Admixture between Archaic and Modern Humans

Alan R. Rogers* Ryan J. Bohlender

2016

Introduction

For many years, anthropologists disagreed about interbreeding between modern humans and archaic forms such as Neanderthals. Some saw evidence of admixture (Bräuer, 1984, 1989; Smith et al., 1989; Trinkaus, 2005; Wolpoff, 1989); others did not (Lahr and Foley, 1998; Stringer and Andrews, 1988). Most population geneticists were until recently among the skeptics. This skepticism rested on the remarkable homogeneity of the mitochondrial DNA (mtDNA) of modern humans (Krings et al., 1997; Serre et al., 2004; Stoneking, 1993). Had archaics contributed mtDNA to the modern human gene pool, we would expect less homogeneity. This inference was weak, however, because mtDNA does not recombine and thus evolves as a single locus. This single locus does not allow strong inferences about archaic admixture. Manderscheid and Rogers (1996, Fig. 2) showed that, to exclude an admixture rate of a few percent, one must assume that the female European population was implausibly large, with a long-term effective size of at least several hundred thousand. Nordborg (1998) reached similar conclusions.

Things began to change near the turn of the century, as nuclear DNA became more important in human population genetics (Harding et al., 1997). Unlinked loci provide essentially independent replicates of the evolutionary process, and this replication greatly strengthens inferences about archaic admixture. More recently, access to ancient genomes has made these inferences even stronger (Green et al., 2010). It is now clear that archaic DNA has introgressed into many populations of modern humans. In some genomic regions, archaic DNA seems to have been purged by natural selection. In others, archaic variants were favored by selection. Thus, archaic ancestors helped shape the genetic adaptations of modern humans.

Detecting archaic admixture

Several ideas underlie the methods that are used to detect archaic admixture. The first involves separation time. When two populations have been separate for a long time and then exchange genes, some loci will carry alleles from both populations. At such loci, the gene tree is deep—at least as deep as the separation time of the populations. On the other hand, loci without admixed alleles have shallower gene trees. Thus, admixture gives us a diversity of gene trees—some long and some short. This variation distorts the site frequency spectrum and related statistics. Eswaran et al. (2005) used this idea a decade ago to argue for archaic admixture.

Admixture also generates a distinctive signature in linkage disequilibrium (LD). In a randomly-mating population, recombination rapidly destroys long blocks of LD. When populations are sep-

*Dept. of Anthropology, 270 S 1400 E Rm 102, University of Utah 84112, Salt Lake City, UT

Table 1: Site patterns and their frequencies in one data set, with 0 and 1 representing the ancestral and derived alleles (Patterson et al., 2010a, p. S138).

	Nucleotide Site Pattern		
	ae	*en*	*an*
Europe (E)	1	1	0
Africa (A)	1	0	1
Neanderthal (N)	0	1	1
Chimpanzee (C)	0	0	0
Count	303,340	103,612	95,347

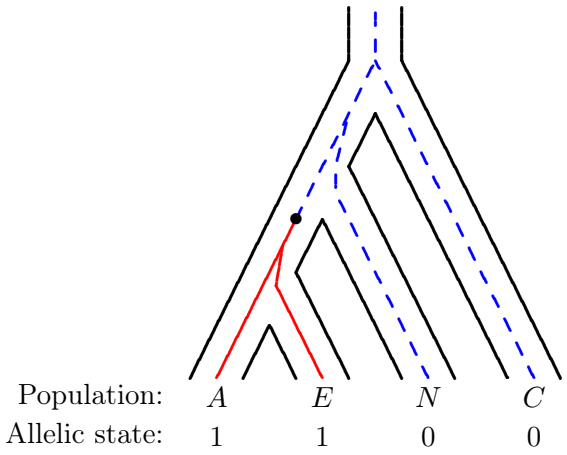

Figure 1: Population tree (bold black lines) with embedded gene tree. Bullet (•) marks mutation from allele 0 to allele 1. Embedded gene tree shows histories of ancestral allele (0, dashed blue lines) and derived allele (1, solid red lines). In this example, the closest relatives (A and E) uniquely share the derived allele.

arate, however, chromosomes from one cannot recombine with those from the other, and between-population LD accumulates. When the isolation finally breaks down, and genes flow from one population into another, the introgressing alleles initially lie on long haplotypes, which are geographically restricted. These haplotypes tend to differ at many nucleotide sites, because of the long separation time. This pattern has often been used to detect archaic admixture (Abi-Rached et al., 2011; Cox et al., 2008; Evans et al., 2005; Garrigan et al., 2005a; Hammer et al., 2011; Hu et al., 2015; Mendez et al., 2012a; Moorjani et al., 2011; Plagnol and Wall, 2006; Reed and Tishkoff, 2006; Wall and Hammer, 2006; Wall, 2000; Wall et al., 2009).

Another class of methods capitalizes on the availability of archaic DNA sequences. These methods infer admixture from the frequency with which derived alleles are shared by pairs of samples. Table 1 illustrates this idea with a data set consisting of four haploid genomes. The populations in this data set are Africa (A), Europe (E), Neanderthal (N), and chimpanzee (C). We use lower-case letters (a, e, n, and c) to represent haploid genomes sampled from those populations. The chimpanzee genome is used only as an indication of which allele is ancestral. If two alleles are present in human samples but only one is present in chimpanzees, the chimpanzee allele is likely to be ancestral. There are three site patterns in which the derived allele is shared by two human samples but not by chimpanzee. We represent these patterns by pairs of letters—*ae*, *en*, and *an*—which indicate which pair of samples shares the derived allele.

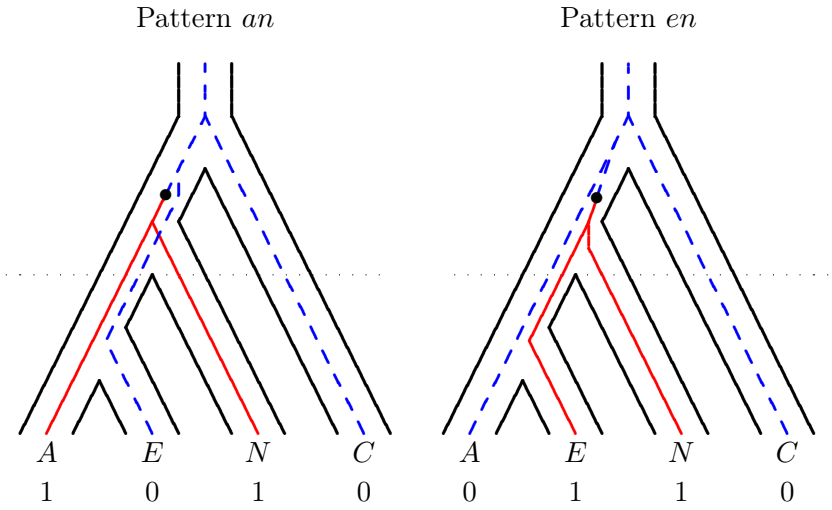

Figure 2: Incomplete lineage sorting: distant relatives may share derived allele.

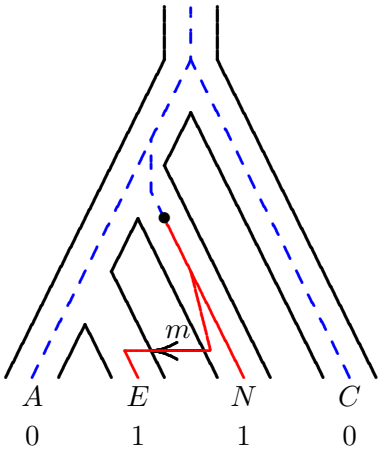

Figure 3: Neanderthal admixture inflates *en* site pattern.

In these data, the *ae* site pattern is much more common than the other two. Fig. 1 shows why. Two samples uniquely share a derived allele only if a mutation occurs in a uniquely shared ancestor. This is most likely to happen in populations such as A and E, which are related and thus share a portion of their evolutionary history.

In the other two site patterns (*en* and *an*), the derived allele links Neanderthal to a modern human sample. In the absence of admixture, these patterns can arise only through *incomplete lineage sorting*, the process is illustrated in Fig. 2. In this figure, the two human lineages remain distinct (just by chance) far back into the past. Above the dotted line, all three lineages (a, e, and n) are part of a single population and can therefore coalesce in any order.

If random mating prevailed within the population ancestral to humans and Neanderthals, the two counter-intuitive patterns (*en* and *an*) ought to occur with equal frequency (Pamilo and Nei, 1988). Yet in table 1 the *en* pattern occurs more often than the *an* pattern. This excess may result from gene flow from Neanderthals into Europeans, as illustrated in Fig. 3.

This pattern is used as a test of archaic admixture in what are referred to as f_3 and f_4-statistics, and D-statistics or ABBA-BABA statistics (Green et al., 2010; Patterson et al., 2012; Reich et al., 2009). D-statistics have been used to detect archaic admixture starting with the draft Neaderthal

genome, while f-statistics were first used to study population history in India. These statistics are distinct but related, and have been used widely to infer the presence or absence of admixture in a set of 3 or 4 populations (Green et al., 2010; Meyer et al., 2012; Prüfer et al., 2014; Reich et al., 2010, 2009; Skoglund and Jakobsson, 2011). The null hypothesis in these methods is no admixture. An excess of either the *en* or *an* site pattern, as in table 1, will result in a deviation of the statistic from 0, and the significance of this deviation can be evaluated at a genome-wide scale using "block-bootstrap" or "block-jackknife" methods.

There is some controversy about results produced by these methods. Initial applications with the low-coverage Denisovan genome found no evidence for Denisovan admixture outside of Melanesia (Meyer et al., 2012; Reich et al., 2010). On the basis of the null results from these D-statistics, estimates of the admixture fraction for Denisova were limited to Melanesian populations. Later analyses using the low-coverage Denisovan genome with additional SNP array data, and using the high-coverage Denisovan genome, have found low levels of Denisovan admixture in East Asian populations (Prüfer et al., 2014; Skoglund and Jakobsson, 2011), some of which appears to have been under positive selection (Huerta-Sánchez et al., 2014).

Several published methods use the site pattern principle to estimate the fraction of archaic genes in modern populations (Durand et al., 2011; Green et al., 2010; Meyer et al., 2012; Patterson et al., 2012; Reich et al., 2010, 2011, 2009; Rogers and Bohlender, 2015).

What fraction of modern DNA derives from archaic admixture?

In the decade before 2010, various geneticists argued that archaics had contributed genes to modern humans. This evidence was based either on the site frequency spectrum (Eswaran et al., 2005) or on LD and deep divergence (Cox et al., 2008; Evans et al., 2005; Garrigan et al., 2005a,b; Harris and Hey, 1999; Hawks et al., 2006; Ziętkiewicz et al., 2003). The archaic origin of at least two of these loci has been confirmed by comparison with the Neanderthal genome (Hammer et al., 2011, p. 15126; Sankararaman et al., 2014; Yotova et al., 2011). These methods are especially valuable because they allow us to detect admixture even in the absence of DNA from archaic fossils (Hammer et al., 2011). Second, LD-based methods allow us to identify introgressed alleles and thus to study how these function in modern humans.

It is also natural to wonder what fraction of modern DNA derives from archaic admixture. This *admixture fraction* can be estimated in several ways. One approach fits simulation models to statistics that measure LD and deep divergence, as discussed just above. This approach suggests that the admixture fraction is about 5% in Europe (Plagnol and Wall, 2006, p. 976) and 2% in Africa (Hammer et al., 2011).

Beginning in 2010, more direct estimates became possible using archaic genome sequences. According to recent estimates, Neanderthals contributed 1.5–2.1% to the genomes of modern Europeans and slightly more to those of eastern Eurasians (Prüfer et al., 2014, p. 45).

Neanderthals were not the only archaic species contributing DNA to modern humans. Exavations at Denisova Cave, in Sibera, produced a 40-thousand-year-old finger bone. The genome of this specimen showed that it derives from an archaic population that was related to Neanderthals, but substantially different (Reich et al., 2010). Between 3% and 6% of the DNA in modern Melanesians seems to derive from this "Denisovan" population (Reich et al., 2010, p. 1053; Meyer et al., 2012, p. 223). There is much less Denisovan admixture in other parts of the world, as shown in Fig. 4.

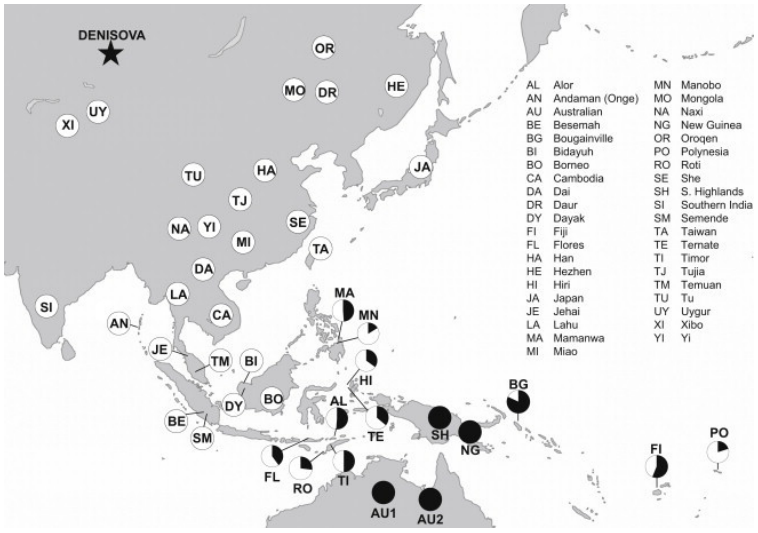

Figure 4: Denisovan admixture most common in Melanesia (Reich et al., 2011, Fig. 1).

How reliable are estimates of the admixture fraction?

Most of these estimates are based on site pattern frequencies, and there are two sources of concern. First, these estimates assume random mating in the ancestral human population and may be biased upwards if the ancestral population was geographically structured, with limited gene flow between regions (Slatkin and Pollack, 2008). This hypothesis of *ancestral subdivision* has been seen as an alternative to that of archaic admixture (Blum and Jakobsson, 2011; Durand et al., 2011; Eriksson and Manica, 2012, 2014; Lowery et al., 2013).

Several lines of evidence argue against this interpretation. First, if the population ancestral to moderns and archaics had been geographically structured, that structure should inflate estimates of the size of that ancestral population. Yet current estimates indicate that moderns and archaics separated during an interval of small effective population size, as shown in Fig. 5.

Second, we can date the introgression of archaic alleles using the length of LD blocks surrounding archaic regions within the genomes of modern humans. These estimates are too young to be consistent with the hypothesis of ancient geographic structure (Sankararaman et al., 2012; Wall et al., 2013).

Third, the genetic differences are rather small between putative Neanderthal segments in modern genomes and the homologous segments in *bona fide* Neanderthals. Under the hypothesis of ancient population structure, these differences should be much larger (Lohse and Frantz, 2014).

Finally, we now have direct evidence from the genome of a 40,000-year-old modern human, who was only a few generations removed from a Neanderthal ancestor (Fu et al., 2015).

These findings make it clear that archaic DNA has introgressed into the genomes of modern humans. They do not, however, imply that existing estimates of the admixture fraction are unbiased. Current estimators rely on detailed assumptions about the tree of populations. Most of them also assume that only one archaic population contributed genes to the modern one. When these assumptions are violated, strong biases can result.

Ghost admixture (Beerli, 2004; Slatkin, 2005; Durand et al., 2011, p. 2240; Harris and Nielsen, 2013) refers to the case in which a modern population receives DNA from several archaics, only one of which is acknowledged by the statistical model. Fig. 6 shows how ghost admixture can inflate the count of the *en* site pattern, leading to an overestimate of Neanderthal admixture.

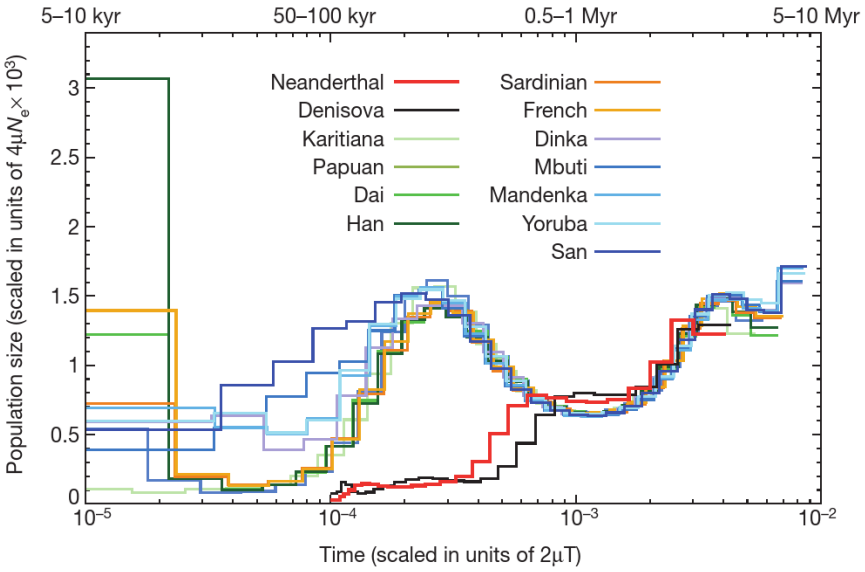

Figure 5: Neanderthal and Denisovan populations were very small (Prüfer et al., 2014, Fig. 4).

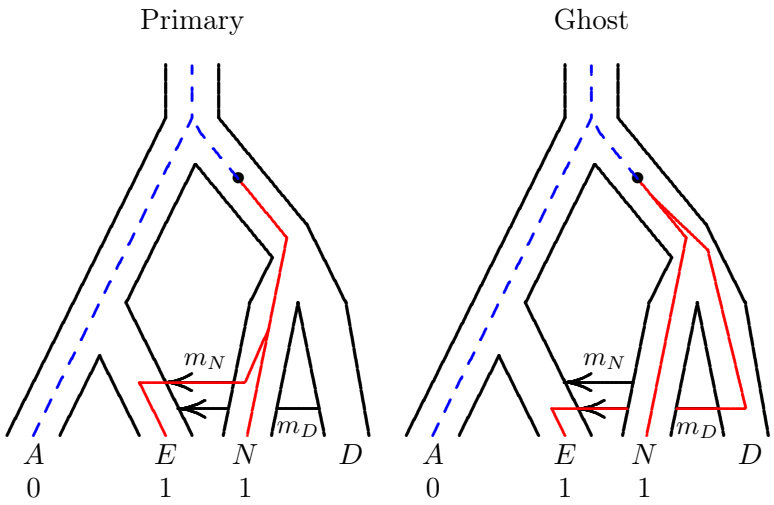

Figure 6: Primary and secondary (ghost) admixture both inflate the *en* site pattern. Key: A, Africa; E, Eurasia; N, Neanderthal; D, Denisova; solid red, derived; dashed blue, ancestral. After Rogers and Bohlender (2015, Fig. 2).

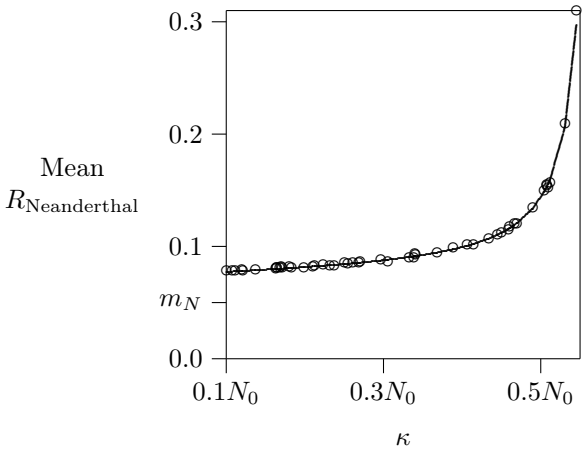

Figure 7: Response of mean $R_{\text{Neanderthal}}$ to the separation time, κ, of Neanderthals and Denisovans. Circles show simulated values and solid line shows expected value. Simulations assume a primary admixture fraction, $m_N = 0.05$, and a ghost admixture fraction, $m_D = 0.025$. The separation time (κ) of Neanderthals and Denisovans is expressed in units of N_0 generations, where N_0 is the size of the population ancestral to moderns, Neanderthals, and Denisovans. After Rogers and Bohlender (2015, Fig. 4).

It is not surprising that specification error could cause bias, yet there are two surprises here. The first is that ghost admixture makes many of these estimators sensitive to ancient population sizes and to the times at which populations separated. These parameters are poorly constrained, so it is difficult to correct these biases. The second surprise is the sheer magnitude of the effect.

Fig. 7 shows shows bias in the estimator $R_{\text{Neanderthal}}$ (Patterson et al., 2010b), which we abbreviate as R_N. The figure considers the case in which there is 5% primary admixture, m_N, and half that much ghost admixture. The horizontal axis shows the separation time, κ, of Neanderthals and Denisovans. If κ is small, R_N is 7.5% and equals the total rate of archaic admixture, including Denisovans as well as Neanderthals. But if κ is large, R_N is 6× the fraction of Neanderthal admixture and 4× the total archaic admixture. These are enormous biases. Different estimators respond to ghost admixture in different ways, but nearly all of them are sensitive.

It is fortunate that the different estimators respond differently to these biases, because this makes

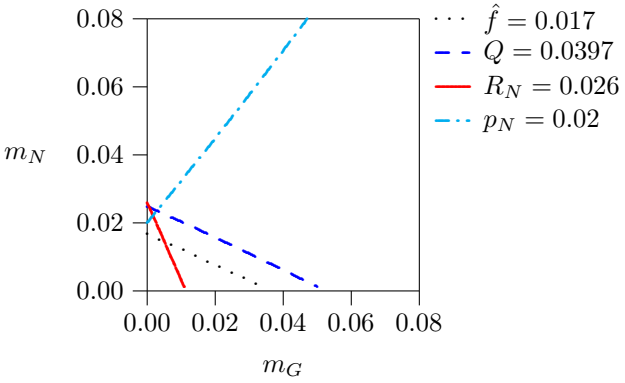

Figure 8: Archaic admixture in France, as implied by various published estimates. m_N is primary (Neanderthal) admixture, as a function of ghost admixture, m_G. After Rogers and Bohlender (2015, Fig. 8).

it possible to construct graphs such as Fig. 8, which describes archaic admixture into the French population. Each estimator plots as a line, which connects the values of Neanderthal admixture (m_N) and ghost admixture (m_G) that are consistent with a given statistic. In an ideal world, these lines would intersect at a point corresponding to the true values of m_G and m_N. They nearly intersect at the left edge of the graph, suggesting that about 2% of French DNA derives from Neanderthals and little or none derives from other archaic populations.

It has not yet been possible to do this sort of analysis in other populations. Yet there is abundant evidence that other populations in Eurasia and Oceania have received gene flow from more than one archaic population. Consequently, the biases discussed above should be important, and there is little basis for confidence in current estimates of the admixture fraction outside of Europe.

The distribution of archaic segments in Eurasian genomes

Fig. 9 shows the distribution of Neanderthal segments in the genomes of modern Europeans (blue lines) and Asians (red lines). Notice that these are not evenly distributed. There are regions with lots of Neanderthal DNA and regions with little or none. This figure is from Vernot and Akey (2014), but Khrameeva et al. (2014) and Sankararaman et al. (2014) documented the same phenomenon.

Why should archaic DNA be so unevenly distributed across the genome? Vernot and Akey (2014) proposed that Neanderthal DNA is absent from some regions because it has been purged by natural selection. Several observations support this hypothesis. Regions depleted of archaic DNA tend to occur in or near genes and are also reduced on the X chromosome, which is often involved in hybrid sterility (Sankararaman et al., 2014, p. 354). Neanderthal segments are least common in genomic regions where the two species (moderns and Neanderthals) are most dissimilar (Vernot and Akey, 2014). Furthermore, Neanderthal segments are rare in genomic regions that are likely to be under strong selection (Sankararaman et al., 2014). Finally, the genomes of several ancient Eurasians, who lived more than 35 ky ago, had more archaic DNA that do those of any modern human (Fu et al., 2015, 2014; Seguin-Orlando et al., 2014).

These results suggest that the process of introgression was not random. Some portions of the Neanderthal genome were removed by natural selection from the genomes of modern humans. Sankararaman et al. (2014, p. 356) estimate that the Neanderthal admixture fraction must have originally been at least half again as high as that estimated from modern samples. All this selection must have been costly: hybrid individuals must have been either less healthy or less fertile than the outbred individuals around them. These are the mechanisms that, in more exaggerated form, separate distinct species—lions from tigers and horses from donkeys. The process of speciation was already well underway when modern humans encountered Neanderthals.

More Neanderthal genes in Asia and America than Europe

Not only are archaic genes unevenly distributed within the human genome; they are also more common in some populations than others. One might expect to find more Neanderthal DNA in regions where Neanderthals once lived, such as Europe and Central Asia. Yet Neanderthal DNA is most common in the genomes of East Asians, as shown in Fig. 10 (Prüfer et al., 2014, p. 45; Green et al., 2010; Skoglund and Jakobsson, 2011; Wall et al., 2013; Vernot and Akey, 2014; Vernot and Akey, 2015; Meyer et al., 2012).

There are several conceivable explanations for this excess. First, East Asians might have received a second pulse of Neanderthal admixture, which did not affect Europeans. Second, Europeans might have received a second pulse of admixture from a population that had not interbred with

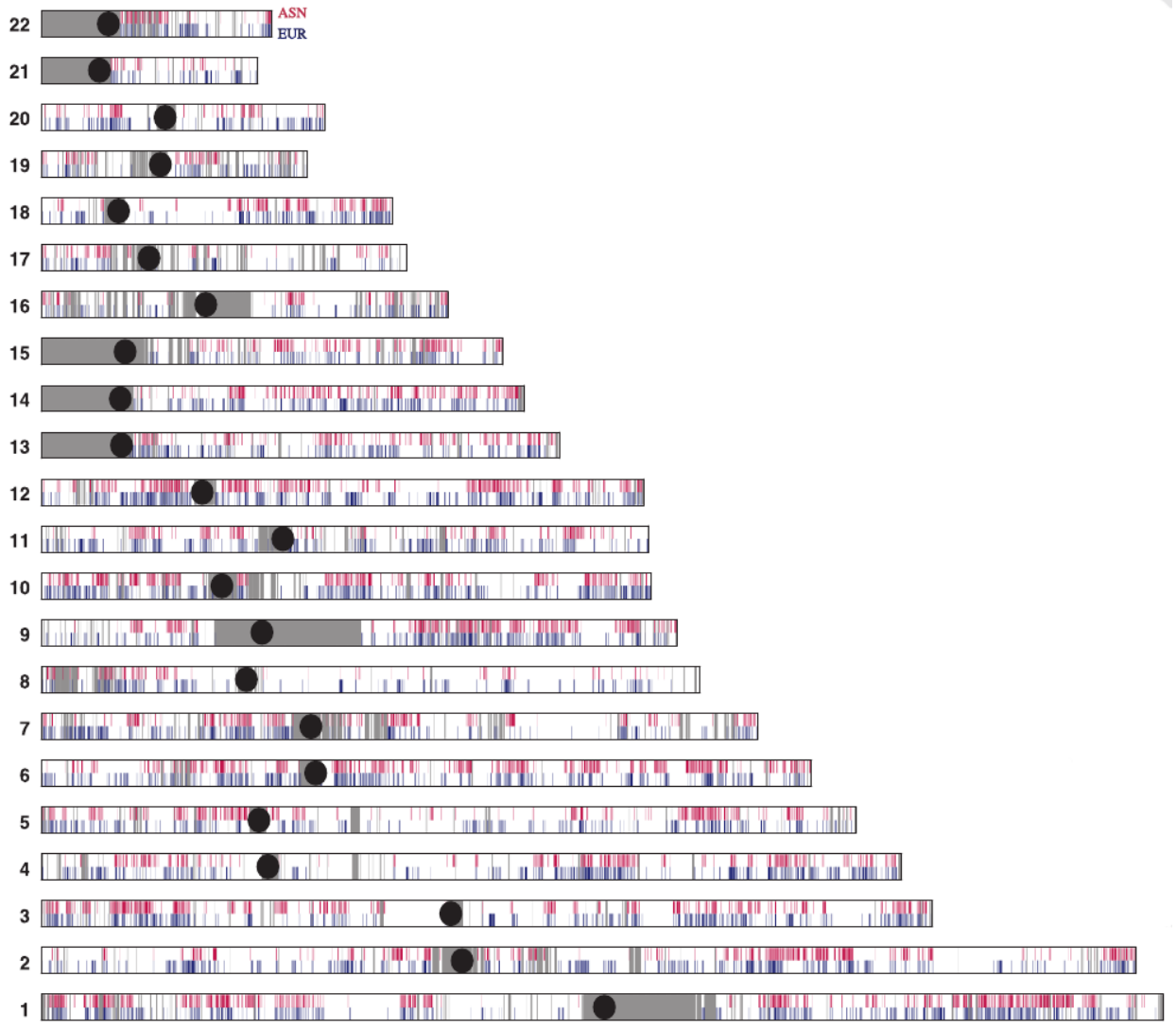

Figure 9: Neanderthal segments in Eurasian genomes. Each row represents a different chromosome. Key: blue, Europe; red, Asia (Vernot and Akey, 2014).

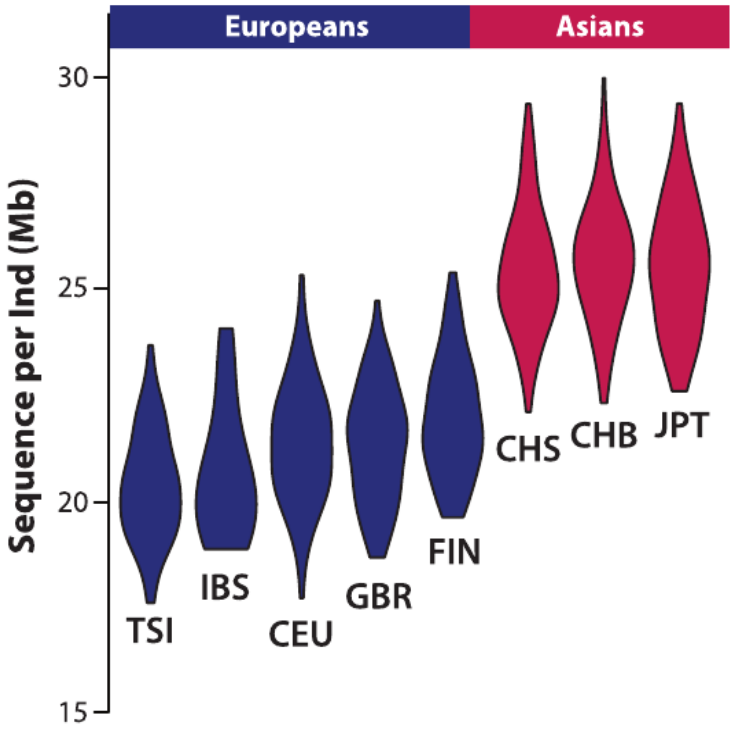

Figure 10: There is more Neanderthal DNA in Asians than in Europeans.(Vernot and Akey, 2014).

Neanderthals. Finally, selection might have been more successful at removing Neanderthal DNA from European genomes than from Asian ones (Sankararaman et al., 2014, p. 356).

This last hypothesis is plausible, because some evidence suggests that the European population was larger than the Asian one for some time after their separation. In the smaller Asian population, purifying selection would have been less effective. Yet this hypothesis has not fared well in recent studies. If selection were responsible for the East Asian excess, we would expect that excess to be largest in genomic regions under strong selection. This is not the case (Kim and Lohmueller, 2015; Vernot and Akey, 2015). Thus, the East Asian excess must reflect either a second pulse of Neanderthal DNA into East Asia or a pulse of Neanderthal-free DNA into Europe.

Archaic adaptations in modern humans

As we emphasized above, some archaic DNA was deleterious and has been purged from the genomes of modern humans. Most of the rest was probably neutral. Yet it is abundantly clear that some archaic DNA was beneficial and underlies modern human adaptations.

For example, life is hard at high altitude, because there is so little oxygen. Yet there are humans who live year-round at very high altitudes. These populations have physiological adaptations that make this possible. The details of these adaptations differ from place to place. One might expect that all high-altituded adaptations would involve an increased supply of hemoglobin—the molecule that transports oxygen. This is indeed the case for most of us. When we move to high altitude, we begin producing more hemoglobin. Yet in Tibetan populations, adaptation to altitude is accomplished by *lowering* the concentration of hemoglobin (Simonson et al., 2010).

This adaptation reflects the actions of several of genes. One of these, EPAS1, is a transcription

factor that becomes active when oxygen is in short supply. At this locus, one haplotype is associated with adaptation to high altitude and is found almost exclusively in Tibet. This high-altitude haplotype differs at many nucleotide sites from all other haplotypes found in modern humans. Yet it is quite similar to the Denisovan haplotype. It seems clear that the adaptation to high altitude in Tibet evolved by adaptive introgression, either from Denisovans or from a related population (Huerta-Sánchez et al., 2014).

In several populations, selection has favored introgressed archaic segments that affect skin and hair (Vernot and Akey, 2014; Sankararaman et al., 2014; Hu et al., 2015, table 2). Europeans may have gotten their freckles from Neanderthals (Racimo et al., 2015, p. 9).

In Europeans, selection has favored Neanderthal-like alleles in genes that break down lipids (Khrameeva et al., 2014).

The most dramatic evidence for adaptive introgression, however, involves the immune system. This makes sense, because our immune systems are in an evolutionary arms race with pathogens. Pathogens evolve much faster than humans and tend to evolve adaptations to whichever immune variants are currently most common. For this reason, rare alleles have an advantage at human immunity loci. Introgressed alleles are initially rare, so those at immunity loci ought to have a selective advantage. Immunity loci ought to be overrepresented in introgressed segments of the human genome.

This hypothesis is hard to test, because of the way this process distorts variation. When the rare allele is always favored, alleles are seldom lost, and polymorphisms can endure for millions of years. Some loci of the Major Histocompatibility Complex (MHC) have alleles that are more similar to alleles in chimpanzees or gorillas than to other human alleles. This is not because of introgression, but because these alleles are older than the human, chimpanzee, and gorilla species. So we cannot infer introgression simply from shared alleles. Fortunately, there is another signal.

The MHC system consists of several loci, one of which is called "HLA-B." This locus is highly polymorphic and has over a thousand alleles. One of these, the HLA-B*73.01 allele, differs at many nucleotide sites from other human alleles. It is thought to have separated about 16 million years ago. Yet all copies of this allele are remarkably similar to each other. This combination of ancient divergence plus modern homogeneity suggests that HLA-B*73.01 introgressed recently from some other population. In addition, HLA-B*73.01 sits on a long block of LD, about 1.3 Mb in length. This extensive LD also suggests that the allele was recently introduced into the human population. Finally, this LD block includes an allele, HLA-C*15, at a different MHC locus, and this other allele is also found in the Denisovan genome (Abi-Rached et al., 2011). All in all, it seems quite likely that this haplotype is of archaic origin. Abi-Rached et al. (2011) estimate that over 50% of Eurasian MHC alleles came from archaics.

Another example involves the OAS1 locus, which is part of the innate immune system. In Melanesia, there are two major alleles: one shared with the rest of the world, and the other unique to Melanesia (or nearly so). The two alleles differ at many nucleotide sites, implying that they have been separate for a long time—perhaps 3.4 million years. Yet there is extensive LD in this region, stretching for 90 kb. This long block of LD implies that the Melanesian allele has been in the modern population for only about 25 thousand years. Furthermore, the Melanesian allele is a close match to the Denisovan genome. This is another clear case of archaic admixture affecting an immunity locus (Mendez et al., 2012a).

Another allele at this locus seems to have come from Neanderthals and is widely distributed in Eurasia (Mendez et al., 2013). Several other introgressed alleles involve other components of the innate immune system (Dannemann et al., 2015; Mendez et al., 2012b). All in all, it seems clear that archaic admixture has had a major effect on the immune systems of modern Eurasians.

Table 2: Neanderthals had low heterozygosity (Castellano et al., 2014).

Species	Population	Relative Heterozygosity†
Neanderthal	El Sidrón	0.28
	Vindija	0.25
	Altai	0.22
Modern	African	1.00
	European	0.76
	Asian	0.71

†Heterozygosity relative to modern Africa.

Health implications of archaic introgression

DNA from archaic populations may also underlie health problems in some human populations. In Oceania, a Neanderthal-like haplotype at the FTO locus predisposes its carriers to obesity. On average, each copy of the obesity-risk allele added 1.6 kg/m^2 to the body-mass index (BMI) of carriers (Naka et al., 2013, p. 1207). A Neanderthal-like haplotype at the SLC16A11 locus predisposes its carriers to type 2 diabetes (The SIGMA Type 2 Diabetes Consortium, 2014). Each copy of this haplotype increases the risk of diabetes by about 20%. This haplotype has a frequency of roughly 50% in Native Americans and 10% in East Asians but is rare or absent elsewhere. It may contribute to the elevated prevalence of type 2 diabetes in Latin America.

There is a common thread here. These deleterious archaic alleles all affect either obesity or the risk of diabetes—disorders of energy storage. It is possible that archaic populations had evolved "thrifty genotypes," which prepare for food shortages by improved energy storage (Neel, 1962). Such genotypes might have been favored by selection during the colonization of Oceania and the New World (Naka et al., 2013, p. 1209).

Consequences of small archaic population size

When a population is reduced in size, it rapidly loses genetic variation. This seems to have happened to both of the archaic populations for which we have data. This effect is seen in table 2, which shows the heterozygosity of several populations relative to that of modern Africans. The heterozygosity of modern Europeans is about 3/4 that of modern Africans, and that of Neanderthals was about 1/4 as large.

It is also seen in Fig. 5, which estimates the size of various populations far back into the past. Prior to about a million years ago, all populations were of similar size. This suggests that the ancestors of all humans were at that time a single population. Shortly after a million years ago, the population ancestral to modern humans began to grow, while the archaic populations entered a steep decline. For several hundred thousand years, until they finally went extinct, the archaic populations were very small.

The genome of the Altai Neanderthal illustrates what life must have been like in these tiny populations. The genome has long stretches of very low heterozygosity, which imply recent close inbreeding. This individual may have been the offspring of a mating between grandfather and granddaughter or uncle and niece. Even when we exclude the effects of this recent inbreeding, the remaining portions of the genome still suggest a long history of small population size (Prüfer et al., 2014).

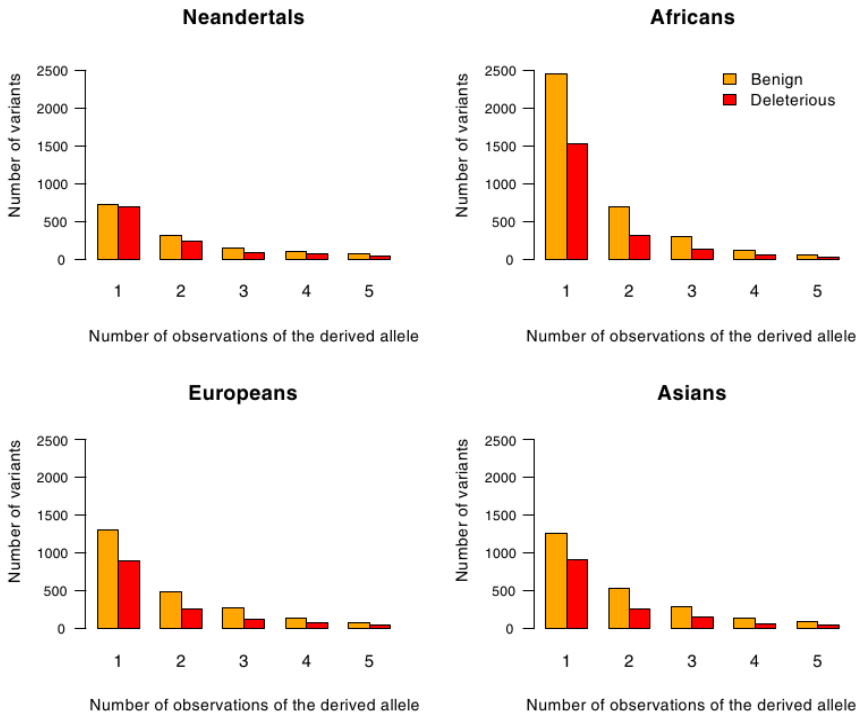

Figure 11: Selection in Neanderthals was less effective at removing deleterious mutations (Castellano et al., 2014).

We can anticipate the consequences of such an extended period of small population size. In large populations, natural selection is constantly weeding out the deleterious alleles that arise each generation by mutation. Most of the mutations that spread through the population have little or no effect on proteins and are selectively neutral. Thus, selection acts primarily as a filter, which prevents the spread of deleterious alleles.

This filter is less effective in small populations, because the random force of genetic drift is strong relative to selection. For this reason, deleterious alleles are more likely to spread in small populations than in large ones. Because the archaic populations were so small, it is likely that deleterious mutations accumulated within them.

This is seen in Fig. 11, which plots the distributions of benign and deleterious alleles within each of several frequency classes. Consider first the upper right panel, which refers to Africans. The two bars at the left side of this plot show the number of polymorphic nucleotide sites at which the rarer of the two alleles is present only once in the sample. The orange bar counts variants that are thought to be benign, and the red bar counts those that are thought to be deleterious. As you can see, the red bar is shorter than the orange bar, indicating that benign variants are more common than deleterious ones. This is *not* because most mutations are benign. Indeed, most mutations are probably harmful. The red bar is not shorter because harmful mutations didn't occur, it's shorter because they've been removed by selection.

The other bars in the African panel refer to different frequency categories—those in which the rare allele is present twice, three times, and so on. In each case, the red bar is shorter than the orange bar, indicating that harmful alleles have been removed by selection.

Now look at the panels for Europeans and Asians. The pattern is qualitatively the same, but the difference between the orange and red bars is not as large. This is because Europe and Asia

had (until recently) smaller populations than Africa, so selection was somewhat less effective at removing harmful alleles.

Now look at the Neanderthal panel. The bars are much shorter, reflecting the fact that there is less genetic variation in small populations. But the big news is in the relative sizes of the red and orange bars. For Neanderthals, these bars are nearly equal in size. This indicates that selection was much less effective at removing harmful alleles. Neanderthals must have been unhealthy even before modern humans arrived in Eurasia. This may have contributed to their demise.

Conclusions

In the past decade, it has become clear that archaic hominins did not go altogether extinct. They interbred with modern humans, and some of their DNA lives on in modern human populations. There are various estimates of the fraction of archaic DNA in our genomes, ranging from roughly 2 to 5%. These estimates, however, are subject to various biases and may be revised in the future. There is little doubt, however, that archaic DNA is widespread within modern humans.

It is also clear that archaic DNA was selected against in many parts of the human genome. Nonetheless, some archaic genes were beneficial and contribute in important ways to modern human adaptations.

The Neanderthals themselves, however, were probably not well adapted. A long history of small population size allowed deleterious alleles to accumulate and may have contributed to their extinction.

References

L. Abi-Rached, M. J. Jobin, S. Kulkarni, A. McWhinnie, K. Dalva, L. Gragert, F. Babrzadeh, B. Gharizadeh, M. Luo, F. A. Plummer, et al. The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science*, 334(6052):89–94, 2011.

Peter Beerli. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, 13(4):827–836, 2004.

Michael GB Blum and Mattias Jakobsson. Deep divergences of human gene trees and models of human origins. *Molecular Biology and Evolution*, 28(2):889–898, 2011.

Günter Bräuer. A craniological approach to the origin of anatomically modern Homo sapiens in Africa and implications for the appearance of modern Europeans. In F.H. Smith and F. Spencer, editors, *The Origins of Modern Humans: A World Survey of the Fossil Evidence*, pages 327–410. Alan R. Liss, New York, 1984.

Günter Bräuer. The evolution of modern humans: A comparison of the African and non-African evidence. In Paul Mellars and Chris Stringer, editors, *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, pages 123–154. Princeton University Press, Princeton, New Jersey, 1989.

Sergi Castellano, Genís Parra, Federico A Sánchez-Quinto, Fernando Racimo, Martin Kuhlwilm, Martin Kircher, Susanna Sawyer, Qiaomei Fu, Anja Heinze, Birgit Nickel, et al. Patterns of coding variation in the complete exomes of three Neandertals. *Proceedings of the National Academy of Sciences, USA*, 111(18):6666–6671, 2014.

Murray P. Cox, Fernando L. Mendez, Tatiana M. Karafet, Maya Metni Pilkington, Sarah B. Kingan, Giovanni Destro-Bisol, Beverly I. Strassmann, and Michael .F Hammer. Testing for archaic hominin admixture on the X chromosome: Model likelihoods for the modern human RRM2P4 region from summaries of genealogical topology under the structured coalescent. *Genetics*, 178 (1):427–437, 2008.

Michael Dannemann, Aida M. Andrés, and Janet Kelso. Adaptive variation in human toll-like receptors is contributed by introgression from both Neandertals and Denisovans. *bioRxiv*, 2015.

Eric Y Durand, Nick Patterson, David Reich, and Montgomery Slatkin. Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, 28(8):2239–2252, 2011.

Anders Eriksson and Andrea Manica. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *Proceedings of the National Academy of Sciences, USA*, 109(35):13956–13960, 2012.

Anders Eriksson and Andrea Manica. The doubly-conditioned frequency spectrum does not distinguish between ancient population structure and hybridisation. *Molecular Biology and Evolution*, 31(6):1618–1621, 2014.

Vinayak Eswaran, Henry Harpending, and Alan R. Rogers. Genomics refutes an exclusively African origin of humans. *Journal of Human Evolution*, 49:1–18, 2005.

Patrick D. Evans, Sandra L. Gilbert, Nitzan Mekel-Bobrov, Eric J. Vallender, Jeffrey R. Anderson, Leila M. Vaez-Azizi, Sarah A. Tishkoff, Richard R. Hudson, and Bruce T. Lahn. Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science*, 309:1717–1720, 2005.

Qiaomei Fu, Mateja Hajdinjak, Oana Teodora Moldovan, Silviu Constantin, Swapan Mallick, Pontus Skoglund, Nick Patterson, Nadin Rohland, Iosif Lazaridis, Birgit Nickel, Bence Viola, Kay Prufer, Matthias Meyer, Janet Kelso, David Reich, and Svante Paabo. An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, 2015.

Qiaomei Fu, Heng Li, Priya Moorjani, Flora Jay, Sergey M Slepchenko, Aleksei A Bondarev, Philip LF Johnson, Ayinuer Aximu-Petri, Kay Prüfer, Cesare de Filippo, Matthias Meyer, Nicolas Zwyns, Domingo C. Salazar-García, Yaroslav V. Kuzmin, Susan G. Keates, Pavel A. Kosintsev, Dmitry I. Razhev, Michael P. Richards, Nikolai V. Peristov, Michael Lachmann, Katerina Douka, Thomas F. G. Higham, Montgomery Slatkin, Jean-Jacques Hublin, David Reich, Janet Kelso, T. Bence Viola, and Svante Pääbo. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*, 514(7523):445–449, 2014.

D. Garrigan, Z. Mobasher, S. B. Kingan, J. A. Wilder, and M. F. Hammer. Deep haplotype divergence and long-range linkage disequilibrium at Xp21.1 provide evidence that humans descend from a structured ancestral population. *Genetics*, 170(4):1849–1856, 2005a.

Daniel Garrigan, Zahra Mobasher, Tesa Severson, Jason A. Wilder, and Michael F. Hammer. Evidence for archaic Asian ancestry on the human X chromosome. *Molecular Biology and Evolution*, 22:189–192, 2005b.

Richard E. Green, Johannes Krause, Adrian W. Briggs, Tomislav Maricic, Udo Stenzel, Martin Kircher, Nick Patterson, Heng Li, Weiwei Zhai, Markus Hsi-Yang Fritz, Nancy F. Hansen,

Eric Y. Durand, Anna-Sapfo Malaspinas, Jeffrey D. Jensen, Tomas Marques-Bonet, Can Alkan, Kay Prüfer, Matthias Meyer, Hernán A. Burbano, Jeffrey M. Good, Rigo Schultz, Ayinuer Aximu-Petri, Anne Butthof, Barbara Höber, Barbara Höffner, Madlen Siegemund, Antje Weihmann, Chad Nusbaum, Eric S. Lander, Carsten Russ, Nathaniel Novod, Jason Affourtit, Michael Egholm, Christine Verna, Pavao Rudan, Dejana Brajkovic, Željko Kucan, Ivan Gušic, Vladimir B. Doronichev, Liubov V. Golovanova, Carles Lalueza-Fox, Marco de la Rasilla, Javier Fortea, Antonio Rosas, Ralf W. Schmitz, Philip L. F. Johnson, Evan E. Eichler, Daniel Falush, Ewan Birney, James C. Mullikin, Montgomery Slatkin, Rasmus Nielsen, Janet Kelso, Michael Lachmann, David Reich, and Svante Pääbo. A draft sequence of the Neandertal genome. *Science*, 328(5979):710–722, 2010.

Michael F. Hammer, August E. Woerner, Fernando L. Mendez, Joseph C. Watkins, and Jeffrey D. Wall. Genetic evidence for archaic admixture in Africa. *Proceedings of the National Academy of Sciences, USA*, 108(37):15123–15128, 2011.

Rosalind M Harding, Stephanie Malia Fullerton, Robert C Griffiths, Jacquelyn Bond, Martin J Cox, Julie A Schneider, Danielle S Moulin, and JB Clegg. Archaic African and Asian lineages in the genetic ancestry of modern humans. *American Journal of Human Genetics*, 60(4):772, 1997.

Eugene E. Harris and Jody Hey. X chromosome evidence for ancient human histories. *Proceedings of the National Academy of Sciences, USA*, 96:3320–3324, 1999. URL `http://www.pnas.org/cgi/content/full/96/6/3320`.

Kelley Harris and Rasmus Nielsen. Inferring demographic history from a spectrum of shared haplotype lengths. *PLoS Genetics*, 9(6):e1003521, 2013.

John Hawks, Gregory Cochran, Henry C. Harpending, and Bruce T. Lahn. A genetic legacy from archaic *Homo*. *Trends in Genetics*, 24(1):19–23, 2006.

Ya Hu, Qiliang Ding, Yi Wang, Shuhua Xu, Yungang He, Minxian Wang, Jiucun Wang, and Li Jin. Investigating the evolutionary importance of Denisovan introgressions in Papua New Guineans and Australians. *bioRxiv*, 2015.

Emilia Huerta-Sánchez, Xin Jin, Asan, Zhuoma Bianba, Benjamin M. Peter, Nicolas Vinckenbosch, Yu Liang, Xin Yi, Mingze He, Mehmet Somel, Peixiang Ni, Bo Wang, Xiaohua Ou, Huasang, Jiangbai Luosang, Zha Xi Ping Cuo, Kui Li, Guoyi Gao, Ye Yin, Wei Wang, Xiuqing Zhang, Xun Xu, Huanming Yang, Yingrui Li, Jian Wang, Jun Wang, and Rasmus Nielsen. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 512(7513):194–197, 2014.

Ekaterina E. Khrameeva, Katarzyna Bozek, Liu He, Zheng Yan, Xi Jiang, Yuning Wei, Kun Tang, Mikhail S. Gelfand, Kay Prufer, Janet Kelso, Svante Paabo, Patrick Giavalisco, Michael Lachmann, and Philipp Khaitovich. Neanderthal ancestry drives evolution of lipid catabolism in contemporary Europeans. *Nature Communications*, 5(3584), Apr 2014.

Bernard Kim and Kirk Lohmueller. Selection and reduced population size cannot explain higher amounts of Neandertal ancestry in East Asian than in European human populations. *American Journal of Human Genetics*, Feb 2015.

M. Krings, A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking, and Svante Pääbo. Neandertal DNA sequences and the origin of modern humans. *Cell*, 90(1):19–30, 1997.

Marta Mirazón Lahr and Robert Foley. Towards a theory of modern human origins: Geography, demography, and diversity in recent human evolution. *Yearbook of Physical Anthropology*, 41: 137–176, 1998.

Konrad Lohse and Laurent AF Frantz. Neandertal admixture in Eurasia confirmed by maximum-likelihood analysis of three genomes. *Genetics*, 196(4):1241–1251, 2014.

Robert K Lowery, Gabriel Uribe, Eric B Jimenez, Mark A Weiss, Kristian J Herrera, Maria Regueiro, and Rene J Herrera. Neanderthal and Denisova genetic affinities with contemporary humans: Introgression versus common ancestral polymorphisms. *Gene*, 530(1):83–94, 2013.

Elisabeth J. Manderscheid and Alan R. Rogers. Genetic admixture in the late Pleistocene. *American Journal of Physical Anthropology*, 100(1):1–5, 1996.

Fernando L Mendez, Joseph C Watkins, and Michael F Hammer. Global genetic variation at OAS1 provides evidence of archaic admixture in Melanesian populations. *Molecular Biology and Evolution*, 29(6):1513–1520, 2012a.

Fernando L Mendez, Joseph C Watkins, and Michael F Hammer. A haplotype at STAT2 introgressed from Neanderthals and serves as a candidate of positive selection in Papua New Guinea. *American Journal of Human Genetics*, 91(2):265–274, 2012b.

Fernando L Mendez, Joseph C Watkins, and Michael F Hammer. Neandertal origin of genetic variation at the cluster of OAS immunity genes. *Molecular Biology and Evolution*, 30(4):798–801, 2013.

Matthias Meyer, Martin Kircher, Marie-Theres Gansauge, Heng Li, Fernando Racimo, Swapan Mallick, Joshua G Schraiber, Flora Jay, Kay Prüfer, Cesare de Filippo, Peter H. Sudmant, Can Alkan, Qiaomei Fu, Ron Do, Nadin Rohland, Arti Tandon, Michael Siebauer, Richard E. Green, Katarzyna Bryc, Adrian W. Briggs, Udo Stenzel, Jesse Dabney, Jay Shendure, Jacob Kitzman, Michael F. Hammer, Michael V. Shunkov, Anatoli P. Derevianko, Nick Patterson, Aida M. Andrés, Evan E. Eichler, Montgomery Slatkin, David Reich, Janet Kelso, and Svante Pääbo. A high-coverage genome sequence from an archaic Denisovan individual. *Science*, 338(6104):222–226, 2012.

Priya Moorjani, Nick Patterson, Joel N Hirschhorn, Alon Keinan, Li Hao, Gil Atzmon, Edward Burns, Harry Ostrer, Alkes L Price, and David Reich. The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genetics*, 7(4):e1001373, 2011.

I. Naka, K. Hikami, K. Nakayama, M. Koga, N. Nishida, R. Kimura, T. Furusawa, K. Natsuhara, T. Yamauchi, M. Nakazawa, Y. Ataka, T. Ishida, T. Inaoka, S. Iwamoto, Y. Matsumura, R. Ohtsuka, N. Tsuchiya, and J. Ohashi. A functional SNP upstream of the beta-2 adrenergic receptor gene (ADRB2) is associated with obesity in Oceanic populations. *International Journal of Obesity*, 37(9):1204–1210, Sep 2013. ISSN 0307-0565.

James V. Neel. Diabetes mellitus: A "thrifty" genotype rendered detrimental by "progress"? *American Journal of Human Genetics*, 14(4):353–362, 1962.

Magnus Nordborg. On the probability of Neanderthal ancestry. *American Journal of Human Genetics*, 63(4):1237–1240, 1998.

Pekka Pamilo and Masatoshi Nei. Relationships between gene trees and species trees. *Molecular Biology and Evolution*, 5(5):568–583, 1988.

Nick Patterson, Nancy Hansen, James Mullikin, and David Reich. Neandertals are more closely related to non-Africans than to Africans. Supplementary Note 15 of Green et al. Green et al. (2010), 2010a.

Nick Patterson, Heng Li, Swapan Mallick, and David Reich. Denisovans share more derived alleles with Melanasians than with other groups. Supplementary Note 8 of Reich et al. Reich et al. (2010), 2010b.

Nick Patterson, Priya Moorjani, Yontao Luo, Swapan Mallick, Nadin Rohland, Yiping Zhan, Teri Genschoreck, Teresa Webster, and David Reich. Ancient admixture in human history. *Genetics*, 192(3):1065–1093, 2012.

Vincent Plagnol and Jeffrey D Wall. Possible ancestral structure in human populations. *PLoS Genetics*, 2(7):e105, 2006.

Kay Prüfer, Fernando Racimo, Nick Patterson, Flora Jay, Sriram Sankararaman, Susanna Sawyer, Anja Heinze, Gabriel Renaud, Peter H Sudmant, Cesare de Filippo, Heng Li, Swapan Mallick, Michael Dannemann, Qiaomei Fu, Martin Kircher, Martin Kuhlwilm, Michael Lachmann, Matthias Meyer, Matthias Ongyerth, Michael Siebauer, Christoph Theunert, Arti Tandon, Priya Moorjani, Joseph Pickrell, James C. Mullikin, Samuel H. Vohr, Richard E. Green, Ines Hellmann, Philip L. F. Johnson, Hélène Blanche, Howard Cann, Jacob O. Kitzman, Jay Shendure, Evan E. Eichler, Ed S. Lein, Trygve E. Bakken, Liubov V. Golovanova, Vladimir B. Doronichev, Michael V. Shunkov, Anatoli P. Derevianko, Bence Viola, Montgomery Slatkin, David Reich, Janet Kelso, and Svante Pääbo. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 505(7481):43–49, 2014.

Fernando Racimo, Sriram Sankararaman, Rasmus Nielsen, and Emilia Huerta-Sanchez. Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, 2015.

Floyd A Reed and Sarah A Tishkoff. African human diversity, origins and migrations. *Current Opinion in Genetics and Development*, 16(6):597–605, 2006.

D. Reich, R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. F. Johnson, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*, 468(7327):1053–1060, 2010.

David Reich, Nick Patterson, Martin Kircher, Frederick Delfin, Madhusudan R Nandineni, Irina Pugach, Albert Min-Shan Ko, Ying-Chin Ko, Timothy A Jinam, Maude E Phipps, Naruya Saitou, Andreas Wollstein, Manfred Kayser, Svante Pääbo, and Mark Stoneking. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *American Journal of Human Genetics*, 89(4):516–528, 2011.

David Reich, Kumarasamy Thangaraj, Nick Patterson, Alkes L Price, and Lalji Singh. Reconstructing Indian population history. *Nature*, 461(7263):489–494, 2009.

Alan R. Rogers and Ryan J. Bohlender. Bias in estimators of archaic admixture. *Theoretical Population Biology*, 100:63–78, March 2015. ISSN 0040-5809.

Sriram Sankararaman, Swapan Mallick, Michael Dannemann, Kay Prüfer, Janet Kelso, Svante Pääbo, Nick Patterson, and David Reich. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*, 507(7492):354–357, 2014.

Sriram Sankararaman, Nick Patterson, Heng Li, Svante Pääbo, and David Reich. The date of interbreeding between Neandertals and modern humans. *PLoS Genetics*, 8(10):e1002947, 2012.

Andaine Seguin-Orlando, Thorfinn S Korneliussen, Martin Sikora, Anna-Sapfo Malaspinas, Andrea Manica, Ida Moltke, Anders Albrechtsen, Amy Ko, Ashot Margaryan, Vyacheslav Moiseyev, et al. Genomic structure in Europeans dating back at least 36,200 years. *Science*, 346(6213):1113–1118, 2014.

David Serre, André Langaney, Mario Chech, Maria Teschler-Nicola, Maja Paunovic, Philippe Mennecier, Göran Possnert, and Svante Pääbo. No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biology*, 2:0313–0317, 2004.

Tatum S Simonson, Yingzhong Yang, Chad D Huff, Haixia Yun, Ga Qin, David J Witherspoon, Zhenzhong Bai, Felipe R Lorenzo, Jinchuan Xing, Lynn B Jorde, Josef T. Prchal, and RiLi Ge. Genetic evidence for high-altitude adaptation in Tibet. *Science*, 329(5987):72–75, 2010.

Pontus Skoglund and Mattias Jakobsson. Archaic human ancestry in East Asia. *Proceedings of the National Academy of Sciences, USA*, 108(45):18301–18306, 2011.

Montgomery Slatkin. Seeing ghosts: The effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology*, 14(1):67–73, 2005.

Montgomery Slatkin and Joshua L Pollack. Subdivision in an ancestral species creates asymmetry in gene trees. *Molecular Biology and Evolution*, 25(10):2241–2246, 2008.

F. H. Smith, A. B. Falsetti, and S. M. Donnelly. Modern human origins. *Yearbook of Physical Anthropology*, 32:35–68, 1989.

Mark Stoneking. DNA and recent human evolution. *Evolutionary Anthropology*, 2(2):60–73, 1993.

C. B. Stringer and Peter Andrews. Genetic and fossil evidence for the origin of modern humans. *Science*, 239:1263–1268, March 1988.

The SIGMA Type 2 Diabetes Consortium. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature*, 506(7486):97–101, Feb 2014. ISSN 0028-0836.

Erik Trinkaus. Early modern humans. *Annual Review of Anthropology*, 34:207–230, 2005.

Benjamin Vernot and Joshua Akey. Complex history of admixture between modern humans and neandertals. *American Journal of Human Genetics*, Feb 2015.

Benjamin Vernot and Joshua M Akey. Resurrecting surviving Neandertal lineages from modern human genomes. *Science*, 343(6174):1017–1021, 2014.

J. D. Wall and M. F. Hammer. Archaic admixture in the human genome. *Current Opinion in Genetics and Development*, 16(6):606–610, 2006.

Jeffrey D. Wall. Detecting ancient admixture in humans using sequence polymorphism data. *Genetics*, 154:1271–1279, 2000.

Jeffrey D Wall, Kirk E Lohmueller, and Vincent Plagnol. Detecting ancient admixture and estimating demographic parameters in multiple human populations. *Molecular Biology and Evolution*, 26(8):1823–1827, 2009.

Jeffrey D Wall, Melinda A Yang, Flora Jay, Sung K Kim, Eric Y Durand, Laurie S Stevison, Christopher Gignoux, August Woerner, Michael F Hammer, and Montgomery Slatkin. Higher levels of Neanderthal ancestry in East Asians than in Europeans. *Genetics*, 194(1):199–209, 2013.

Milford H. Wolpoff. Multiregional evolution: The fossil alternative to Eden. In Paul Mellars and Chris Stringer, editors, *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, pages 62–108. Princeton University Press, Princeton, New Jersey, 1989.

Vania Yotova, Jean-Francois Lefebvre, Claudia Moreau, Elias Gbeha, Kristine Hovhannesyan, Stephane Bourgeois, Sandra Bédarida, Luisa Azevedo, Antonio Amorim, Tamara Sarkisian, Patrice Hodonou Avogbe, Nicodeme Chabi, Mamoudou Hama Dicko, Emile Sabiba Kou' Santa Amouzou, Ambaliou Sanni, June Roberts-Thomson, Barry Boettcher, Rodney J. Scott, and Damian Labuda. An X-linked haplotype of Neandertal origin is present among all non-African populations. *Molecular Biology and Evolution*, 28(7):1957–1962, 2011.

Ewa Ziętkiewicz, Vania Yotova, Dominik Gehl, Tina Wambach, Isabel Arrieta, Mark Batzer, David EC Cole, Peter Hechtman, Feige Kaplan, David Modiano, Jean-Paul Moisan, Roman Michalski, and Damian Labuda. Haplotypes in the dystrophin DNA segment point to a mosaic origin of modern human diversity. *American Journal of Human Genetics*, 73(5):994–1015, 2003.