BIO Web of Conferences **4**, 00005 (2015)
DOI: 10.1051/bioconf/20150400005
© Owned by the authors, published by EDP Sciences, 2015

Tracing the origin of our species through palaeogenomics

Eva-Maria Geigl, E. Andrew Bennett, and Thierry Grange

Institut Jacques Monod, UMR 7592 CNRS and University Paris Diderot, Paris, France

Abstract. The recent breathtaking progress in whole genome sequencing technology allows access to the genomes both of ancient organisms and populations, including those now extinct. Despite the heavy degradation and the extremely low quantities of ancient DNA, it is sometimes possible to sequence an entire genome from a fossil. This enterprise has been successful in the case of fossilized remains from Neanderthals, a lineage of hominids that lived in Europe for 200,000 years and disappeared 30,000 years ago. An even greater surprise was the genome that has been obtained from a small finger bone preserved in a cave in the Siberian Altai Mountains. This genome revealed the existence of a human lineage previously unknown from the fossil record. The corresponding population mixed with the Neanderthals and the ancestors of the present day populations of South-East Asia. These hybridization events left different traces in the non-African human populations emphasizing the fact that we are genomic mosaics. The comparison of the different genomes also gives hints to how the genome of present-day populations was shaped and helps us to better understand which parts of our genetic make-up are responsible for the biological features of *H. sapiens*.

1. Introduction

The last years have seen a huge advance in the field of palaeoanthropology through the deciphering of the genetic code of extinct human populations, such as the Neanderthals who disappeared about 30,000 years ago. These successes were the result of major breakthroughs in sequencing technologies during the early 2000s that allowed the rapid sequencing of whole genomes. Next generation sequencing generates massive amounts of sequence data. The informatics revolution made the development of bioinformatic tools for the analysis of these vast amounts of data possible. The team of Svante Pääbo (Max-Planck Institute, Leipzig, Germany) united both excellent scientists and a huge research budget to construct first a low-coverage draft of the Neanderthal genome from three individuals from Croatia [1] and then decipher the genomes at a higher coverage than many modern genomes of a young girl and a Neanderthal woman, both found in a cave in the Altai mountains in Southern Siberia known as Denisova cave [2, 3]. Only very small bones were available, a toe of the Neanderthal woman and the base of a finger phalanx of the Denisova girl. The genome of this girl revealed that she belonged to a so far unknown lineage contemporary to the Neanderthals, now called the Denisovans. Currently there is no skeleton associated to this lineage. Only three teeth from the same cave, the morphology of which is closer to an *H. erectus* or Australopithecines tooth than to a Neanderthal or *H. sapiens* tooth, were

This is an Open Access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

found to share the same genetic signature [4]. Her genome, however, has already delivered substantial information about this mysterious population's phylogeny, geographical range, and contribution to the present-day human populations, as well as about some aspects of its phenotype and metabolism.

The sequences of the ancient genomes are a rich information source that is being analyzed to shed light on the recent evolution of hominin genomes and in particular that of modern humans. Before archaic human genomes were available; our view of the genome evolution of our species was limited to the comparison with the genome of our closest living relative, the chimpanzee. These ancient genomes allow us to fill some gaps in the picture that we have of our evolution since the divergence from our last common ancestors with the chimpanzees. They give access to the intermediate steps of the evolutionary branch that led to present-day humans, which opens up completely new horizons. The comparison between the genomes of Neanderthals, a lineage that is one of the last to have split from that which evolved into *H. sapiens*, and of present-day and ancient modern humans give us an insight into the changes that became fixed in the human genome over the last 300,000 to 600,000 years, since the most recent common ancestor (MRCA) of both lineages. Some of those changes are likely responsible for our unique biological features.

The first Neanderthal genome that had been sequenced was the mitochondrial genome. The sequences were obtained from five different bones originating from Spain, Germany, Croatia and the Caucasus [5]. Their analysis showed that around 49 to 38,000 years ago, i.e., about 10,000 years before their extinction, the Neanderthal populations had a small effective population size, which means that only few individuals reproduced [5]. Moreover, the mitochondrial genomes of two bones from the same archaeological site (Vindija) that had lived at a 6,000 years interval were identical, indicating a continuity of the maternal lineage in this geographical area. Even at a larger scale, the comparison of the mitochondrial genomes of Neanderthals in El Sidron in Spain, Feldhofer in Germany and Vindija in Croatia, showed only few differences despite their geographical distance and did not reveal a significant population structure [5]. Finally, the sequence of the mitochondrial genome of the Neanderthal child from the Caucasian cave of Mezmaiskaya, dated from 60,000 to 70,000 years ago, reveals a more pronounced difference, suggesting that either the Neanderthal populations went through a bottleneck towards the end of their existence or that some populations went extinct locally [5].

The quality of the two nuclear genomes of the Denisova cave, the one of the young Denisovan and the one of the Altai-Neanderthal, was such that it was possible to obtain almost a complete catalog of all the point mutations that distinguish Neanderthals, Denisovans and modern humans.

2. Methodological approaches

2.1 The sequencing of the European Neanderthal genome

The most recent progress of massively parallel sequencing technologies allows the deciphering of entire genomes of complex organisms in a very short time. This stands in sharp contrast to the sequencing of the first human genome through the international scientific community that was achieved in 2000 after 15 years of work and at a cost of 3 billion dollars. While by now sequencing of whole genomes has become a routine exercise, this is still not the case for ancient genomes, and sequencing ancient genomes is still a "tour de force". The DNA molecules are fragmented into very small pieces of only 30–100 base pairs (bp) on average, with larger molecules becoming very rare, and show chemically transformed nucleotide bases which are enriched at the ends of the molecules. These transformed bases are characteristic of ancient DNA. They change the coding information and are called miscoding lesions. Only tiny quantities of these ancient fragmented DNA molecules "survive" long time periods in calcified fossilizing tissue.

The sequencing of the genomes of the aforementioned Neanderthal bones was performed with the Illumina sequencing method producing several tens of millions of reads of approximately 100 bp in

length. These sequences had then to be mapped to the phylogenetically closest genome available, in this case the one of present-day humans, since a *de novo* assemblage of an ancient genome based on these short molecules is not possible. The sequences may contain mistakes due to the damages that occur in DNA over time, therefore a higher coverage is required. Ideally each base should be sequenced about fifty times on average for accurate calling of most single nucleotide polymorphisms (SNPs), or single sequence differences between the two genomes [6]. The mapping of very small Neanderthal DNA fragments to the human genome is based on the assumption that the chromosomes of the two lineages shared a similar organizational structure. This strategy allows for the detection of SNPs, and of regions with varying copy number (copy number variation: CNV) between the two genomes, but not the detection of potential chromosomal rearrangements that would change the order of genes on the chromosomes and that could be responsible for changes in the regulation of gene expression.

The fact that only traces of DNA molecules "survive" microbial and chemical degradation over time increases the risk that either the bones and/or the extracts are contaminated with modern human DNA. Even though the initial amount of contaminating modern DNA may be low, it may still outnumber these trace amounts of endogenous DNA, and would be incorporated in the DNA libraries, amplified and sequenced. This is one of the major problems of ancient DNA research. Therefore, before even analyzing in detail the Neanderthal genome sequences, the level of contamination with modern DNA had to be evaluated. To do so, the authors have analyzed three independent genetic markers, i.e., the mitochondrial genome, which had been shown previously to lie outside of the diversity of H. sapiens [5], some Y chromosome markers (knowing that the analyzed Neanderthal bones belonged to females, Y chromosome sequences must be contaminants), and some autosomal markers known to be specific for Neanderthals, thus allowing discrimination with potentially contaminating *H. sapiens* DNA. This analysis yielded a contamination level of less than 1% and confirmed the authenticity and reliability of the genome sequences. Thus, their analysis ensured robust interpretations.

Another difficulty had to yet to be tackled. The traces of short aDNA molecules are diluted with a much larger amount of environmental DNA, mostly of microbial provenance. This environmental DNA constitutes the vast majority of the DNA extracted from most fossils. Thus, if DNA is preserved in Middle or Upper Pleistocene or Holocene fossils, it exists only as a tiny fraction of the DNA that can be extracted from them, generally below 1 %. This means that on one hand one has to screen a large number of fossils to find those few in which DNA is well preserved and on the other hand, that one has to develop enrichment methods to increase the fraction of ancient endogenous DNA. Enrichment also reduces sequencing costs since it reduces the proportion of the sequences that have to be discarded. This research is underway and some promising solutions have been proposed [7–10]. The first enrichment method was based on the hybridization of small synthetic DNA molecules that "capture" the complementary DNA strands from the fossil extracts from the Neanderthal bones of the caves of Vindija and El Sidron [7]. This gave solid insights into the coding regions of the two specimens.

2.2 The sequencing of the Hominins of the Denisova cave

A tiny fragment of a pinky bone from the Denisova cave in the Altai Mountains in Siberia, whose DNA was exceptionally well preserved, yielded a mitochondrial genome that was different from both Neanderthals and modern humans, and which had diverged from the human mitochondrial lineage about one million years ago [11]. Further methodological progress allowed for the sequencing of the whole genome of the pinky bone with a 31 fold coverage [2]. This improvement was based on the observation that DNA degradation over time occurs mainly through depurination followed by spontaneous hydrolysis of abasic sites. As a consequence the ends of the DNA molecules are single-stranded. Moreover, a considerable portion of the molecules has a phosphate on its 3′ end, which hampers the construction of libraries for sequencing. In the case of aDNA this leads to both the loss of DNA molecules with this particular ends when not repaired, and the loss of a portion of the

single-stranded ends. The new approach brings about a solution to both problems in that it eliminates the phosphate groups at the end of the molecules and constructs the libraries with single-stranded DNA molecules. This procedure first denatures double-stranded DNA molecules, dephosphorylates them, ligates adapters with a chemical modification at one of the ends and immobilizes the single-stranded molecules on magnetic beads coupled to a protein that recognizes this modification. These immobilized molecules are then converted into double-stranded molecules, a new adaptor is added to the other end, and the molecules are amplified with a DNA polymerase using primers that correspond to the ligated adapters. This method increases significantly the yield of ancient DNA amenable to sequencing in the majority of ancient samples [12]. It was used to sequence the whole genome of the Denisovan from 10 mg of the pinky bone and to cover 99.9% of all nucleotide positions at least once, and 99.4% at least six times [2]. The 31-fold coverage of the 1.86 gigabases (10^9 nucleotides) of non-duplicated sequences aligned to the human genome is exceptional for an ancient genome. This was possible because about 70% of the total DNA purified from the bone corresponded to the ancient hominin. This rare and spectacular value is due to both exceptional endogenous DNA preservation and to low environmental DNA content. The cause for such an exceptional preservation cannot be explained at present. Indeed, no other fossil of this age outside of a permafrost environment has so far been sequenced that contained only 30% environmental DNA. This exceptional preservation and the efficiency of the new method has resulted in a high quality genome that can be considered as authentic and reliable. In 2013 another bone from the same cave has been sequenced using the same single-stranded method and a 52-fold coverage of the genome was achieved. This was from a toe bone of a Neanderthal with a similar exceptional DNA preservation as the Denisovan finger bone. Thus, the Denisova cave appears to provide a suitable environment for exceptional DNA preservation and low environmental DNA contamination.

3. Insights into the evolution of the genomes of modern humans and into the relationships between Homo sapiens – Neanderthals – Denisovans

3.1 Estimation of the divergence date of lineages and populations

First of all, the sequence of the Neanderthal genome has been used to appreciate the global extent of sequence differences across the whole genome in order to date the most recent common ancestor of the two populations. The dating approach is based on both the molecular clock hypothesis, i.e., the approximation that sequence evolution proceeds at constant speed over time, but also on the date of separation of another, more ancient branch that has been estimated independently. These models have serious limits since the dates can fluctuate depending on the references used and on the genomic regions considered. More so, the older the date of separation of the reference branch, the less reliable are the models used. The models have been recently refined based on the intergenerational mutation rate that has been determined through the direct sequencing of several genomes of parents and their children [13, 14]. These studies have shown that the dates vary by a factor of up to two. Thus, the calculated divergence dates should be taken with caution. The number of mutations specific for each lineage has been measured by comparing the genomes of the Neanderthal and the Denisovan with the genomes of present-day humans, chimpanzees and the modeled common ancestor of humans and chimps. Based on the estimation of a divergence date between humans and chimpanzees of 6.6 million years ago, and also integrating the intergenerational mutation rate, one obtains large windows: the divergence between the human lineage and the one of the ancestors of Neanderthals and Denisovans would have taken place 300,000 to 750,000 years ago. The Neanderthal and Denisovan lineages would have separated more recently, about 200,000 to 470,000 years ago [3].

When the authors compared the number of substitutions that would have occurred between the common ancestor of Homo and Pan and the human genome on one side and the Denisovan genome

on the other, 0.81% fewer substitutions were found in the Denisovan and 1.02% in the Neanderthal genome [3]. This difference was interpreted as the difference in age of the analyzed bones since the associated individuals had less time to accumulate the mutations than the present-day individuals. In the phylogenetic reconstruction, the branch of the Neanderthal lineage was 20% shorter than the one of the Denisovan lineage, suggesting that the phalanx of the foot is older than the phalanx of the hand, which is in agreement with the stratigraphy of the cave. The dates that are proposed depend on the quality of the ancient and modern genomes. The uncertainties linked to both the estimation of the number of substitutions that occurred on the lineage leading to present-day humans and the mutation rate per generation do not enable us yet to determine the absolute age on the two phalanges of the Denisova cave. In the future, however, it might be possible to infer from its genomic sequence the absolute age of a fossil in the exceptional cases where the DNA is very well preserved.

The exceptional quality of the two Altai genomes made it possible to measure the heterozygosity, i.e., the fraction of nucleotide sites that are different between the maternal and the paternal genome of an individual. Different calculations indicate that the heterozygosity of the Denisovan girl was about 0.022% [3]. This presents only a small fraction of the heterozygosity that exists in present-day humans, i.e., about 20% of the heterozygosity of Africans, 26–33% of the heterozygosity of the Eurasians and 36% of the one of the Karitianas, a population from Southern America whose heterozygosity is very much reduced. This level of heterozygosity indicates that the genetic diversity of the population to which the Denisovan girl belonged was much lower than all known present-day human populations. The alternative hypothesis of interbreeding between blood relations in the direct ancestors of this individual can be excluded since such a situation would have led to long genomic homozygous regions, which was not observed. The result rather suggests that the Denisovan population was very small throughout time. This is surprising when one considers the large area over which this population is believed to have expanded, a region spanning Siberia and Southeast Asia (see below). It will be necessary to identify new Denisovan specimens to clarify the spatio-temporal distribution of this archaic lineage.

The heterozygosity of the genome of the Altai Neanderthal proved to be even lower, with only 84% of the number of heterozygous sites in the Denisovan genome, 22–30% of that found in the genomes of present-day non-Africans and 16–18% of that found in the genomes of present-day Africans [3]. Large homozygous regions were found that could only be explained as the product of interfamilial relations such as grandfather and his granddaughter or between uncle and niece [3]. This was not an exception, however, since the genetic diversity of the late Neanderthals across their entire habitat range was lower than in present-day humans [15].

Moreover, the degree of kinship of non-homozygous regions allowed the reconstruction of the demographic history of the two archaic populations based on a model of coalescence, and to compare them with the one of present-day humans. This has shown that the present-day populations of H. sapiens have all undergone similar changes in population size, in particular a doubling between 125,000 and 250,000 years ago, which points to a period before their migration out-of-Africa. In contrast, the size of the Denisovan and the Altai-Neanderthal populations appears to have drastically declined at about the same time when the expansion of the present-day human populations has started. This does not mean that the two events are necessarily linked. Previous population genetic analyses showed that in the smallest populations slightly deleterious mutations are not as efficiently subjected to purifying selection, leading to a higher ratio between non-synonymous substitutions (those which have an effect on the function of proteins) and synonymous substitutions (those that do not change the amino acids, and so are free to mutate without the restrictions of selection). This is what was found in the two genomes analyzed, with a ratio of 1.5 to 2.5 fold higher in the Denisovan and Neanderthal genomes than in the genomes of present-day humans. This suggests that the Denisovans and Neanderthals had a smaller population size than present-day humans and that the deleterious mutations were less efficiently removed during the evolution of these smaller populations. The analysis of coding regions in the genomes of the two other Neanderthals from Croatia and Spain also confirmed that (i) the Neanderthals have accumulated

more deleterious mutations than modern humans, (ii) all of these populations had a small size, and (iii) the different Neanderthal populations of Europe had little genetic exchange between them. Despite the fact that the evidence for the third conclusion is weak since the bones used for analysis were not contemporaneous, one could speculate that these populations may have disappeared due to the accumulation of deleterious mutations as a consequence of mating in small, isolated populations [15]. In any case, they would have been more sensitive to adverse living conditions and thus to extinction.

3.2 Identification of genes that have evolved

The sequences of the Neanderthal and Denisovan genomes were analyzed to try to identify changes that are unique to the human genome which have a phenotype. The difficulty of finding sense in the genetic variation is that a large number of accumulated mutations are neutral, i.e., they do not cause changes of characters. They can also have slightly deleterious effects that were not counter-selected due to the very small size of the populations. First, only 1.5% of the human genome codes for proteins. It is the genetic variation in these coding regions, however, that would have the highest probability to have a phenotypic effect. The mutations in the remaining 98.5% can also have a phenotypic effect in that they can modify the regulation of gene expression, which is potentially very important since the evolution of proteins is very constrained and the changes in regulation can allow more easily for the gradual emergence of new characters [16]. In contrast, it is more difficult to ensure that a mutation in a non-coding region affects a regulatory region since these regions are more difficult to define than the coding regions, despite the considerable progress in the functional identification of regulatory sequences that has recently been achieved.

The strategy to identify human-specific genomic changes consisted in searching in the genomes of present-day humans for the derived mutations that appeared specifically since the separation with the chimpanzee. These positions are the ones where the human genome differs from the genomes of chimpanzees, orangutans and macaques since a sequence that is shared between these latter species must be ancestral. Next, the state of the corresponding alleles in the Neanderthal and Denisovan genomes was characterized. A catalog of sites distributed over the whole genome and identical among all present-day humans but different in the Neanderthals, the Denisovan and the great apes was established [3]. In this way 31,389 nucleotide substitutions and 4,113 short insertions and deletions were found that are shared among the analyzed present-day humans but different in the Neanderthals, Denisovans and big apes. This corresponds to a rather short list of changes that distinguish present-day humans from other hominins. All the other sites are identical among present-day humans, Neanderthals and Denisovans. This means that the human alleles that are different from the ancestral alleles shared by all the great apes were fixed before the divergence of these populations. 3,117 of these genetic variations specific for modern humans are localized in characterized regulatory sequences and only 96 in coding regions, corresponding to 87 different proteins. Thus, only few changes of amino acids have been fixed in all present-day humans since they had diverged from the Neanderthals. Five proteins were modified at more than one position: one codes for a protein that is involved in the mobility of spermatozoids, one could be involved in wound healing, one is particularly present in the sweat glands and in hair, one has an unknown function but considered to be essential, and another one has also an unknown function [3]. Two genes coding for proteins carry modifications that affect protein length, a dramatic change. Apart from the protein mentioned above that is expressed in hair, a protein involved in the pigmentation of the skin is also affected by a change in protein length. Indeed, this was found to be true for two genes *GPR143* and *LYST* that are involved in the melanosome function. One notices that there is an amazing enrichment of genes involved in the function and appearance of the skin that may have been subject to important modifications in the final phases of our evolution. It is not clear, however, whether this is due to a variation that was fixed because of the small size of the initial human populations or whether it was selected because these modifications conferred a reproductive advantage. Among these 87 genes

whose coding capacity seems to have specifically evolved in modern humans, there is a high number of genes expressed in the ventricular zone of the cerebral cortex, which could have played a role in the development of cognitive capacities that are specific to the lineage of modern humans. There are five of these latter genes, three of which are associated with the mitotic spindle. One can therefore speculate that they may play a role in the specification of the identity of certain neurons after cell division. Another one of these genes is involved in the maintenance of neuronal stem cells in this region of the cortex. Finally, there is an overrepresentation of amino acid-changing substitutions in genes related to behavioral traits that are associated with *de novo* purine biosynthesis (*ADSL*), mitochondrial glycine degradation (*GLDC*), and neurite outgrowth (*SL1TRK1*) [15]. Mutations in the gene *ADSL* are known to lead to psychomotor retardation, autism and muscle wasting in present-day humans, whereas mutations in *GLDC* lead to glycine encephalopathy and in *SL1TRK1* to a syndrome involving motoric and vocal ticks [15].

The group of Svante Pääbo suspected a gene coding for a transcription factor named *FOXP2* to be involved in the emergence of novel cognitive capacities in the course of the evolution of modern humans. *FOXP2* could play a role in the development of the capacity to speak and to master a language. The gene *FOXP2* evolves specifically in the human lineage and has two point mutations that are present in the genomes of humans, Neanderthals and Denisovans, and that distinguish them from the other primates [17]. When the *FOXP2* gene of mice was humanized, i.e., these two human-specific mutations were introduced in the mouse gene, effects on the neuronal plasticity of the basal ganglions were observed, suggesting that these mutations could have played a role in the evolution of language in the human lineage [18]. On the other hand, mutations affecting non-coding regions of the gene *FOXP2* are different in the genomes of humans, Denisovans and Neanderthals [19]. One of these mutations modifies the properties of the sequence that regulates the expression of this gene, which suggests a new evolution of the linguistic capacities determined by *FOXP2* specifically in the human lineage after its divergence from the Neanderthals/Denisovans. Nevertheless, one has to keep in mind that the search for genes responsible for the cognitive capacities that are specifically human can easily lead to speculations concerning the involvement of genetic variations affecting genes that play a role in the brain development. It might be appropriate therefore to keep a critical sense with respect to such interpretations. The identity of candidate genes, however, has evolved with the improvement of the quality of the data produced. It may be difficult to ensure that these genetic variations really played a key role in the emergence of new cognitive capacities unless one identifies modern humans that happen to carry mutations which restore the ancestral sequences.

3.3 Genetic exchanges between archaic populations

The analysis of the genomes of two archaic human lineages has nourished the debate about the diversity of present-day humans. Two opposing hypotheses had been previously proposed, one being the hypothesis of an African origin of present-day human populations, the so-called “out-of-Africa” hypothesis, and the second being the “multiregional” hypothesis. The first one postulates that the ancestors of present-day humans had evolved in Africa and then replaced all other human populations when colonizing the planet (e.g., [20, 21]. At the end of the 1980s, this hypothesis obtained an important support through the genetic studies of mitochondrial DNA and the Y chromosome in present-day human populations [22–24]. These results were interpreted according to the formulation of an “African Mitochondrial Eve” who would have lived 200,000 years ago, and an “African Adam” who would have lived 130,000 to 150,000 years ago. Their descendants would have populated the world and replaced the autochthonous archaic human populations. In its simplest form, this hypothesis doesn’t allow for any hybridization between these migrants and the autochthonous populations they encountered. The second hypothesis argues for a contemporaneous evolution of modern human beings in several parts of the world [25]. This model postulates that it was the gene flow between the ancient African, European

and Asian people that led to the evolution of modern *H. sapiens*. The modern genetic data indicate clearly that the "Out of Africa" model with substantial or total dilution of the genetic contribution of more ancient autochthonous populations is the most likely. Since the Khoisans of Southern Africa show the highest genetic diversity [26], this population could be the closest to the one that is at the origin of modern humans. This diversity would place its most recent common ancestor in the middle of the distance that separates us from the common ancestor with the Neanderthals. The genetic distances between the different African and non-African populations indicate that it was a population from Eastern Africa that would have migrated "out of Africa", about 100,000 years ago. At that time, before the availability of actual genetic data from archaic populations, population geneticists did not agree on the genetic models that would indicate a genetic contribution from more ancient autochthonous populations. The genome fraction proposed to have a Neanderthal or even *H. erectus* origin was either more or less large, or zero. Ten years ago, the analysis of a small number of genetic markers concluded in favor of a model postulating several successive migrations out of Africa over the last two million years followed by admixture that would have allowed the reintroduction of some ancestral alleles in the genomes of population descending from the most recent migration wave of the population. The latter would have been the major contributor to the genetic material of modern humans [27]. This proposition was controversially discussed.

Although it is clear that the contribution of the African population to the genetic heritage 200,000 years ago is major, it is difficult based on mathematical models to determine the degree of gene flow coming from other populations that had left Africa before, because the conclusions depend on the chosen parameters. The data from ancient genomes give essential material support to test the hypotheses. These recent Neanderthal and Denisovan genomes have finally allowed us to firmly establish the concept of admixture between ancient and modern human populations. Indeed, the sequencing of the two palaeogenomes revealed gene flow between Neanderthals and non-African *H. sapiens* and between Denisovans and Asian *H. sapiens*. This result invalidates the results obtained from some morphological studies that did not find evidence for hybridization (e.g., [20] and confirms others that argued in favor of hybridization [28, 29]). This result was not expected, since the mitochondrial DNA sequence of Neanderthals and Denisovans that had been obtained previously showed two distinct mitochondrial lineages, both extinct and outside the variation of present-day humans [5, 11]. This had led to the initial conclusion that Neanderthals did not contribute to our genome. The nuclear genome, however, tells a more complex story.

When the genetic distance between the first version of the Neanderthal genome was compared with those of five present-day humans of diverse geographical origins, it was shown that the single nucleotide polymorphisms (SNPs) that were different in the present-day human genomes compared to the Neanderthal genome were more often shared between a French, a Han Chinese and a Papuan than with a San or a Yoruban [1]. The calculation of the distance between Neanderthals and any other present-day human also showed that the gene flow went from Neanderthals towards the ancestors of the non-African modern humans. Moreover, the genomic regions that could have a Neanderthal origin could also be the ones where the non-African populations had a very high diversity since some individual could have carried the Neanderthal sequences and others those of the modern humans of African origin. The portion of the genome proposed to be of Neanderthal origin was between 1 and 2% [3, 30].

When and where did Neanderthals and anatomically modern humans interbreed? The fossil archives have not yielded any witnesses of the presence of Neanderthals in East Asia and Papua New Guinea. On the other hand, the fact that the contribution of the Neanderthals to the genomes of Europeans, Asians and Papuans is similar suggests that this contribution occurred in the ancestor common to all of these populations (but see below). The admixture between Neanderthals and ancestors of the present-day humans must have taken place before the migration of modern humans into the Asian continent. Neanderthals and ancestors of present-day humans might have coexisted in the Near and Middle East between 100,000 and 50,000 years ago. It is at this time and at this location where the admixture could

have taken place, which is compatible with an east African origin of the non-African populations. Indeed, the DNA of Neanderthal origin in the present-day non-Africans is closer to the genome of the Neanderthal from the Caucasus than to that of the Altai and the Croatian Neanderthals [3]. The analysis of the linkage disequilibrium in the genomes of the present-day European populations, which determine whether there is a preferential association between two alleles, indicates that the last gene flow between Neanderthals and ancestors of the Europeans took place between 65,000 and 37,000 years ago, most likely between 65,000 and 47,000 years ago [31]. According to this, they would have met in Southwest Asia where the archeological remains and fossils indicate that *H. sapiens* arrived around more than 100,000 years ago (as witnessed by the *H. sapiens* skeletons of the Skhul and the Qafzeh caves), the Neanderthals 70,000 years ago (as witnessed by the Neanderthal skeletons of the Tabun cave). The anatomically modern humans would have arrived 50,000 years ago. The genetic estimation of 65,000 to 47,000 years is too recent to correspond to the first fossil trace of modern humans outside of Africa (see also below). This implies that the genomes of the present-day humans cannot be exclusively the result of a single event of gene flow between the anatomically modern humans of Skhul/Qafzeh and the Neanderthals, but rather indicates that there was gene flow as well (or exclusively?) during a more recent period, maybe during the Upper Paleolithic when modern humans left Africa with their specific technologies. Nevertheless, Neanderthal traces were found in the Maasai genome, an East African population, but this has been explained with a secondary contact in Africa of a non-African population who were already carrying admixed Neanderthal DNA [32]. This interpretation is also compatible with genetic traces of admixture between African populations and those of the Middle East and Southern Europe [33] as well as with a higher similarity between the genomes of non-African individuals and east Africans rather than with West Africans [34]. This would mean that there were not just several "out of Africa", but also some "back to Africa" migrations. Finally, the three populations of the Dai, Han (East Asia) and Karitiana (South America) share more alleles with the Denisovans than the Sardinians and the present-day French. The corresponding genomic regions being closer to the ones of the Neanderthals than to the Denisovans, however, these data argue in favor of a larger genetic contribution from the Neanderthals to the genomes of the Asian and Amerindian populations than to the European ones [3]. A second study confirms that observation. Indeed, a comparison of the genomes of Neanderthals, Denisovans and 42 genomes of various present-day populations proposes that the Neanderthal admixture rate into East Asians (Han Chinese and Japanese) is 40% higher than in Europeans [32]. Other analyses that took advantage of the excellent quality of the Altai Neanderthal genome together with the data from the 1,000 genome project propose only 1 to 1.2% of the genome being of Neanderthal origin in the European genomes, depending on the population considered, and 1.4% in Asians, which corroborates also the interpretation of several events of gene flow between Neanderthals and modern humans in different populations [30, 35]. There would have been at least two Neanderthal admixture events that left their traces in modern human genomes, one concerning the common ancestors of Europeans and Asians, the second concerning only the ancestors of Asian populations. These admixture events would have taken place between 80,000 and 50,000 years [36, 37]. The separation between Europeans and East Asians would have taken place before the extinction of the Neanderthals, i.e., 30,000 years ago, which is in agreement with the discovery in the cave of Tianyuan in China of a modern human fossil of 40,000 [38]. Different fractions of the Neanderthal genome would have been dispersed in the individuals of European or Asian origin and up to 20% of the Neanderthal genome would still be present collectively in the genomes of modern humans [35].

The most recent report of the genome sequence of a 45,000 year-old *H. sapiens* from Ust'-Ishim in western Siberia indicates that Neanderthal gene flow occurred 232-430 generations before the Ust'-Ishim individual lived, as deduced from the size of the genome portions that have a Neanderthal origin [39]. The estimation of the time of admixture between the ancestors of the Ust'-Ishim individual and Neanderthals, based on the simplifying assumption that the gene flow occurred as a single event and assuming a generation time of 29 years, yielded approximately 60,000 to 50,000 years BP, which

is in agreement with the estimated time of the major expansion of modern humans out of Africa and the Middle East. Thus, the bulk of the Neanderthal contribution of present-day people outside Africa does not go back to mixture between Neanderthals and the anatomically modern humans who lived in the Middle East at earlier times, such as those whose remains have been found at Skhul and Qafzeh [39]. The presence of some longer Neanderthal-derived DNA fragments in the Ust'-Ishim genome may indicate, however, that additional admixture occurred even later [39].

Other admixture events between archaic and modern humans were brought to light. The analysis of the Denisovan genome and the comparison with the genomes of eleven present-day humans showed that the Denisovans have contributed up to 8% of their genome to the Melanesians, the Papuans of New Guinea, the Australian Aborigenes and other inhabitants of the islands of southeast Asia [2-4, 40, 41]. In contrast, the genetic contribution of the Denisovans to the present-day northern Asians as well as to the autochthonous Amerindian populations is about 0.2%, and thus about 25 times smaller than that to the Melanesian populations. The same seems to be true for the ancient populations of Northern Asia as deduced from the analysis of the mitochondrial genome and portions of the nuclear genome of the Tianyuan individual from the region of Beijing that was dated to 40,000 years [38]. The distribution in Asia of the populations that carry a portion of the Denisovan genome suggests that the admixture event took place in Southeast Asia and thus that the geographical region inhabited by Denisovans was much larger than the small area where their only remains have been found [40]. A sophisticated statistical analysis developed for the analysis of ancient genomes and carried out on the alleles that are not present in 35 African genomes confirmed these results [2]. This means that a subgroup of the *sapiens* population that interbred with the Neanderthals during their westwards migration would have subsequently interbred with the Denisovan population in Southeast Asia.

The proportion of the alleles that are shared between Denisovans and Melanesians is not distributed homogeneously over all the chromosomes since there are less on the X chromosome of the Papuans [2]. This could indicate that gene flow between Denisovans and the ancestors of the Papuans involved mainly Denisovan men. Alternatively, this result could also be the consequence of a substructure of the population combined with migration that involved mainly *sapiens* women. The fact that some autosomal chromosomes contain less Denisovan alleles could however indicate that other factors came into play. It could be that some alleles were disadvantageous for the hybrids, and were removed by selection. Such a hypothesis has also been put forward to explain the distribution of the Neanderthal contribution to modern human genomes. Indeed, large parts of the present-day non-African genome are depleted of Neanderthal DNA whereas in other parts the Neanderthal heritage is much richer, where some of the Neanderthal alleles are present at a frequency of up to 60% in European and Asian populations [35 Sankararaman, 2014 #1694]. These results indicate that some regions of the Neanderthal genome were counter-selected in the hybrid offspring, whereas others conveyed a selective advantage and therefore became dominant in present-day human populations. The genomic regions of the Neanderthal that were lost due to counter-selection in the hybrids are strongly enriched in genes that are expressed in the testes, suggesting that the hybrid males were less fertile [2, 30].

In contrast, it seems as if the contribution of the ancestral genomes also conveyed selective advantages to modern humans. In particular, several of the genes involved in the development of keratinocytes and the formation of the keratin fiber have alleles of Neanderthal origin that are present at a frequency of up to 60% in European and Asian populations [30, 35]. These genes are responsible for the resistance of skin, hair and nails and guarantee a better protection in colder environments suggesting that the Neanderthals were already adapted to non-African latitudes and transmitted this genetic adaptation to modern humans. Furthermore, the interbreeding between different human lineages also left traces in the immune system. The ancestors of present-day humans acquired different alleles of class I of the major histocompatibility complex HLA, which is very polymorphic and subject to pressure through balanced selection [42]. These alleles represent more than half of the HLA alleles in present-day Eurasians and seem to have been introduced later into the African populations [42]. This

overrepresentation indicates that the alleles must be under positive selection and therefore convey a selective advantage, at least in their corresponding ecological niches. This suggests that while *H. sapiens* migrated outside of Africa, they adapted through interbreeding favorable alleles that led to the improved fitness of archaic populations against pathogens that were present in their new environments. Similarly, genetic adaptation of the Tibetans to high altitudes appears to originate from an allele transmitted from Denisovans or Denisovan-related individuals [43]. This adaptation results from an allele of the hypoxia pathway gene EPAS2 that has the most extreme signature of positive selection in Tibetans, while it is found only at very low frequency in Han Chinese. EPAS2 is associated with differences in hemoglobin concentration at high altitude. This allele has not been found in Melanesians nor in the Altai Neanderthal. Since Melanesians are the present-day populations with the highest genomic portions of Denisovan origin whereas Han Chinese and Mongolians have very little Denisovan DNA despite their geographical proximity to the Tibetan region, this suggests that there has been several instances of admixture with Denisovans with natural selection allowing fixation of the EPAS2 Denisovan allele only in Tibetans [43]. In contrast, some Neanderthal alleles that occur at high frequency in some populations and that could have conferred a selective advantage in the past can be less advantageous in the modern world. For example, an allele of the gene *SLC16A11* associated to a higher predisposition to type 2 diabetes in autochthonous populations of the Americas is present at a frequency of 50% in these populations and results from a Neanderthal introgression in the ancestors of the North Asiatic populations [44]. Presumably this allele conferred an advantage during times of irregular food supply, which explains its spreading among the Asian populations that migrated into the Americas whereas currently, at times of high food supply, it is no longer beneficial.

Admixture between archaic populations is not limited to populations that had migrated « out of Africa » but was also observed in present-day African populations. The discovery of a Y chromosome lineage present in 6.3% of tested individuals from the population of the “Mbo” in Cameroon, and very different from the rest of the present-day human population including the Khoisans and the Pygmies has shattered the hypothesis that the ancestor of all present-day men had lived 140,000 years ago in Africa. Indeed, the calculations based on this Y lineage implied that the ancestor of the Y chromosome of all present-day men and the Y chromosome Mbo had lived in Africa 340,000 years ago [45]. This predates largely the date for the “mitochondrial Eve”, the mother of all present-day maternal lineages, and the most ancient fossils of *Homo sapiens*.Even if the date is more likely to be closer to 208,000, as proposed by a later analysis [46] using a Y mutation rate in line with more recent calculations [39], this is still a surprising result. How can it be explained? It could be the result of either an ancient population structure of *H. sapiens* in Africa or of the introgression of an archaic ancestor of *H. sapiens* [45]. Could the human skeletons found at Iwo Eleru, an archeological site in Niger that is 800 km away from the Mbo population, that are dated to 13,000 years BP and show modern and archaic morphological features be the descendants of an admixture between modern and archaic humans [47]? The sequencing of 15 genomes of three hunter-gatherer populations, the pygmies from Cameroon, as well as the Hadza and the Sandawe from Tanzania, revealed the admixture between ancestors that are common to these three populations with an archaic population that was separated from the modern lineage 1.2 to 1.3 million years ago [48]. Thus, Africa was also a theater of admixture between different archaic populations. Future paleogenomic research may allow us to identify the archaic population(s) involved.

In addition to the evidence of gene flow from ancient populations into modern humans, the analysis of the Altai Neanderthal and Denisovan genomes also show signatures of gene flow toward the Denisovans from Neanderthals from the Altai, but not those from Croatia and the Caucasus, since the derived alleles in the Denisovan genome are closer to those of the Altai Neanderthal than to the other analyzed Neanderthals [3]. At least 0.5% of the Denisovan genome seems to stem from a Neanderthal population related to the Altai Neanderthal. The signal for gene flow is particularly high for the HLA alleles and the CRISP cluster that are involved in immunity and the functioning of the sperm, respectively. At the moment, there is no indication for reciprocity of the gene flow. The admixture

of Denisovan populations with a more archaic population was also postulated following comparative analyses of the Denisovan and Neanderthal genomes of the Altai with numerous modern genomes [3]. Indeed, while it is clear that the common ancestor of the Neanderthals and the Denisovans first diverged from the ones of the modern humans before the Neanderthal and Denisovan lineages diverged, the African genomes share up to 7% more alleles with the Neanderthal genome than with the Denisovan genome. The scenario that accounts best for the exact structure of the data is that of gene flow into the Denisovans from an archaic population that diverged from the ancestors of modern humans 1 to 4 million years ago and that would have contributed 1 to 8% to the Denisovan genome [3]. A complex phylogeny of the Denisovans was already predicted by the contradiction between the divergence date of the Denisovan mitochondrial genome from the one of the modern human lineage [11] and the more recent date estimated from the nuclear genome [2, 3]. Moreover, a mitochondrial lineage relatively close to the one of the Denisovan has been identified in a 400,000 to 300,000 year-old femur preserved in the cave of "Sima de los Huesos", located on the site of Atapuerca in Spain, a bone which, based on its morphology, is contentiously assigned to *H. heidelbergensis*, the predecessor of the Neanderthals. [49]. Was this mitochondrial lineage the signature of an archaic European population with African provenance that would have preceded the ancestors of Neanderthals and Denisovans? Or local ancestral populations belonging to *H. erectus*, such as the 800,000 year-old bones from another site in Atapuerca that are designated *H. antecessor*? Has the admixture of the ancestors of the Denisovans with other *H. erectus* populations in Asia left these traces in the Denisovan genome? The conclusion has to wait for the nuclear genome of the individuals of Sima de los Huesos.

These data indicate as well that one can consider that Neanderthals and Denisovans still belonged to the same species as the modern humans shortly before their extinction: the individuals of the three populations were capable of reproducing with each other under natural conditions, which corresponds to the definition of the term "species" according to Wallace and backed by Ernst Mayr [50, 51], even though this definition is probably too restricted (see [52, 53]). Nevertheless, one needs to take into account the possibility that the fertility of the hybrids was not optimal, and in particular of the males, as discussed before. Although it affects a large part of the present-day population, the extent of the gene flow remained limited. This can be related to the relative size of the archaic and modern populations in expansion. It is not necessary to assume that the interbreeding events between Neanderthals and modern humans have occurred often to explain that parts of the Neanderthal genome could have survived such a long time. Indeed, the descendant human populations expanded very efficiently following their migration "Out of Africa". This could have been sufficient to allow for fixation in the population of these introgressed alleles. Such a low number of interbreeding events could explain the difficulty to find hybrid fossil traces between these populations, especially since the phenotypic characters of the Neanderthals became rapidly diluted through successive interbreeding with other modern humans. It is also possible that the first-generation Neanderthal-*sapiens* hybrids were less hardy and could not reproduce as well with full-blooded *sapiens*, but the contribution of some advantageous alleles coming from archaic populations would have given a selective advantage to their offspring, while successive interbreeding would have diluted the contributions of the Neanderthal and Denisovan genomes that were less advantageous.

What about admixture between Neanderthals and modern humans in Europe? So far in the genomes of Europeans, no clear additional traces of a genetic contribution of Neanderthals have been found. Thus, if there was admixture, the offspring would not have expanded, at least not as much as those from the Near/Middle East and the hybrid ancestors of the Asian populations. This population of modern Europeans that was contemporaneous to the Neanderthals could also have been reduced severely for other reasons, such as climatic ones. Moreover, no traces of gene flow from *sapiens* into Neanderthals have been found. In contrast, the Neanderthal population being in its extinction rather than expansion phase, rare interbreeding events could have been masked. Some paleoanthropological data suggest that such interbreeding could have taken place also in Europe [28].

In conclusion, there were at least four to five admixture events between modern and archaic humans known to date, one or two in Africa involving one of the archaic populations that are not yet described paleogenomically, two in Southwest Asia with the Neanderthals and one in Southeast Asia with the Denisovans, and at least two other admixture events between archaic humans, with Neanderthals and Denisovans in the Altai on one hand and between Denisovans and another archaic human population that could correspond to *H. erectus* on the other. Thus, it seems that the evolution of the human population that inhabited the earth alone for the past 30,000 years was much more complex than previously thought. Indeed, a number of admixture events with other archaic populations took place when modern humans were not quite alone. Thus, these data allow a more nuanced rather than a dichotomic view of human evolution.

References

[1] R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H. Fritz, N. F. Hansen, E. Y. Durand, A. S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prufer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Hober, B. Hoffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Paabo, Science **328**, 710–722 (2010)

[2] M. Meyer, M. Kircher, M. T. Gansauge, H. Li, F. Racimo, S. Mallick, J. G. Schraiber, F. Jay, K. Prufer, C. de Filippo, P. H. Sudmant, C. Alkan, Q. Fu, R. Do, N. Rohland, A. Tandon, M. Siebauer, R. E. Green, K. Bryc, A. W. Briggs, U. Stenzel, J. Dabney, J. Shendure, J. Kitzman, M. F. Hammer, M. V. Shunkov, A. P. Derevianko, N. Patterson, A. M. Andres, E. E. Eichler, M. Slatkin, D. Reich, J. Kelso, S. Paabo, Science **338**, 222–226 (2012)

[3] K. Prufer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinze, G. Renaud, P. H. Sudmant, C. de Filippo, H. Li, S. Mallick, M. Dannemann, Q. Fu, M. Kircher, M. Kuhlwilm, M. Lachmann, M. Meyer, M. Ongyerth, M. Siebauer, C. Theunert, A. Tandon, P. Moorjani, J. Pickrell, J. C. Mullikin, S. H. Vohr, R. E. Green, I. Hellmann, P. L. Johnson, H. Blanche, H. Cann, J. O. Kitzman, J. Shendure, E. E. Eichler, E. S. Lein, T. E. Bakken, L. V. Golovanova, V. B. Doronichev, M. V. Shunkov, A. P. Derevianko, B. Viola, M. Slatkin, D. Reich, J. Kelso, S. Paabo, Nature **505**, 43–49 (2014)

[4] D. Reich, R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. Johnson, T. Maricic, J. M. Good, T. Marques-Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V. Shunkov, A. P. Derevianko, J. J. Hublin, J. Kelso, M. Slatkin, S. Paabo, Nature **468**, 1053–1060 (2010)

[5] A. W. Briggs, J. M. Good, R. E. Green, J. Krause, T. Maricic, U. Stenzel, C. Lalueza-Fox, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, R. Schmitz, V. B. Doronichev, L. V. Golovanova, M. de la Rasilla, J. Fortea, A. Rosas, S. Paabo, Science **325**, 318–321 (2009)

[6] S. S. Ajay, S. C. Parker, H. O. Abaan, K. V. Fajardo, E. H. Margulies, Genome Res **21**, 1498–1505 (2011)

[7] H. A. Burbano, E. Hodges, R. E. Green, A. W. Briggs, J. Krause, M. Meyer, J. M. Good, T. Maricic, P. L. Johnson, Z. Xuan, M. Rooks, A. Bhattacharjee, L. Brizuela, F. W. Albert, M. de la Rasilla, J. Fortea, A. Rosas, M. Lachmann, G. J. Hannon, S. Paabo, Science **328**, 723–725 (2010)

[8] T. Maricic, M. Whitten, S. Paabo, PLoS one **5**, e14004 (2010)

[9] M. L. Carpenter, J. D. Buenrostro, C. Valdiosera, H. Schroeder, M. E. Allentoft, M. Sikora, M. Rasmussen, S. Gravel, S. Guillen, G. Nekhrizov, K. Leshtakov, D. Dimitrova, N. Theodossiev, D. Pettener, D. Luiselli, K. Sandoval, A. Moreno-Estrada, Y. Li, J. Wang, M. T. Gilbert, E. Willerslev, W. J. Greenleaf, C. D. Bustamante, Am J Hum Genet **93**, 852–864 (2013)

[10] J. M. Enk, A. M. Devault, M. Kuch, Y. E. Murgha, J. M. Rouillard, H. N. Poinar, Mol Biol Evol **31**, 1292–1294 (2014)
[11] J. Krause, Q. Fu, J. M. Good, B. Viola, M. V. Shunkov, A. P. Derevianko, S. Paabo, Nature **464**, 894–897 (2010)
[12] E. A. Bennett, D. Massilani, G. Lizzo, J. Daligault, E. M. Geigl, T. Grange, BioTechniques **56**, 289–290, 292-286, 298 (2014)
[13] A. Kong, M. L. Frigge, G. Masson, S. Besenbacher, P. Sulem, G. Magnusson, S. A. Gudjonsson, A. Sigurdsson, A. Jonasdottir, A. Jonasdottir, W. S. Wong, G. Sigurdsson, G. B. Walters, S. Steinberg, H. Helgason, G. Thorleifsson, D. F. Gudbjartsson, A. Helgason, O. T. Magnusson, U. Thorsteinsdottir, K. Stefansson, Nature **488**, 471–475 (2012)
[14] J. C. Roach, G. Glusman, A. F. Smit, C. D. Huff, R. Hubley, P. T. Shannon, L. Rowen, K. P. Pant, N. Goodman, M. Bamshad, J. Shendure, R. Drmanac, L. B. Jorde, L. Hood, D. J. Galas, Science **328**, 636–639 (2010)
[15] S. Castellano, G. Parra, F. A. Sanchez-Quinto, F. Racimo, M. Kuhlwilm, M. Kircher, S. Sawyer, Q. Fu, A. Heinze, B. Nickel, J. Dabney, M. Siebauer, L. White, H. A. Burbano, G. Renaud, U. Stenzel, C. Lalueza-Fox, M. de la Rasilla, A. Rosas, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, M. V. Shunkov, A. P. Derevianko, B. Viola, M. Meyer, J. Kelso, A. M. Andres, S. Paabo, Proc Natl Acad Sci USA **111**, 6666–6671 (2014)
[16] D. L. Stern, V. Orgogozo, Evolution **62**, 2155–2177 (2008)
[17] J. Krause, C. Lalueza-Fox, L. Orlando, W. Enard, R. E. Green, H. A. Burbano, J.-J. Hublin, C. Hänni, J. Fortea, M. de la Rasilla, J. Bertranpetit, A. Rosas, S. Pääbo, Current Biology **17**, 1908–1912 (2007)
[18] W. Enard, S. Gehre, K. Hammerschmidt, S. M. Holter, T. Blass, M. Somel, M. K. Bruckner, C. Schreiweis, C. Winter, R. Sohr, L. Becker, V. Wiebe, B. Nickel, T. Giger, U. Muller, M. Groszer, T. Adler, A. Aguilar, I. Bolle, J. Calzada-Wack, C. Dalke, N. Ehrhardt, J. Favor, H. Fuchs, V. Gailus-Durner, W. Hans, G. Holzlwimmer, A. Javaheri, S. Kalaydjiev, M. Kallnik, E. Kling, S. Kunder, I. Mossbrugger, B. Naton, I. Racz, B. Rathkolb, J. Rozman, A. Schrewe, D. H. Busch, J. Graw, B. Ivandic, M. Klingenspor, T. Klopstock, M. Ollert, L. Quintanilla-Martinez, H. Schulz, E. Wolf, W. Wurst, A. Zimmer, S. E. Fisher, R. Morgenstern, T. Arendt, M. H. de Angelis, J. Fischer, J. Schwarz, S. Paabo, Cell **137**, 961–971 (2009)
[19] T. Maricic, V. Gunther, O. Georgiev, S. Gehre, M. Curlin, C. Schreiweis, R. Naumann, H. A. Burbano, M. Meyer, C. Lalueza-Fox, M. de la Rasilla, A. Rosas, S. Gajovic, J. Kelso, W. Enard, W. Schaffner, S. Paabo, Mol Biol Evol **30**, 844–852 (2013)
[20] C. B. Stringer, P. Andrews, Science **239**, 1263–1268 (1988)
[21] C. Stringer, Nature **405**, 24–25, 27 (2000)
[22] R. L. Cann, M. Stoneking, A. C. Wilson, Nature **325**, 31–36 (1987)
[23] F. R. Santos, N. O. Bianchi, S. D. Pena, Genome Res **6**, 601–611 (1996)
[24] R. Thomson, J. K. Pritchard, P. Shen, P. J. Oefner, M. W. Feldman, Proc Natl Acad Sci USA **97**, 7360–7365 (2000)
[25] M. H. Wolpoff, J. N. Spuhler, F. H. Smith, J. Radovcic, G. Pope, D. W. Frayer, R. Eckhardt, G. Clark, Science **241**, 772–774 (1988)
[26] S. C. Schuster, W. Miller, A. Ratan, L. P. Tomsho, B. Giardine, L. R. Kasson, R. S. Harris, D. C. Petersen, F. Zhao, J. Qi, C. Alkan, J. M. Kidd, Y. Sun, D. I. Drautz, P. Bouffard, D. M. Muzny, J. G. Reid, L. V. Nazareth, Q. Wang, R. Burhans, C. Riemer, N. E. Wittekindt, P. Moorjani, E. A. Tindall, C. G. Danko, W. S. Teo, A. M. Buboltz, Z. Zhang, Q. Ma, A. Oosthuysen, A. W. Steenkamp, H. Oostuisen, P. Venter, J. Gajewski, Y. Zhang, B. F. Pugh, K. D. Makova, A. Nekrutenko, E. R. Mardis, N. Patterson, T. H. Pringle, F. Chiaromonte, J. C. Mullikin, E. E. Eichler, R. C. Hardison, R. A. Gibbs, T. T. Harkins, V. M. Hayes, Nature **463**, 943–947 (2010)
[27] A. Templeton, Nature **416**, 45–51 (2002)

[28] C. Duarte, J. Mauricio, P. B. Pettitt, P. Souto, E. Trinkaus, H. van der Plicht, J. Zilhao, Proc Natl Acad Sci USA **96**, 7604–7609 (1999)
[29] E. Trinkaus, Proc Natl Acad Sci USA **104**, 7367–7372 (2007)
[30] S. Sankararaman, S. Mallick, M. Dannemann, K. Prufer, J. Kelso, S. Paabo, N. Patterson, D. Reich, Nature **507**, 354–357 (2014)
[31] S. Sankararaman, N. Patterson, H. Li, S. Paabo, D. Reich, PLoS Genet **8**, e1002947 (2012)
[32] J. D. Wall, M. A. Yang, F. Jay, S. K. Kim, E. Y. Durand, S. Stevison, C. Gignoux, A. Woerner, M. F. Hammer, M. Slatkin, Genetics **194**, 199–209 (2013)
[33] P. Moorjani, N. Patterson, J. N. Hirschhorn, A. Keinan, L. Hao, G. Atzmon, E. Burns, H. Ostrer, A. L. Price, D. Reich, PLoS Genet **7**, e1001373 (2011)
[34] S. A. Tishkoff, F. A. Reed, F. R. Friedlaender, C. Ehret, A. Ranciaro, A. Froment, J. B. Hirbo, A. A. Awomoyi, J. M. Bodo, O. Doumbo, M. Ibrahim, A. T. Juma, M. J. Kotze, G. Lema, J. H. Moore, H. Mortensen, T. B. Nyambo, S. A. Omar, K. Powell, G. S. Pretorius, M. W. Smith, M. A. Thera, C. Wambebe, J. L. Weber, S. M. Williams, Science **324**, 1035–1044 (2009)
[35] B. Vernot, J. M. Akey, Science **343**, 1017–1021 (2014)
[36] P. Skoglund, M. Jakobsson, Proc Natl Acad Sci USA **108**, 18301–18306 (2011)
[37] M. Currat, L. Excoffier, Proc Natl Acad Sci USA **108**, 15129–15134 (2011)
[38] Q. Fu, M. Meyer, X. Gao, U. Stenzel, H. A. Burbano, J. Kelso, S. Paabo, Proc Natl Acad Sci USA **110**, 2223–2227 (2013)
[39] Q. Fu, H. Li, P. Moorjani, F. Jay, S. M. Slepchenko, A. A. Bondarev, P. L. Johnson, A. Aximu-Petri, K. Prufer, C. de Filippo, M. Meyer, N. Zwyns, D. C. Salazar-Garcia, Y. V. Kuzmin, S. G. Keates, P. A. Kosintsev, D. I. Razhev, M. P. Richards, N. V. Peristov, M. Lachmann, K. Douka, T. F. Higham, M. Slatkin, J. J. Hublin, D. Reich, J. Kelso, T. B. Viola, S. Paabo, Nature **514**, 445–449 (2014)
[40] D. Reich, N. Patterson, M. Kircher, F. Delfin, M. R. Nandineni, I. Pugach, A. M. Ko, Y. C. Ko, T. A. Jinam, M. E. Phipps, N. Saitou, A. Wollstein, M. Kayser, S. Paabo, M. Stoneking, Am J Hum Genet **89**, 516–528 (2011)
[41] M. Rasmussen, Y. Li, S. Lindgreen, J. S. Pedersen, A. Albrechtsen, I. Moltke, M. Metspalu, E. Metspalu, T. Kivisild, R. Gupta, M. Bertalan, K. Nielsen, M. T. Gilbert, Y. Wang, M. Raghavan, P. F. Campos, H. M. Kamp, A. S. Wilson, A. Gledhill, S. Tridico, M. Bunce, E. D. Lorenzen, J. Binladen, X. Guo, J. Zhao, X. Zhang, H. Zhang, Z. Li, M. Chen, L. Orlando, K. Kristiansen, M. Bak, N. Tommerup, C. Bendixen, T. L. Pierre, B. Gronnow, M. Meldgaard, C. Andreasen, S. A. Fedorova, L. P. Osipova, T. F. Higham, C. B. Ramsey, T. V. Hansen, F. C. Nielsen, M. H. Crawford, S. Brunak, T. Sicheritz-Ponten, R. Villems, R. Nielsen, A. Krogh, J. Wang, E. Willerslev, Nature **463**, 757–762 (2010)
[42] L. Abi-Rached, M. J. Jobin, S. Kulkarni, A. McWhinnie, K. Dalva, L. Gragert, F. Babrzadeh, B. Gharizadeh, M. Luo, F. A. Plummer, J. Kimani, M. Carrington, D. Middleton, R. Rajalingam, M. Beksac, S. G. Marsh, M. Maiers, L. A. Guethlein, S. Tavoularis, A. M. Little, R. E. Green, P. J. Norman, P. Parham, Science **334**, 89–94 (2011)
[43] E. Huerta-Sanchez, X. Jin, Asan, Z. Bianba, B. M. Peter, N. Vinckenbosch, Y. Liang, X. Yi, M. He, M. Somel, P. Ni, B. Wang, X. Ou, Huasang, J. Luosang, Z. X. Cuo, K. Li, G. Gao, Y. Yin, W. Wang, X. Zhang, X. Xu, H. Yang, Y. Li, J. Wang, R. Nielsen, Nature **512**, 194–197 (2014)
[44] S. T. D. Consortium, A. L. Williams, S. B. Jacobs, H. Moreno-Macias, A. Huerta-Chagoya, C. Churchhouse, C. Marquez-Luna, H. Garcia-Ortiz, M. J. Gomez-Vazquez, N. P. Burtt, C. A. Aguilar-Salinas, C. Gonzalez-Villalpando, J. C. Florez, L. Orozco, C. A. Haiman, T. Tusie-Luna, D. Altshuler, Nature **506**, 97–101 (2014)
[45] F. L. Mendez, T. Krahn, B. Schrack, A. M. Krahn, K. R. Veeramah, A. E. Woerner, F. L. Fomine, N. Bradman, M. G. Thomas, T. M. Karafet, M. F. Hammer, Am J Hum Genet **92**, 454–459 (2013)
[46] E. Elhaik, T. V. Tatarinova, A. A. Klyosov, D. Graur, Eur J Hum Genet (2014)

[47] K. Harvati, C. Stringer, R. Grun, M. Aubert, P. Allsworth-Jones, C. A. Folorunso, PLoS one **6**, e24024 (2011)
[48] J. Lachance, B. Vernot, C. C. Elbers, B. Ferwerda, A. Froment, J. M. Bodo, G. Lema, W. Fu, T. B. Nyambo, T. R. Rebbeck, K. Zhang, J. M. Akey, S. A. Tishkoff, Cell **150**, 457–469 (2012)
[49] M. Meyer, Q. Fu, A. Aximu-Petri, I. Glocke, B. Nickel, J. L. Arsuaga, I. Martinez, A. Gracia, J. M. de Castro, E. Carbonell, S. Paabo, Nature **505**, 403–406 (2014)
[50] A. R. Wallace, Transactions of the Linnean Society of London **25**, 1–71 (1865)
[51] E. Mayr, Science **141**, 765 (1963)
[52] K. de Queiroz, Proc. Natl. Acad. Sci. USA **102** Suppl 1, 6600–6607 (2005)
[53] K. de Queiroz, Syst. Biol. **56**, 879–886 (2007)