

OLLI 497: Ancient DNA

Session 1: September 21st

Summary and Observations

Introduction

Many authors of science books often take the opportunity of the “Introduction” to tell their readers what to expect in the book. David Reich eventually does so in this “Introduction” but he first jumps with two feet right into a discussion of two previous efforts to use genetics to inform us of humankind’s past. In this case, he discusses the work of Luca Cavalli-Sforza, a scientist that Reich much admires, and the work of Svante Pääbo and his colleagues, especially Matthias Meyer and Qiaomei Fu, in sequencing DNA from Neanderthals. He points to the limitations that Cavalli-Sforza worked under, limitations that ultimately doomed his efforts. Reich clearly sees himself as fulfilling Cavalli-Sforza’s vision, winning his bet, and carrying his insight to the intended conclusion. He also builds on Meyer and Fu’s work, tweaking it to make their technique more robust and productive. Standing on the shoulders of giants perhaps?

Reich begins with “... Luca Cavalli-Sforza, the founder of genetic studies of our past”, who “... made a grand bet in 1960 that would drive his entire career. He bet that it would be possible to reconstruct the great migrations of the **past** based entirely on the genetic differences among **present-day** peoples.”

Reich points out that at the time Cavalli-Sforza began his research the only possibility open to him was to “... measure proteins in the blood, using variations like the A, B, and O blood types.... By the 1990s, he and his colleagues had assembled data from more than one hundred such variations in diverse populations. Using these data they were able to reliably cluster individuals by continent based on how often they matched each other at these variations: for example, Europeans have a high rate of matching to other Europeans, East Asians to East Asians, and Africans to Africans.”

Beginning in the 1990s, Cavalli-Sforza’s team was able to directly examine DNA. They analyzed about “... one thousand individuals from around fifty populations spread across the planet, examining variation at more than three hundred positions in the genome.” They were able to cluster the individuals into five groups, and “... the results corresponded uncannily well to commonly held intuitions about the deep ancestral divisions among humans (West Eurasians, East Asians, Native Americans, New Guineans, and Africans).”

Reich next describes more of their technique. “Cavalli-Sforza was especially interested in interpreting the genetic clusters among present-day people in terms of population history. He and his colleagues analyzed their blood group data by using a technique that identifies **combinations of biological variations** that are most efficient at summarizing differences across individuals.” [Unfortunately, Reich does not tell us what “biological variations” were used, nor their combinations. We are left in the dark.] “Plotting these combinations of blood group types onto a map of West Eurasia, they found that the one summarizing the **most variation** across individuals reached its extreme value in the **Near East**, and declined along a southeast-to-northwest gradient into Europe. They interpreted this as a genetic footprint of the **migration of farmers** into Europe from the Near East, known from archaeology to have occurred after nine thousand years ago. The declining intensity suggested to them that after arriving in Europe, the first farmers mixed with local hunter-gatherers, accumulating more hunter-gatherer ancestry as they expanded, a process they called “**demic diffusion**.”

Reich notes that his conclusion was wrong. In 2008, “... John Novembre and colleagues demonstrated that gradients like those observed in Europe can arise without migration. They

OLLI 497: Ancient DNA

then showed that a Near Eastern farming expansion into Europe might counter-intuitively cause the **mathematical technique** that Cavalli-Sforza used to produce a gradient perpendicular to the direction of migration, not parallel to it as had been seen in the real data.”

It’s important to keep in mind that, at least according to Reich, Cavalli-Sforza’s failure resulted from **limited data and faulty mathematical technique**; this issue will be relevant when we discuss Reich’s own mathematical technique. In any event, this leads Reich to introduce the topic of this book, ancient DNA. “It took the revolution wrought by the ability to extract DNA from ancient bones—the “ancient DNA revolution”—to drive a nail into the coffin of the demic diffusion model. The ancient DNA revolution documented that the first farmers even in the most remote reaches of Europe—Britain, Scandinavia, and Iberia—had very little hunter-gatherer-related ancestry.” In fact: “The highest proportion of early farmer ancestry in Europe is today not in Southeast Europe, the place where Cavalli-Sforza thought it was most common based on the blood group data, but instead is in the Mediterranean island of Sardinia.”

Reich admits that Cavalli-Sforza was right in his assumption that the genetic makeup of current populations can point to events in the distant human past. “For example, the lower genetic diversity of non-Africans compared to Africans reflects the reduced diversity of the modern human population that expanded out of Africa and the Near East after around fifty thousand years ago. **But the present-day structure of human populations cannot recover the fine details of ancient events.** The problem is not just that people have mixed with their neighbors, blurring the genetic signatures of past events. It is actually far more difficult, in that we now know, from ancient DNA, that the people who live in a particular place today almost never exclusively descend from the people who lived in the same place far in the past.”

Reich notes: “His bet was that the past was a much simpler place than the present, and that by focusing on populations today that are not affected by major migrations in their recorded history, he might be studying **direct descendants** of people who lived in the same places long before. But what the study of ancient DNA has now shown is that the past was no less complicated than the present. **Human populations have repeatedly turned over.**”

Reich sums up Cavalli-Sforza’s contribution: “Cavalli-Sforza saw before anyone else the full potential of genetics for revealing the human past, but his vision predated the technology needed to fulfill it. Today, however, things are very different. We have several hundred thousand times more data, and in addition we have access to the rich lode of information contained in ancient DNA, which has become a **more definitive source** of information about past population movements than the traditional tools of archaeology and linguistics.” [I wonder how archeologists and linguists would react to that claim!]

And that “more definitive source” is growing rapidly. He notes that in “...2015, whole-genome analysis of ancient DNA went into hyperdrive.” His graph shows the steep curve from 2010 of the number of samples of whole genome data for ancient DNA.

This explosive growth in whole genome ancient DNA analysis resulted from techniques first developed in Svante Pääbo’s lab. “Much of the technology for the genome-wide ancient DNA revolution was invented by Svante Pääbo and his colleagues at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who developed it to study extremely old samples such as archaic Neanderthals and Denisovans.” Reich’s next notes that his contribution was to ramp up the process, that Svante Pääbo helped him to establish his own lab in the U.S., and that...“My partner in this effort has been Nadin Rohland, who did her own seven-year apprenticeship in Pääbo’s laboratory before she came to mine. Our idea was to make ancient DNA industrial—to build an American-style genomics factory out of the techniques developed in Europe to study individual samples”

OLLI 497: Ancient DNA

Reich and Rohland realized that "... a technique developed by Matthias Meyer and Qiaomei Fu in Pääbo's laboratory could be the key to the industrial-scale study of ancient DNA. Meyer and Fu's invention was born of necessity: the need to extract DNA from an approximately forty-thousand-year-old early modern human from Tianyuan Cave in China. When Meyer and Fu extracted DNA from Tianyuan's leg bones, they found that only about 0.02 percent of it was from the man himself. The rest came from microbes that had colonized his bones after he died. This made direct sequencing too expensive, even using the hundred-thousand-times cheaper technology that had become available after around 2006. To get around this challenge, Meyer and Fu borrowed a page from the playbook of methods developed by medical geneticists. Just as medical geneticists had developed methods to isolate DNA from the 2 percent of the genome that is most interesting and to discard the other 98 percent, **Meyer and Fu isolated a tiny fraction of sequences from the Tianyuan bone that were human and discarded the rest**"

Reich spells out more of this technique: "The method of DNA isolation that Meyer and Fu developed has been central to the success of the ancient DNA revolution. In the 1990s, molecular biologists learned how to adapt laser-etching techniques invented for printing electronic circuits to attach millions of DNA sequences of their choice to silicon or glass wafers. These sequences could then be cut off the wafers using molecular scissors (enzymes) and released into a watery mix. Meyer and Fu took advantage of this method to synthesize fifty-two-letter-long sequences of DNA that, overlapping like shingles on a roof, covered much of human chromosome 21. **Exploiting DNA's tendency to bind to highly similar sequences**, they "fished" out the DNA sequences from Tianyuan that they were interested in by using as "bait" the sequences they had artificially synthesized. [Although Reich does not spell this out here, they must be using single stranded DNA, not the iconic double helix. A single strand would attract the complementary strand, "fishing out" the human DNA from the bacterial. Right?] They found that a large fraction of the DNA they obtained was from Tianyuan's genome. Not only that, but it was from the parts of Tianyuan's genome that they wanted to study. They analyzed the data to show that Tianyuan was an early modern human, part of the lineage leading to present-day East Asians. He did not have a particularly large amount of ancestry from archaic human lineages that were diverged by hundreds of thousands of years from modern human lineages, contradicting earlier claims based on the shape of his skeleton."

"Rohland and I adapted this technique to study the whole genome. We worked with our colleagues in Germany to synthesize fifty-two-letter-long DNA sequences covering more than a million positions at which people are known to vary. We used these bait sequences to **enrich** for human compared to microbial DNA, which in some cases increased the fraction of DNA that was of interest to us by more than a hundredfold. We gained another approximately tenfold jump in efficiency because we only **targeted** informative positions in the genome."

Reich goes on to note the gains in productivity realized through the use of robots in processing the DNA, and the computer systems used to analyze the data.

"These advances mean that whole-genome study of ancient DNA no longer requires screening large numbers of skeletal remains before it is possible to find a few individuals whose DNA can be analyzed. Instead, a substantial fraction of screened samples dating to the last ten thousand years can now be converted to working genome-wide data. The new methods have made it possible to analyze hundreds of samples in a single study. With such data, it is possible to reconstruct population changes in exquisite detail, transforming our understanding of the past."

OLLI 497: Ancient DNA

And transform they have: “We discovered that the population of northern Europe was largely replaced by a mass migration from the eastern European steppe after five thousand years ago; that farming developed in the Near East more than ten thousand years ago among multiple highly differentiated human populations that then expanded in all directions and mixed with each other along with the spread of agriculture; and that the first human migrants into the remote Pacific islands beginning around three thousand years ago were not the sole ancestors of the present-day inhabitants. In parallel, I initiated a project to survey the diversity of the world’s present-day populations, using a microchip for analyzing human variation that my collaborators and I designed specifically for the purpose of studying the human past. We used the chip to study more than ten thousand individuals from more than a thousand populations worldwide—a dataset that has become a mainstay of studies of human variation not just in my laboratory but also in other laboratories around the world.”

Reich believes that we now have the data and toolsets to achieve a “... detailed reconstruction of deep relationships among ancient human populations.”

In spite of the fact that he relies on archeologists to identify significant features of the skeletal remains he analyses, such as: was this person a farmer, a nomad, a warrior; how old is this skeleton; where precisely was it located; were there artifacts buried with it that would identify societal relationships; etc.etc., Reich asserts that: “These days, human genome variation has **surpassed** the traditional toolkit of archaeology—the study of the artifacts left behind by past societies—in what it can reveal of changes in human populations in the deep past.” I’m sure it does, but...

Reich finishes this “Introduction” in a more traditional way: “This book is about the genome revolution in the study of the human past. This revolution consists of the avalanche of discoveries based on data taken from the whole genome—meaning, the entire genome analyzed at once instead of just small stretches of it such as mitochondrial DNA. The revolution has been made far more powerful by the new technologies for extracting whole genomes’ worth of DNA from ancient humans.” Since 2009, “... whole-genome data have begun to challenge long-held views in archaeology, history, anthropology, and even linguistics—and to resolve controversies in those fields.”

“The ancient DNA revolution is rapidly disrupting our assumptions about the past. Yet there is at present no book by a working geneticist that lays out the impact of the new science and explains how it can be used to establish compelling new facts.”

But the field is in its infancy, and rapidly evolving. Since publication of this book, new findings have arisen which challenge some of those in this book. So a caveat is in order: “I hope that readers will take the topics I discuss as examples of the disruptive power of whole-genome studies, not as a **definitive** summary of the state of the science.” I agree and I’m glad he said this.

From this point to the end of the “Introduction”, Reich more or less gives us a cook’s tour of the book. I want only to highlight a few of his statements which I think are relevant and which we may come back to in subsequent chapters.

“I try to make a virtue of my laboratory’s central role in the ancient DNA revolution by telling the story of my own work where it is relevant—as this is a subject on which I can speak with great authority—while also discussing work in which I was not involved when it is critical to the story. Because I take this approach, the book disproportionately highlights the work from my laboratory.”

OLLI 497: Ancient DNA

“I also highlight some of the great themes that are emerging, especially the finding that **mixture between highly differentiated populations** is a recurrent process in the human past.”

And this “mixture” upsets the notion of “primeval groups” corresponding to “races”: “... this long-held view about “race” has just in the last few years been proven wrong—and the critique of concepts of race that the new data provide is very different from the classic one that has been developed by anthropologists over the last hundred years. A great surprise that emerges from the genome revolution is that in the relatively recent past, human populations were just as different from each other as they are today, but that the fault lines across populations were almost unrecognizably different from today. DNA extracted from remains of people who lived, say, ten thousand years ago shows that the structure of human populations at that time was qualitatively different. **Present-day populations are blends of past populations, which were blends themselves.**”

“Chapter 1, “How the Genome Explains Who We Are,” argues that the genome revolution has taught us about who we are as humans not by revealing the distinctive features of our biology compared to other animals but by uncovering the **history of migrations and population mixtures** that formed us.”

“The sequencing of the Denisovan genome unleashed a cavalcade of discoveries of additional archaic populations and mixtures, and demonstrated unequivocally that **population mixture** is central to human nature.”

And so, for this book, “... population **mixture** as a unifying theme.”

“Ancient DNA and the genome revolution can now answer a previously unresolvable question about the deep past: the question of **what happened**—how ancient peoples related to each other and how migrations contributed to the changes evident in the archaeological record.”

Stayed tuned. We will be in for an interesting and “moving” experience!