A genetic perspective on human origins

Matthew Williams and João Teixeira (University of Adelaide, Australia)

Who are we? Where did we come from? Why are we here? These fundamental questions have been widespread throughout human history, shared across different cultures from distant epochs and geographical locations. The search has been as much a philosophical as an empirical one, capturing the imagination of the philosopher, the theologian, the artist and the scientist alike. Hence, the quest for unveiling our origins is probably as old as humanity itself. From a scientific point of view, which we address in the present article, the question of human origins became deeply intertwined with Charles Darwin’s theory of evolution in the late 19th century. This led to the development of scientific fields such as palaeoanthropology, which analyses fossil remains, stone tools and cultural artefacts to piece together our past. Recently, however, the possibility to assess genetic information from thousands of individuals across the world and, more importantly, to obtain DNA from specimens that lived thousands of years in the past (so-called ancient DNA [aDNA] analyses) is rapidly transforming long-held beliefs about our origins. As such, we have never been in a better position to ask what do our genomes have to tell us about where we came from. Ultimately, however, can they tell us who we are?

Within most scientific disciplines a complex relationship exists between Occam’s razor, which states that the simplest explanation is best, and Bonini’s paradox, elegantly put by the philosopher and poet Paul Valéry: ‘Everything simple is false. Everything complex is unusable.’ This is particularly true for aDNA research. As researchers, we are not simply trying to infer often immeasurably complex demographic events from relatively small amounts of incomplete data, but likewise historical events. Thus, to understand the history of our species, aDNA studies rely on synthesizing population genetics theory and modelling with archaeological and palaeoanthropological fossil record reconstructions.

Of necessity, the genetic models are often overtly simple caricatures of human demography and do not fully capture the complexity of human societal structure through time. Therefore, it is the duty of the researcher, and the broader scientific community, to understand how violations of these models impact the interpretation of genetic results. Moreover, it is of growing importance to effectively communicate that, whilst we reconstruct the past using simplistic genetic models, these models should not be interpreted simplistically.

Models of human evolution

The prevailing view amongst palaeoanthropologists well into the mid-20th century was dominated by an imperialistic interpretation of the fossil record, whereby phylogenetic classifications of human fossil variation followed prevailing racial theories about human populations. Hence, this period was dominated by an evolutionary polygenist view of human origins, which asserted, at the time, that geographically dispersed human populations had separately evolved from different species at different evolutionary times. However, when the concept of human races lost its scientific validity during the second half of the 20th century, evolutionary polygenism was abandoned as a valid scientific hypothesis and replaced by two contrasting (and competing) models of human evolution: multiregional evolution and Out of Africa. The development of these two models has historically been associated with the analysis and interpretation of the fossil record, with a particular focus on geographical variation and continuity (or lack thereof) of morphological traits through time.

Multiregional evolution states that geographical variation observed in the fossil record starting in the early Pleistocene should be interpreted as intraspecific diversity existing within the genus Homo, which defines the human species. The theory builds strongly upon Franz Weidenreich’s model of an intricate network of human evolution across a wide geographical area spanning the Old World. Multiregionalism maintains that while most of the human population lived in Africa throughout the Pleistocene, there was a continuous (in time and space) network of gene flow between human populations that lived in relative isolation from one another. This process allowed a relative, but never complete, isolation of human groups in different parts of the Old World,
Ancient DNA

Figure 1. Human evolution during the late Pleistocene. Before human populations migrated out of Africa and into different parts of the world, Eurasia was occupied by several ‘archaic’ human groups, such as Neanderthals in the West (blue) and Denisovans in the East (red). Genetic studies have shown that an early group of humans moved out of Africa ~100,000 years ago and admixed with Neanderthals; Denisovans and Neanderthals met and interbred in Siberia ~90,000 years ago; an unknown divergent group admixed with Denisovans, probably in East Asia/Siberia. These events are depicted by pink arrows and occurred before the migration of human populations out of Africa ~60–50,000 years ago. The potential routes taken by migrating human populations are shown in black arrows and should be interpreted cautiously, as they remain speculative. Before leaving Africa, human populations probably admixed with an unknown ‘archaic’ human population (purple U circle). Further admixture events occurred soon after humans left Africa: with Neanderthals ~60–50,000 years ago in the Middle East (blue N circle); with another ‘archaic’ group in South Asia (purple U circle); and with Denisovans in East Asia, Island Southeast Asia and the Philippines ~50,000 years ago (red D circles).

whereby humans were a polytypic species. The process of gene flow mediated by continuous migrations across the landscape, mostly from Africa to different parts of Eurasia in a process named centre-and-edge, promoted the spread and fixation of advantageous traits across the entire network, whereby no geographical place for modernity can be easily ascribed.

In sharp contrast, the Out of Africa theory considers that the human species, Homo sapiens, emerged ~250,000 years ago somewhere in eastern/southern Africa, much more recently than proposed by multiregional evolution. A key stance/position of the Out of Africa model is that features of modernity first arose in Africa around this time and then spread across the rest of the world ~60–50,000 years ago through a major expansion of anatomically modern humans into Eurasia. The expansion of modern humans across the planet resulted in the replacement of existing (so-called) archaic forms in Eurasia, such as the Neanderthals of Europe.

**Population genetics and personal ancestry**

The debates regarding modern human origins in the light of these two competing models coincided with the growing availability of genetic information across populations around the world. The advantage of using DNA to trace the history of our species through time is that we know, since the rediscovery of the work of...
Ancient DNA

Gregor Mendel, how genetic information is transmitted from parents to offspring. More importantly, we can also estimate the mutation rate of DNA, i.e., how many errors occur on average in the copying of DNA in each generation. Hence, by contrasting the number of differences between any two DNA copies we can estimate the time to the most recent common ancestor between them.

The first genetic studies tackling the question of human origins made use of mitochondrial DNA (mtDNA), a small fragment of DNA that is present in mitochondria, cellular organelles that function as energy powerhouses for each of the cells in our body. As there are thousands of mitochondria in each cell, mtDNA is highly abundant and easy to obtain. Importantly, each of us inherits mtDNA through our mother, whereby mtDNA is a direct source to unveil the history of our maternal ancestry. In a highly impactful article, published in 1987, Cann, Stoneking and Wilson demonstrated that all the mtDNA diversity found in contemporary human populations emerged from a single mtDNA lineage that arose in Africa ~200,000 years ago. This finding strongly supports the Out of Africa theory, and even an extreme variation of the model, known as Eve theory, as an allusion to Biblical accounts of creation. Importantly, because mtDNA is transmitted only through the maternal side, it contains no information regarding the paternal ancestry of human populations. However, there is another genetic marker, the Y-chromosome, which is only transmitted from father to son and contains the genes responsible for the genetic determination of males. Since the Y-chromosome has an opposite pattern of inheritance to mtDNA, it offers the possibility to investigate the history of our paternal ancestry. Interestingly, the results obtained for the Y-chromosome corroborate a recent African origin for the human species, and similarly strongly favour the Out of Africa theory.

Nonetheless, things get more complicated, as both mtDNA and the Y-chromosome represent a minute fraction of genetic ancestry in each individual human. If we consider only two generations in the past, neither of these markers offers any information on the maternal grandfather nor the paternal grandmother of a given individual and, as a consequence, of none of their ancestors. In fact, the further back we go in time, the more genetic ancestral information we fail to include from our own family tree, at an exponential rate.

So, how do we assess the genetic information passed on for millennia from our vast family tree? The answer is in the nucleus of the cell. As readers of The Biochemist will be aware, humans have 22 pairs of autosomal chromosomes, one of each inherited from each of our parents, plus the sex chromosomes. This nuclear genome contains the overwhelming majority of genes that encode most of the biological functions in our body, from growth to metabolism and immunity. Due to the cut and paste nature of recombination during gametogenesis, this genetic information is inherited from our entire family tree in a mosaic-like fashion. For example, the copy of chromosome 1 that person X inherits from her mother (i.e., the maternal copy, but the same logic applies to the paternal copy) is a completely new assortment of the two copies of chromosome 1 that she had inherited from her own parents (i.e., the grandparents of person X). Indeed, this recombination process was repeated across all the maternal ancestors of person X, such that their maternal chromosome 1 copy captures an exponentially decreasing amount of the chromosome 1 from each of their increasingly distant ancestors.

aDNA and our extended family tree

So, what does the information contained in our nuclear genome say about the origin of our species? The analyses of genomic data across contemporary human populations around the world confirm that the majority of our genetic ancestry traces back to Africa. Importantly, populations currently living outside of Africa are the descendants of a migration of people from Africa ~60–50,000 years ago, and collectively represent only a subset of the genetic variation observed in Africa. Again, these observations favour the Out of Africa model of recent human evolution.

However, a recent revolution in the field of evolutionary biology, beginning in the 21st century, has made it possible to obtain DNA sequences from ancient remains and established the field of aDNA. The availability of ancient genomes has allowed us to uncover previously hidden chapters in the history of our species, namely remarkable population migrations and admixture events. In a groundbreaking article published in 2010, Svante Pääbo and colleagues sequenced the first complete Neanderthal genome. To the astonishment of many palaeoanthropologists, by comparing the
Neanderthal genome to those of contemporary human populations, Pääbo and colleagues showed that Neanderthals had contributed ~2% to the genetic ancestry of present-day human populations living outside of Africa. Together with the observation that the vast majority of our ancestry traces back to Africa, these findings unveiled that human populations who left the continent ~60–50,000 years ago met and interbred with Neanderthals shortly after arriving in Eurasia, probably somewhere in the Middle East. Importantly, this observation directly includes Neanderthals in the collective human family tree. But the incredible findings from Pääbo’s group in 2010 went even further. The team obtained DNA from a finger phalanx uncovered at Denisova Cave, in Siberia, which belonged to a completely different human group unknown to anthropology, that the team named Denisova man. Amazingly, the research found that Denisovans contributed ~4% to the genetic ancestry of Papuan and Aboriginal Australian populations, and also smaller amounts in Asian and New World populations. Further work over the last decade has demonstrated that Neanderthals and Denisovans interbred in Siberia ~90,000 years ago; that Denisovans likely interbred with an even more divergent human group living in Asia; and that potential multiple interbreeding events between Denisovans and humans occurred in Island Southeast Asia, as Out-of-Africa migrants were travelling to New Guinea and Australia ~50,000 years ago. Hence, aDNA is providing incredible insights into recent human population history and, while it is certain that most of our genetic ancestry derives from Africa ~200,000 years ago, a non-negligible multiregional contribution exists from Eurasian archaic groups to the contemporary genetic makeup of human populations (see Figure 1).

**Population genetics and personal ancestry**

Advancements in genetics research have revolutionized our conceptions of the history of our species; introducing us to unknown groups such as the Denisovans and revealing the place of Neanderthals in our family tree. Thus, we can expect that genetic research, coupled with the palaeo and archaeo sciences, will take us far in answering the second of the three questions we posed at the beginning of this article – ‘where did we come from?’ However, the question of human origins is as much personal as it is academic, and answering the remaining two questions: ‘who are we?’ and ‘why are we here?’ – will require more than robust statistical models about human demography. In a time where over 26 million at-home ancestry DNA kits have been sold (as per MIT Technology Review report) many are still left wondering, ‘where does my genetic heritage fit within our broader human history?’ Except for in very specific circumstances, it is unnecessary to use high-powered, sophisticated statistical genetics to answer questions about our immediate ancestry, and a cursory discussion at a family gathering will likely reveal an approximate answer. Rather, our desire to discover our distant genetic heritage (both temporally and geographically), linked with a curiosity to know with whom one shares it with is what ‘consumer genomics’ companies have tapped into. However, when trying to disentangle our deeper ancestries, things become a little unintuitive.

As discussed above, the process of recombination shuffles up chromosomes. Individuals have two copies of each autosomal chromosome but only pass on a single mixed copy to their progeny, leaving one of the homologous genomic regions behind. As a consequence, for each generation in your family tree, the amount of genetic information you inherit from that particular ancestor is halved. In other words, the further back you go the less genetic information any given ancestor directly contributed to you. This brings us to the somewhat surprising fact that there are genealogical ancestors in your family tree who are not your genetic ancestors. You only have to go back on average nine generations to find direct biological ancestors from whom you have inherited no genetic information, as it was lost in the lottery of genetic assortment before it could make it to you.
Alongside this rather paradoxical concept is an equally obscure thought. During the 1st century BCE it is estimated that the global population numbered approximately 300 million people. If we take a generation time of 25 years, there are 80 or so generations between the present day and the 1st century BCE. If we then assume that for someone living today their ancestors were all unique individual people extending back to the 1st century BCE, they would have 1.2 × 10^{15} ancestors; notably that is 4 × 10^{15} more ancestors than there were people alive at that time. This over-representation of ancestors occurs because whilst in theory the number of ancestors for any one individual grows exponentially into the past, we know there is a certain amount of ‘doubling over’ of everyone’s ancestors, meaning that some appear more than once in your family history. Thus, it is a mathematical certainty that we all have ‘double over’ of everyone’s ancestors, meaning that some appear more than once in your family history. Thus, it is a mathematical certainty that we all have some levels of ‘inbreeding’ in our genealogy. The vastness of our genealogical family tree, both temporally and geographically, has led to surprising results suggesting that the most recent genealogical ancestor of all present-day humans lived only a few thousand years ago. Further work by Ralph and Coop in 2013 looking at present-day European genetic ancestry found that within only 1500 years, individuals from across Europe have hundreds of genetic ancestors in common.

So, on these seemingly contradictory paradoxes; the exponential increase in the number of ancestors with the simultaneous reduction in genetic contribution by any single ancestor, emerges one of the most powerful and unifying propositions of population genetics, that “no matter the languages we speak or the colour of our skin, we share (genealogical) ancestors who planted rice on the banks of the Yangtze, who first domesticated horses on the steppes of the Ukraine, who hunted giant sloths in the forests of North and South America, and who laboured to build the Great Pyramid of Khufu” (Rohde, Olson and Chang, 2004).

Further reading

- Graham Coop blog GCbias https://gcbias.org/ (Accessed 26/11/19)