

DNAdots™

Simple explanations
of modern genetic technologies

Personal genetic ancestry testing

What it is:

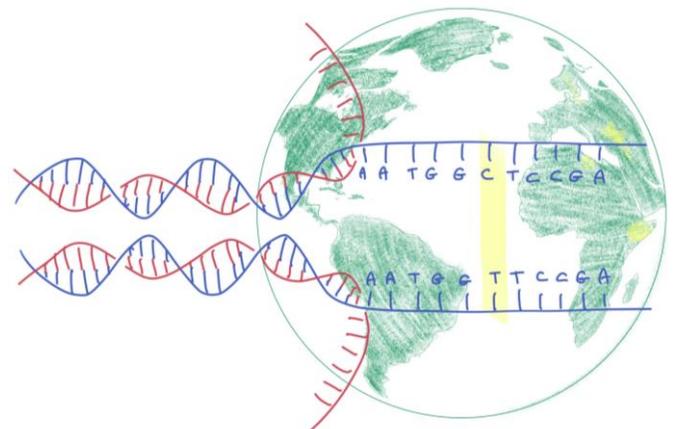
Using your DNA to track where your ancestors are from

You've certainly seen the ads; spit in a tube, put it in the mail, and find out who you are and where you are from. For people interested in their ancestry, the promise of a precise breakdown of what regions of the world and what ethnic groups their ancestors are from is enticing. And today, with price tags usually under one hundred dollars, individual genetic testing is directly accessible to most consumers. There are several companies that provide such testing services, and they all use the same basic approach: analyze your DNA for specific bits of variation and then match those variants to regions and populations in the world where the variants are most common.

How it works:

Comparing SNPs from around the world

If you were to compare any two random people's DNA, you would find that they are remarkably similar – around 99.9% of the DNA would be identical at the base pair level. Where they differ, however, can be incredibly informative, and it is that 0.1% of differences that makes a person genetically unique. But while your DNA sequence as a whole is unique, it is very unlikely that any particular stretch of your DNA is. Your DNA is put together from many different pieces that have been mixed and matched and recombined over generations. Each individual stretch of sequence is actually shared by many, many other people. You share these sequences with other people because we are all related, however distantly. Ancestry testing relies on the fact that people from certain regions of the world, or from certain ethnic groups, tend to be more closely related to each other than they are to other people.



The most common technique that ancestry testing companies use is to analyze so-called single nucleotide polymorphisms, or SNPs (pronounced "snips"). Your DNA is made up of long stretches of the four nitrogenous bases, adenine (A), thymine (T), guanine (G), and cytosine (C) laid out in a very particular order. When comparing two people's DNA sequences, this order is largely identical, but when a single nucleotide is different – when, for example, you have an A where another person has a G – that's a SNP. The vast majority of these SNPs have little to no effect on the person carrying them and become more or less common in specific populations over time through mostly random chance. By finding many of these variations throughout the genome and testing people from all over the world, certain patterns start to emerge. At one site, C may be common in Southern Europe, while everywhere else a T is common. At another site, a G may be common in Western Africa, and so on. Testing companies analyze hundreds of thousands of these SNPs in many regions of the genome and in many different populations around the world. By doing so, it is possible to produce profiles of the types of variations that are common or uncommon across all the different

stretches of DNA. Testing companies will then compare your DNA at the same SNPs to these different population profiles and use complex algorithms to try to reconstruct where your particular combinations of SNPs seem to match best. They report these matches in terms of percentages; for example, 50% of your DNA seems to match DNA that is common in Southeast Asia, 25% Ireland, 20% Europe broadly, and 5% is unknown.

Does it work?

Theoretically, yes. Practically?

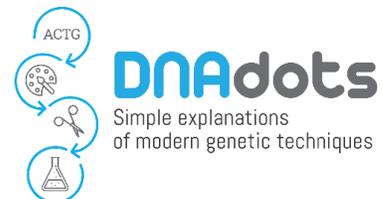
While identifying which SNPs are particularly common or uncommon and developing genetic profiles for different regions or ethnic groups is fairly straightforward, the reverse, accurately matching stretches of a single person's DNA to one specific group, can be much more difficult. Perhaps most important is the fact that throughout history people have migrated, and people from different ethnicities have intermarried. This means that there is really no such thing as distinct "groups" of people with clear boundaries to which we belong. Those stretches of DNA that have been mixed and matched to make you, have been mixed and matched in your ancestors for thousands of years. Because of this, there are very few DNA sequences that belong exclusively to one region or ethnic group. When a SNP is common in one area, it tends to spread, and just because a SNP is rare in a certain region, it doesn't mean that you couldn't have inherited it from there. Some companies have addressed some of these concerns by adding confidence scores to their testing results, indicating which parts of your ancestry they are most sure about. When looking at results, if large portions of the genome match a particular geographic region with high confidence, it is likely the results can be trusted. However, the higher the confidence that is reported, the less specific geographically the results will be, and smaller bits of variation will not be detectable. While a standard report may indicate many different specific regions that you match, high confidence reports tend to give much broader descriptions, like "Northern European" or "Central Asian." For many users, that lack of precision may be less satisfying.

Compounding these problems is the fact that the accuracy with which your ancestry can be determined greatly depends on the quality of the data your sample is being compared to. Testing companies invest heavily to make sure that they have large, quality sets of sequence data, but unfortunately, it is not possible to independently verify that quality, as companies do not make their data publicly available. We do know that there is a heavy bias in genetic testing databases towards European samples. This means that people with African, Asian, Native American, or other heritages are likely to find the tests less accurate and less informative, though testing companies are aware of this problem and are constantly working to make their data sets better.

Ancestry testing often reports your results in very specific percentages, but it is important to remember that genealogical ancestry and genetic ancestry are not necessarily the same thing. DNA is a physical molecule that is passed through generations. Each generation, only half of a person's DNA can be passed, and which bits are passed on is a mostly random process. This means that you don't carry DNA from all of your ancestors, and that the amount of your DNA tracing back to a specific region won't exactly match your actual percentage of ancestors from there. While one DNA testing website advertises: "Discover what makes you uniquely you," it may be important to remember that what makes you unique and what makes your history your own is much more than a DNA sequence.

Learn more:

- "Special Report: Genetic testing goes mainstream." Science News [Link](#).
- "What is Direct to Consumer Genetic Testing?" NIH [Link](#).



Personal genetic ancestry testing

Questions

Review:

1. Explain how your DNA is unique to you, but all of your DNA is shared with other people.
2. Explain what a SNP is.
3. Why are some SNPs more common in some area and less common in others?
4. The text says that there is no such thing as distinct groups of people. What are some factors that make defining clear groups of people impossible?
5. It is entirely possible for two siblings to get accurate ancestry results from a testing company that gives them different percentages of ancestry from different regions, even though they share all the same ancestors. How could this be?

Critical thinking:

1. Each individual SNP that exists originally came from a mutation that occurred in a single person's DNA, changing a single nucleotide to something different from what other people had at that site. Knowing that:
 - a) If any two people have the same nucleotide at a particular SNP, what does that say about their ancestry?
 - b) If you compared your DNA to any random person on Earth, you would share at least several SNPs. What does that say about human ancestry in general?
2. The number of ancestors you have theoretically doubles every generation you trace your lineage back in time. How many ancestors would you expect to have if you trace your ancestry back 5 generations? 10 generations? 50 generations? Do those numbers seem possible? Explain.

Discussion:

1. The majority of SNPs that testing companies test for give little information about the health of the person that was tested. But many *are* associated with health and disease. Your SNP results can also be used to identify you and people who are related to you, reliably out to 3rd or 4th cousins. When you give a DNA sample to a testing company, they own the data they get from that sample and use it to improve their ancestry testing, medical tests and other tests that they develop. How do you feel about giving a private company your genetic information? Should there be limits on what they can do with it? What if the information can be used to develop new health treatments or tests? What is their responsibility to your privacy?