ISSN 1019-3316, Herald of the Russian Academy of Sciences, 2015, Vol. 85, No. 5, pp. 392–396. © Pleiades Publishing, Ltd., 2015.
Original Russian Text © S. Pääbo, 2015, published in Vestnik Rossiiskoi Akademii Nauk, 2015, Vol. 85, No. 10, pp. 879–884.

Papers by Laureates of the 2014 Lomonosov Grand Gold Medal of the Russian Academy of Sciences

DOI: 10.1134/S1019331615050081

The Contribution of Ancient Hominin Genomes from Siberia to Our Understanding of Human Evolution[1]

S. Pääbo*

The human genome is contained in chromosomes that are present in almost every cell in our bodies. It is composed of approximately 3.2 billion nucleotides. When cells replicate to form germ cells that will contribute to the next generation, mutations occur. As a result of these mutations, about 50 to 200 new substitutions exist in every new individual that is born. These substitutions accumulate in the genome over time to the extent that roughly one nucleotide in a thousand differs between two human genomes today, whereas roughly one nucleotide in a hundred differs between a human and a chimpanzee genome. In addition, duplicated DNA sequences differ both between individuals and between species.

Each particular nucleotide site in the genome has its own history that could, in principle, be traced back through past generations. Such a history can be depicted in the form of a tree showing common ancestors shared with the same site seen in other individuals today. However, in reality, it is impossible to trace the history of a single nucleotide site. Therefore, one generally traces the average history of a segment of the genome, or the entire genome, and depicts that in the form of a tree that thus represents an average picture of how most sites in the DNA segments or genomes whose nucleotide sequences have been determined are related.

Until recently, it was only possible to determine DNA sequences and whole genome sequences from present-day individuals from which DNA can be isolated in good condition from fresh tissues, such as blood. To evolutionary scientists, this is somewhat frustrating because it represents an indirect way to study the past: one studies DNA sequences that exist today, uses the best models we have for how mutations accumulate, and estimates what common ancestors in the past may have looked like. This is frustrating because what we have are estimates subject to many uncertainties, for example, as a result of the mutational models used. For over 30 years, our laboratory has worked on methods to overcome this "time trap" by going back in time and retrieving DNA sequences from archaeological and paleontological remains. This is possible only on rare occasions when well-preserved tissues can be found. Direct ancestors of present-day organisms are also almost never available. However, this approach, nevertheless, opens up new possibilities in that it allows DNA sequences from past populations and extinct species to be determined.

Of particular interest to us is the closest extinct relative of all present-day humans: the Neanderthals. This robust form of hominins emerged in Europe and western Asia approximately 300000 to 400000 years ago and disappeared between 30000 and 40000 years ago. The debate concerning the relationships between Neanderthals and modern humans, and about what happened when they met, lasted for decades. One idea was that modern humans replaced Neanderthals without interbreeding, in which case the Neanderthal contribution to present-day human genetic variation would be zero. Another idea was that Neanderthals were the direct ancestors of Europeans. In this case, the Neanderthal genetic contribution to present-day people in Europe would approach 100%. Obviously, all levels of contribution between 0 and 100% are also possible, and different levels of Neanderthal contribution to present-day Europeans have been argued for on the basis of archaeological and paleontological data.

We got the first chance to test these hypotheses directly in the mid-1990s, when we were allowed to analyze the Neanderthal bones that had been discovered in the Neandertal Valley in Germany in 1856 and gave its name to this hominin group. At the time, we were able to draw on over ten years of experience with the development of techniques to extract and amplify small amounts of DNA from ancient remains of cave bears, mammoths, and other late Pleistocene mammals [1]. We focused on the mitochondrial DNA (mtDNA), because every cell contains hundreds or

[1] The text was submitted by the author in English.

* Svante Pääbo is a professor and director of the Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
e-mail: paabo@eva.mpg.de

even thousands of mtDNA copies, making it easier to retrieve mtDNA than any particular part of the nuclear genome. We reconstructed the most variable part of the mtDNA and estimated phylogenetic trees to reconstruct the history of the mtDNAs of Neanderthals and present-day people. In contrast to the nuclear genome, the mtDNA is inherited as one single unit from mothers to offspring without recombination so a phylogenetic tree for mtDNA reflects not the average history but the exact maternal lineages that relate the mtDNA analyzed.

On the one hand, these trees showed what was already known: that the mtDNAs of all people inhabiting the Earth today trace their ancestry back to a common ancestor about 100000–200000 years ago. But they also showed that the mtDNA lineage of the Neanderthal type specimen went much further back in time and shared a common ancestor with present-day mtDNAs on the order of half a million years ago [2]. Subsequently, we and others have determined several other Neanderthal mtDNA sequences. They all fall together outside the variation of the mtDNAs of present-day people. Thus, in 1997, it was clear that for the mtDNA, the complete replacement model held: no person today carries an mtDNA derived from a Neanderthal.

However, the mtDNA represents only a tiny part of our total genome. The full picture of our genetic history can only be obtained by studying the nuclear genome. In the early years of this millennium, it became feasible to consider sequencing genomes from ancient organisms thanks to new techniques that made it possible to sequence millions of DNA molecules rapidly and inexpensively. We were fortunate to receive funding from the Max Planck Society for a five-year effort to improve the technique of extraction of DNA from ancient bones and to make DNA libraries that could be used for high-throughput DNA sequencing. We also analyzed a large number of bones from many sites in Europe to find those bones that contained the largest relative proportion of Neanderthal DNA.

We settled on a site in Croatia, from which we used three bones from different Neanderthal individuals and sequenced more than one billion short DNA fragments extracted from the bones. We developed computer algorithms to match these short DNA sequences to the human genome while accounting for errors induced by chemical process that have affected them over tens of thousands of years. Only a few percent of all sequences were derived from the Neanderthal individuals. Nevertheless, in 2010 we were able to present about 3 billion nucleotides of Neanderthal DNA that had been mapped to the human genome. Together these DNA fragments covered about 55% of the parts of the Neanderthal genome to which short fragments can be mapped [3]. This was enough to ask if any genetic interaction had occurred when modern humans encountered Neanderthals.

If Neanderthals had made no genetic contribution to modern humans, the Neanderthal genome would be equally far from Africans, Europeans, and any other present-day populations. In contrast, if present-day Europeans carried DNA that they had inherited from Neanderthals, European genomes would carry fewer differences from Neanderthals than African genomes, since Neanderthals were never in Africa; therefore, they would not be expected to have contributed to genomes there. To test this, we sequenced the genomes of five present-day people and identified the positions where two of these differed from each other. We then asked how often at these positions the Neanderthal genome carried the variant seen in one present-day person and how often it carried the variant seen in the other present-day person. This approach of counting matches to pairs of present-day genomes was necessary since the quality of the Neanderthal genome was so low that we could not trust sequence variants that were seen only in the Neanderthal genome and not in one of the present-day genomes. When we compared two African genomes in this way, the Neanderthal genome matched variants in the two genomes equally often. This is to be expected since there was no reason to expect that Neanderthals would have contributed DNA to the ancestors of any of the Africans. Intriguingly, when we compared a European and an African to the Neanderthal genome, we detected statistically significantly greater matching to the European genome, suggesting that Neanderthals had contributed DNA to the ancestor of the Europeans. Even more surprising was that, when we compared a person from China to an African and a person from Papua New Guinea to an African, we always found that the non-African matched the Neanderthal genome more often than the African genome. This was surprising to us since Neanderthals had probably never been in China and surely never in New Guinea. How could this be?

The explanation that we suggested and that has since been borne out by work in our own and other groups was that Neanderthals met modern humans and mixed with them probably in the Middle East. If these modern humans later became the ancestors of everybody that today lives outside Africa, these early modern humans can, so to speak, have carried with them the Neanderthal genetic contribution also to geographical areas where Neanderthals never existed. As a result, between 1 and 2% of the genomes of every person whose roots are non-African is of Neanderthal origin. The Neanderthal component in the genomes of present-day people has since been dated by studies of the extent to which Neanderthal-like DNA segments have been broken down to smaller pieces by recombination that happens in each generation [4]. It has also

been confirmed by subsequent studies of a modern human from about 40000 years ago who carried much larger segments of Neanderthal DNA than present-day people, since they lived much closer to the time of mixture [5].

Of course, it is unlikely that mixing between Neanderthals and modern humans happened only in one population and exclusively in the Middle East, but, given the data at hand in 2010, this was the simplest explanation of our findings. Further insights were largely limited by the comparatively low quality of the Neanderthal genome. This was to be changed thanks to our collaboration with Anatolii Panteleevich Derevyanko.

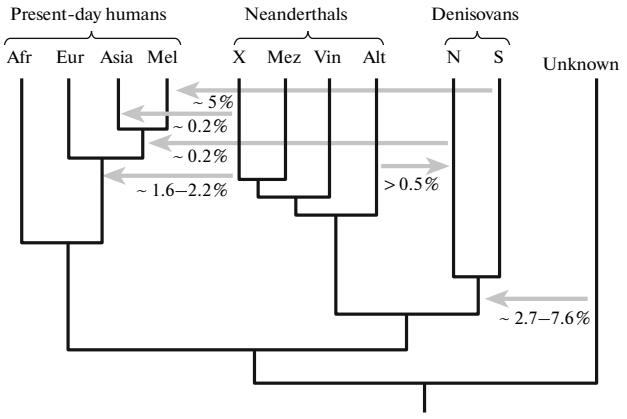
The excavations at Denisova Cave, led by Academician Derevyanko and Professor M.V. Shun'kov of the Institute of Archaeology and Ethnography of the Siberian Branch of the Russian Academy of Sciences have generated many fundamental and novel insights into human evolution. One of their crucial finds is a hominin toe bone discovered in 2010. When we applied new, ultrasensitive methods that my laboratory developed to extract DNA and produce DNA libraries to this bone, we were able to sequence almost 50-fold more endogenous DNA from this single small bone than from the three bones from Croatia that had been used to produce the first Neanderthal genome a few years earlier. This individual turned out to be a Neanderthal, and its genome was sequenced to a quality higher than most genomes determined from present-day living people [6].

Using such high-quality genomic information, it is possible to observe differences between the two genomes that the individual inherited from her father and from her mother. One can thus gauge the extent of variation in the population where the parents of the individual lived. One can also estimate how closely related the mother and the father of the individual were to each other. In the case of the Neanderthal from Denisova Cave, this yielded an unexpected result. The paternal and maternal genomes had long segments of DNA that were identical. This means that the parents of this individual were closely related. One can estimate that they must have been related at the level of half siblings. When in the future further Neanderthal genomes are sequenced to the same high quality as the one from Denisova Cave, it will be interesting to see if this was an unusual situation among Neanderthals or if it reflects a social pattern typical of Neanderthals.

The high quality of the Neanderthal genome from Denisova Cave can also be used to estimate what parts of the genomes of present-day people were inherited from Neanderthals. This confirms that everybody outside Sub-Saharan Africa carries between 1 to 2% of Neanderthal DNA. This proportion is slightly larger in East Asia than in Europe, suggesting that additional admixture between Neanderthals and modern humans may have happened during the colonization of Asia [7]. To get a perspective on this, you may recall that we all have one-half of our DNA from each of our parents, about 25% from each grandparent, about 12% from our great grandparents, and so on. From an ancestor six generations back, we have, on average, inherited about 1.5% of our DNA. Thus, from the point of view of the total amount of DNA people today have inherited from Neanderthals, it is as if they had a Neanderthal ancestor six generations back. However, due to recombination that occurs when new germ cells are formed in each generation, the Neanderthal DNA is distributed in much smaller fragments than the DNA you have inherited from your ancestors six generations back. You may also ask how much of the total Neanderthal genome exists distributed among people living today. This estimate is still very approximate, but it would seem that at least about 40% of the Neanderthal genome can be found in people today.

Amazingly, the high-quality Neanderthal genome is not the only great gift that Denisova Cave has given the world. In 2008 a tiny piece of the phalanx of a fifth finger of a child was discovered in the East gallery of the cave. We were privileged to work on this find and were happy to be able to generate first a low quality genome [8] and then, as our techniques improved, a high-quality genome from it. In this genome, each position in the part of the genome amenable to mapping short pieces of DNA was covered over 30 times [9]. When we compared this genome to other genomes, we were surprised to find that it was neither a modern human nor a Neanderthal. It shared a common ancestor with Neanderthals, but this ancestral population lived about four times further back in time than the oldest ancestral population shared among present-day human populations. After discussions with Academician Derevyanko and his team in Novosibirsk, it was decided to name this new hominin group "Denisovans." It is the first hominin group described on the basis of a genome sequence rather than a morphological description. Although remains of Denisovans have yet to be found outside of Denisova Cave, we can learn about their history and the history of other hominins by studies of their genome.

Interestingly, on the order of 5% of the genomes of people that today live in the Pacific, for example, Aboriginal Australians and Papuans, come from Denisovans [10], suggesting the ancestors of these populations met Denisovans and sired offspring with them. In addition, about 0.2% of the genomes of people in Mainland Asia come from Denisovans [6, 11]. By comparing the two high-quality genomes of a Neanderthal and a Denisovan that have been determined from Denisova Cave, we can also discern gene flow events that have occurred between these two groups and other gene flow events that have affected these two



A schematic illustration of some archaic and modern groups and their genetic interactions.

Modern humans are represented by African (Afr), European (Eur), Asian (Asia), and Melanesian (Mel) populations; Neanderthals are represented by an unknown population (X) contributing to non-Africans, and by Neanderthal genomes from the Russian Caucasus (Mez), Croatia (Vin), and Denisova Cave (Alt); Denisovans are represented by an unknown population (S) contributing to people in the Pacific and by the population in the Altai Mountains (N). "Unknown" represents a hominin that diverged between one and four million years ago and contributed to the Denisovan genome. For each of the six genetic contributions detected to date, the approximate percentages of the genome contributed are indicated. For details, see [6].

groups differently. A minimum of two additional instances of gene flow can be detected by these comparisons: one from eastern Neanderthals into Denisovans, and one from an unknown hominin that diverged a million years ago or more from the human lineage into Denisovans [6]. In addition, recent work shows that early modern humans in Europe mixed with Neanderthals when they first arrived there [12].

The emerging picture is thus a complicated one (figure) in which many different hominin groups exchanged genes with one another on what must have been many occasions. Often this exchange was of limited magnitude, but it shows that the gene pools of most or even all hominin groups in the Late Pleistocene were open systems that allowed genetic variants to spread from one group to another. An interesting question if any of these variants were of functional importance. As yet, we do not know much about this, but I want to bring up a few examples of what has emerged from studies by several groups in the last two years.

One way to ask what functional role Neanderthal genetic variants may play in present-day genomes is to ask which genes that carry Neanderthal variants have risen to high frequency. The fact that these variants have become frequent today may suggest that they were positively selected in the past. The group of genes that is statistically overrepresented in such genomic segments are keratins, i.e., the structural protein present in skin and hair [13, 14]. Thus, it is thus likely that in the future we will find that some aspects of the morphology or function of skin and hair that is present in some people in Europe and Asia derived from Neanderthals.

There are also aspects of metabolism that are affected by Neanderthal variants. For example, Europeans but not Asians carry more Neanderthal variants than statistically expected in genes involved in catabolism [15]. It is not yet known what these variants do, but it will hopefully be discovered in the next few years. Interestingly, a variant of the gene that encodes the protein that transports lipids across cell membranes and that has been inherited from Neanderthals has risen to a frequency of up to 35% in East Asia and Native Americans. This variant is associated with an increased risk to develop type 2 diabetes [16]. It may seem surprising that a Neanderthal gene variant that confers risk of disease has become frequent in the population. One may speculate that a variant that causes diabetes today in people who enjoy ample nutrition throughout life may have represented an advantage in a situation of food shortages. Thus, this gene variant may represent a Neanderthal adaptation to starvation, which in the past was advantageous also for modern humans.

Have Denisovans, like the Neanderthals, contributed functionally to present-day people? Recent work suggests that this is the case. The population in Tibet carries genetic adaptations to life where the partial pressure of oxygen is low as is the case in the high altitudes on the Tibetan High Plateau. The major gene variant responsible for this adaptation affects the number of red blood cells and occurs at a frequency of about 80% in Tibet, but it is very rare elsewhere in Asia. Last year it was shown that this gene variant is likely to have been inherited from Denisovans [17]. Thus, it seems that gene flow from Denisovans contributed to making life on the high plateau in Tibet possible. Similarly, there are indications that gene variants important for how the immune system deals with infectious diseases may have been acquired both from Denisovans and from Neanderthals [18].

There is thus a picture emerging in which Denisovans, Neanderthals, and possibly other archaic groups who had lived in Eurasia for hundreds of thousands of years and had adapted to local environments met and mixed with modern humans on many occasions. This gave modern humans the opportunity to acquire locally advantageous gene variants from these groups. This is a phenomenon often referred to as "adaptive introgression" in other species [19], which may have been of some importance for modern humans as they colonized new environments throughout Eurasia [20].

In summary, the fact that gene flow has been detected not only from Denisovans and Neanderthals into modern humans but also between various other

hominin groups shows that these were not closed genetic systems. They may best be regarded as a "metapopulation"—a web of populations that included Neanderthals, Denisovans, modern humans, and other groups, which were linked by intermittent or sometimes perhaps even persistent gene flow [21]. In this metapopulation, gene variants spread directly but also potentially indirectly between groups who were in contact with each other over other groups.

These results support the idea expressed by Academician Derevyanko already in 2005 when he said, "Dear colleagues, please do not offend Neanderthals. They are among our ancestors!" [22, p. 507]. The analyses of the genomes from Denisova Cave have shown that this generous attitude was correct and should be extended to Denisovans and perhaps also other hominin forms.

ACKNOWLEDGMENTS

Our work is made possible by the long-standing, fruitful, and stimulating collaboration with Academician Derevyanko and Prof. Shun'kov who lead the work at Denisova Cave. I am also indebted to past and present members of our research team and collaborators who have made the analyses of archaic genomes possible, to the Max Planck Society for continuous support, and to the Russian Academy of Sciences for its support of the important work in Siberia and for the honor bestowed upon me today.

This work was supported by the Russian Science Foundation, project no. 14-50-00036.

REFERENCES

1. S. Pääbo, *Neanderthal Man: In Search of Lost Genomes* (Basic Books, New York, 2014; Corpus, Moscow, 2016).
2. M. Krings, A. Stone, R. W. Schmitz, et al., "Neanderthal DNA sequence and the origin of modern humans," Cell **90**, 19 (1997).
3. R. E. Green, J. Krause, A. W. Briggs, et al., "A draft sequence Neanderthal genome," Science **328**, 710 (2010).
4. S. Sankararaman, N. Patterson, H. Li, S. Pääbo, D. Reich, "The date of interbreeding between Neanderthals and modern humans," PLoS Genet., No. 8 (2012).
5. Q. Fu, Heng Li, P. Moorjani, et al., "Genome sequence of a 45.000-year-old modern human from western Siberia," Nature **514**, 445 (2014).
6. K. Prüfer, F. Racimo, N. Patterson, et al., "The complete genome sequence of a Neanderthal from the Altai Mountains," Nature **505**, 43 (2014).
7. B. Vernot and J. M. Akey, "Complex history of admixture between modern humans and Neanderthals," Am. J. Hum. Genet. **56**, 448 (2015).
8. D. Reich, R. E. Green, M. Kircher, et al., "Genetic history of an archaic hominin group from Denisova cave in Siberia," Nature **468**, 1053 (2010).
9. M. Meyer, M. Kircher, T. Gansauge, et al., "A high-coverage genome sequence from an archaic Denisovan individual," Science **338**, 222 (2012).
10. D. Reich, N. Patterson, M. Kircher, et al., "Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania," Am. J. Hum. Genet. **89**, 516 (2011).
11. P. Skoglund and M. Jakobsson, "Archaic human ancestry in East Asia," Proc. Natl. Acad. Sci. U. S. A. **108**, 18301 (2011).
12. Q. Fu, M. Hajdinjak, O. A. Moldovan, et al., "An early modern human from Romania with a recent Neanderthal ancestor," Nature **524**, 216 (2015).
13. B. Vernot and J. M. Akey, "Resurrecting surviving Neanderthal lineages from modern human genomes," Science **343**, 1017 (2014).
14. S. Sankararaman, S. Mallick, M. Dannemann, et al., "The genomic landscape of Neanderthal ancestry in present-day humans," Nature **507**, 354 (2014).
15. E. E. Khrameeva, K. Bozek, L. He, et al., "Neanderthal ancestry drives evolution of lipid catabolism in contemporary Europeans," Nature Commun., No. 4 (2014).
16. A. L. Williams, S. B. Jacobs, H. Moreno-Macias, et al., "The SIGMA Type 2 Diabetes Consortium," Nature **506**, 97 (2014).
17. E. Huerta-Sánchez, X. Jin, A. Asan, et al., "Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA," Nature **512**, 194 (2014).
18. L. Abi-Rached, M. J. Jobin, S. Kulkarni, et al., "The shaping of modern human immune systems by multiregional admixture with archaic humans," Science **334**, 89 (2011).
19. P. W. Hedrick, "Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation," Mol. Ecol. **22**, 4606 (2013).
20. F. Racimo, S. Sankararaman, R. Nielsen, and E. Huerta-Sánchez, "Evidence for archaic adaptive introgression in humans," Nat. Rev. Genet. **16**, 359 (2015).
21. S. Pääbo, "The diverse origins of the human gene pool (commentary)," Nat. Rev. Genet. **16**, 313 (2015).
22. A. P. Derevyanko, *The Middle to Upper Paleolithic Transition in Eurasia: Hypotheses and Facts* (Izd. IAET SB RAS, Novosibirsk, 2005).