

Modern genes yield atlas of ancient inter-ethnic sex

Encounters among populations left marks in the DNA of living people.

Daniel Cossins

13 February 2014



James L. Stanfield/National Geographic Creative

The Kalash people who live in the Hindu Kush Mountains of modern Pakistan carry genes that probably originated in Europe and might have been carried East by the Macedonian Army of Alexander the Great.

From invasions and migrations to slavery and trade, history is embroidered with events that led to interactions between previously separate populations — and, in many cases, hanky-panky. By mining genetic data from living people, researchers have now created a historical atlas of instances of such mixing.

The study, published today in *Science*¹, uses statistical methods to make inferences about which populations interbred, and when, over the past 4,000 years. The findings are also presented as an [interactive map](#). The authors sequenced DNA from 1,490 people from 95 genetically distinct populations around the world, and tested almost 500,000 single-letter variations in DNA known as single nucleotide polymorphisms, or SNPs.

Eighty of the populations showed evidence of genetic mixing. Many of the mixing events identified are consistent with historical records, validating the approach. For instance, records suggest that the Hazara people of Afghanistan and Pakistan are descended partially from Mongol warriors, and the study confirmed this, showing that Mongol DNA entered the Hazara around the time that the Mongol Empire was expanding.

Other mixing events inferred by the study were previously unknown and do not show up in historical records, but may be plausible. For instance, sequences in the DNA of the Tu people from modern China indicate that Europeans similar to modern Greeks mixed with an East Asian population around 1200 bc. The source of this European DNA might have been merchants travelling the Silk Road.

In another example, the DNA of the Kalash people, a population isolated in a remote valley in Pakistan, showed evidence of input from Europe or the Middle East (the researchers could not pin down a precise geographic location) between 990 and 210 bc — a period that overlaps with that of Alexander the Great. Local Kalash tradition holds that they are descended from Alexander the Great's army, although there is no historical record of such a mixture.

“They’ve developed a sophisticated statistical method for making inferences about human history,” says Joshua Akey, a geneticist at

the University of Washington in Seattle who was not involved in the work. “Such advances are critical if we are going to read the stories of our past written in our genomes,” he adds. Akey recently led a similar study² on European genomes, the results of which were published in January (see 'Modern human genomes reveal our inner Neanderthal'). Although he has not reviewed the latest paper¹ in detail, he says that the authors' statistical method could be quite powerful and is a significant improvement on earlier approaches.

Mixed signals

Such genetic mapping is based on a genetic re-shuffling process called recombination. With the exception of the sex-determining chromosomes, every chromosome type in human cells occurs in two similar, but not identical, versions — one from each parent. But when it comes to creating the next generation, those two versions are merged into one chromosome by piecing together segments from each, and the resulting 'recombined' chromosomes are what goes into sperm and egg cells. The length of uninterrupted segments from each source decreases over successive generations, so the size of those segments serves as a timer: the longer the ancestral segments remain in modern-day genomes, the more recently the genetic mixing took place.

But using the DNA of living people has its limitations. For one, modern genomes will not always carry DNA that is a good approximation of that from the original source groups, especially for older mixing events where one of the ancestral populations might have died out without leaving any direct descendants. “Hopefully in future, we can find them by adding data from ancient DNA samples,” says Simon Myers, a statistician who specializes in bioinformatics at the University of Oxford, UK, and co-author of the latest study. It is also hard to precisely define sources of mixing when it occurred between genetically similar groups, and scenarios involving multiple waves of mixing over time or between multiple groups can be difficult to tease apart.

But Myers expects that larger sample sizes will improve resolution. “That will give us a deeper understanding of human history,” he says, “but it could also help us to understand how rare genetic variants, which can spread through mixing events, influence health and disease in different populations.”

Nature | doi:10.1038/nature.2014.14718

References

1. Hellenthal, G. *et al.* *Science* **343**, 747–751 (2014).
2. Vernot, B. & Akey, J. M. *Science* <http://dx.doi.org/10.1126/science.1245938> (2014).