Major transitions in human evolution revisited: A tribute to ancient DNA

Luca Ermini, Clio Der Sarkissian, Eske Willerslev, Ludovic Orlando*

Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 14 January 2014
Accepted 19 June 2014
Available online 19 December 2014

Keywords:
Paleogenomics
Archaic hominins
Neolithic transition
Admixture
Migration
Paleodemography
Selection
Epidemics
Epigenomics
Metagenomics

ABSTRACT

The origin and diversification of modern humans have been characterized by major evolutionary transitions and demographic changes. Patterns of genetic variation within modern populations can help with reconstructing this ~200 thousand year-long population history. However, by combining this information with genomic data from ancient remains, one can now directly access our evolutionary past and reveal our population history in much greater detail. This review outlines the main recent achievements in ancient DNA research and illustrates how the field recently moved from the polymerase chain reaction (PCR) amplification of short mitochondrial fragments to whole-genome sequencing and thereby revisited our own history. Ancient DNA research has revealed the routes that our ancestors took when colonizing the planet, whom they admixed with, how they domesticated plant and animal species, how they genetically responded to changes in lifestyle, and also, which pathogens decimated their populations. These approaches promise to soon solve many pending controversies about our own origins that are indecipherable from modern patterns of genetic variation alone, and therefore provide an extremely powerful toolkit for a new generation of molecular anthropologists.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

The first DNA sequence isolated from the remains of an extinct organism was reported in November 1984 (Higuchi et al., 1984). It derived from the dried muscle of a quagga zebra curated at the Münich Museum since the early 1900s. This kicked off the field of ancient DNA (aDNA), which represents a unique field in evolutionary biology and aims at retrieving genetic information from a broad range of archaeological remains such as bones (Hagelberg et al., 1989), teeth (Drancourt et al., 1998), hair (Gilbert et al., 2004a), coprolites (Poinar et al., 1998) and even from sediments (Willerslev et al., 2003). Despite many controversies (Marota and Rollo, 2002) and a colourful history punctuated by early spectacular claims (i.e., DNA from dinosaurs) known now to have resulted from contamination (Zischler et al., 1995), aDNA research is now celebrating its thirtieth anniversary with promise of a bright future.

In those thirty years, much progress has been made, and nowadays we know that kilobases-long DNA fragments barely survive in fossils, except in the extraordinarily cold conditions of Antarctica where up to 1600 base pairs (bp) of penguin ancient mitochondrial DNA (mtDNA) could be PCR-amplified (Lambert et al., 2002). Ancient DNA generally comes in a bulk of short molecules not longer than 50–100 bp characterized by typical patterns of molecular damage (Pääbo et al., 2004). Hydrolytic attacks severely degrade the DNA backbone (Dabney et al., 2013a) and introduce a series of chemical modifications to nucleotidic bases. The most prominent of such degradations consists of the deamination of Cytosines into Uracils, which leads to characteristic GC/AT nucleotide mis-incorporations (Hofreiter et al., 2001a). Over the years, dedicated analytical procedures, ranging from the best-of-three rule in the early days of cloning/sequencing (Hofreiter et al., 2001b) to the down-scaling of base quality scores in the recent days of next-generation sequencing (Jönsson et al., 2013), have been proposed to eliminate such spurious mutations and reconstruct genuine sequence information. A number of other modifications (Heyn et al., 2010) preclude PCR amplification and limit our ability to manipulate aDNA fragments. As a result, fresh DNA contaminants, which can be preferentially amplified, represent a constant threat to molecular analyses, especially for human archaeological remains where many possible contamination sources exist, from the excavation itself to laboratory reagents (Leonard et al., 2007). Despite this,
recent in silico procedures exploiting the mis-incorporation patterns introduced by post-mortem DNA damage have proved efficient in distinguishing aDNA sequences from modern contaminants. (Skoglund et al., 2014). This new procedure has shown great success in recovering genuine information even from heavily contaminated anthropological material (Meyer et al., 2014).

Ancient DNA research has now come of age with replicable and stringent procedures (Fig. 1), and in 2010, i.e., no later than a decade after the reference human genome was characterized (Lander et al., 2001; Venter et al., 2001), the first ancient human genome sequence was released (Rasmussen et al., 2010). With the massive throughput of so-called next-generation sequencing (NGS) platforms, the complete genomes of at least eight ancient humans have been characterized at 1X to up to 20X coverage (average number of times a genomic position is sequenced from independent templates) (Rasmussen et al., 2010, 2011, 2014; Keller et al., 2012; Raghavan et al., 2013; Olalde et al., 2014; Skoglund et al., 2014). The genome of our known closest relatives, the Neanderthals, has also been characterized (Green et al., 2010; Prüfer et al., 2014) together with that of other archaic hominins, the Denisovans (Reich et al., 2010; Krause et al., 2010a; Meyer et al., 2012). Their quality even competes with that achieved in sequencing the genome of living individuals. Ancient DNA researchers have also gathered genome-wide sequence data of a few more ancient individuals (Sánchez-Quinto et al., 2012; Skoglund et al., 2012; Fu et al., 2013a; Olalde et al., 2014) and many new complete ancient human genomes are expected in the forthcoming months. The recent success in characterizing a first draft of the genome from a 560–780 kyr (thousands of years) old horse (Orlando et al., 2013) and the almost complete mitochondrial genome of Homo heidelbergensis (Meyer et al., 2014) augurs for the genome sequencing of archaic hominins who lived in the Middle Pleistocene (Millar and Lambert, 2013). Besides nuclear genomes, many ancient mtDNA markers have illuminated our understanding of past population migration dynamics (Krause et al., 2010b; Fu et al., 2012; Brotherton et al., 2013), often revealing important features that were difficult to reconstruct from modern genetic data alone.

In this review, we will present how recent achievements in aDNA research have revisited major transitions in human evolution. We will describe when we dispersed out of Africa, the routes we followed to colonize the planet, and whom we met and potentially admixed with. We will also report what pathogens infected us and caused major epidemics, and illustrate the way we genetically responded to such prominent changes in lifestyle.

Recent technological developments in ancient genomics and aDNA research have revisited major transitions in human evolution. Their genomics of archaic hominins: Neanderthals

Archaic hominins such as Homo erectus, and its near relative, Homo ergaster, started migrating out of Africa prior to 1.8 million years ago (Grine et al., 2009) and colonized Eurasia multiple times independently. Although the exact phylogenetic relationships amongst archaic humans are still controversial, it is recognized that they underwent local evolution and adaptation (Grine et al., 2009). Our present knowledge of the biology of archaic hominins mainly derives from those physical features directly observable on fossilized remains, but aDNA can now complete the picture and help to resolve complex evolutionary questions regarding their origins.

Neanderthals were the first archaic hominins ever characterized at the genetic level by the sequencing of short hypervariable mitochondrial fragments (Krings et al., 1997). Seventeen years later, partial hypervariable mitochondrial regions of at least 13 individuals (Krings et al., 1997, 1999; Ovchinnikov et al., 2000; Serre et al., 2004; Lalueza-Fox et al., 2006; Orlando et al., 2006; Krause et al., 2007b; Dalen et al., 2012) and the complete mitochondrial genomes of at least eight individuals (Green et al., 2008; Briggs et al., 2009; Prüfer et al., 2014) have been characterized. A first draft of the nuclear genome was characterized in 2010 at 1.2X-coverage by shotgun sequencing three ca. 40 kyr-old Neanderthal specimens originating from the Vindija cave, Croatia (Green et al., 2010). With shotgun sequencing, DNA library inserts are sequenced at random and a large fraction of the reads generated can originate from microbes that colonize archaeological remains after death. However, the same year, the exome capture technology, which preferentially selects human coding templates from the diversity of DNA library inserts, was applied to a 49 kyr Neanderthal bone and revealed the complete sequence of about 14,000 protein-coding positions (Burbano et al., 2010). More recently, a high-quality genome sequence (ca. 52X coverage) of a Neanderthal woman from the Siberian Altai mountains has been determined together with a low coverage Neanderthal neonate genome from the Caucasus (Prüfer et al., 2014). The availability of such a vast sequence of datasets is hitherto unprecedented for an extinct hominin, and makes Neanderthals the most widely known extinct hominin at the genetic level.

Comparisons of the Neanderthal and modern human genomes revealed four important traits in Neanderthal evolution. First, assuming a genome-wide mutation rate of 0.5 × 10⁻⁶ mutation per site per year (Scally and Durbin, 2012), Neanderthal and modern human populations split ca. 550–765 kyr ago (Green et al., 2010; Prüfer et al., 2014). Such population split times would be halved assuming faster mutation rates, using those calibrated on the 6 million year (myr) divergence between humans and chimpanzees, in agreement with earlier estimates (Green et al., 2010). For consistency, population split times provided in this review are based on the former rate.

Second, although Neanderthals were believed to have mainly spread into Europe and the Levant, their genetic presence was
found as far East as the Okladnikov and Denisova caves in the Altai mountains, Western Siberia (Krause et al., 2007b; Prüfer et al., 2014). How much Neanderthal populations were structured over their full geographic range is still contentious, although the existence of distinct groups has been suggested based on mitochondrial DNA data (Briggs et al., 2009).

Third, despite their wide geographic range, Neanderthals appear to have lived in small groups of extremely limited sizes. This feature was first suggested by evidence based on relaxed purifying selection of a subset of mitochondrial genes (Green et al., 2008; Briggs et al., 2009) and has been recently confirmed by Pairwise Markov Sequential Chain (PSMC) analyses of a high-quality diploid genome sequence (Prüfer et al., 2014). As a whole, the effective size of Neanderthal populations is estimated to represent about one-tenth that of modern humans.

Fourth, Neanderthals showed patrilocal (Lalueza-Fox et al., 2011) and/or intra-familial mating behaviours (Prüfer et al., 2014). The inbreeding coefficient calculated from the high-quality Neanderthal genome from the Altai mountains (Prüfer et al., 2014) is indeed compatible with inter-marriage between half-siblings, uncle and niece, grand-father and grand-daughter, or double first cousins, who share both sets of grandparents.

The characterization of complete Neanderthal genomes has also allowed the compilation of a genome-wide catalogue of the genetic differences between archaic hominins and modern humans. Of particular importance are genomic regions where Neanderthals show ancestral alleles (i.e., chimpanzee-like) while most modern humans carry derived alleles at the same loci, significantly enriched for non-synonymous mutations. Such cases might represent regions of our genome that made us human by acquiring very early in our evolutionary history advantageous mutations that further increased in frequency until their (near-) fixation. One such region includes the gene RUNX2, encoding for a transcription factor essential for osteoblastic differentiation and endochondral ossification (Mundlos et al., 1997). Allelic variants in this gene are associated with cleidocranial dysplasia, a hereditary congenital disorder, characterized by cranial malformations (i.e., protruding frontal bone), bell-shaped rib cages, under-developed or absent collarbones and abnormal dentition, which altogether, represent major anatomical differences between Neanderthal and modern humans.

Derived mutations around exon 7 of the FOXP2 transcription factor are fixed in modern humans and were first hypothesized as potential drivers for our unique language capabilities (Enard et al., 2002). However, since these derived mutations were also found in Neanderthals (Krause et al., 2007a; Green et al., 2010), they were probably inherited from a common ancestor of both species (Maricic et al., 2013). Therefore, it has been proposed that the mutations around exon 7 could be not adaptive or that the cognitive features potentially driven by the derived FOXP2 allele were shared between Neanderthals and modern humans.

Recent target enrichment experiments have revealed a position in intron 8 located within a 14-bp long binding motif of the transcription factor POU3F2 (Maricic et al., 2013), where a unique and nearly-fixed mutation is found in modern humans. This position is highly conserved in all tetrapods, including three Neanderthal individuals from El-Sidrón (Maricic et al., 2013), and both chromosomes from a Denisovan individual (Meyer et al., 2012). Functional assays revealed reduced POU3F2 binding for the modern human allele, suggesting a lower FOXP2 expression in modern humans. Further functional work is still required before the respective role of FOXP2 and POU3F2 in mediating the appearance of language and speech in the human lineage can be fully understood.

The catalogue of genetic differences between archaic and modern human groups extends way beyond the genes listed above, and currently includes 87 proteins, and a handful of miRNAs and segmental gene duplications (Prüfer et al., 2014). We are only beginning to understand their respective functional consequences and we will need to develop functional assays similar to those used for FOXP2 (Maricic et al., 2013), MC1R (which revealed light-skin pigmentation mutations segregating at low frequencies in Neanderthal populations) (Lalueza-Fox et al., 2007), and mir-1304 (a miRNA gene involved in enamel formation that is proposed to represent one essential factor for the difference in the dental morphology of modern humans and Neanderthals) (Lopez-Valenzuela et al., 2012).

The genomics of archaic hominins: Denisovans

A distal manual phalanx of a juvenile individual from the Denisova cave revealed the presence of a yet genetically unknown archaic hominid population. The mtDNA sequence retrieved from the bone indeed showed no genetic affinity with Neanderthals and anatomically modern humans (AMHs) (Krause et al., 2010a), and depicted a new mitochondrial lineage (the so-called Denisovan branch), which diverged before the Neanderthal/modern human split about one million years ago. This new branch was therefore proposed to correspond to the descent of the H. erectus lineage. The nuclear genome, which was characterized soon after at 1.9X coverage, revealed a completely different picture, where Denisovans and Neanderthals are sister groups (Reich et al., 2010) (Fig. 2). The discovery of an upper molar (Reich et al., 2010) from a young Denisovan adult provided some information about Denisovan dental morphological characteristics, which differed in several aspects from those of Neanderthals, but could match in size teeth from H. erectus and Homo habilis.

The development of a highly sensitive DNA sequencing library building method, together with the exceptional endogenous DNA content of the Denisovan finger bone, allowed the characterization of a high-quality genome sequence (Meyer et al., 2012). Branch shortening analyses, by which the phylogenetic distance separating ancient individuals from the chimpanzee outgroup is shorter than those separating the latter and modern humans, revealed that the Denisovan individual lived 50–110 kyr ago, depending on how molecular clocks are calibrated (Meyer et al., 2012; Prüfer et al., 2014). Denisovan and Neanderthal populations most likely split 381–473 kyr ago (Prüfer et al., 2014).

The Denisovan genome does not show particularly long homozygosity tracts, which contrasts with the inbreeding detected in the Altai Neanderthal excavated from the same cave. The PSMC demographic reconstructions revealed a remarkably low effective population size (Meyer et al., 2012), with a heterozygosity level about one-fifth that of present-day African populations.

Comparative genomics revealed 260 non-synonymous mutations that are (nearly)-fixed amongst modern humans (Meyer et al., 2012). Interestingly, those mutations occurring at the 23 most

![Figure 2. Mitochondrial and nuclear genomes depict conflicting phylogenetic histories in human evolution.](image-url)
conserved positions in primates preferentially affect genes associated with brain function or nervous system development, axonal and dendritic growth, synaptic transmission, autism and susceptibility to language disorders. Some of those mutations occurred in a total of 34 genes associated with human diseases, particularly with skin and eye disorders. Another mutation occurred in EVc2, a gene involved in Ellis-van Creveld syndrome, which is notably characterized by wider dental pulp cavities and fusion of tooth roots (Kamal et al., 2013). Such changes in EVc2 might explain some of the major dental differences that characterize hominin evolution (Meyer et al., 2012).

The first nearly complete mitochondrial genome sequence from a 400,000-year-old hominin was recently characterized (Meyer et al., 2014). The hominin was excavated at Sima de los Huesos cave site in Spain, where different skeletons bearing morphological features characterizing the pre-Neanderthal H. heidelbergensis lineage were found. This followed a methodological tour-de-force, which combined DNA extraction methods tailored to ultrashort DNA fragments (Dabney et al., 2013b), target enrichment methods (Maricic et al., 2010; Fu et al., 2012) and in silico approaches filtering for contamination originating from modern humans. Surprisingly, phylogenetic analyses revealed the archaic hominin as a sister group to Denisovans and not Neanderthals. Different scenarios have been proposed to explain this unexpected result. First, recalling the mitochondrial and nuclear placement of Denisovans in the hominin phylogeny, the mitochondrial tree might not reflect the true population history (Orlando, 2014), a phenomenon known as incomplete lineage sorting. The close proximity between Denisovans and H. heidelbergensis might then be misleading. Alternatively, it could reflect the introgression of a mitochondrial lineage from another, yet undetermined, archaic hominin population in any, or both, archaic lineages, or even the introgression of a mitochondrial haplotype related to H. heidelbergensis in the Denisovan branch. Testing which scenarios are more plausible will require genome-wide sequence information from the nuclear genome, which appears challenging but feasible given recent advances in paleogenomics and the recent characterization of the nuclear genome of an even older horse (Orlando et al., 2013).

Out of Africa

According to the so-called multiregional model, which is mostly based on paleo-anthropological data rather than genetic evidence, a local transition from H. erectus to AMHs would have occurred in different regions of the Old World, where regional populations slowly evolved into modern humans. Some level of gene flow maintained a relative genetic homogeneity across the whole geographic range, although it also allowed the preservation of some patterns of regional morphological differentiation (Ash and Robinson, 2011). The vast majority of paleoanthropological and genetic data currently available supports, however, a different model for our origins. In this model, AMHs would have originated in Africa and later dispersed in Eurasia and Oceania, finally replacing pre-existing archaic hominin populations descending from earlier migrations (Jobling et al., 2013). The date for the AMH expansion out of Africa has been estimated to about 50–62 kyr ago, starting from a small population of effective size equivalent to ~1000 individuals (Liu et al., 2006). The tempo and mode of dispersal in Eurasia and Oceania is however still controversial with different models competing. The single dispersal model supports a unique dispersion into Eurasia about 50 kyr ago (HUGO Pan-Asian SNP Consortium, 2009) followed by a series of founder events and separate migrations into Asia (55–40 kyr ago) and Europe (40–25 kyr ago). According to this model, a further expansion from Asia into Australia would have given rise to the ancestors of Aboriginal Australians, as early as 50–40 kyr ago (O’Connell and Allen, 2004; Summerhayes et al., 2010). The single dispersal hypothesis is mainly supported by mtDNA and Y-chromosomal studies showing, for each marker, that descendants of all non-African lineages come from a single ancestor who lived 55–75 kyr ago. Phylogenetic studies based on mtDNA indicated that almost all African mtDNA lineages belong to the L clades, while all non-African lineages (clades M, N, and R) arose from one of its sub-clades called L3. It is thus likely that L3 appeared in Africa before the out of Africa single wave, while M, N, and R, emerged and diversified outside Africa (Behar et al., 2012). The Y-chromosome data show a similar phylogeographic pattern (Karafet et al., 2008; Wei et al., 2013).

In contrast to the single dispersal model, the multiple dispersal hypothesis (Cavalli-Sforza and Piazza, 1994) assumes separate successive migrations from Africa to the rest of the world, with Australia being populated first by an early (60–50 kyr ago) and possibly independent out of Africa dispersal. The first complete genome of an ancient Aboriginal Australian was sequenced in 2011 (Rasmussen et al., 2011) and provided new insights into the debate on human out of Africa dispersal. The Aboriginal Australian genome was sequenced from a hair tuft that was voluntarily donated to the famous ethnologist Dr A.C. Had- don when exploring Australia in the early 1900s. It showed that the ancestors of Aboriginal Australians likely started spreading into eastern Asia about 62–75 kyr ago, admixing with Denisovans (see below), and arriving in Australia about 50 kyr ago. The ances- tors of present-day mainland Asians would have then followed a second and independent wave of expansion taking place ca. 25–38 kyr ago. These findings are in line with previous reports dating the dispersal of the first humans at about 50–40 kyr ago (O’Connell and Allen, 2004; Summerhayes et al., 2010; Reich et al., 2011).

Interestingly, a recent study found evidence of archaic introgression in the genome of three African populations (pygmies from Cameroon, and Khoisan-speaking Hadza and Sandawe from Tanzania), suggesting that ancestors of the African hunter-gatherers admixed with one or more archaic human populations (Lachance et al., 2012). This echoes the findings from Pickrell et al. (2014), who found significant genome-wide signatures of a West Eurasian ancestry in modern hunter-gatherer populations from southern Africa. Altogether, this suggests that a recent migration of people from West Eurasia into Ethiopia occurred around 3.0 kyr ago, thereby bringing archaic (e.g., Neanderthal-like) alleles into Africa, which could have dispersed 1.5 kyr later into southern Africa.

Population admixture: Neanderthals

The question of a possible admixture between AMHs and archaic hominins on their way out of Africa has nurtured fascinating anthropological debates for the last 150 years. The characterization of complete genomes from archaic hominins has now provided strong evidence for such admixture events. The Nean- derthal draft genome was first found to share more derived alleles with Eurasian than with African populations, suggesting that 1–4% of the non-African genomes derived from Neandertals (Green et al., 2010). The high-quality Neanderthal genome recently characterized refined this estimate to ~2%. Alternative models of population history have been proposed, where this closer genetic proximity did not reflect recent admixture but the presence of an older population structure within Africa, with AMHs and Nean- derthals dispersing from the same population background (Eriksson and Manica, 2012). However, analyses based on the rate
of decay of Linkage Disequilibrium further supported the admixture scenario by estimating that the last gene flow between Eurasians and Neanderthals occurred outside of Africa 37–86 kyr ago (most likely 47–65 kyr ago) (Sankararaman et al., 2012). Neanderthals appear to have contributed more DNA to modern East Asians than to modern Europeans (Meyer et al., 2012; Wall et al., 2013). As a consequence, simple population models where Neanderthals and AMHs admixed when they cohabited in the Levant before the latter colonized Asia, Oceania and Europe, should probably be dismissed and more complex models preferred, where additional gene flow from Neanderthals into East Asians took place after they diverged from Europeans. It is possible that a minimum of two admixture events occurred between Neanderthals and the ancestors of Europeans and East Asians (Vernot and Akey, 2014). In addition, admixture signals between Eurasian and Neanderthals were found to be stronger in Mezmaiskaya Neanderthals from the Caucasus than in Croatian and Siberian Neanderthals (Prüfer et al., 2014), suggesting some apparent structure within the Neanderthal populations.

Interestingly, some of the alleles that we inherited from Neanderthals were advantageous and likely helped our immune system to fight the new range of pathogens that we encountered on our way out of Africa (see below). However, some of the archaic alleles probably also caused male hybrid sterility when transferred into an AMH genetic background (Sankararaman et al., 2014), and have been puriﬁed from our genome. As the strength of natural selection is inversely related to the effective population size (Meyer et al., 2014), and Asian populations have experienced a demographic bottleneck (Keinan et al., 2007), the resulting reduced efﬁciency of purifying selection could at least partly explain the presence of a higher proportion of Neanderthal ancestry in Asian populations (Sankararaman et al., 2014).

Demographic modelling of admixture combined with explicit spatial expansion simulations indicated that the observed levels of Neanderthal ancestry in the genomes of modern Europeans are only compatible with very rare successful admixture occurrences, representing at best a few cases every century occurring along the whole hybridization area (Currat and Excoffier, 2011). A similar conclusions was achieved by Neves and Serva (2012) using a stochastic simulation model. This suggests that interbreeding was highly unsuccessful, and whether it was due to active pre-zygotic isolation (e.g., inter-marriage avoidance), or to the action of negative selection on incompatible alleles in hybrids, or to a combination of both still remains to be determined.

Population admixture: Denisovans

Contrary to Neanderthals, Denisovans showed no evidence of admixture with most present-day Eurasian populations, with the exception of Melanesians (Reich et al., 2010, 2011) and a small contribution into mainland Asian and Native American populations (Skoglund and Jakobsson, 2011; Prüfer et al., 2014). The overall level of Denisovan ancestry detected in Melanesian genomes is ca. 4–6%, which added to the ca. 2% of Neanderthal ancestry, makes them the modern human population with the highest levels of archaic hominin ancestry identiﬁed to date. Interbreeding of Denisovans with AMHs originally seemed to have mainly occurred in remote islands of Southeast Asia located east of Wallace’s Line, a natural biogeographic barrier crossed only by rodents and AMHs, among known terrestrial mammals. Therefore, it was proposed that the ﬁrst AMHs crossing the Wallace Line encountered a Denisovian population already established there since the Middle Pleistocene (Cooper and Stringer, 2013). This model has been recently challenged by new genetic evidence showing a contribution of Denisovans to populations not located east of the Wallace’s Line, such as mainland Asians and Native Americans, although 25-fold smaller (Prüfer et al., 2014). Interestingly, the fraction of Denisovan derived polymorphisms present in Melanesian Papuans is smaller for the X chromosome than autosomes. This suggests that either the gene flow preferentially involved Denisovan males and AMH females, or that hybrid incompatibility alleles within the X chromosome were removed by purifying selection. The fact that this pattern is also shared with Neanderthal ancestry is compatible with still untested scenarios, such as the existence of an increased positive selection on the X chromosome in AMHs.

The advantageous contribution of archaic hominins to our genome

Admixture between archaic hominins and AMHs has probably inﬂuenced our capacity of adaptation to our new environment out of Africa. Signatures of adaptive introgression from Neanderthals in the human genome are enriched for genes involved in the makeup of the integumentary system, which comprises the skin, hair and nails (Vernot and Akey, 2014). For example, an adaptive haplotype in the BNC2 gene associated with skin pigmentation (Jacobs et al., 2013) was found to be virtually absent in modern East Asians, but with a frequency of ~70% in Europeans. Conversely, another adaptive haplotype at POU2F3 shows a frequency of ~66% in East Asians and less than 1% in Europeans (Vernot and Akey, 2014). This gene is a homeobox transcription factor regulating keratinocyte proliferation and differentiation, and is primarily expressed in the epidermis (Cabral et al., 2003; Takemoto et al., 2010). More strikingly, some of the HLA class I alleles of the human Major Histocompatibility Complex, such as the HLA-B*73 allele, segregating in modern western Asian populations, might have been acquired from Denisovans (Abi-Rached et al., 2011). Two recent studies also showed that two genes involved in our immune system, STAT2 and OAS1, also carry signatures of archaic introgression (Mendez et al., 2012a, b), with the Eurasian STAT2 allele, absent in sub-Saharan Africans, closely matching the Neanderthal haplotype and an OAS1 polymorphism, unique to eastern Indonesians and Melanesians, matching the Denisovian allele (Mendez et al., 2012a). Therefore, it seems plausible that some of the genetic information that we inherited from archaic hominins facilitated our expansion out of Africa, acclimatizing us to an environment to which archaic hominins had already become adapted hundreds of thousand years before us.

Gene ﬂow did not only concern AMHs, but also occurred between archaic hominin populations. The gene ﬂow between Neanderthals to Denisovians was likely asymmetric with at least 0.5% of the Denisovan genome contributed from Neanderthals (Prüfer et al., 2014). Perhaps even more surprising, gene ﬂow did not necessarily involve a known archaic hominin population either. As a matter of fact, present-day African genomes show an excess of shared alleles with Neanderthals as compared with Denisovians (Prüfer et al., 2014). This is unexpected in the absence of external gene ﬂow because the ancestors of both Neanderthals and Deni-
started colonizing the American continent (Swisher et al., 1994; Gabunia et al., 2000). The time and routes followed by the first humans colonizing the New World is one of the most debated topics in archaeology (Barton et al., 2004). The most prominent archaeological model suggests that humans associated with the Clovis culture 11.5–10.9 $^{14}$C kyr ago were the first inhabitants of the Americas, as no human osteological remains have been directly dated prior to Clovis (Adovasio and Pedler, 2004). This dogma was however challenged by the discovery of pre-Clovis sites with evidence of human occupation (Meltzer, 1991) and by aDNA evidence from human coprolites found at Paisley Caves in south-central Oregon (USA) and radiocarbon dated to approximately 12.3 kyr ago (Gilbert et al., 2008a). Despite many criticisms (Goldberg et al., 2009; Poinar et al., 2009), these aDNA findings have been confirmed some years later by Jenkins et al. (2012), who first demonstrated the validity of the chronostratigraphic model at Paisley Cave and replicated the successful characterization of human DNA from pre-Clovis coprolites. In addition, further aDNA analyses made on the tip of a bone projectile point implanted in a rib of a single disarticulated mastodon unearthed at the Manis site (Wisconsin, USA) and radiocarbon dated to about 11.9 kyr ago, brought additional evidence that humans already hunted proboscideans before Clovis (Waters et al., 2011). Altogether, this supports a pre-Clovis colonization of the New World (Jenkins et al., 2012).

The genome sequence of a 24 kyr old individual from the Mal’ta site in south-central Siberia (Russian Federation) revealed new insights on the origins of Native Americans (Raghavan et al., 2013). The Mal’ta mitochondrial genome belonged to haplogroup U, which is found at high frequency in present-day Europeans (Richards et al., 2000) and in Upper Paleolithic and Mesolithic European hunter-gatherers (Bramanti et al., 2009; Malmström et al., 2009; Krause et al., 2010b: Der Sarkissian et al., 2013). The Mal’ta Y chromosome haplotype instead corresponded to a basal lineage of haplogroup R, widespread in modern-day western Eurasians and sister lineage to the haplogroup Q, the most common haplogroup in Native Americans. These western Eurasian affinities of the Mal’ta specimen were then confirmed by tree-based analyses of population splits and admixture tests based on autosomal single-nucleotide polymorphisms (SNPs), which also supported gene flow from the Mal’ta specimen to all Native Americans. This gene flow likely occurred prior to the diversification of Native American populations in the New World, and could represent as much as 14–38% of the Native American ancestry. It also appears to have occurred after the divergence of the ancestors of Native Americans and eastern Asians, as no particular eastern Asian signature was identified in the Mal’ta specimen. This contrasts with the significant eastern Asian genetic component found amongst all present-day Native Americans (Schurr and Sherry, 2004) and suggests that different population backgrounds may have participated in the early colonization of the New World. Whether this occurred as a single wave with mixed ancestries or different successive waves still remains to be determined. For now, the western Eurasian ancestry signature found in Mal’ta suggests that the European component present in modern-day Native Americans does not only originate from post-Columbian admixture, but may have derived from a mixed population ancestry of the first Americans, which included a European contribution. It also provides valid explanation for the presence of mtDNA haplogroup X among both Europeans and Native Americans, and challenges extreme migration scenarios such as the Solutrean Theory, which, based on stone tool characteristics, hypothesized that Clovis people originated from a migration across the Atlantic of Pleistocene Europeans (Bradley and Stanford, 2004).

The complete genome of a child (Anzick-1) buried approximately 12.6 kyr ago with ochre-covered Clovis artifacts at the Anzick site (Montana, USA) revealed additional insights into the origins of Native Americans and provided further support for the pre-Clovis colonization of the Americas (Rasmussen et al., 2014). The genome of the Anzick-1 individual showed a significantly greater genetic affinity with Native American than with any extant Eurasian population, and amongst American populations, Central and South Native Americans (SA) were found to be more closely related to the Anzick-1 individual than to northern Native Americans (NA). This suggests that the North American and South American lineages diverged early, with Anzick-1 belonging to the NA lineage and potentially ancestral to all SA groups. Alternatively, Anzick-1 would be ancestral to all Native Americans, but further gene flow into the NA lineage following a later admixture from a more basal lineage would have artificially inflated its apparent genetic affinity with SA groups.

Either way, the Anzick-1 individual belonged to a population who used Clovis tools and was closely related to all indigenous American populations. Since Anzick-1 and the above Mal’ta individual shared the same relative affinity to western and eastern Eurasians, the gene flow from the Siberian Upper Paleolithic Mal’ta population into Native American populations occurred before 12.6 kyr ago.

The northern extremes of Canada and Greenland were not colonized until 4.5 kyr ago with the so-called Early Paleo-Eskimo Independence I-Saqqaq culture (McGhee, 1996). The Independence II-Dorset culture succeeded this early Paleo-Eskimo culture and lasted until the Neo-Eskimo Thule culture arose some 2000 years ago. Present-day Yupik and Inupiat from Alaska and Inuit from Canada and Greenland represent the descent of Thule people (Gilbert et al., 2008b). Genetic analyses of hair remains from a Saqqaq individual who lived about 4000 years ago along the southwestern coast of Greenland revealed a genetic discontinuity in the peopling of the New World Arctic. First, the complete sequence of the mitochondrial genome was found to belong to an extinct haplotype within haplogroup D2a1, which is common among modern Aleuts and Siberian Sireniku Yuits but not found in Greenlandic Inuit (Gilbert et al., 2008b). The high-quality nuclear genome sequence that was further characterized refined the population affinities by showing a much closer similarity to some contemporary northeast Siberia modern populations (Chukchis and Koryak) than with present-day Greenlanders. Population divergence time calculations showed that Saqqaq ancestors moved into Greenland some 5.5 kyr ago, well before the arrival of Inuit that later replaced the Saqqaq population.

**We, megafauna killers?**

The human exodus from Africa and our colonization of the Americas and Australia, coincided with a period of climate change and downfall of large-bodied mammals (called megafauna) (Haynes, 2008). For instance, at the end of the Late Quaternary (~12 kyr ago), about one-third of large-bodied mammals became extinct in Eurasia and about two-thirds in North America (Barnosky et al., 2004). The arrival of humans to new geographical regions certainly had a profound impact on local ecosystems, and the presence of new human groups has long been proposed as a trigger for the demise of large herbivore populations and their predators, possibly even for their total extinction (Barnosky et al., 2004; Johnson, 2009).

These mass extinctions were neither synchronous nor universal, and paleontologists and archaeologists have debated the causes of these extinctions for decades with different human-driven extinction scenarios being proposed, including, among many others, the overkill of large mammals by human hunters and the spread of new diseases perhaps introduced by new human
immigrants (paleodisease hypothesis) (Martin and Klein, 1989; Surovell and Waguespack, 2009). According to the former hypothesis, upon their arrivals into the Americas, AMHs encountered a large number of species, which had never been in contact with humans and did not recognize them as a threat. Highly skilled human hunters therefore would have had no difficulty in rapidly exterminating large proportions of species. Because North American archaeological evidence suggested that the extinction might have been very rapid, this overkill model was further refined into a blitzkrieg or rapid overkill scenario (Martin, 1973). Note that the Americas is not an exception and that everywhere AMHs encountered naïve animals, mass extinction has followed, except in Africa where megafauna and hominids co-evolved for millions of years and probably mutually adapted to gradually increasing hunting skills. In the paleodisease hypothesis, anthroponoses introduced by new arriving humans are also proposed as the trigger to extinction (Diamond, 1989). This model, however, lacks epidemiological support as no known disease can eradicate completely unrelated genera (Lyons et al., 2004).

In addition to humans, climate and the major changes occurring during the Late Quaternary have been proposed as prominent factors driving massive extinctions (Airoy, 2001). In particular, radiocarbon dates of Late Pleistocene fossils from Alaska and the Yukon Territory indicated that a number of megafauna populations expanded at the time of human arrival in these regions (Guthrie, 2006).

Over the years, ancient DNA studies have been carried out to disentangle the respective roles of climate and humans on megafauna extinctions. The overall methodology mostly consisted of inferring a time for major population declines and testing whether this time better fit with data on paleoe climatic changes or human arrival. For example, in bison, which lived throughout Beringia during the Late Pleistocene (125.0–11.7 kyr ago), the genetic diversity was found to be larger than in present-day American populations (Shapiro et al., 2004). This pattern was proposed to result from a global population expansion over the last 100 kyr, before a first demographic collapse occurred immediately prior to the period corresponding to the Last Glacial Maximum (26.5–19.0 kyr ago). The estimated time of the transition from population expansion to decline suggested climate change as the main driver of the bison population collapse (Shapiro et al., 2004). A second collapse was further detected at the time humans settled in Alaska (Drummond et al., 2005), suggesting an additional potential role of humans in shaping the demographic trajectory of bison populations. More recent analyses, based on climate niche modelling, have shown however that human impact became only significant following the European introduction of firearms in the last three centuries (Metcalf et al., 2014). The population dynamics of the Pleistocene musk ox also showed several expansions and contractions over the past 60 kyr, mainly due to major climatic changes rather than to anthropogenic causes (Campos et al., 2010).

Lorenzen et al. (2011) generalized aDNA-based demographic reconstructions to four additional megafauna species, the extinct woolly rhinoceros, mammoths, horses and reindeer. In order to assess the respective impacts of climate and humans, megafauna demographic histories were compared with the potential distribution range of the megafauna species (as estimated using climate niche modelling), and to their geographical overlap with humans (as estimated using the dates of faunal remains and sites occupied by humans). Results showed that neither humans nor climate alone affected the megafaunal extinctions. Instead, each species responded differently to the impact of climate change, habitat redistribution and human colonization. Although humans appear to have played no part in the extinction of the woolly rhino or the musk ox in Eurasia, a combination of anthropogenic and climatic effects appears to be responsible for the extinction of others mammals, including Eurasian steppe bison and wild horses. Reindeer were relatively unaffected by human impact and climate change, while the causes of the extinction of mammoths was left unresolved.

In a follow-up study, the authors obtained ancient plant and animal DNA directly from more than 200 permafrost samples around the Arctic covering the past 50 kyr and from stomach contents and faeces of Pleistocene megafauna species. Comparing the vegetation history with megafauna diet, it appeared that the disappearance of forbs during the Pleistocene/Holocene transition was devastating for many big-bodied mammals (Willerslev et al., 2014), which undermines the role of humans as drivers for megafauna extinctions. This was also the case for woolly mammoths, the demography of which was reconstructed over the last 200 kyr (Palkopoulou et al., 2013). Woolly mammoth populations experienced bottlenecks during warmer climatic conditions, to which they were not adapted, followed by expansions in population size and geographical range when the climate became cooler and drier. Woolly mammoths did not survive the last warming period initiated at the end of the Pleistocene around 15 kyr ago, which led to a fatal population size drop and some 10 kyr later to extinction. Future research aiming at unravelling the impact of our ancestors and climate change (Larson et al., 2013) on the extinctions will most likely rely on ecological niche models and demographic reconstructions based on ancient unlinked SNP markers, which provide stronger statistical power for inferring the magnitude and the timing of past population changes (Mourier et al., 2012).

**Change in lifestyle: domestication**

For most of their history, humans relied on a foraging economy based on hunting, fishing and gathering (Bellwood et al., 2007). Starting from ~12 kyr ago, some human groups began to establish sedentary settlements (de Laet, 1994), and to gradually domesticate wild plants and animals (Childe, 1950). Such major changes have been of utmost importance in the history of humans and are part of the cultural and technological transition called the ‘Neolithic’ (Ammerman and Cavalli-Sforza, 1984; Bellwood et al., 2007). Genetic studies suggested that this transition from hunting to stock-raising occurred more often and in more places than those advocated by traditional archaeological evidence (Zeder et al., 2006). Ancient DNA has provided direct evidence for the geographical origin of domesticates, their modes of domestication, and the genetic variants that humans selected to propagate traits that best served their interests. Ancient DNA has also revealed the limits of our understanding of the domestication process when based on modern genetic information alone. For example, in chickens, two genes supposedly selected in the early stages of domestication and currently showing no variation were shown to be extremely variable 500 years ago (Girdland Flink et al., 2014), suggesting that modern alleles only reflected the recent history of intensive selective breeding.

Pig domestication appears to have occurred throughout separate processes acting on differentiated subspecies of European and Asian wild boar, which were either domesticated or at least incorporated into domestic stocks descending from regionally differentiated wild populations (Larson et al., 2005; Wu et al., 2007; Tanaka et al., 2008; Ottoni et al., 2013). European domestic pigs were proposed to have been introduced first by early farmers from the Near East before they were completely replaced by autochthonous domesticated pigs (Larson et al., 2007a). Early European farmers may have thus exchanged domesticated animals, as suggested by aDNA evidence from the pre-Neolithic Ertebølle culture of South Scandinavia, where people, who primarily hunted for survival, likely obtained domestic pigs by trading, exchange, or by hunting. 
and capturing escaped animals from neighbouring Neolithic communities (Krause-Kyora et al., 2013). In Asia, pigs appear to have been domesticated in multiple centres in China, India and in peninsular Southeast Asia (Larson et al., 2010). Domesticated pigs were then brought by humans while migrating south and east to New Guinea, and eventually reached the remote Pacific as far east as Hawaii (Larson et al., 2010). The genetic signatures, mainly mitochondrial, of ancient domesticated (and also commensal) animals have been utilized as a direct proxy to reconstruct the timing and routes of ancient human dispersal. For example, this evidence supported an Austronesian human dispersal route connecting Southeast Asia to Oceania (Larson et al., 2007b, 2010), in agreement with studies based on human mtDNA and Y-chromosome diversity (Su et al., 2000; Hage and Marck, 2003; Kayser et al., 2006).

In addition to pigs, ancient domesticated horses have been analyzed for aDNA. Modern domestic horses carry large mtDNA diversity, but almost no Y-chromosomal variation (Jansen et al., 2002; Lindgren et al., 2004). This pattern suggested that horse domestication mostly involved the reproduction of a limited number of stallions and constant mare restocking from different population backgrounds. This potentially reflects differences in the management of mares and stallions during domestication, with females perhaps easier to handle for propagating livestock (Lira et al., 2010; Warmuth et al., 2012). As a result, a large fraction of the genetic diversity that used to be present in wild horse populations (up to 73% according to Lippold et al., 2011) has been successfully incorporated into the domestic pool through mares. The persistence of a haplotype now absent from the modern pool of Y-chromosomes in a 2800 year-old Scythian horse suggested that this sex bias in favour of females was not extreme in the early stages of domestication some ~5.5 kyr ago (Outram et al., 2009), but only intensified recently. Although no mitochondrial haplogroup could be associated with a particular geographic population or domesticated breed in horses, haplogroup D1 was proposed to reflect the outcome of a local domestication in Iberia given its relatively elevated frequencies in Iberian breeds and North-African Barb horses (Jansen et al., 2002). However, aDNA revealed its virtual absence from the Iberian Neolithic and Bronze Age (Lira et al., 2010). It therefore most likely reflects the descent of horse populations that were introduced into Iberia post-Bronze Age. Some mtDNA haplotypes that populated Iberia in the Bronze Age are still occasionally found in Iberian breeds (Lira et al., 2010) and could either reflect local domestication episodes and/or restocking from wild mares, therefore leaving open the question of an Iberian domestication center for horses.

Through selective breeding, our ancestors probably targeted easily selectable phenotypic traits, such as the coat colour in horses. The genetic identification of spotted horses in the Late Pleistocene whose phenotype was maintained for at least 17 kyr (Bellone et al., 2013) has suggested that the famous cave paintings of the ‘dappled horses of Pech-Merle’ could be simple representations of living animals, and not necessarily symbolic expressions (Pruvost et al., 2011). In contrast to present-day domesticated horses, pre-domestic horses were characterized by reduced coat colour variability, mainly with bay, black, and leopard complex spotting (Pruvost et al., 2011). During and following the Iron Age in Europe (~2–3 kyr ago), horses were selectively bred in part for their variability in coat colour and pattern, which resulted in the explosion of colour patterns observed in horses today (Ludwig et al., 2009).

The use of aDNA has proven essential for understanding other complex domestication processes, such as those of cattle and dogs. A recent aDNA analysis (Zhang et al., 2013) shows the presence of early attempts to domesticate taurine cattle in China ca. 10.5 kyr ago. This occurred independently from the almost contemporary domestication from the Near East and from the two millennia younger domestication of the zebu in Asia (Meadow, 1993). For dogs, a recent aDNA analysis demonstrated that European hunter-gatherers started to tame wolves more than 18 kyr ago, before they gradually evolved into dogs (Thalmann et al., 2013). Such an old date makes wolves the first wild animals ever tamed. Further examples include studies on barley (Palmer et al., 2009), grapes (Cappellini et al., 2010), maize (Jaenick-Després et al., 2003), chicken (Storey et al., 2012), sheep (Niemi et al., 2013) and goats (Fernandez et al., 2006).

It is becoming evident that studies aimed at unravelling domestication processes and pinpointing wild ancestors of domestic species need to rely, where possible, on DNA recovered from ancient specimens (Larson and Burger, 2013). Future analyses based on full genome sequencing will reveal which genetic pathways were selected to transform wild organisms into the diversity of domesticated forms that we know today, but also to time when, and in which cultural contexts, particular traits arose.

The Neolithic transition and the gene pool of Europeans

Assessing the impact of the Neolithic transition on the European gene pool is critical for understanding the genetic history of Europeans. According to the proponents of the ‘demic diffusion’ model, the Neolithic transition was accompanied by significant migrations of early farmers from the Near East (Ammerman and Cavalli-Sforza, 1984; Boyle and Renfrew, 2000). This model suggests a genetic discontinuity between pre-Neolithic and Neolithic European populations, as well as a significant Near-Eastern Neolithic ancestry in present-day Europeans. Conversely, the ‘cultural diffusion’ model proposes that ideas and technologies, rather than humans, spread into Europe (e.g., Whittle, 1996), and implies a genetic continuity between hunter-gatherers and early farmers in Europe. Note that the ‘demic diffusion’ and ‘cultural diffusion’ models have crystallized most of the debate as both result in very distinct predictions for European ancestry. These models represent two extremes of a long continuum, and intermediate scenarios have been proposed in order to account for the archaeological record.

The ‘demic diffusion’ and ‘cultural diffusion’ models have been directly tested using ancient human mitochondrial sequences (De Benedetto et al., 2000; Caramelli et al., 2003; Chandler et al., 2005; Haak et al., 2005; Sampietro et al., 2007; Bramanti et al., 2009; Malmstrom et al., 2009; Krause et al., 2010b; Deguilloux et al., 2011; Gamba et al., 2012; Hervella et al., 2012; Bollongino et al., 2013; Brandt et al., 2013; Brotheron et al., 2013; Der Sarkissian et al., 2013; Fu et al., 2013b) and Y-chromosome markers (Haak et al., 2010; Lacan et al., 2011a,b) from a variety of European remains. Genome-wide data were also obtained for nine hunter-gatherers from Spain (~7 kyr old; Sánchez-Quinto et al., 2012; Olalde et al., 2014) and Sweden (~4.1–7.5 kyr ago; Skoglund et al., 2012, 2014), and five early farmers from Italy and Sweden (~4.8–5.3 kyr ago) (Skoglund et al., 2012). This knowledge of the ancient genetic diversity in Europe allows the circumvention of potential biases associated with pre- and post-Neolithic migrations from the Near East and errors associated with dating the most recent common ancestor of the major groups of mitochondrial/Y-chromosome sequences, also called ‘haplogroups’.

Mitochondrial data recovered from hunter-gatherers of the Paleolithic/Mesolithic from Spain to Western Russia over a period ranging from ~4.1 to 30.0 kyr ago (Bramanti et al., 2009; Malmstrom et al., 2009; Krause et al., 2010b; Hervella et al., 2012; Sánchez-Quinto et al., 2012; Der Sarkissian et al., 2013) suggested a relative homogeneity in the mitochondrial gene-pool of hunter-gatherers across Europe, with haplogroup U as the dominant haplogroup. This is consistent with results of coalescent time
estimates for haplogroup U, which was described as the oldest haplogroup in Europe based on modern data (Richards et al., 2000). At the genomic level, hunter-gatherers from Spain (Sánchez-Quinto et al., 2012; Olalde et al., 2014) and Scandinavia (Skoglund et al., 2012, 2014) were found to carry similar genomic signatures, thus giving support to a common genetic background of foraging populations across pre-Neolithic Europe. European hunter-gatherers also displayed a lower diversity of mitochondrial haplogroups than present-day Europeans (e.g., Richards et al., 2000) and no particular genetic affinity with present-day European populations at the genome level (Sánchez-Quinto et al., 2012; Skoglund et al., 2012, 2014; Olalde et al., 2014). In addition, differences in phenotype-associated variants were observed between pre-Neolithic, Neolithic and present-day Europeans. This is the case, for example, at the SLC24A5 locus (rs1426654), where two hunter-gatherers from Spain (Olalde et al., 2014) and Sweden (Skoglund et al., 2014) showed an ancestral variant, whereas the corresponding derived allele associated with light pigmentation was found in two prehistoric farmers of Italy (Keller et al., 2012) and Sweden (Skoglund et al., 2014), as well as in almost all Europeans today. All together these results suggest a genetic discontinuity due to post-Mesolithic migrations and/or adaptations. Genetic data from post-Mesolithic prehistoric individuals has revealed the Neolithic as the main event responsible for the genetic discontinuity observed between hunter-gatherers and present-day populations. Early Neolithic farmers of the Linearbandkeramik culture (LBK) in Germany (~7.5 kyr ago) were indeed found to be genetically different from European hunter-gatherers (Bramanti et al., 2009; Haak et al., 2010; Brandt et al., 2013; Fu et al., 2013b). Patterns of interactions between hunter-gatherers and farmers at the onset of the Neolithic appear, however, complex and spatially heterogeneous. Mitochondrial and genomic data showed that, despite ~4 kyr of parallel existence in South Scandinavia, little admixture occurred between farmers of the Funnel Beaker culture and local hunter-gatherers of the Pitted-Ware culture who had preserved their foraging lifestyle after the arrival of the Neolithic (Malmström et al., 2009; Skoglund et al., 2012). Whereas Scandinavian hunter-gatherers showed no sign of genetic introgression from agricultural populations, local contemporary early farmers showed evidence of admixture with hunting-gathering groups, which probably had occurred earlier during the migration associated with the spread of the Neolithic to Scandinavia (Skoglund et al., 2014). In Late Neolithic Germany (~5–6 kyr ago), local genetic continuity was observed between Mesolithic hunter-gatherers and a Neolithic group who had retained a foraging diet even 2 kyr after the onset of the Neolithic in this area. However, a co-existing group with an agriculturalist diet showed admixture with Mesolithic hunter-gatherers (Bollongino et al., 2013). Genetic signatures from the Near East, Anatolia and the Caucasus were detected in Early Neolithic LBK farmers, as exemplified by the presence of the Y-chromosome haplogroup G2a and the relatively high occurrence of the mitochondrial haplogroup N1a (Haak et al., 2010). The Y-chromosome haplogroup G2 was found in Early Neolithic individuals from Spain (~7 kyr ago) (Lacan et al., 2011b), as well as in later Neolithic individuals from Southern France (~5 kyr ago) (Lacan et al., 2011a) and Northern Italy (Keller et al., 2012). Mitochondrial haplogroup N1a was also identified in other Neolithic individuals from Hungary (Haak et al., 2005) and France (Deguilloux et al., 2011), but was not detected in Early Neolithic groups of the Iberian Peninsula (Chandler et al., 2005; Sampietro et al., 2007; Lacan et al., 2011b; Gamba et al., 2012; Hervella et al., 2012) for which an independent Neolithic route from the Near East was proposed. The genetic affinity between Early Neolithic individuals and present-day populations of the Near East provides evidence for a model involving substantial demic diffusion. It also reveals a supplemental genetic discontinuity between Early Neolithic people and today’s Europeans, suggesting subsequent population movements. Genome-wide data from the Late Neolithic Alpine mummy popularly known as the Tyrolean Iceman (~5 kyr ago), showed genetic affinities with contemporary Funnel Beaker Culture hunter-gatherers of Scandinavia (Skoglund et al., 2012, 2014), suggesting that the genetic component characterizing those Scandinavian groups was geographically widespread at the time. This genetic signature appears to have persisted to some extent, as it was found in an Iron Age individual from Bulgaria (2.8 kyr ago) and in present-day Sardinians (Sikora et al., 2014). This unexpected link between Neolithic farmers and present-day Sardinians was explained by a significant gene flow into the Sardinian gene pool associated with the spread of agriculture into Europe, followed by relative genetic isolation (Sikora et al., 2014). In contrast, mainland Europeans underwent a different population history after the arrival of the Neolithic.

The 2500 years following the establishment of the Neolithic in Germany were characterized by relative genetic continuity throughout the LBK, Rössen, Schönengen, Baalberg and Salzmünde cultures of the Early and Middle Neolithic (Brandt et al., 2013). Middle Neolithic cultural groups of Germany (Brandt et al., 2013) showed mitochondrial affinities with other Middle Neolithic groups of the Funnel Beaker culture from South Scandinavia (~5 kyr ago) (Malmström et al., 2009; Skoglund et al., 2012) and from France (~5 kyr ago) (Lacan et al., 2011a). Later, the genetic shift associated with the emergence of the Middle Neolithic Bernburg culture (~3 kyr ago) was characterized by an increase in typical hunter-gatherer mitochondrial lineages in Germany. It was proposed to reflect the input from South Scandinavian groups that had retained hunter-gatherer lineages thousands of years after the arrival of farming in the region (Brandt et al., 2013).

In the Late Neolithic, the emergence of the Corded Ware culture in Germany (~4.8 kyr ago) was also accompanied by genetic changes, following in particular the introduction of the mitochondrial lineages I and U2 from an eastern European origin (Brandt et al., 2013). The Corded Ware culture co-existed in Germany with the Bell Beaker culture (BBC), which is thought to have left a substantial legacy on the mitochondrial gene pool of central Europeans, as genetic continuity between BBC individuals and present-day Europeans could not have been rejected (Brandt et al., 2013). The BBC diffusion from the Iberian Peninsula to the rest of Europe is thought to be responsible for the spread of most mitochondrial haplogroup H lineages (Brandt et al., 2013; Brotherton et al., 2013). Haplogroup H was found in Mesolithic populations of Iberia (Hervella et al., 2012) and today is the most common European haplogroup with frequencies greater than 40% (Brotherton et al., 2013). As a support for the Iberian origin of the BBC established on the basis of the archaeological record, the closest maternal relatives to the BBC group in Germany were found in Early Neolithic Northern Spain (Hervella et al., 2012) and Portugal (Chandler et al., 2005). In these groups, despite signs of introgression of Near East mitochondrial lineages, the contribution of hunter-gatherer lineages appears to be relevant, suggesting more genetic continuity in these populations than in Central Europe, and thus a different process for Neolithization. On the basis of the mitochondrial data available from the temporal transect in Germany, the relative contributions of the different prehistoric cultures to the present-day gene pool of Europeans was estimated to be 16% for the Paleolithic/Mesolithic, 31.2% for the Early/Middle Neolithic, 5.8% for the Late Neolithic, leaving 47% for other lineages of undetermined origin (Brandt et al., 2013).

The heterochronous dataset available for prehistoric populations in Germany allowed the detection of genetic changes with
a rather high resolution, but admittedly reflects only the maternal genetic history of Europeans. Y-chromosome and/or genomic data will be necessary to reconstruct the full population history of Europeans and to detect discrepancies between male and female contributions. A wider geographical and temporal sampling is also still necessary. New aDNA data from Europe could for instance help to investigate the structure of hunter-gatherer populations or to shed light on the back migration of BBC carriers into the Iberian Peninsula, which may be the genetic ‘sink’ of the north-eastern European component detected in the present-day Spanish populations (Patterson et al., 2012). New aDNA data could also help determining when and where the ‘Ancient North Eurasian’ component of hypothetically North East Eurasian origin detected in present-day Europeans, came to contribute to the European gene pool (Patterson et al., 2012; Lipson et al., 2013).

Impacts of Neolithic-associated dietary shifts

The Neolithic transition has reshaped the genetic makeup of Europeans, but also introduced important physiological and metabolic changes. The Neolithic lifestyle transition from hunting-gathering to farming was, for instance, accompanied by a shift to a carbohydrate-rich diet (Braidwood et al., 1961; Oelze et al., 2012). Dairy farming and the life span capability to digest lactose (the main sugar present in milk) provided a selective advantage to Europeans carrying the dominant allele T located 13,910 bp upstream from the lactase gene (Enattah et al., 2002). While this allele is nearly fixed in some present-day European groups (Ingram et al., 2009), it was not as frequent within our Neolithic ancestors. It appears indeed to be absent from nine Neolithic individuals from Central Europe (~7.5 kyr ago) (Burger et al., 2007) and 26 Late Neolithic individuals from Southern France (~5.0 kyr ago) (Lacan et al., 2011a), and to be present at low frequencies amongst ten Middle Neolithic individuals from Scandinavia (~4.5 kyr ago) (Malmstrom et al., 2010). The frequency of the T allele further increased and reached 27% in the Late Neolithic in the Iberian Peninsula (~4.5–5.0 kyr ago) (Plantinga et al., 2012), evidence of the strong selective advantage conferred by this allele (Simoons, 1970; Tishkoff et al., 2007). The distribution of lactase persistence in Europe may have been aided further by the dispersal of Neolithic groups such as those associated with the LBK culture (Itan et al., 2009).

Ancient DNA revealed that the Neolithic dietary shift has not solely introduced changes to our genomes but also to our microbiome (Adler et al., 2013). The microbial diversity present in the dental plaque of six Mesolithic hunter-gatherers and six Neolithic farmers could be characterized and compared in terms of taxa and their relative abundances through 16S ribosomal DNA ‘mini-barcodes’ (very short genetic sequences employed as genetic markers to identify species) deep-sequencing. Clostridiales and non-pathogenic oral bacteria of the Ruminococcaceae family were found to be predictive of hunter-gatherers, whereas decay-associated Veillonellaceae were found to characterize Neolithic farmers. In addition, Neolithic farmers showed more taxa associated with periodontal diseases, in line with the skeletal evidence showing an increase of periodontal diseases associated with carbohydrate-rich diet in Neolithic groups (Keri, 1998). Further comparison with four Bronze Age, 18 medieval and 1413 present-day individuals revealed that another important change in our oral microbiome probably occurred with the advent of the Industrial Revolution in the nineteenth century. This change represents the second most important nutritional shift in Europe after the Neolithic transition, with the post-industrialization diet characterized by high concentration of beet or cane sugars fermentable by oral bacteria responsible also for caries (Adler et al., 2013).

What has aDNA told us about human pathogens?

A significant number of infectious diseases have emerged and evolved as a consequence of the Neolithic transition, following the settlement of ever denser human groups and the close proximity with domesticated animals (Dobson and Carper, 1996; Diamond, 1997). Our recent history is also rich in massive epidemics and pandemics, such as the one caused by the 1917 Spanish flu, which killed more people than the First World War (Johnson and Mueller, 2002). Important information regarding such episodes can be tracked in historical records, but most often, with not enough details to enable a precise identification of their etiologic agents, since most records predate the development of modern medicine. Yet, genetic traces from pathogens can be identified in human archaeological remains and aDNA has successfully identified the agents responsible for the flu (Taubenberger et al., 1997; Reid et al., 1999, 2000, 2002, 2004; Basler et al., 2001), the plague (Morelli et al., 2010; Schuenemann et al., 2011; Bos et al., 2011, 2012; Harbeck et al., 2013), leprosy (Taylor et al., 1999; Economou et al., 2013; Schuenemann et al., 2013), smallpox (Biagini et al., 2012), tuberculosis (Saló et al., 1994; Taylor et al., 1999, 2005, 2007; Bouwman et al., 2012; Müller et al., 2013) and cholera (Devault et al., 2014). The validity of early results has been largely questioned due to a lack of appropriate controls and high risks of contamination by microbes from the depository, storage, and/or laboratory environments (Gilbert et al., 2004b; Willerslev and Cooper, 2005; Achtman, 2008; Wilbur et al., 2009). With more rigorous controls and recent advances in high-throughput sequencing, aDNA researchers have now gone beyond the sole molecular diagnostic of the presence/absence of a given pathogen, and nearly complete genomes of bacterial pathogens from historical specimens (1.5 kyr—150 years ago) have been reconstructed (Bos et al., 2011; Schuenemann et al., 2013). Skeletal remains and museum collections of historical pathological specimens represent valuable resources for investigating the genetic basis of pathogenicity and virulence, and the evolutionary history and phylogeography of pathogens. The potentiality of archived specimens was demonstrated in a recent study where the draft genome of a Vibrio cholerae strain responsible for the 1849 cholera outbreak in Philadelphia was reconstructed from a victim’s intestine preserved in the Mütter Museum collections (Philadelphia, USA) (Devault et al., 2014).

Historical plague epidemics provide good examples of what aDNA can achieve for understanding past outbreaks. First, genetic analyses identified Yersinia pestis as the bacterial organism responsible for two of the most deadly plague epidemics in history, namely the Justinian Plague, which struck Europe 542–740 Anno Domini (A.D.), and the Black Death, which decimated 30–50% of the European population between 1347 and 1351 A.D. (Benedictow, 2004). This ended a long controversy about possible etiologic agents of these tragic epidemics (Scott and Duncan, 2001; Cohn, 2003).

The complete sequence of the bacterial chromosome, the pMT1 and pPCP1 plasmids of Y. pestis was characterized from remains excavated in a graveyard dedicated to victims of the Black Death in London (Bos et al., 2011; Schuenemann et al., 2011). Sequence comparison with modern strains currently infecting human populations revealed no non-synonymous changes in known virulence-associated genes, leaving no satisfactory explanation for the extreme pathogenicity of the Black Death. The Black Death genome was even found identical to two modern Y. pestis strains across 636 SNP positions. Even though full genome information is not yet available for the latter two strains, this finding suggests that genetic variants highly similar to the Black Death are still segregating in Y. pestis populations. The authors proposed that external factors, such as the genetic background of the host population, but
also climatic and social conditions, could have been responsible for the severity of the Black Death. Alternatively, the latter might have been due to major chromosomal rearrangements and/or the presence of extra plasmids/operons potentially missed in the ancient strain where the genetic information was retrieved following target enrichment against the genetic background of a modern bacterial strain. The extremely low amounts of Y. pestis traces preserved in DNA extracts were indeed incompatible with shotgun sequencing, making the target enrichment approach a prerequisite and precluding de novo assembly (even though in silico algorithms are emerging for scaffolding ancient DNA contigs and recovering insights about the evolution of the genome structure; Rajaraman et al., 2013).

Phylogenetic trees for Y. pestis were first constructed on the subset of strains for which full genome information was available (Bos et al., 2011) but were further re-evaluated using genome-wide SNP information characterized for 289 additional strains (Bos et al., 2012). In such analyses, the Black Death strain fell at the root of all Y. pestis strains commonly associated with human infection (branches 1 and 2). Using the age of the Medieval samples for tip-calibration, the most recent common ancestor for all pathogenic strains was estimated to have emerged around 1282–1343 A.D. (Bos et al., 2011). A small group of strains present in China today appears to have branched off from the previous cluster right before the emergence of the Black Death. Those strains likely represent the descent of an early Asian diversification episode that gave rise to the Black Death agent and further pathogenic strains.

Recent aDNA studies on SNP data (Harbeck et al., 2013) and Y. pestis genomes (Wagner et al., 2014) recovered from Justinian Plague victims indicated that the etiological strain responsible for that pandemic was different from the one causing the Black Death. The strain involved in the Justinian Plague was therefore found to have no modern descendants, suggesting extinction of the strain or a still insufficient dataset of modern data.

Conclusion

Ancient DNA has provided invaluable information about the origins of humans, their past migrations, their impact on animals and plants, their metabolic adaptations, and their diseases. Given recent methodological developments, we can confidently predict that this is only the beginning. Not even a decade ago, no technology could sequence a complete genome from an ancient human (branches 1 and 2). Using the age of the Medieval samples for tip-calibration, the most recent common ancestor for all pathogenic strains was estimated to have emerged around 1282–1343 A.D. (Bos et al., 2011). A small group of strains present in China today appears to have branched off from the previous cluster right before the emergence of the Black Death. Those strains likely represent the descent of an early Asian diversification episode that gave rise to the Black Death agent and further pathogenic strains.

Recent aDNA studies on SNP data (Harbeck et al., 2013) and Y. pestis genomes (Wagner et al., 2014) recovered from Justinian Plague victims indicated that the etiological strain responsible for that pandemic was different from the one causing the Black Death. The strain involved in the Justinian Plague was therefore found to have no modern descendants, suggesting extinction of the strain or a still insufficient dataset of modern data.

Acknowledgements

This work was supported by the Danish National Research Foundation (DNRF94), a Marie-Curie Intra-European Fellowships IEF (302617) supporting LE, a Marie-Curie Career Integration Grant CIG-293845 and a Danish National Research Foundation FNU grant attributed to LO. Authors declare no financial interests.

References


