they are tracing the inheritance of particular sections of DNA or genes. And we know that different sections of our DNA have different patterns of genetic ancestry. This means that researchers can get very different estimates of how recently we share ancestors, depending on what they are looking at:

- Researchers look at mtDNA to follow ancestry passed along the female line. For mtDNA, everyone alive today shares a common ancestor who lived between 160,000 and 200,000 years ago.[1]
- When researchers look at Y chromosome DNA to follow ancestry through the male line, the most recent estimate is of a common ancestor who lived between 240,000 and 580,000 years ago.[2]
- If we look at sections of DNA from other parts of the genome (autosomal DNA), the date of a 'common ancestral section of DNA' (that is, a section of DNA that everyone alive today has inherited) varies from gene to gene, but has been estimated to average around 1 million years ago.[3]

Sense About Genetic Ancestry Testing

There are some things genetic ancestry tests can tell you quite accurately

There are credible ways to use the genetic data from mtDNA or Y chromosomes in individual ancestry testing, such as to supplement independent, historical studies of genealogy. If, for example, two men have identified – through historical research, possibly involving surnames – a common male-line ancestor in the sixteenth century, it would be reasonable to use their Y chromosome data to test this. There are some ancestry testing companies that offer this service.

To answer a specific question about individual ancestry with any degree of confidence requires a combination of historical records and genetic information.

- If, however, you look for the most recent person that everyone alive today is descended from, the best current estimate is that the individual lived only 3,500 years ago[4] – which is much more recently than you might imagine.

Can genetic ancestry testing tell you that...

...you are related to a historical figure?

A company might tell you that you are related to the Queen of Sheba or Napoleon. The short response to this is, yes, you probably are! We could say this for many people alive today in connection with many people from the past without having to do any genetic test at all. We are all related, it's a matter of degree. Not only is our common ancestor estimated to have lived 3,500 years ago, but reasonable estimates show that every individual alive around 5,000 years ago was either a common ancestor of everyone alive today, or of no one alive today[4]. So at that point in the past we all have *exactly* the same set of ancestors.

If you are told that you are genetically related (share a genetic marker) to someone who lived a long time in the past, it may well be true but is not very meaningful. In reality, we all share the vast majority of our DNA through remote common ancestors – and we may have little DNA that is directly inherited from an ancestor who lived even just a few generations ago.

...your ancestors came from a particular group of people or part of the world?

Genetic ancestry testing presents a simplified view of the world where everyone belongs to a group with a label, such as 'Viking' or 'Zulu'. But people's genetics don't reflect discrete groups. Even strong cultural boundaries, such as between the Germanic and Romance language groups in Europe, do not have very noticeable genetic differences. The more remote and less-populated parts of the UK, such as the Scottish Highlands, do have some genetic differences from the bulk of the population, but they are not big. There is no such thing as a 'Scottish gene'[5]. Instead groups show a story of gradual genetic change and mixing.

Where we can make a connection between a tribal group and a particular section of DNA, for example, we could say that if you carry it today there is a possibility that some of your ancestors were in that group. But that is all. Human history is complicated and involves a lot of migration – so your ancestors are probably from many different places. Each of your genes has its own history, which means there are thousands of possible versions of your ancestry. If nothing is done to test which of the set of possible histories is the most likely explanation for your DNA, it is simply storytelling.

...your ancestors moved to a particular part of the world at a specific time?

Researchers use the genetic differences between Y chromosomes or mtDNA among a set of individuals to infer possible trees of relatedness. We can estimate the times of common ancestors on those trees, although these estimates lack precision. But it is not reasonable to make a leap from these DNA trees to mapping your ancestors onto geographical locations or past migrations. For example, a man in the UK might have a type of Y chromosome that has been found more often in North Africa than elsewhere. However, this is based on populations in North Africa now, not in the past, and people have moved over the centuries. And, the same Y type may be found in other parts of the world – he could equally have inherited it from one of these. And even if some of his ancestors did come from North Africa, it does not show when they came to the UK or how many of his millions of ancestors came from that region.

So if a genetic ancestry company talks confidently about your genetic ancestors moving from one location to another at a specific time, that is either made-up, or it is so true as to be equally meaningless i.e. any individual alive today would have a similar genetic connection.

Patterns of genetic ancestry are messy

Why is the business of working out ancestry so tricky? At its simplest it is because ancestry is a lot more complicated than it appears. Patterns in our DNA are shaped by the human history of moving around. Looking back through time it takes only a small amount of migration to 'pull in' ancestry from a wide geographic range. There is also a lot of interference from other, random processes: DNA mutation, who has children with whom and which genes pass from one generation to another. All these noisy processes limit what any test will be able to tell you because:

- The same patterns of genetic differences between us can result from a range of very different human histories.
- One version of history could produce lots of very different patterns of genetic differences between us.
- There are sometimes unexpected changes in DNA patterns, so that what looks to us like a reflection of one history in fact results from a completely different one.

Sense About Genetic Ancestry Testing

Tests using DNA information about a single lineage (Y or mtDNA) cannot tell you detailed genealogical information

Because Y chromosome and mtDNA follow one line of inheritance (patrilineal or matrilineal), they can inform us about a single line of ancestors. But that is only one of your many ancestral lineages, and it becomes increasingly irrelevant to your total ancestry the further back in time you go: beyond more than a few generations, most people's ancestry is complex and becomes difficult to determine from genetics. Researchers choose to look at Y chromosome or mtDNA precisely because they follow a simplified pattern of inheritance that it is possible to study. The problem comes when commercial genetic ancestry tests treat this kind of data as though it is the whole story.

Genetic ancestry research into human history

The commercial genetic ancestry tests borrow selectively and misleadingly from the research field looking at population genetic ancestry, but they should not be confused with it!

There are three main approaches used in research looking at the human past: phylogeography, descriptive statistics, and population-genetic modelling. There are researchers using all three of these, and we can take some insights from each of them. However, some of the methods are more robust than others and researchers in this field disagree about when and how it is appropriate to use them, particularly in relation to asking questions about human history.

Phylogeography

This approach describes geographic patterns in genetic or other types of data (e.g. languages), that can be related through a tree diagram to make suggestions about migration histories of individuals, populations or species. This is almost always the approach that ancestry testing companies use with Y chromosome or mtDNA data. It can generate hypothetical histories, but does not tell us with any confidence that they are likely to be true accounts of how a DNA pattern came about. One major criticism of using phylogeographic methods to infer population history is that they are usually interpretive, and so guided by subjective biases. A second is that the methods have not been validated by showing that they work with examples where we already know the history.

Descriptive summary patterns

Some statistics can be used to describe the main characteristics of a set of genetic data. If the scenario is kept simple, these statistics can represent processes such as migration and changes in population size, which influence the movement of genes and the amount of genetic variation in human populations. When ancestry testing companies analyse autosomal data they will usually use one or a number of these statistics. However, this approach cannot represent the complex influences on genes over time, or the role of randomness in DNA mutation and in the way genes are passed to later generations.

Population-genetic modelling

Many population geneticists are interested in being able to say something about the origins, past migrations, and past mixing between individuals within human populations. This is a very difficult task but by looking at many individuals from different regions, and in particular by modelling possible population histories and calculating how likely they are to explain a DNA pattern, different ideas about population history can be tested in robust scientific ways. This approach helps researchers to work out which of the many possible histories is most likely to have led to the pattern in the DNA being studied. Because of this, the approach is considered to be more useful and scientific. However, while it can tell us something about populations, it tells us very little about specific individuals.

References

1. Ingman M, Kaessmann H, Pääbo S & Gyllensten U. 2000. **Mitochondrial genome variation and the origin of modern humans.** *Nature* 408, 708-713
2. Mendez FL et al. 2013. **An African American Paternal Lineage Adds an Extremely Ancient Root to the Human Y Chromosome Phylogenetic Tree.** *American Journal of Human Genetics http://dx.doi.org/10.1016/j.ajhg.2013.02.002*
3. Garrigan D & Hammer MF. 2006. **Reconstructing human origins in the genomic era.** *Nature Reviews Genetics* 7, 669-680
4. Rohde DLT, Olson S & Chang JT. 2004. **Modelling the recent common ancestry of all living humans.** *Nature* 431, 562-566
5. Nielsen R & Beaumont M. 2009. **Statistical inferences in phylogeography.** *Molecular Ecology* 18, 1034-1047

This Sense About... note was researched and drafted in February 2013 by Professor David Balding, Professor Mark Thomas and Tabitha Innocent, Sense About Science; with assistance from Dr Turi King, Dr Lounès Chikhi, Dr Rosalind Harding, Professor Mark Jobling & Professor Guido Barbujani. This document is prepared and published by Sense About Science, which has final responsibility for the content.

Designed by Stephen Chorlton and Sean McMahon.