A high-coverage Neandertal genome from Chagyrskaya Cave

Fabrizio Mafessoni, Steffi Grote, Cesare de Filippo, Viviane Slon, Kseniya A. Kolobova, Benoit Ciliberti, Sergey V. Markin, Manjuasha Chintalapati, Stephane Peyrégne, Laurits Skov, Pontus Skoglund, Andrey I. Krivoshapkin, Anatoly P. Derevianko, Matthias Meyer, Janet Kelso, Benjamin Peter, Kay Prüfer, and Svante Pääbo

*Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany; †Institute of Archaeology and Ethnography, Russian Academy of Sciences, 630090 Novosibirsk, Russia; ‡Department of Anthropology, University of Toronto, ON M5S 2S2, Canada; ‡Ancient Genomics Laboratory, Francis Crick Institute, NW1 1AT London, United Kingdom; and §Human Evolutionary Genomics Unit, Okinawa Institute of Science and Technology, Onna-son, 904-0495 Okinawa, Japan

Contributed by Svante Pääbo, May 13, 2020 (sent for review March 16, 2020; reviewed by Kelley Harris and Roberto Macchiarelli)

We sequenced the genome of a Neandertal from Chagyrskaya Cave in the Altai Mountains, Russia, to 27-fold genomic coverage. We show that this Neandertal was a female and that she was more related to Neandertals in western Eurasia [Prüfer et al., Science 358, 655–658 (2017); Hajdínjak et al., Nature 555, 652–656 (2018)] than to Neandertals who lived earlier in Denisova Cave [Prüfer et al., Nature 505, 43–49 (2014)], which is located about 100 km away. About 12.9% of the Chagyrskaya genome is spanned by homozygous regions that are between 2.5 and 10 centiMorgans (cM) long. This is consistent with the fact that Siberian Neandertals lived in relatively isolated populations of less than 60 individuals. In contrast, a Neandertal from Europe, a Denisovan from the Altai Mountains, and ancient modern humans seem to have lived in populations of larger sizes. The availability of three Neandertal genomes of high quality allows a view of genetic features that were unique to Neandertals and that are likely to have been at high frequency among them. We find that genes highly expressed in the striatum in the basal ganglia of the brain carry more amino-acid-changing substitutions than genes expressed elsewhere in the brain, suggesting that the striatum may have evolved unique functions in Neandertals.

Neandertals and Denisovans are the closest evolutionary relatives of present-day humans. Analyses of their genomes showed that they contributed genetically to present-day people outside sub-Saharan Africa [1, 2]. However, to date, the genomes of only two Neandertals and one Denisovan have been sequenced to high quality. One of these Neandertal genomes (Vindija 33.19) comes from an individual found in Vindija Cave in Croatia [3], whereas the other Neandertal genome (Denisova 5 or the “Altai Neandertal”) [4] and the Denisovan genome (Denisova 3) [5] both come from specimens discovered in Denisova Cave, in the Altai mountains in Siberia.

A number of archaic genomes of moderate quality (one- to threefold genomic coverage) have yielded additional insights into Neandertal history. For example, genome sequences from five late Neandertals from Europe have shown that they carried little genetic variation (6, 7) and were more closely related to the Vindija 33.19 than to the Denisova 5 Neandertal. A genome sequence from a morphologically undiagnostic bone from Denisova Cave, Denisova 11, belonged to the direct offspring of a Neandertal mother and a Denisovan father [8], indicating that the two groups met in the Altai region. The Neandertal mother of Denisova 11 was more closely related to Vindija 33.19 than to Denisova 5, indicating that a replacement of Neandertal populations in the Altai Mountains occurred [8]. Here, we present the high-coverage genome sequence of a Neandertal from Chagyrskaya Cave, located 106 km to the west of Denisova Cave (9–12) (Fig. 1). This genome provides insights into Neandertal population structure and history and allows the identification of genomic features unique to Neandertals.

Significance

We present the third high-quality genome to be determined from a Neandertal. Patterns of variation in the genome suggest that her ancestors lived in relatively isolated populations of less than 60 individuals. When we analyze this genome together with two previously sequenced Neandertal genomes, we find that genes expressed in the striatum of the brain may have changed especially much, suggesting that the striatum may have evolved unique functions in Neandertals.


Reviewers: K.H., University of Washington; and R.M., Universite de Poitiers.
The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The genome sequence of Chagyrskaya 8 can be downloaded from http://ftp.eva.mpg.de/neandertal (accession ID MK388903).

1To whom correspondence may be addressed. Email: paabo@eva.mpg.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2004944117/-/DCSupplemental.

when compared to Vindija 33.19 (Fig. 2), Chagyrskaya 8 shares fewer derived alleles than Denisova 5.

When compared to Vindija 33.19 (Fig. 2), Chagyrskaya 8 shares fewer derived alleles with other Neandertals that lived in Europe ~50 kya, i.e., approximately at the same time as Vindija 33.19. However, Chagyrskaya 8 shares more derived alleles than Vindija 33.19 with Denisova 11, a first-generation Neandertal–Denisovan offspring (8). Since Vindija 33.19 and Chagyrskaya 8 do not differ in their sharing of derived alleles with the Denisovan Denisova 3, this indicates that Chagyrskaya 8 is most closely related among currently known Neandertal to the mother of Denisova 11 found at Denisova Cave (Fig. 1 and SI Appendix 7).

Relationship with Modern Humans. Non-African present-day humans carry ~2% (1, 3) Neandertal ancestry as a result of gene flow from Neandertals that occurred between 50 and 90 kya (1, 17). Genomewide, Chagyrskaya 8 shares more alleles with present-day human populations outside Africa than Denisova 5 does, and a similar proportion of alleles as Vindija 33.19 (SI Appendix 7). However, if the analysis is restricted to previously detected Neandertal haplotypes introgressed in present-day humans (18), or to derived alleles that occur at low frequencies in present-day non-African populations and that are therefore more likely to be introgressed from Neandertals, Vindija 33.19 shares more alleles with present-day populations than Chagyrskaya 8. This indicates that Vindija 33.19 is more closely related than Chagyrskaya 8 to Neandertal populations that contributed the majority of the DNA to present-day populations.

To test if any modern human population carries an additional genetic contribution from Neandertals more closely related to Chagyrskaya 8 than to Vindija 33.19, we made use of previously published haplotypes inferred to come from Neandertals that are found today exclusively in East Asia, Europe, India, or Oceania. Among 300 genomes from the Simons Genome Diversity Panel (19) and 89 Papuan genomes (20, 21), the proportions of alleles shared with Chagyrskaya 8 and with Vindija 33.19 are similar (SI Appendix 7), giving no indication that Neandertals closer related to Chagyrskaya 8 than to Vindija 33.19 contributed to the populations tested. Thus, within the limits of the resolution of these analyses, we conclude that if several Neandertal populations contributed to the genomes of different modern human populations in different proportions (22, 23), these Neandertal populations were similarly related to Chagyrskaya 8 and to Vindija 33.19.

Small Population Size and Inbreeding. The Neandertal genome of Denisova Cave, Denisova 5, carries a high proportion of long tracts that are homozygous by descent (HBD tracts) (4). Whereas tracts that are over 10 cM long indicate that the parents of Denisova 5 were closely related, HBD tracts between 2.5 and 10 cM indicate that the population from which Denisova 5 comes was of small size over ~100 generations before the individual lived (SI Appendix 8).

Fig. 1. The Chagyrskaya 8 Neandertal and its relationship with other archaic individuals. (Top) Locations of Chagyrskaya Cave and other sites where archaic specimens analyzed here were found are indicated. (Bottom) Schematic illustration of the relationship among archaic genomes. The gray outline describes relationships between high-coverage genomes (SI Appendix 6). Population split times estimated using the F(A) statistics (3, 4, 8) (SI Appendix 7) between high-coverage and low-coverage genomes are indicated by horizontal dashed line. Within the gray outlines, colored silhouettes schematically indicate population sizes (Ne) over time as estimated in SI Appendix 6. Neandertals are indicated in blue, Denisovans in red. Genomes determined to high genomic coverage (~27-fold) are indicated by large circles. The genomes from Spy, Goyet, and Les Cottés are reported in ref. 6, Scaldina and Holenstein-Stadel in ref. 13, other genomes in refs. 2–6, 8.

Fig. 2. Relative sharing of derived allele among the Chagyrskaya 8 and other archaic genomes. Positive values in the D statistic indicate more allele sharing with Chagyrskaya 8 than with Denisova 5 (4) or than with Vindija 33.19 (8). Error bars indicate one SE. One asterisk indicates Z > 2; two asterisks indicate Z > 3. Late Neandertals of low genomic coverage refer to the genomes in ref. 6.
Compared to Denisova 5, the Chagyrskaya 8 genome carries fewer HBD tracts longer than 10 cM, but more HBD tracts of intermediate length than almost all present-day and prehistoric modern human genomes, as well as the Denisovan genome (Denisova 3) (Fig. 3A). We show by coalescent simulations that this cannot be explained by an overall small but panmictic population. Rather, it suggests that Neandertal populations were subdivided (SI Appendix 8).

By coalescent modeling, we infer that Chagyrskaya 8 and Denisova 5 may have lived in subpopulations of 60 or fewer individuals. In contrast, current and past modern human populations as well as Denisovans (based on the Denisova 3 genome) lived in subpopulations of more than 100 individuals, assuming a migration rate between subpopulations of 1% or less (Fig. 3B and SI Appendix 8). Interestingly, the Vindija 33.19 Neandertal seems to have lived in a subpopulation of larger size than the two Siberian Neandertals, although this difference is only marginally statistically significant when the proportion of the genomes covered by all HBD tracts longer than 2.5 cM is considered (likelihood-ratio test, $P = 0.05$).

**Derived Genomic Features in Neandertals.** We used the three high-coverage Neandertal genomes to identify biological pathways where protein-coding genes show more derived nonsynonymous substitutions fixed in the three Neandertals than expected from the silent and polymorphic changes. We identified 993 fixed nonsynonymous substitutions among 889 genes, and 2,952 polymorphic nonsynonymous substitutions in the three Neandertals.

No groups of genes associated with a specific known biological functions or phenotypes (26) show a higher ratio of nonsynonymous to synonymous fixed changes relative to the ratio of nonsynonymous to synonymous polymorphic changes (MacDonald–Kreitman ratio) (27), compared to other groups of genes (family-wise error rate $> 0.1$) (28 and SI Appendix 9).

However, when analyzing genes preferentially expressed in different brain regions according to the Allen Brain Atlas (29, 30), we find that genes expressed in the striatum in individuals 12–19 y of age show a higher MacDonald-Kreitman ratio (1.02, family-wise error rate $= 0.029$) than genes expressed in other brain regions and at other ages (0.53–0.83) (SI Appendix 9). This may indicate that negative selection in genes expressed in the striatum was relaxed in Neandertals. Alternatively, some of the proteins encoded by these genes might have been the targets of positive selection. In addition, genes expressed in the prenatal striatum carry more substitutions in their untranslated regions than genes expressed elsewhere (family-wise error rate $= 0.049$) and at other times. Among genes expressed in the striatum, those carrying fixed nonsynonymous changes in Neandertals are more often present in genomic regions that carry little or no DNA introgressed from Neandertals in present-day humans than striatal A

**Fig. 3.** Homozygous tracts in archaic genomes. (A) Proportion of the genome spanned by HBD tracts of size 2.5–10 cM (Top, dark) and all HBD tracts (Bottom, light) for archaic genomes, ancient modern humans and present-day humans from the Simons Genome Diversity Project (18). Values for Chagyrskaya 8 are indicated by a triangle. The genetic size of the tracts was estimated assuming a uniform recombination rate of $1.3 \times 10^{-8}$ recombinations/bp. (B) Estimated group size $(x$ axis) and migration rate $(y$ axis) for three Neandertals, one Denisovan, and ancient modern human genomes based on HBD tracts $> 2.5$ cM long. The colored areas delimit 95% confidence intervals of likelihood ratios. Ancient modern humans give estimates for the genomes of four individuals dated to between $\sim 45$ and $\sim 8$ kya [“Ust’ishim” (15), “Sunghir 2” (24), “Loschbour” (25), and “LBK-Stuttgart” (25)].
genes not carrying such changes (Fisher’s exact test $P = 0.026$). This pattern is not observed for all genes carrying fixed non-synonymous substitutions in Neandertals ($P > 0.1$), suggesting that some substitutions in Neandertal striatal genes might have been negatively selected in modern humans. Besides the striatum, genes expressed prenatally in the posterior parietal cortex, in the ventrolateral prefrontal cortex, and in the primary somatosensory cortex carry more fixed substitutions in their regulatory regions in Neandertals than genes expressed in other brain regions and at other times (SI Appendix 9).

In a further attempt to detect positive selection along the Neandertal lineage, we performed the Hudson, Kreitman, Aguadé (HKA) test (31) and the population branch statistic (PBS) (32) in 25-kb sliding windows across the genome (SI Appendix 10). We estimated the probability of obtaining the observed values of the different statistics by coalescent simulations, and retained windows with a false discovery rate $< 5\%$ (Fig. 4). On the autosomes, we identify a total of 35 separate candidate regions. One of these candidates, a 75-kb-long region identified by PBS on chromosome 5, overlaps two separate windows identified by HKA. This overlap is lower than if expected by change ($P$ value $< 4 \times 10^{-4}$ and SI Appendix 10). The candidate regions identified overlap genes involved in neural development ($\text{EXOC6B}$), immunity, and wound healing ($\text{HTN1}, \text{EVPLL}$) and mitochondrial functions ($\text{NSUN3}, \text{TIMM29}$). For both statistics, we find an overlap with genomic regions previously identified as positively selected in modern humans (33) (enrichment test, $P$ values = $0.010$ and $0.056$ for HKA and PBS, respectively).

**Discussion**

*Chagyrskaya* 8 is more closely related to *Vindija* 33.19 and other late Neandertals in western Eurasia than to the Denisova 5 Neandertal who lived earlier in the Altai Mountains (Fig. 2). *Chagyrskaya* 8 is thus related to Neandertal populations that moved east sometime between 120 and 80 kya (13). Interestingly, the artifacts found in Chagyrskaya Cave show similarities to artifact assemblages in central and eastern Europe (10 and SI Appendix 1), suggesting that Neandertal populations coming from western Eurasia to Siberia may have brought their material culture with them (10, 34). Some of these incoming Neandertal encountered local Denisovan populations, as shown by Denisova 11, who had a Denisovan father and a Neandertal mother related to the population in which *Chagyrskaya* 8 lived.

In this regard, it is interesting that the *Chagyrskaya* 8 and Denisova 5 Neandertals lived in smaller populations than the *Vindija* 33.19 Neandertal in Croatia, the Denisovan *Denisova* 3, and modern humans (24) (Fig. 3B). Neandertals in the Altai region may have lived in smaller and more isolated populations than Neandertals elsewhere as that region represented the periphery of their geographical distribution and may have been an area where Denisovans were more continuously present. However, more detailed studies of the population history of Denisova Cave and other sites will be necessary to clarify this.

When analyzing genetic changes on the Neandertal lineage using the three Neandertal genomes, the number of changes in genes expressed in the striatum during adolescence stands out. One possibility is that these changes accumulated in Neandertals due to their small population size, perhaps in combination with a relaxation of selection on genes expressed in the striatum. Interestingly, genes expressed in the striatum overlap more frequently than expected with genomic regions where Neandertal introgressed fragments in modern human genomes are rare. We speculate that striatal genes may carry Neandertal-specific changes that were disadvantageous when introduced into modern humans. This, as well as positive selection for derived changes in the modern human lineage, may underlie so-called Neandertal deserts in present-day human genomes, i.e., regions that are depleted of Neandertal ancestry (20, 35).

![Fig. 4](image-url) Positive selection on the Neandertal lineage. Manhattan plots of HKA (Top) and PBS (Bottom) tests for positive selection applied to three Neandertal genomes. Different colors indicate different autosomes, from 1 (Left) to 22 (Right). The $y$ axis indicates $-\log_{10}$ of false discovery rates (FDR). The gray line indicates $FDR = 5\%$. Candidate windows that overlap exons are indicated by the name of the corresponding gene. The regions that are significant for both tests are indicated by an asterisk.
As more high-quality Neandertal genomes become available, it will be possible to more comprehensively explore genes and groups of genes that carried functionally relevant changes in Neandertals. Currently, there is suggestive evidence that such findings may be forthcoming. For example, in addition to genes expressed in the striatum, untranslated regions and promoters of genes expressed in the posteroventral (inferior) parietal cortex, a brain region that has been associated with speech and mathematical cognition, carry more changes in the three Neandertals than expected by chance. In addition, among the top phenotypes associated with changes in regulatory regions in Neandertals are abnormalities in parts of the skeleton where Neandertal morphology stand out, such as the nasal bridge and the rib cage (e.g., ref. 37).

**Methods**

DNA was extracted from ∼100 mg of bone power after sodium hypochlorite treatment (SI Appendix 2). Sequencing, filtering of the data, and genotyping were carried out as previously described (ref. 3 and SI Appendix 3). Present-day human DNA contamination was estimated using mitochondrial, autosomal data, and from the proportion of DNA fragment mapping on the Y chromosome (SI Appendixes 4 and 5). Demographic histories and split times were estimated using the Pairwise Sequentially Markovian Coalescent (38 and SI Appendix 6) and the F(A/B) method (refs. 35 and SI Appendix 7), respectively, and coestimated with momi2 (39 and SI Appendix 7). HBD tracts were computed as previously published (refs. 35 and SI Appendix 8). Selection analyses were performed using the R packages GOfuncR (https://github.com/sgrote/GOfuncR), ABAEnrichment (ref. 29 and SI Appendix 9), and by computing HKA and a modified PBS statistic in 25-kb windows (SI Appendix 10).

**Data Availability Statement.** The genome sequence of Chagyrskaya 8 can be downloaded from ftp://eva.mpg.de/neandertal/Chagyrskaya/VCF/. Its mitochondrial DNA genome sequence has been deposited in GenBank (accession ID MK388903).

**ACKNOWLEDGMENTS.** This study was funded by the Max Planck Society; the European Research Council (Grant Agreement 694707 to S.P.); and the Russian Science Foundation (Project 19-48-04107 to K.A.K.). We thank Sarah Nagel, Birgit Nickel, Barbara Schellbach, and Antje Weihmann for laboratory work; and Heiko Temming for computed tomography scans.

3. K. Prüfer et al., The complete genome sequence of a Neandertal from the Altai Mountains. Nature 505, 43–49 (2014).