HUMAN EVOLUTION

A high-coverage Neandertal genome from Vindija Cave in Croatia

Kay Prüfer,[1]* Cesare de Filippo,[1]† Steffi Grote,[1]† Fabrizio Mafessoni,[1]† Petra Korlević,[1]
Mateja Hajdinjak,[1] Benjamin Vernot,[1] Laurits Skov,[2] Pinghsun Hsieh,[3]
Stéphane Peyrégne,[1] David Reher,[1] Charlotte Hopfe,[1] Sarah Nagel,[1] Tomislav Maricic,[1]
Qiaomei Fu,[4] Christoph Theunert,[1,8] Rebekah Rogers,[8] Pontus Skoglund,[5]
Manjusha Chintalapati,[1] Michael Dannemann,[1] Bradley J. Nelson,[3] Felix M. Key,[1]
Pavao Rudan,[6] Željko Kućan,[6] Ivan Gušić,[6] Liubov V. Golovanova,[7]
Vladimir B. Doronichev,[7] Nick Patterson,[5] David Reich,[5,9,10] Evan E. Eichler,[3,11]
Montgomery Slatkin,[8] Mikkel H. Schierup,[2] Aida M. Andrés,[1] Janet Kelso,[1]
Matthias Meyer,[1] Svante Pääbo[1]*

To date, the only Neandertal genome that has been sequenced to high quality is from an individual found in Southern Siberia. We sequenced the genome of a female Neandertal from ~50,000 years ago from Vindija Cave, Croatia, to ~30-fold genomic coverage. She carried 1.6 differences per 10,000 base pairs between the two copies of her genome, fewer than present-day humans, suggesting that Neandertal populations were of small size. Our analyses indicate that she was more closely related to the Neandertals that mixed with the ancestors of present-day humans living outside of sub-Saharan Africa than the previously sequenced Neandertal from Siberia, allowing 10 to 20% more Neandertal DNA to be identified in present-day humans, including variants involved in low-density lipoprotein cholesterol concentrations, schizophrenia, and other diseases.

Neandertals are the closest evolutionary relatives identified to date of all present-day humans and therefore provide a unique perspective on human biology and history. In particular, comparisons of genome sequences from Neandertals with those of present-day humans have allowed genetic features specific to modern humans to be identified (1, 2) and have shown that Neandertals mixed with the ancestors of present-day people living outside sub-Saharan Africa (3). Many of the DNA sequences acquired by non-Africans from Neandertals were likely detrimental and were purged from the human genome via negative selection (4–8), but some appear to have been beneficial and were positively selected (9); among people today, alleles derived from Neandertals are associated with both susceptibility and resistance to diseases (7, 10–12).

However, our knowledge about the genetic variation among Neandertals is still limited. To date, genome-wide DNA sequences of five Neandertals have been determined. One of these, the "Altai Neandertal," found in Denisova Cave in the Altai Mountains in southern Siberia, the eastern-most known reach of the Neandertal range, yielded a high-quality genome sequence (~50-fold genomic coverage) (2). In addition, a composite genome sequence from three Neandertal individuals has been generated from Vindija

Cave in Croatia in southern Europe but is of low quality (~1.2-fold total coverage) (3), while a Neandertal genome from Mezmaiskaya Cave in the Caucasus (2) is of even lower quality (~0.5-fold coverage). In addition, chromosome 21 (13) and exome sequences (14) have been generated from a different individual from Vindija Cave and one from Sidron Cave in Spain. The lack of high-quality Neandertal genome sequences, especially from the center of their geographical range and from the time close to when they were estimated to have mixed with modern humans, limits our ability to reconstruct their history and the extent of their genetic contribution to present-day humans.

Neandertals lived in Vindija Cave in Croatia until relatively late in their history (3, 15). The cave has yielded Neandertal and animal bones, many of them too fragmentary to determine from their morphology from what species they derive. Notably, DNA preservation in Vindija Cave is relatively good and allowed the determination of Pleistocene nuclear DNA from a cave bear (16), a Neandertal genome (3), and exome and chromosome 21 sequences (13, 14).

To generate DNA suitable for deep sequencing, we extracted DNA (17) and generated DNA libraries (18) from 12 samples from Vindija 33.19, one of 19 bone fragments from Vindija Cave determined to be of Neandertal origin by mitochondrial (mt) DNA analyses (19). In addition, 567 mg were removed for radiocarbon dating and yielded a date of greater than 45,500 years before present (OxA 32,278). One of the DNA extracts, generated from 41 mg of bone material, contained more hominin DNA than the other extracts. We created additional libraries from this extract, but to maximize the number of molecules retrieved from the specimen, we omitted the uracil-DNA-glycosylase (UDG) treatment (20, 21). A total of 24 billion DNA fragments were sequenced, and ~10% of these could be mapped to the human genome. Their average

[1]Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. [2]Bioinformatics Research Centre, Aarhus University, DK-8000 Aarhus C, Denmark. [3]Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA 98195, USA. [4]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China. [5]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [6]Anthropology Center of the Croatian Academy of Sciences and Arts, 10000 Zagreb, Croatia. [7]ANO Laboratory of Prehistory 14 Linia 3-11, St. Petersburg 1990 34, Russia. [8]Department of Integrative Biology, University of California, Berkeley, CA 94720-3140, USA. [9]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [10]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA. [11]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.
*Corresponding author. Email: pruefer@eva.mpg.de (K.P.); paabo@eva.mpg.de (S.P.) †These authors contributed equally to this work.
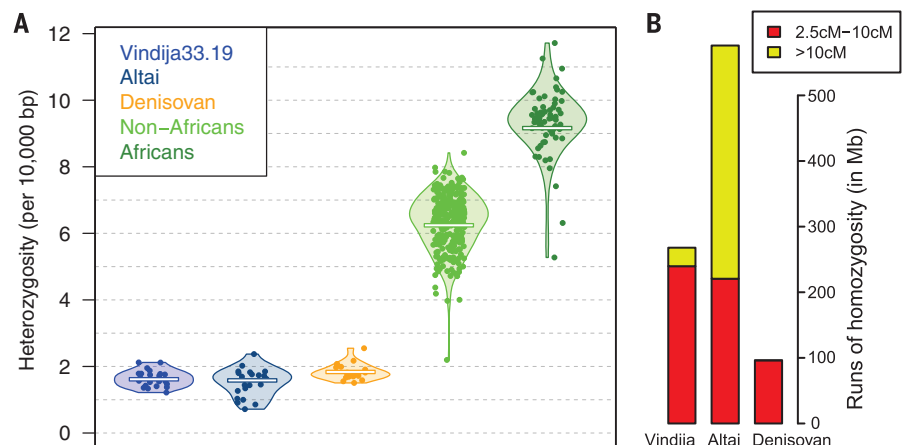
Fig. 1. Heterozygosity and inbreeding in the Vindija Neandertal. (A) Distribution of heterozygosity over all autosomes in the three archaic hominins, 12 non-Africans, and 3 Africans. Each dot represents the heterozygosity measured for one autosome. The center bar indicates the mean heterozygosity across the autosomal genome(s). (B) Genome covered by shorter (2.5 to 10 cM, red) and longer (>10 cM, yellow) runs of homozygosity in the three archaic hominins.

length was 53 base pairs (bp), and they yielded 30-fold coverage of the ~1.8 billion bases of the genome to which such short fragments can be confidently mapped.

We estimated present-day human DNA contamination among the DNA fragments (20). First, using positions in the mtDNA where present-day humans differ from Neandertals, we estimated an mtDNA contamination rate of 1.4 to 1.7%. Similarly, using positions in the autosomal genome where all present-day humans carry derived variants whereas all archaic genomes studied to date carry ancestral variants, we estimated a nuclear contamination rate of 0.17 to 0.48%. Because the coverage of the X chromosome is similar to that of the autosomes, we inferred that the Vindija 33.19 individual is a female, allowing us to use DNA fragments that map to the Y chromosome to estimate a male DNA contamination of 0.74% (between 0.70 and 0.78% for each of the nine sequencing libraries). Finally, using a likelihood method (2, 3), we estimated the autosomal contamination of 0.18 to 0.23%. We conclude that the nuclear DNA contamination rate among the DNA fragments sequenced is less than 1%. After genotyping, this will result in contamination that is much lower than 1%.

Because ~76% of the DNA fragments were not UDG-treated, they carry C to T substitutions throughout their lengths. This causes standard genotyping software to generate false heterozygous calls. To overcome this, we implemented snpAD, a genotyping software that incorporates a position-dependent error profile to estimate the most likely genotype for each position in the genome. This results in genotypes of quality comparable to that of UDG-treated ancient DNA given our genomic coverage (20). The high coverage of the Vindija genome also allowed for characterization of longer structural variants and segmental duplications (20).

To gauge whether the Vindija 33.19 bone might stem from a previously sequenced individual from Vindija Cave, we compared heterozygous sites in the Vindija 33.19 genome to DNA fragments sequenced from the other bones. The three bones from which a low-coverage composite genome has been generated (Vindija 33.16, 33.25, and 33.26) do not share variants with Vindija 33.19 at a level compatible with their being derived from the same individual. By contrast, more than 99% of heterozygous sites in the chromosome 21 sequence from Vindija 33.15 (13) are shared with Vindija 33.19, indicating that they come from the same individual (20). Additionally, two of the other three bones may come from individuals that shared a maternal ancestor to Vindija 33.19 relatively recently in their family history because all carry identical mtDNAs.

In addition to the Altai Neandertal genome, a genome from a Denisovan, an Asian relative of Neandertals, has been sequenced to high coverage (~30-fold) from Denisova Cave. These two genomes are similar in that their heterozygosity is about one-fifth of that of present-day Africans and about one-third of that of
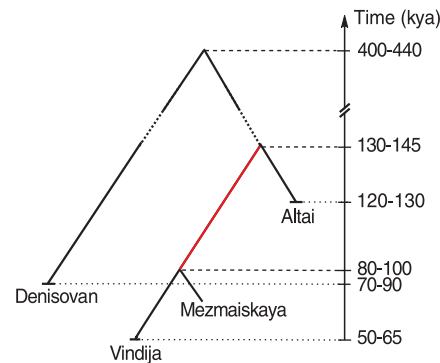

Fig. 2. Approximate ages of specimens and population split times. Age estimates for the genomes estimated from branch shortening (i.e., the absence of mutations in the archaic genomes) are indicated by dotted lines. Population split time estimates are indicated by dashed lines. The majority of Neandertal DNA in present-day people comes from a population that split from the branch indicated in red. All reported ages assume a human-chimpanzee divergence of 13 million years. Numbers show ranges over point estimates (split times), or ranges over different data filters (branch shortening).

present-day Eurasians. We estimated the heterozygosity of the Vindija 33.19 autosomal genome to 1.6×10^{-5}; similar to the heterozygosity of the Altai Neandertal genome and slightly lower than that of the Denisovan genome (1.8×10^{-5}) (Fig. 1A). Thus, low heterozygosity may be a feature typical of archaic hominins, suggesting that they lived in small and isolated populations with an effective population size of around 3000 individuals (20). In addition to low overall heterozygosity, the Altai Neandertal genome carried segments of many megabases (Mb) [>10 centimorgans (cM)] without any differences between its two chromosomes, indicating that the parents of that individual were related at the level of half-sibs (2). Such segments are almost totally absent in the Vindija genome (Fig. 1B), suggesting that the extreme inbreeding between the parents of the Altai Neandertal was not ubiquitous among Neandertals. We note, however, that the Vindija genome carries extended homozygous segments (>2.5 cM) comparable to what is seen in some isolated Native American populations today (20).

The high quality of the three archaic genome sequences allows their approximate ages to be estimated from the number of new nucleotide substitutions that they carry relative to present-day humans when compared to the inferred ancestor shared with apes (1). Using this approach, we estimate that the Vindija 33.19 individual lived 52 thousand years ago (ka), the Altai Neandertal individual 122 ka, and the Denisovan individual 72 ka (Fig. 2) (20). Many factors make such absolute age estimates tentative. Among these are uncertainty in generation times and mutation rates. Nevertheless, these results indicate that the Altai Neandertal

lived about twice as far back in time as the Vindija 33.19 Neandertal, whereas the Denisovan individual lived after the Altai but before the Vindija Neandertal.

We next estimated when ancestral populations that gave rise to the three archaic genomes and to modern humans split from each other based on the extent to which they share genetic variants (1–3, 20). The estimated population split time between the Vindija Neandertal and the Denisovan is 390 to 440 ka and that between the Vindija Neandertal and modern humans 520 to 630 ka, in agreement with previous estimates using the Altai Neandertal (2). The split time between the Vindija and the Altai Neandertals is estimated to be 130 to 145 ka. To estimate the population split time to the Mezmaiskaya 1 Neandertal previously sequenced to 0.5-fold coverage, we prepared and sequenced libraries yielding an additional 1.4-fold coverage. Because the present-day human DNA contamination of these libraries is on the order of 2 to 3% (20), we estimated the population split time to the Vindija 33.19 individual with and without restricting the analysis of the Mezmaiskaya 1 individual to fragments that show evidence of deamination. The resulting split time estimates are 100 ka for the deaminated fragments and 80 ka for all fragments (Fig. 2).

It has been suggested that Denisovans received gene flow from a hominin lineage that diverged prior to the common ancestor of modern humans, Neandertals, and Denisovans (2). In addition, it has been suggested that the ancestors of the Altai Neandertal received gene flow from early modern humans that may not have affected the ancestors of European Neandertals (13). In agreement with these studies, we find that the Denisovan genome carries fewer derived alleles that are fixed in Africans, and thus tend to be older, than the Altai Neandertal genome, whereas the Altai genome carries more derived alleles that are of lower frequency in Africa, and thus younger, than the Denisovan genome (20). However, the Vindija and Altai genomes do not differ significantly in the sharing of derived alleles with Africans, indicating that they may not differ with respect to their putative interactions with early modern humans (Fig. 3, A and B). Thus, in contrast to earlier analyses of chromosome 21 data for the European Neandertals (13), analyses of the full genomes suggest that the putative early modern human gene flow into Neandertals occurred before the divergence of the populations ancestral to the Vindija and Altai Neandertals ~130 to 145 ka (Fig. 2). Coalescent simulations show that a model with only gene flow from a deeply diverged hominin into Denisovan ancestors explains the data better than one with only gene flow from early modern humans into Neandertal ancestors, but that a model involving both gene flows explains the data even better. It is likely that gene flow occurred between many or even most hominin groups in the late Pleistocene and that
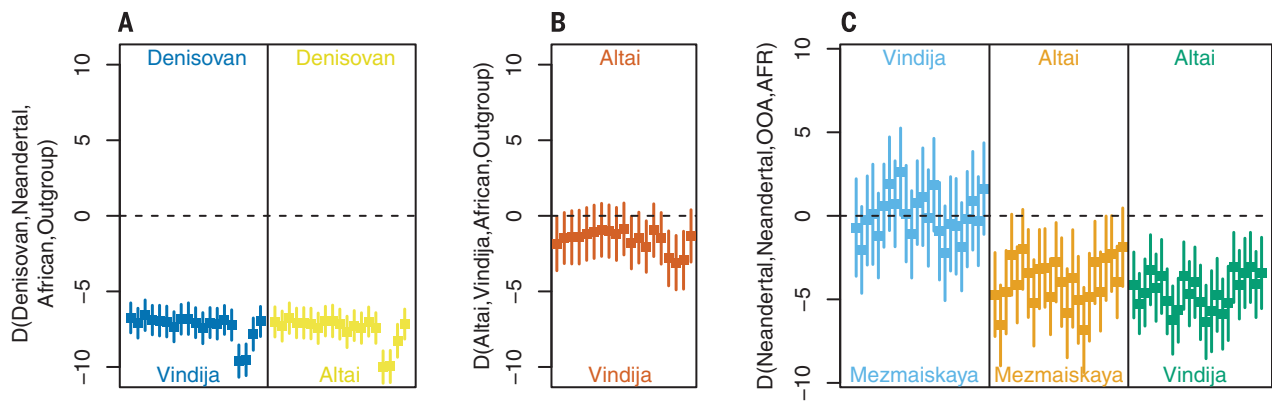
Fig. 3. Allele sharing between archaic and modern humans. (**A**) Derived allele sharing (in percent) of 19 African populations with the Altai and Denisovan, and Vindija and Denisovan genomes, respectively. (**B**) Sharing of derived alleles in each of the 19 African populations with the Vindija and Altai genomes. (**C**) Allele sharing of Neandertals with non-Africans and Africans. Points show derived allele sharing (in percent) for all pairwise comparisons between non-Africans (OOA: French, Sardinian, Han, Dai, Karitiana, Mixe, Australian, Papuan) and Africans (AFR: San, Mbuti, Yoruba). Mezmaiskaya 1 data were restricted to sequences showing evidence of deamination to reduce the influence of present-day human DNA contamination. Bars show two standard errors from the mean in all plots.
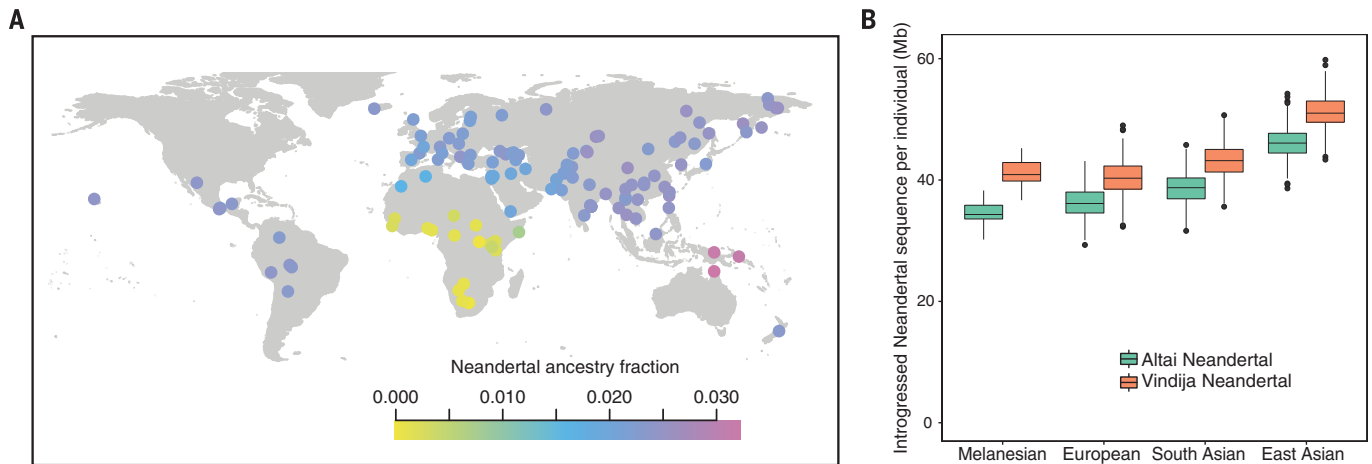
Fig. 4. Estimates of fraction of Neandertal DNA for present-day populations. (**A**) Colors indicate Neandertal ancestry estimates (20). Oceanian populations show high estimates due to Denisovan ancestry that is difficult to distinguish from Neandertal ancestry. (**B**) Amount of Neandertal sequence in present-day Europeans, South Asians, and East Asians (20).

more such events will be detected as more ancient genomes of high quality become available.

A proportion of the genomes of all present-day people whose roots are outside Africa derives from Neandertals (2, 3, 22). We tested if any of the three sequenced Neandertals falls closer to the lineage that contributed DNA to present-day non-Africans by asking if any of them shares more alleles with present-day non-Africans than the others (20, 23). The Vindija 33.19 and Mezmaiskaya 1 genomes share more alleles with non-Africans than the Altai Neandertal, and there is no indication that the former two genomes differ in the extent of their allele sharing with present-day people (Fig. 3C). Using a likelihood approach, we estimate the proportion of Neandertal DNA in present-day populations that is closer to the Vindija than the Altai genomes to be 99 to 100% (20). Thus, the majority of Neandertal DNA in present-day populations appears to come from

Neandertal populations that diverged from the Vindija and Mezmaiskaya 1 Neandertals before their divergence from each other some 80 to 100 ka.

The two high-coverage Neandertal genomes allow us to estimate the proportion of the genomes of present-day people that derive from Neandertals with greater accuracy than was hitherto possible. We asked how many derived alleles non-Africans share with the Altai Neandertal relative to how many derived alleles the Vindija Neandertal shares with the Altai Neandertal—essentially asking how close non-Africans are to being 100% Neandertal (24). We find that non-African populations outside Oceania carry between 1.8 and 2.6% Neandertal DNA (Fig. 4A), higher than previous estimates of 1.5 to 2.1% (2). As described (25), East Asians carry somewhat more Neandertal DNA (2.3 to 2.6%) than people in Western Eurasia (1.8 to 2.4%).

We also identified regions of Neandertal ancestry in present-day Europeans and Asians using the Vindija and the Altai Neandertal genomes (8, 20). The Vindija genome allows us to identify ~10% more Neandertal DNA sequences per individual than the Altai Neandertal genome (e.g., 40.4 Mb versus 36.3 Mb in Europeans) because of the closer relationship between the Vindija genome and the introgressing Neandertal populations. In Melanesians, the increased power to distinguish between Denisovan and Neandertal DNA sequences results in the identification of 20% more Neandertal DNA (Fig. 4B).

Many Neandertal variants associated with phenotypes and susceptibility to diseases have been identified in present-day non-Africans (6, 7, 10–12). That the Vindija Neandertal genome is more closely related to the introgressing Neandertals allows ~15% more such variants to be identified (20). Among these are variants associated with

plasma concentrations of low-density lipoprotein (LDL) cholesterol (rs10490626) and vitamin D (rs6730714), eating disorders (rs74566133), visceral fat accumulation (rs2059397), rheumatoid arthritis (45475795), schizophrenia (rs16977195), and the response to antipsychotic drugs (rs1459148). This adds to mounting evidence that Neandertal ancestry influences disease risk in present-day humans, particularly with respect to neurological, psychiatric, immunological, and dermatological phenotypes (7).

REFERENCES AND NOTES

1. M. Meyer *et al.*, *Science* **338**, 222–226 (2012).
2. K. Prüfer *et al.*, *Nature* **505**, 43–49 (2014).
3. R. E. Green *et al.*, *Science* **328**, 710–722 (2010).
4. K. Harris, R. Nielsen, *Genetics* **203**, 881–891 (2016).
5. I. Juric, S. Aeschbacher, G. Coop, *PLOS Genet.* **12**, e1006340 (2016).
6. S. Sankararaman *et al.*, *Nature* **507**, 354–357 (2014).
7. C. N. Simonti *et al.*, *Science* **351**, 737–741 (2016).
8. B. Vernot, J. M. Akey, *Science* **343**, 1017–1021 (2014).
9. F. Racimo, S. Sankararaman, R. Nielsen, E. Huerta-Sánchez, *Nat. Rev. Genet.* **16**, 359–371 (2015).
10. A. L. Williams *et al.*, *Nature* **506**, 97–101 (2014).
11. M. Dannemann, A. M. Andrés, J. Kelso, *Am. J. Hum. Genet.* **98**, 22–33 (2016).
12. H. Quach *et al.*, *Cell* **167**, 643–656.e17 (2016).
13. M. Kuhlwilm *et al.*, *Nature* **530**, 429–433 (2016).
14. S. Castellano *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6666–6671 (2014).
15. T. Devièse *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 201709235 (2017).
16. A. D. Greenwood, S. Pääbo, *Mol. Ecol.* **8**, 133–137 (1999).
17. J. Dabney *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15758–15763 (2013).
18. M. T. Gansauge, M. Meyer, *Nat. Protoc.* **8**, 737–748 (2013).
19. M. T. Gansauge, M. Meyer, *Genome Res.* **24**, 1543–1549 (2014).
20. Materials and methods are available as supplementary materials.
21. A. W. Briggs *et al.*, *Nucleic Acids Res.* **38**, e87 (2010).
22. S. Mallick *et al.*, *Nature* **538**, 201–206 (2016).
23. N. Patterson *et al.*, *Genetics* **192**, 1065–1093 (2012).
24. D. Reich *et al.*, *Am. J. Hum. Genet.* **89**, 516–528 (2011).
25. J. D. Wall *et al.*, *Genetics* **194**, 199–209 (2013).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6363/655/suppl/DC1
Materials and Methods
Supplementary text
Figs. S1 to S103
Tables S1 to S52
References (*26–118*)

Science

A high-coverage Neandertal genome from Vindija Cave in Croatia

Kay Prüfer, Cesare de Filippo, Steffi Grote, Fabrizio Mafessoni, Petra Korlevic, Mateja Hajdinjak, Benjamin Vernot, Laurits Skov, Pinghsun Hsieh, Stéphane Peyrégne, David Reher, Charlotte Hopfe, Sarah Nagel, Tomislav Maricic, Qiaomei Fu, Christoph Theunert, Rebekah Rogers, Pontus Skoglund, Manjusha Chintalapati, Michael Dannemann, Bradley J. Nelson, Felix M. Key, Pavao Rudan, Zeljko Kucan, Ivan Gusic, Liubov V. Golovanova, Vladimir B. Doronichev, Nick Patterson, David Reich, Evan E. Eichler, Montgomery Slatkin, Mikkel H. Schierup, Aida M. Andrés, Janet Kelso, Matthias Meyer and Svante Pääbo

Revelations from a Vindija Neandertal genome

Neandertals clearly interbred with the ancestors of non-African modern humans, but many questions remain about our closest ancient relatives. Prüfer *et al.* present a 30-fold-coverage genome sequence from 50,000- to 65,000-year-old samples from a Neandertal woman found in Vindija, Croatia, and compared this sequence with genomes obtained from the Altai Neandertal, the Denisovans, and ancient and modern humans (see the Perspective by Bergström and Tyler-Smith). Neandertals likely lived in small groups and had lower genetic diversity than modern humans. The findings increase the number of Neandertal variants identified within populations of modern humans, and they suggest that a larger number of phenotypic and diseaserelated variants with Neandertal ancestry remain in the modern Eurasian gene pool than previously thought.

Science, this issue p. 655; see also p. 586

ARTICLE TOOLS	http://science.sciencemag.org/content/358/6363/655
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2017/10/04/science.aao1887.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/358/6363/586.full http://science.sciencemag.org/content/sci/358/6363/659.full
REFERENCES	This article cites 111 articles, 25 of which you can access for free http://science.sciencemag.org/content/358/6363/655#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service