Nuclear DNA from two early Neandertals reveals 80,000 years of genetic continuity in Europe

Stéphane Peyrégne1*, Viviane Slon1, Fabrizio Mafessoni1, Cesare de Filippo1, Mateja Hajdinjak1, Sarah Nagel1, Birgit Nickel1, Elena Essel1, Adeline Le Cabc2, Kurt Wehrberger3, Nicholas J. Conard4, Claus Joachim Kind2, Cosimo Posth6, Johannes Krause6, Grégory Abrams7, Dominique Bonjean7, Kévin Di Modica7, Michel Toussaïnt7, Janet Kelso1, Matthias Meyer1, Svante Pääbo1, Kay Prüfer1,6*

Little is known about the population history of Neandertals over the hundreds of thousands of years of their existence. We retrieved nuclear genomic sequences from two Neandertals, one from Hohlenstein-Stadel Cave in Germany and the other from Scladina Cave in Belgium, who lived around 120,000 years ago. Despite the deeply divergent mitochondrial lineage present in the former individual, both Neandertals are genetically closer to later Neandertals from Europe than to a roughly contemporaneous individual from Siberia. That the Hohlenstein-Stadel and Scladina individuals lived around the time of their most recent common ancestor with later Neandertals suggests that all later Neandertals trace at least part of their ancestry back to these early European Neandertals.

**INTRODUCTION**

Neandertals lived in western Eurasia for hundreds of thousands of years before modern humans spread outside Africa. The earliest morphological and genetic evidence of Neandertals reaches back approximately 430 thousand years ago (1, 2), while the last Neandertals disappeared around 40 ka ago (3). Denisovans, a sister group of Neandertals discovered by genetic analysis of remains from Denisova Cave (Altai Mountains, Russia; Fig. 1) (4), may have been widespread in Asia (5).

Recent analyses of nuclear genome sequences from Neandertals have shown that, toward the end of their existence, Neandertals across their entire geographic range from Europe to Central Asia belonged to a single group sharing a most recent common ancestor less than 97 ka ago (6, 7). However, population discontinuity has been observed in Denisova Cave, Russia, further back in time, where the Neandertal component in the genome of a ~90-ka-old Neandertal-Denisovan offspring (7) shows stronger affinities to late Neandertals in Europe than to the Altai Neandertal, another individual found in the same cave (8). The latter lived 120 ka ago according to the number of missing mutations in her genome compared to present-day human genomes. Thus, a population replacement likely occurred in the easternmost part of the Neandertal territory between 90 and 120 ka ago.

Without nuclear genome sequences from early European Neandertals, it has not been possible to determine the origin of the replacement and whether it was limited to the east. To learn more about the early population history of European Neandertals, we studied the nuclear genomes of two individuals from Western Europe that are dated to approximately 120 ka ago and from which only mitochondrial DNA (mtDNA) was previously recovered. The first, a femur from Hohlenstein-Stadel Cave (HST) in Germany (9), carries an mtDNA genome that falls basal to all other known Neandertal mtDNAs and was dated to ~124 ka ago based on its branch length in the mitochondrial tree (95% highest posterior density interval [HPDI], 62 to 183 ka ago; associated faunal remains suggest a date between 80 and 115 ka ago) (10). The second, a maxillary bone from Scladina Cave [Scladina 1-4A, here referred to as Scladina (11)], yielded the hypervariable region of the mtDNA genome (12) and was dated to 127 ka ago based on uranium and thorium isotopic ratios [1 SD, 95 to 173 ka ago (13)].

**RESULTS**

Because of the great age of the specimens and their extensive handling in the decades after their discovery, obtaining DNA of sufficient quantity for genome-wide analyses is challenging. We thus used the most efficient DNA extraction and library preparation methods currently available (14–16) and coupled them with pre-treatment methods for the removal of human and microbial contamination (note S1) (17). We then characterized the libraries prepared from both specimens by hybridization capture of mtDNA and shallow shotgun sequencing to identify those libraries that were most suitable for further analysis (Materials and Methods; notes S2 and S3).

On the basis of 450- and 107-fold coverage of the mtDNA genome, respectively, we were able to verify the published mtDNA sequence from HST (10) and reconstruct the complete mtDNA of Scladina (notes S5 and S6). Scladina dates to ~120 ka ago according to the branch length in the mtDNA tree (95% HPDI, 76 to 168 ka ago; note S7), consistent with the aforementioned date. Confirming previous results from the hypervariable region (10), we find that the complete Scladina mtDNA is most similar to the Altai Neandertal mtDNA (note S7). On the basis of only the mtDNA, it thus appears that both individuals fall outside the variation of later European Neandertals. However, mtDNA is only a single maternally inherited locus and of limited value for reconstructing the relationships among Neandertals and other archaic humans (1).
We generated a total of 168 and 78 million base pairs (Mb) of nuclear DNA sequence from the two individuals, respectively (note S3). Ancient DNA sequences often carry cytosine to thymine substitutions that are caused by cytosine deamination accumulating in DNA fragments over time, most often at the ends of the fragments (18). The frequency of these substitutions on both molecule ends (1), confirms that ancient nuclear DNA is present but that a large proportion of the HST and Scladina sequences are contaminants from present-day humans (note S8). At positions that are derived only in the Altai Neandertal [ancestral in the genomes of a Denisovan (19) and an Mbuti (19)], 57.8 and 31.1% of HST and Scladina sequences, respectively, show the Neandertal allele (note S9). However, sequences also match the derived allele in an Mbuti genome (19) more often than the high-coverage genome of the Altai Neandertal does (HST, 8.8%; Scladina, 22.3%, Altai Neandertal, 1.4%; note S8). This excess of sharing suggests that 23 and 65% of the HST and Scladina sequences, respectively, are modern human contaminants (note S8). To reduce contamination, we restricted all further analyses to sequences that show evidence for deamination (Materials and Methods), leaving us with 51 Mb of the HST genome and 12 Mb of the Scladina genome (note S3). This procedure reduces the estimated contamination to 2% for HST and 5.5% for Scladina and results in a coverage on the X chromosome and autosomes that shows that HST was male, whereas Scladina was female, in agreement with the morphological assessments (notes S4 and S8) (9, 13).

To investigate the relationship of HST and Scladina to Neandertals, we compared their nuclear sequences to two high-coverage Neandertal genomes. The genome of a ~50-ka-old Neandertal from Vindija Cave in Croatia [Vindija 33.19, referred to as Vindija (20)] is a representative of the group of later Neandertals that inhabited Eurasia after 90 ka ago (6, 7), whereas the Altai Neandertal represents the earlier group of Neandertals in the east. We identified Vindija-specific- and Altai-specific-derived variants by randomly sampling an allele from these two Neandertal genomes and retaining only those variants that differ from the other high-coverage Neandertal genome and from the Denisovan (19), one Mbuti (19), and several ape outgroup genomes (Materials and Methods) (21–24). At these sites, HST shares Vindija-specific alleles more often than Altai-specific alleles (531 versus 466; two-sided binomial test, P = 0.043), while no significant difference was observed for Scladina (110 versus 106; P = 0.838; Fig. 2 and note S9). Since the number of DNA sequences with putative deamination-induced substitutions is small for Scladina, we repeated this analysis including all sequences and found that Scladina then shows more Vindija-specific alleles than Altai-specific alleles (Scladina, 443 versus 321; P < 10^{-4}; HST, 1676 versus 1326; P < 10^{-5}; note S9). This cannot be accounted for by contamination with present-day human DNA, since the proportion of Neandertal ancestry in present-day humans is, on average, smaller than 3% (note S9). Thus, these results indicate that both HST and Scladina are more closely related to Vindija than they are to the Altai Neandertal.

If HST and Scladina truly have a most recent common ancestor with Vindija more recently than with the Altai Neandertal, then their genomes are expected to share derived alleles with the Altai Neandertal genome as often as the Vindija genome does. However, the genomes of Vindija and the Altai Neandertal share more derived alleles with each other than the HST or Scladina genomes share with either of them (Fig. 2 and note S9). This imbalance in allele sharing can largely be accounted for by a reference bias that favors the alignment of HST and Scladina sequences that carry a modern human reference allele over those carrying a Neandertal allele (note S9). By aligning to an alternative reference genome containing alleles seen in the high-coverage Neandertals, we recover further sequences that we combine with the original set of alignments and compensate for this bias (Fig. 2, Materials and Methods, and note S9). The remaining imbalance in allele sharing can be explained by contamination and sequencing errors in Scladina and HST (Fig. 2 and note S9).

Using the reference bias–corrected alignments and two methods, we estimated split times between the populations represented by HST and Scladina and the Vindija population (note S10). Our first estimates are based on the sharing of derived alleles by HST/Scladina at sites where the high-coverage Vindija genome is heterozygous [F(A|B) statistic (8, 20)]. The estimated split times of HST and Scladina from the ancestor with Vindija are 101 ka ago [confidence interval (CI), 80 to 123 ka ago] and 100 ka ago (CI, 66 to 153 ka ago), respectively. The second estimates are based on a coalescent divergence
model (25) and suggest, for both Neandertals, a ~10-ka-long shared history with Vindija after the split of the latter from the Altai Neandertal population (i.e., 122 to 141 ka ago, assuming 130 to 145 ka ago for the Vindija-Altai split time; note S10). The estimates from both methods are close to the estimated age of ~120 ka ago for these individuals (10, 13). Therefore, HST and Scladina could be members of an ancestral Neandertal population that gave rise to all Neandertals sequenced to date with the exception of the Altai Neandertal, who did not leave any descendants among sequenced Neandertals. This ancestral Neandertal population was established in the west by ~120 ka ago, and later descendants may have migrated east and replaced at least partially the eastern population of Neandertals represented by the Altai Neandertal.

It seems unexpected that HST carries an mtDNA lineage that diverged ~270 ka ago from other mtDNAs, given the recent population split times from the Vindija ancestors and the low levels of genetic diversity in the nuclear genomes of Neandertals (8, 20). To test whether such a deeply diverging mtDNA lineage could be maintained in the Neandertal population by chance, we used coalescent simulations with a demography estimated from the high-coverage Neandertal genomes (20), which was scaled to match the lower effective population size of the mtDNA, taking into account the difference in effective population size between the two sexes (8). We find that population split times between HST and other Neandertals of less than 150 ka ago make the occurrence of a mitochondrial time to the most recent common ancestor (TMRCA) of 270 ka ago unlikely (1.2% of all simulated loci have such a deep TMRCA; note S11). We note that this result is robust to uncertainties in the estimates of the Neandertal population size and of the mitochondrial TMRCA (note S11). The presence of this deeply divergent mtDNA in HST thus suggests a more complex scenario in which HST carries some ancestry from a genetically distant population.

**DISCUSSION**

What scenarios could explain the deeply divergent mtDNA in HST? An explanation could be related to a replacement of mtDNAs in Neandertals that has been suggested to explain the discrepancy between the mtDNA divergence time (<470 ka ago) (10) and the population split times based on nuclear DNA (>520 ka ago) (20) between modern humans and Neandertals. The Sima de los Huesos hominins, and perhaps other early Neandertals, carried mtDNAs that shared a common ancestor with Denisovan mtDNAs more recently than with those of modern humans, whereas later Neandertals carried mtDNAs that shared a more recent common ancestor with the mtDNAs of modern humans. Admixture between Neandertals and ancestors or relatives of modern humans could explain the origin of this later Neandertal mtDNA (1, 10). If several mtDNAs were introduced into the Neandertal population by such a putative gene flow, then the deeply divergent mtDNA in HST may represent the remnants of the mitochondrial diversity of this introgressing population (Fig. 3) (10). This would imply that this admixture into Neandertals occurred later than the previously suggested lower boundary of 270 ka ago (219 to 316 ka ago) (10). We estimate that
the probability for this late mtDNA replacement is nearly identical to the admixture rate, i.e., more than 5% admixture is required to reach a probability of 5% for such an event to occur (note S12) (10).

An alternative source for the deeply divergent mtDNA in HST could be an isolated Neandertal population, for example, a population that separated from other Neandertals before the glacial period preceding HST and Scladina (~130 to 190 ka ago; Fig. 3). Such an isolated population may have preserved the mtDNA that was later re-introduced during a warmer period between 115 and 130 ka ago (the “Eemian” period) when these populations met again and gene flow resumed. We note that the contact may have been a result of a recolonization from the Middle East or Southern Europe (26, 27).

Our analysis shows that late Neandertals that lived in Europe at around 40 ka ago trace at least part of their ancestry back to Neandertals that lived there approximately 80,000 years earlier. The latter became widespread, appearing in the east at least 90 ka ago. The genetic continuity seen in Europe contrasts, however, with the deeply divergent mtDNA in HST, which hints at a more complex history that affected at least some of the European Neandertals before ~120 ka ago. DNA sequences from even older Neandertals are needed to clarify whether Neandertal substructure, gene flow from relatives of modern humans, or both are the explanation for HST’s peculiar mtDNA.

**MATERIALS AND METHODS**

**DNA extraction and library preparation**

Bone or tooth powder was sampled from the HST and Scladina specimens using a sterile dentistry drill after removing the external surface of the specimen at the sampling site (note S1). For the initial assessment of ancient DNA preservation in the specimens, DNA was extracted using a silica-based method (14), as implemented in (17), either from untreated powder or following one of three decontamination procedures described in the note S1. The treatment of the bone powder with 0.5% sodium hypochlorite yielded the highest proportion of fragments mapping to the human reference genome for HST and resulted in the lowest estimates of contamination by present-day human mtDNA for both HST and Scladina (note S2). For the subsequent generation of additional sequencing data, the bone or tooth powder was therefore incubated in 0.5% sodium hypochlorite solution before DNA extraction (17). Single-stranded DNA libraries were prepared from these DNA extracts (15, 16). Each library was tagged with two unique indexes, amplified into plateau, and purified (17, 28) before shotgun sequencing. In addition, an aliquot of each indexed DNA library was enriched for human mtDNA fragments using a hybridization capture method (29).

**Sequencing and raw data processing**

Libraries were sequenced on an Illumina MiSeq and HiSeq 2500 platforms in 76-cycle paired-end runs (28). For a detailed description of the read processing, see note S3. When analyzing the relationship of HST and Scladina to Vindija and the Altai Neandertal, further processing was necessary to avoid a reference bias of the alignments. First, we aligned DNA sequences to both the human reference genome (GRCh37/hg19) and a modified (“Neandertalized”) version of the reference genome that includes the alternative alleles seen in Vindija and/or the Altai Neandertal. If there was more than one alternative base at a given site (i.e., a triallelic site), then a random
base was chosen. We then merged sequences that aligned to either reference genome and removed one duplicate of the sequences that mapped to both. If a sequence aligned to the two references at different positions, then both alignments were excluded (representing 522 and 332 such sequences for HST and Scladina, respectively). We developed an algorithm called bam-mergeRef to perform these merging steps, wrote it in C++, and made it available on GitHub (https://github.com/StephanePeyregne/bam-mergeRef). For a description of the reference bias and the effects of this processing, see note S9. Sequences from libraries enriched for mtDNA fragments were aligned to the revised Cambridge reference sequence (30) or the Altai Neandertal mtDNA with the same parameters as those applied to nuclear sequences (note S3).

Analysis of the mitochondrial genomes
Mitochondrial genome sequences were reconstructed from a consensus call at each position where at least two-thirds of the fragments aligning to the Altai Neandertal mtDNA carried the same base and if the position was covered by at least three fragments. Further details about the mtDNA reconstruction and the estimated proportion of contamination by present-day human mtDNA for both specimens, as well as the phylogenetic analyses, are described in notes S5 to S7.

Analysis of the relationship to other archaic and modern humans
We determined lineage-specific derived alleles by comparing the high-quality genomes of Vindija and the Altai Neandertal (8, 20), Denisova 3 (19), and a present-day human from Africa [Mbuti, HGDP00456 (19)]. At sites where one of these individuals was heterozygous, we randomly picked an allele. An allele was regarded as ancestral when it matched at least three of four aligned great ape reference genome assemblies [chimpanzee (panTro4) (21), bonobo (panPan1.1) (22), gorilla (gorGor3) (23), and orangutan (ponAbe2) (24)]. LASTZ alignments to the human genome GRCh37/hg19 prepared in-house and by the University of California, Santa Cruz, genome browser (31)]. The fourth ape was allowed to carry a third allele or have missing data but not to carry the alternative allele. To avoid errors from ancient DNA damage on HST and Scladina sequences at these positions, we only considered sequences that aligned in forward orientation when the ancestral or derived allele at the position was a G or in reverse orientation when either allele was a C and excluded sequences with a third allele. Only positions passing the published map35_100 filter for Denisova 3, Vindija, and the Altai Neandertal genotypes (20) were retained. A correction for the level of present-day human DNA contamination was applied in this analysis and is described in note S9.

Assessment of present-day human nuclear DNA contamination
We estimated contamination from the proportion $p$ of sites where the Neandertal (HST or Scladina) carries a derived allele seen in the genome of a present-day Mbuti individual [HGDP00456 (19)] but not in Denisova 3 and a Neandertal high-coverage genome (either Vindija or the Altai Neandertal). This proportion $p$ is the result of a mixture of present-day human DNA contamination and DNA endogenous to the ancient specimens as follows: $c \times p_e + (1 - c) \times p_c = p$, with $p_c$ and $p_e$ being the expected proportions of derived alleles for the contaminant and endogenous molecules, respectively, and $c$ is the contamination rate. The proportions $p_c$ and $p_e$ are unknown but can be approximated by the observed proportion of shared alleles between the Mbuti genome and another present-day human genome [33.2% for either a French, HGDP00521 (19) or a Han, HGDP00775 (8)] or a Neandertal high-coverage genome (1.4% for the Altai Neandertal and 1.5% for Vindija), respectively. To compute $p_c$ and $p_e$, we used the genotypes from the high-coverage genomes (randomly sampling one allele at heterozygous positions) under the assumption that these are unaffected by sequencing errors or present-day human DNA contamination. CIs were calculated from the bounds of the binomial CIs of $p$. Assuming that $p$ is the parameter of a binomial distribution (instead of the expected success rate in Poisson trials) is a conservative approximation for calculating CIs, as the variance for Poisson trials is lower or equal to the variance of the binomial distribution with parameter $p$.

Coalescent simulations of the mitochondrial common ancestor
Coalescent simulations using scrm (32) were used to compute the expected distribution of times to TMRCAs for the mitochondrial genomes, given different population split times (from 100 to 200 ka ago, with a step of 10 ka). To be able to compare these to the estimated date for the common mitochondrial ancestor of HST and Vindija, the simulations followed the Vindija demographic history estimated from the Pairwise Sequentially Markovian Coalescent model (PSMC) (33) [that assumed a mutation rate of $1.45 \times 10^{-8}$ per base pair per generation and a generation time of 29 years (20)]. The scaling for the mitochondrial effective population size was calculated according to the females-to-males ratio, previously estimated to be 1.54 for the Altai Neandertal population (note S11) (8).

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/6/eaaw5873/DC1
Note S1. Ancient DNA recovery and treatment.
Note S2. Decontamination methods and initial screening.
Note S3. Data generation and data processing.
Note S4. Sex determination.
Note S5. Mitochondrial contamination estimates.
Note S6. Reconstruction of the mitochondrial genomes.
Note S7. Phylogenetic analysis of the mitochondrial genomes.
Note S8. Characterization of present-day human DNA contamination in the nuclear genome.
Note S9. Genetic relationships and effect of present-day human DNA contamination, sequencing errors, and reference bias.
Note S10. Split time estimates.
Note S11. Discordance between the nuclear and mitochondrial divergence of HST to other Neandertals.
Note S12. Likelihood of a recent mitochondrial replacement in Neandertals.
Table S1. Overview of DNA extracts and libraries prepared from the HST femur.
Table S2. Overview of DNA extracts and libraries prepared for Scladina I-4A.
Table S3. DNA content in the libraries prepared from HST extracts prepared following different decontamination methods (set 1 in table S1).
Table S4. DNA content in the libraries prepared from the bone powder treated with sodium hypochlorite.
Table S5. DNA content in the initial libraries prepared from the untreated extracts from Scladina I-4A.
Table S6. Present-day human DNA contamination estimates after three decontamination methods applied to bone powder from the HST femur.
Table S7. Present-day human DNA contamination estimates from Scladina I-4A mtDNA based on differences between Neandertals and modern humans.
Table S8. Sequencing summary statistics for HST with the following filters: length $\geq 35$ bp and mapping quality $\geq 23$.
Table S9. Sequencing summary statistics for HST with the following filters: length $\geq 30$ bp and mapping quality $\geq 25$. 

Peyrégne et al., Sci. Adv. 2019; 5 : eaaw5873 26 June 2019
Table S10. Sequencing summary statistics for Scladina I-4A with the following filters: length (≥35 bp) and mapping quality (≥25).
Table S11. Sequencing summary statistics for Scladina I-4A with the following filters: length (≥30 bp) and mapping quality (≥25).
Table S12. Sequencing statistics of the negative controls for HST (see table S11).
Table S13. Sequencing statistics of the negative controls for Scladina I-4A (see table S12).
Table S14. Summary of HST mtDNA sequencing.
Table S15. Summary of Scladina I-4A mtDNA sequencing.
Table S16. Coverage statistics for all sequences from HST within the alignability track, map35_L100.
Table S17. Coverage statistics for HST sequences with a C-to-T substitution within the three first or last positions of either ends.
Table S18. Coverage statistics for all sequences from Scladina I-4A within the alignability track, map35_L100.
Table S19. Coverage statistics for Scladina I-4A sequences with a C-to-T substitution within the three first or last positions of either ends.
Table S20. Present-day human DNA contamination estimates from HST mtDNA.
Table S21. Present-day human DNA contamination estimates from Scladina I-4A mtDNA based on differences between Neandertals and modern humans.
Table S22. Present-day human DNA contamination estimates from Scladina I-4A mtDNA based on admixture rate as a function of the admixture rate.
Table S23. Present-day human DNA contamination estimates on mtDNA in the blank libraries of HST based on differences between HST and modern humans.
Table S24. Present-day human DNA contamination estimates on mtDNA in the blank libraries of Scladina I-4A based on differences between Neandertals and modern humans.
Table S25. Best substitution models according to the three model selection measures computed by jModelTest 2.1.10.
Table S26. Marginal likelihoods of the different tested clock and tree models obtained from a path sampling approach using only the coding region of the mitochondrial sequences.
Table S27. Marginal likelihoods of the different tested clock and tree models obtained from a path sampling approach using the full mitochondrial genome sequences.
Table S28. Estimates of molecular age and divergence times.
Table S29. Present-day human DNA contamination estimates for HST nuclear DNA based on deamination rates on the last positions of the molecules.
Table S30. Present-day human DNA contamination estimates for Scladina I-4A nuclear DNA based on deamination rates on the last positions of the molecules.
Table S31. Relationship between sequence length and present-day human DNA contamination estimate based on deamination rates in HST nuclear DNA sequences.
Table S32. Present-day human DNA contamination estimates based on the sharing of derived alleles with a modern human.
Table S33. Genome-wide counts of the three possible allelic configurations informative about the underlying topologies relating Vindija 33.19 and the Altai Neandertal to HST and Scladina I-4A before correcting for reference bias or contamination (see tables S40 and S41 for corrected results and fig. S17 for a description of these allelic configurations).
Table S34. Comparison of alignments to hg19 and panTro4.
Table S35. Excess of ancestral alleles in Late Neandertals compared to Vindija 33.19 at sites that are derived in the Altai Neandertal genome but ancestral in the genomes of an Mbuti and a Denisovan.
Table S36. Effect of the modified alignment procedure on the allele sharing with the Altai Neandertal.
Table S37. Alleles seen in Vindija 87 at positions that are heterozygous in Vindija 33.19.
Table S38. Sequencing and alignment errors of Vindija 87 sequences at positions where Vindija 33.19 is homozygous different from the Altai Neandertal, comparing the original alignments to hg19 with our modified alignment procedure.
Table S39. Summary of the alignments to the two references.
Table S40. Applying different sequence length cutoffs does not affect the allele sharing with the Altai Neandertal after realignments.
Table S41. Genome-wide counts of the three possible allelic configurations informative about the underlying topologies relating Vindija 33.19 and the Altai Neandertal to HST and Scladina I-4A after correcting for reference bias (see table S33 to compare with uncorrected results and table S42 for results corrected for contamination).
Table S42. Counts of the three possible allelic configurations informative about the underlying topologies relating Vindija 33.19 and the Altai Neandertal to HST and Scladina I-4A after correcting for both reference bias and contamination.
Table S43. Summary statistics about the physical distance between the positions used to infer the genetic relationship of HST to Vindija 33.19 and the Altai Neandertal.
Table S44. Summary statistics about the physical distance between the positions used to infer the genetic relationship of Scladina I-4A to Vindija 33.19 and the Altai Neandertal.
Table S45. Effective number of independent positions.
Table S46. Comparison between split time estimates from the Vindija population based on a constant divergence model and the F(A|B) statistic for five low-coverage Neandertal genomes.
Table S47. Split time estimates from the Vindija population based on a coalescent divergence model.
Table S48. Age estimate for individual B (branch shortening) used to convert the F(A|B) values shown in table S47 into time before present.
Table S49. Summary of the number of sites and blocks used to compute the F(A|B) statistic and CIs.
Table S50. Split time estimates between HST or Scladina I-4A and different populations (population B) based on the calibration of the F(A|B) statistic.
Table S51. Predictions of the mitochondrial TMRCA given different split times between the populations of HST and Vindija 33.19.
Table S52. Predictions of the mitochondrial TMRCA given different split times between the Vindija 33.19 population and a hypothetical isolated Neandertal population.
Table S53. Predictions of the mitochondrial TMRCA as done for table S51 but using either the upper or the lower estimates of the Neandertal population size.
Fig. S1. Length distribution of unique DNA fragments aligned to the human reference genome hg19 with a mapping quality of 25 or above (average length = 33 bp for HST and 25 bp for Scladina I-4A) and mapping uniquely (alignability track, map35_L100).
Fig. S2. Proportion of spurious alignment for different sequence lengths in the three libraries of HST that represent ~80% of the generated sequences for this specimen.
Fig. S3. Proportion of spurious alignment in the libraries of Scladina I-4A (same as for HST in fig. S2).
Fig. S4. Bivariate plot of root length against labio-lingual crown diameter (in millimeter) for the permanent mandibular canine.
Fig. S5. Bivariate plot of root length against labio-lingual crown diameter (in millimeter) for the permanent maxillary central incisor.
Fig. S6. Bivariate plot of root pulp volume against total root volume (in cubic millimeter) for the permanent maxillary central incisor.
Fig. S7. Ratio of sequences aligning to the X chromosome and autosomes.
Fig. S8. Number of sequences mapping to each chromosome normalized by chromosome length.
Fig. S9. Deamination patterns from the mtDNA.
Fig. S10. Maximum parsimony tree built with MEGA6 (Molecular Evolutionary Genetics Analysis, program version 6).
Fig. S11. Phylogenetic relationship of currently available archaic human mitochondrial genomes reconstructed from a Bayesian analysis with BEAST 2 (Bayesian Evolutionary Analysis Sampling Trees, program version 2).
Fig. S12. C-to-T substitution frequencies at the end of nuclear DNA sequences (dashed lines), including frequencies conditioned on a C-to-T substitution at the other end (solid lines).
Fig. S13. Proportion of alleles that are derived in the Altai Neandertal but ancestral in the Vindija 33.19 Neandertal and Denisovan genomes stratified by the allele frequency in the Luhya and Yoruba populations (AFR) of the 1000 genomes dataset.
Fig. S14. Deamination frequencies on sequences from HST that carry a modern human allele absent from the currently available Neandertal genomes.
Fig. S15. Deamination frequencies on sequences from Scladina I-4A that carry a modern human allele absent from the currently available Neandertal genomes.
Fig. S16. Lineage assignment before correcting for the reference bias.
Fig. S17. Expectations for the genetic relationship of HST and Scladina I-4A to Vindija 33.19 and the Altai Neandertal.
Fig. S18. Lineage assignment after correcting for the reference bias.
Fig. S19. Comparison of the expected and observed mitochondrial TMRCA of HST with other European Neandertals.
Fig. S20. Probability that all sampled Neandertal mtDNAs come from an early modern human population as a function of the admixture rate.
Fig. S21. Probability that all sampled Neandertal mtDNAs come from an early modern human population as a function of the admixture rate.
References (28–114)

REFERENCES AND NOTES
1. M. Meyer, J.-L. Arsuaga, C. de Filippo, S. Nagel, A. Aximu-Petri, B. Nickel, I. Martínez, A. Gracia, J. M. B. de Castro, E. Carboneil, B. Viola, J. Kelso, K. Prüfer, S. Pääbo, Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. Nature 531, 504–507 (2016).
2. J. L. Arsuaga, I. Martínez, J. L. Arnold, A. Aranburu, A. Gracia-Téllez, W. D. Sharp, R. M. Quam, C. Falguères, A. Pantoja-Pérez, J. Bischoff, E. Pozza-Rey, J. M. Parés, J. M. Carretero, M. Demuro, C. Lorenzo, N. Sala, M. Martín-Torres, N. García, A. Alcalz de Velasco, G. Cuerva-Bescós, A. G. Méndez-Ortega, R. García, R. A. Bonnietti, J. M. Bermúdez de Castro, E. Carboneil, Neandertal roots: Cranial and chronological evidence from Sima de los Huesos. Science 344, 1358–1363 (2014).
3. T. Hingham, D. Douka, R. Wood, C. B. Ramsey, F. Brock, L. Baseli, M. Camps, A. Arizabala, A. J. P. Álvarez, C. Baena, C. Barroso-Rizo, C. Bergman, C. Boitard, P. Boscardin, M. Caparros, N. J. Conard, C. Drai, A. Froment, B. Galván, P. Gambassini, A. Garcia-Moreno, S. Grimagi, P. Haesaerts, B. Holt, M.-J. Istrate-Chiapusso, A. Jenik, J. F. Jordà-Pardo, L.-M. Maillo-Fernández, A. Marro, J. Maroto, M. Menéndez, L. Metz, E. Morin, A. Moroni, F. Negrino, E. Panagopoulou, M. Peresani, S. Pioner, M. de la Rasilla, J. Riel-Salvatore, A. Ronchitelli, A. Salvadori, R. Salas, A. Villalonga, A. P. Pinhasi, R. Jacobí,
11. S. Guši, N. Nagel, S. Skoglund, M. Hsieh, S. an Meyer, Single-stranded DNA library preparation from the archaic Denisovan individual. V. Peyrégne, D. E. Racimo, N. a79 (2017).

12. K. Prüfer, K. Munch, I. Hellmann, K. Akagi, J. R. Miller, B. Walzen, S. Koren, G. Sutton, C. Koudelka, K. W. Winer, J. R. Knight, C. J. Mulkinn, S. J. Meader, C. P. Ponting, G. Lunter, S. Higashino, A. Hobolt, J. Dutheil, E. Karaok, C. Alkan, S. Sajjadian, C. R. Catachizia, M. Ventura, T. Marques-Bonet, E. E. Eicher, C. Andre, R. Atenica, L. Mugisha, J. Zunkel, P. 2019; 5 : eaaw5873 26 June 2019

13. S. Nagel, B. Douka, T. Higham, M. B. Kozlkin, V. M. Shunov, A. P. Derveiriako, J. Kels, M. Skatlin, S. Paabo, Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468, 1035–1060 (2010).

14. S. R. Browning, L. Browning, Y. Zhou, S. Tucci, J. M. Akey, Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. Cell 173, 53–61 (2018).

15. M. Henxian, Q. Fu, A. Huber, M. Pet, F. Mafessoni, S. Grote, P. Skolgund, V. Narasimham, R. Hugouer, J. Crevecoeur, P. Semal, M. Seroset, T. Sammo, J-J. Rublin, I. Gilis, Z. Kuca, P. Rudan, L. V. Golovanova, V. B. Dornorichv, C. Posth, J. Krause, P. Korlevi, S. Nagel, B. Nickel, M. Skatlin, N. Patterson, D. Reich, K. Prüfer, M. Meyer, S. Paabo, J. Kelso, Reconstructing the genetic history of late Neanderthals. Nature 555, 652–656 (2018).

16. V. Sion, F. Mafessoni, B. Verrot, C. de Filippo, S. Grote, B. Viola, M. Hajdinka, P. Peyrigne, S. Nagel, S. Brown, K. Douka, T. Higham, M. B. Kozlkin, V. M. Shunov, A. P. Derveiriako, J. Kels, M. Meyer, K. Prüfer, S. Paabo, The genome of the offspring of a Neanderthal mother and a Denisovan father. Nature 561, 113–116 (2018).

17. K. Prüfer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinez, G. Renaud, P. H. Sudmant, C. de Filippo, H. Li, S. Mallic, M. Danneizman, Q. Fu, M. Kircher, M. Kuhlwil, M. Lachmann, M. Meyer, M. Ongheret, M. Siebaur, C. Theunert, A. Tandon, P. Moranour, J. Pickrell, J. C. Millikin, S. V. Hohr, R. E. Green, I. Hellmann, P. L. F. Johnson, H. Blanche, H. Cian, J. O. Kitzman, J. Shendure, E. E. Eicher, E. S. Leen, T. E. Bakken, L. V. Golovanova, V. B. Dornorich, M. V. Shunov, A. P. Derveiriako, B. Viola, M. Skatlin, D. Reich, J. Kelso, S. Paabo, The complete genome sequence of a Neanderthal from the Altai Mountains. Nature 505, 43–49 (2014).

18. M. Muntener, J. Wahl, Das Felsennfragment eines Neandertalers aus der Stadelhoehle des Hohlensteins im Lonetal. Fundberichte aus Baden-Wuerttemberg 17, 111–124 (1992).

19. C. Posth, W. Cilling, K. Kitagawa, L. Pagani, L. van Holstein, F. Racimo, K. Wehrberger, N. J. Cordon, C. J. Kind, H. Bocherens, J. Krause, Depth divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neandertals. Nat. Commun. 8, 16046 (2017).

20. M. Toussaint, M. Otte, D. Bonjean, H. Bocherens, C. Falguieres, Y. Yokokaya, Les restes humains neandertaliens immediatiss de la couche 4A de la grotte Saldina (Andenne, Belgique). Comptes Rendus de l’Academie des Sciences–Series IIA-Earth and Planetary Science 326, 737–742 (1998).

21. L. Orlando, P. Darlu, M. Toussaint, D. Bonjean, M. Otte, C. Hanny, Revisiting Neandertal diversity with a 100,000 year old mtDNA sequence. Curr. Biol. 16, R400–R402 (2006).

22. M. Toussaint, D. Bonjean, The Saldina-I44 Juvenile Neandertal, Andenne, Belgium: Palaeoanthropology and Context (ERAUL Editions, 2014).

23. J. Deyney, M. Knapp, I. Glodp, M. T. Gansauge, A. Wehmann, B. Nickel, C. Valdiosera, N. Garcia, S. Paabo, J. L. Arsuaga, M. Meyer, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. U. S. A. 110, 15758–15763 (2013).

24. M.-T. Gansauge, T. Gerber, I. Glode, P. Kerlevi, L. Lippik, S. Nagel, L. M. Riel, A. Schmidt, M. Meyer, Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. Nucleic Acids Res. 45, e79 (2017).

25. M.-T. Gansauge, M. Meyer, Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protoc. 8, 737–748 (2013).

26. P. Kerlevi, T. Gerber, M.-T. Gansauge, M. Hajdinka, S. Nagel, A. Axiuni-Petri, M. Meyer, Reducing microbial and human contamination in DNA extraction from ancient bones and teeth. Biotechniques 59, 89–93 (2015).

27. A. W. Briggs, U. Stenzel, P. L. F. Johnson, R. E. Green, J. Kelso, K. Prüfer, M. Meyer, J. Krause, T. M. Ronan, M. Lachmann, S. Paabo, Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. U. S. A. 104, 14616–14621 (2007).

28. M. Meyer, M. Kircher, M.-T. Gansauge, H. Li, F. Racimo, M. Salkic, J. G. Schrader, F. Jay, K. Prüfer, C. de Filippo, P. H. Sudmant, C. Alkan, Q. Fu, R. Do, N. Rohland, A. Tandon, M. Siebaur, E. Green, K. Brey, A. W. Briggs, U. Stenzel, J. Dabney, J. Shendure, J. Kitzman, M. F. Hammer, M. V. Shunov, A. P. Derveiriako, N. Patterson, A. M. Andrs, E. E. Eicher, M. Skatlin, D. Reich, J. Kelso, S. Paabo, A high-coverage genome sequence from an archaic Denisovan individual. Science 338, 222–226 (2012).

29. K. Prüfer, C. de Filippo, S. Grote, F. Mafessoni, P. Kerlevi, M. Hajdinka, B. Verrot, L. Skov, P. Hsieh, S. Peyrigné, D. Reher, C. Hopfe, S. Nagel, T. Maricic, Q. Fu, C. Theunert, R. Rogers, P. Skolgund, M. Chinalatilapa, M. Dannemann, B. J. Nelson, F. M. Key, P. Rudan, Z. Kučan, I. Guliš, L. V. Golovanova, V. B. Dornorichv, N. Patterson, D. Reich, E. Eicher, M. Skatlin, M. H. Schierup, A. M. Andrs, J. Kelso, M. Meyer, S. Paabo, A high-coverage Neandertal genome sequence from the Zhokhov Cave. Nature 437, 655–658 (2017).

30. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437, 69–87 (2005).
Peyrégne et al., Sci. Adv. 2019; 5 : eaaw5873  26 June 2019
Nuclear DNA from two early Neandertals reveals 80,000 years of genetic continuity in Europe

Stéphane Peyrégné, Viviane Slon, Fabrizio Mafessoni, Cesare de Filippo, Mateja Hajdinjak, Sarah Nagel, Birgit Nickel, Elena Essel, Adeline Le Cabc, Kurt Wehrberger, Nicholas J. Conard, Claus Joachim Kind, Cosimo Posth, Johannes Krause, Grégory Abrams, Dominique Bonjean, Kévin Di Modica, Michel Toussaint, Janet Kelso, Matthias Meyer, Svante Pääbo and Kay Prüfer

Sci Adv 5 (6), eaaw5873.
DOI: 10.1126/sciadv.aaw5873