

# Nuclear and mitochondrial DNA sequences from two Denisovan individuals

Susanna Sawyer<sup>a,1</sup>, Gabriel Renaud<sup>a,1</sup>, Bence Viola<sup>b,c,d</sup>, Jean-Jacques Hublin<sup>c</sup>, Marie-Theres Gansauge<sup>a</sup>, Michael V. Shunkov<sup>d,e</sup>, Anatoly P. Derevianko<sup>d,f</sup>, Kay Prüfer<sup>a</sup>, Janet Kelso<sup>a</sup>, and Svante Pääbo<sup>a,2</sup>

<sup>a</sup>Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany; <sup>b</sup>Department of Anthropology, University of Toronto, Toronto, ON M5S 2S2, Canada; <sup>c</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany; <sup>d</sup>Institute of Archaeology and Ethnography, Russian Academy of Sciences, Novosibirsk, RU-630090, Russia; <sup>e</sup>Novosibirsk National Research State University, Novosibirsk, RU-630090, Russia; and <sup>f</sup>Altai State University, Barnaul, RU-656049, Russia

Contributed by Svante Pääbo, October 13, 2015 (sent for review April 16, 2015; reviewed by Hendrik N. Poinar, Fred H. Smith, and Chris B. Stringer)

Denisovans, a sister group of Neandertals, have been described on the basis of a nuclear genome sequence from a finger phalanx (*Denisova 3*) found in Denisova Cave in the Altai Mountains. The only other Denisovan specimen described to date is a molar (*Denisova 4*) found at the same site. This tooth carries a mtDNA sequence similar to that of *Denisova 3*. Here we present nuclear DNA sequences from *Denisova 4* and a morphological description, as well as mitochondrial and nuclear DNA sequence data, from another molar (*Denisova 8*) found in Denisova Cave in 2010. This new molar is similar to *Denisova 4* in being very large and lacking traits typical of Neandertals and modern humans. Nuclear DNA sequences from the two molars form a clade with *Denisova 3*. The mtDNA of *Denisova 8* is more diverged and has accumulated fewer substitutions than the mtDNAs of the other two specimens, suggesting Denisovans were present in the region over an extended period. The nuclear DNA sequence diversity among the three Denisovans is comparable to that among six Neandertals, but lower than that among present-day humans.

Denisovans | ancient DNA | Neandertals

In 2008, a finger phalanx from a child (*Denisova 3*) was found in Denisova Cave in the Altai Mountains in southern Siberia. The mitochondrial genome shared a common ancestor with present-day human and Neandertal mtDNAs about 1 million years ago (1), or about twice as long ago as the shared ancestor of present-day human and Neandertal mtDNAs. However, the nuclear genome revealed that this individual belonged to a sister group of Neandertals. This group was named Denisovans after the site where the bone was discovered (2, 3). Analysis of the Denisovan genome showed that Denisovans have contributed on the order of 5% of the DNA to the genomes of present-day people in Oceania (2–4), and about 0.2% to the genomes of Native Americans and mainland Asians (5).

In 2010, continued archaeological work in Denisova Cave resulted in the discovery of a toe phalanx (*Denisova 5*), identified on the basis of its genome sequence as Neandertal. The genome sequence allowed detailed analyses of the relationship of Denisovans and Neandertals to each other and to present-day humans. Although divergence times in terms of calendar years are unsure because of uncertainty about the human mutation rate (6), the bone showed that Denisovan and Neandertal populations split from each other on the order of four times further back in time than the deepest divergence among present-day human populations occurred; the ancestors of the two archaic groups split from the ancestors of present-day humans on the order of six times as long ago as present-day populations (5). In addition, a minimum of 0.5% of the genome of the *Denisova 3* individual was derived from a Neandertal population more closely related to the Neandertal from Denisova Cave than to Neandertals from more western locations (5).

Although Denisovan remains have, to date, only been recognized in Denisova Cave, the fact that Denisovans contributed

DNA to the ancestors of present-day populations across Asia and Oceania suggests that in addition to the Altai Mountains, they may have lived in other parts of Asia. In addition to the finger phalanx, a molar (*Denisova 4*) was found in the cave in 2000. Although less than 0.2% of the DNA in the tooth derives from a hominin source, the mtDNA was sequenced and differed from the finger phalanx mtDNA at only two positions, suggesting it too may be from a Denisovan (2, 3). This molar has several primitive morphological traits different from both late Neandertals and modern humans. In 2010, another molar (*Denisova 8*) was found in Denisova Cave. Here we describe the morphology and mtDNA of *Denisova 8* and present nuclear DNA sequences from both molars.

## Results

**Denisova 8.** The *Denisova 8* molar (Fig. 1) was found at the interface between layers 11.4 and 12 in the East gallery of Denisova Cave, slightly below the Neandertal toe phalanx (*Denisova 5*, layer 11.4) and the Denisovan finger (*Denisova 3*, layer 11.2). Radiocarbon dates for layer 11.2, as well as for the underlying 11.3 layer, yield ages more than ~50,000 y (OxA-V-2359-16 and OxA-V-2359-14) (2). *Denisova 8* is thus older than *Denisova 3*, which is at least 50,000 y

## Significance

Denisovans are a sister group of Neandertals that were identified on the basis of a nuclear genome sequence from a bone from Denisova Cave (Siberia). The only other Denisovan specimen described to date is a molar from the same site. We present here nuclear DNA sequences from this molar and a morphological description, as well as mitochondrial and nuclear DNA sequences from another molar from Denisova Cave, thus extending the number of Denisovan individuals known to three. The nuclear DNA sequence diversity among the Denisovans is higher than among Neandertals, but lower than among present-day humans. The mtDNA of one molar has accumulated fewer substitutions than the mtDNAs of the other two specimens, suggesting Denisovans were present in the region over several millennia.

Author contributions: S.S. and S.P. designed research; S.S., G.R., B.V., and M.-T.G. performed research; M.V.S. and A.P.D. contributed new reagents/analytic tools; S.S., G.R., B.V., J.-J.H., K.P., and J.K. analyzed data; and S.S., G.R., B.V., J.K., and S.P. wrote the paper.

Reviewers: H.N.P., McMaster University; F.H.S., Illinois State University; and C.B.S., The Natural History Museum, London.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the European Nucleotide Archive database (accession no. PRJEB10828). The mitochondrial assembly of *Denisova 8* has been deposited in the GenBank database (accession no. KT780370).

<sup>1</sup>S.S. and G.R. contributed equally to this work.

<sup>2</sup>To whom correspondence should be addressed. Email: paabo@eva.mpg.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519905112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519905112/-DCSupplemental).

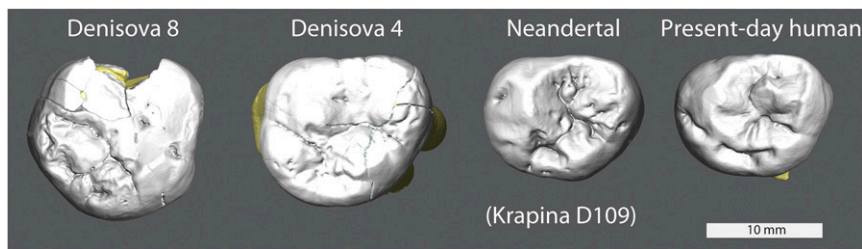


Fig. 1. Occlusal surfaces of the *Denisova 4* and *Denisova 8* molars and third molars of a Neandertal and a present-day European.

old. It is reassembled from four fragments that fit well together, although a piece of enamel and most of the root is missing (*SI Appendix, Fig. S1B*).

The *Denisova 4* molar was found in layer 11.1 in the South gallery, a different part of the cave. Radiocarbon dates for layer 11.2 of the South gallery are more than 50,000 y (OxA-V-2359-17 and OxA-V-2359-18) and  $48.6 \pm 2.3$  thousand years before present (KIA 25285) (2). Although the lack of direct stratigraphic connection between the different parts of the cave makes relative ages difficult to assess, it is likely that *Denisova 4* is younger than *Denisova 8*.

On the basis of crown shape and the presence of a marked crista obliqua, a feature unique to maxillary molars, we identify *Denisova 8* as an upper molar, despite it having five major cusps. The mesial half of the crown is worn, with a small dentine exposure on the protocone, whereas there is no wear on the distal part. The lack of a distal interproximal facet indicates that the tooth is a third molar, or a second molar without the eruption of the  $M^3$ . Usually, when Neandertal and *Homo heidelbergensis* upper  $M^2$ s reach wear levels to the extent seen here, the adjacent  $M^3$  is already erupted and an interproximal facet is visible. One possibility is that the *Denisova 8* is a second molar of an individual with  $M^3$  agenesis. Despite being common in modern humans, this is rare in archaic hominins, but it does occur in Asian late *Homo erectus* and Middle Pleistocene hominins. We analyze *Denisova 8* as an  $M^3$  in the following comparisons, but see *SI Appendix* for discussion of alternative possibilities.

The previously described *Denisova 4* molar is characterized by its large size, flaring buccal and lingual sides, strong distal tapering, and massive and strongly diverging roots (2). Not all of these characteristics can be assessed in *Denisova 8*, but it is clear that it lacks the strong flare of the lingual and buccal surfaces and distal tapering of *Denisova 4*.

The length of *Denisova 8* is more than three SDs larger than the means of Neandertal and modern human molars, and in the range of Pliocene hominins (Fig. 1 and *SI Appendix, Fig. S1A*). Both *Denisova 8* and *Denisova 4* are very large compared with Neandertal and early modern human molars, and *Denisova 8* is even larger than *Denisova 4*. Only two Late Pleistocene third molars are comparable in size: those of the inferred early Upper Paleolithic modern human *Oase 2* in Romania and those of *Obi-Rakhmat 1* in Uzbekistan (7, 8).

The morphology of third molars is variable, and thus not very diagnostic. Nevertheless, Neandertal third molars differ from *Denisova 8*, in that they frequently show a reduction or absence of the hypocone, reduction of the metacone, and generally lack a continuous *Crista obliqua* (8, 9). This applies also to Middle Pleistocene European hominins, who also only rarely show a cusp 5 (9). The massive and diverging roots of *Denisova 4* are very unlike the root morphology of Neandertals and Middle Pleistocene hominins in Europe. East Asian *H. erectus* and Middle Pleistocene *Homo* frequently show massive roots similar to *Denisova 4*, but in these groups, crown size become strongly reduced starting around 1 million years ago (10). The recently described Xujiayao teeth from China (11) have massively flaring roots and relatively large

and complex crowns, similar to the *Denisova* teeth, but they also have reduced hypocones and metacones.

Early and recent modern humans show the most morphological variability of third molars, and there are specimens that have large hypocones, metacones, or continuous *cristae obliquae* (9). The combination of an unreduced metacone and hypocone, continuous crista obliqua, a large fifth cusp, and large overall size is reminiscent of earlier *Homo*, but *Denisova 8* lacks the multiple distal accessory cusps frequently seen in early *Homo* and Australopithecines.

**DNA Isolation and Sequencing.** DNA was extracted from 36 mg dentine from *Denisova 8* in our clean room facility (12), and DNA libraries from this specimen, as well as from a previously prepared extract of *Denisova 4*, were prepared as described (3, 13) (*SI Appendix, Table S2*). From both teeth, random DNA fragments were sequenced and mapped to the human reference genome (hg19). In addition, mtDNA fragments were isolated from the libraries (14) and sequenced.

Of the DNA fragments sequenced from *Denisova 4* and *Denisova 8*, 0.05% and 0.9%, respectively, could be confidently mapped to the human genome sequence, yielding 54.6 and 265 million base pairs (Mb) of nuclear DNA sequences for *Denisova 4* and *Denisova 8*, respectively (see Table 1 for overview). MtDNA sequences from the two specimens were aligned to the mtDNA of *Denisova3* (NC\_013993.1). For *Denisova 4*, the average mtDNA coverage is 72.1-fold. The lowest support for the majority base at any position is 89% (*SI Appendix, Fig. S4*), and the consensus sequence is identical to the previously published mtDNA sequence from this specimen (2). For *Denisova 8*, the mtDNA coverage is 118.9-fold, and the lowest support for the majority base is 86% (*SI Appendix, Fig. S4*).

**DNA Sequence Authenticity.** We used three approaches to estimate present-day human DNA contamination in the two libraries. First, for each library, we used all unique DNA fragments that aligned to the present-day human reference mtDNA (15) and counted as contaminating those that carried a nucleotide different from the majority mtDNA sequence determined from the molar at positions where the endogenous majority consensus differed from all of 311 present-day human mtDNAs. The mtDNA contamination thus estimated was 5.2% [95% confidence interval (CI), 4.5–6.0%] for *Denisova 4* and 3.2% (95% CI, 2.9–3.6%) for *Denisova 8*.

Table 1. Overview of DNA sequences produced, contamination estimates, and amount of nuclear sequences used for analyses

Data amount and quality	<i>Denisova 4</i>	<i>Denisova 8</i>
Amount of mapped sequences	54.6 Mb	265 Mb
mtDNA coverage	72-fold	119-fold
Autosomal contamination	~66%	~15%
mtDNA contamination	~5.2%	~3.2%
X chromosome contamination	~28%	~9%
Nuclear sequences used	1 Mb	24 Mb

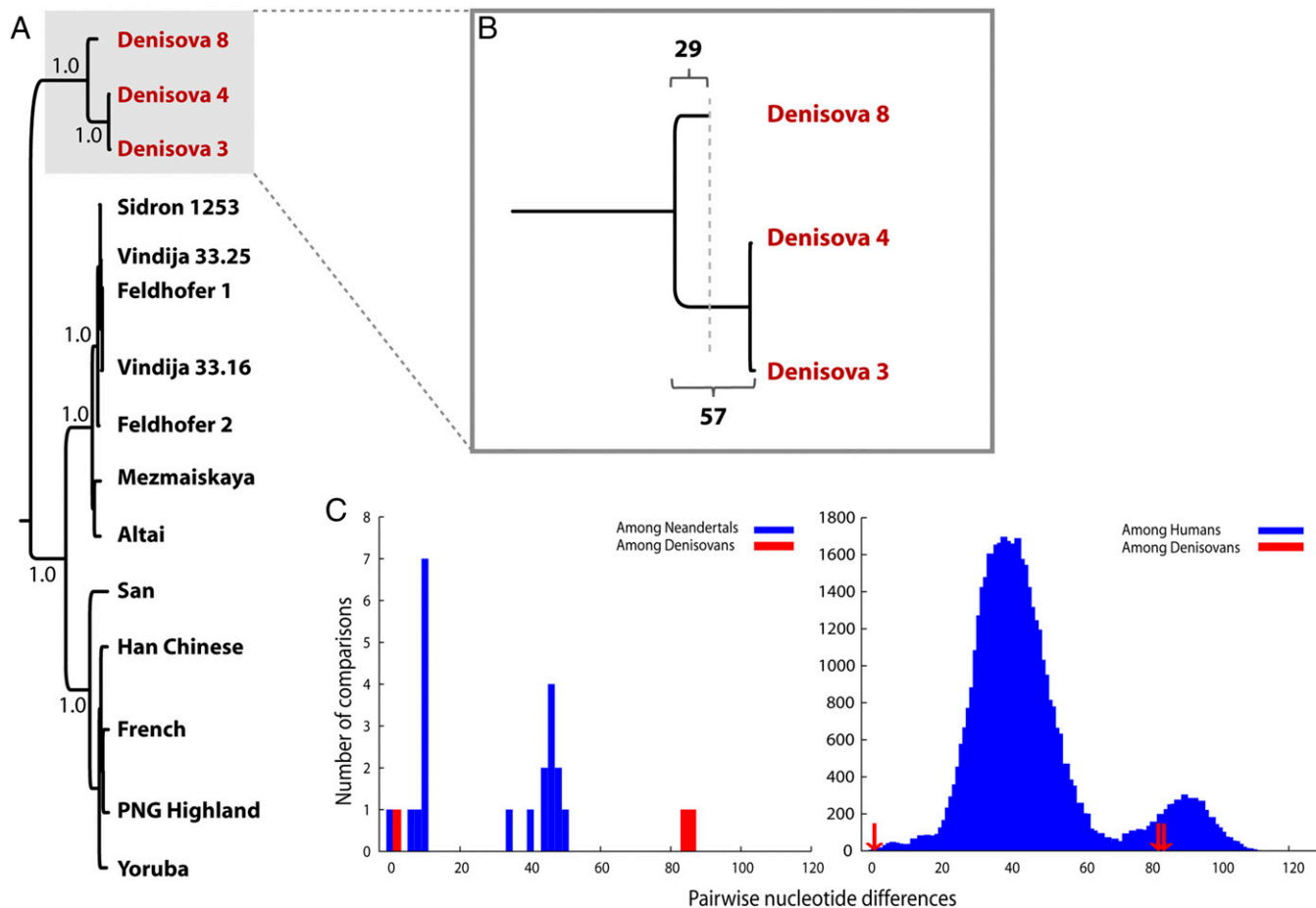
Second, we estimate contamination by present-day nuclear DNA by estimating DNA sequence divergence (as described below) of the two molars from present-day humans. We assume that the divergence of two present-day European individuals from each other represents 100% contamination, whereas the divergence of the high-quality genome determined from *Denisova 3* from present-day humans represents 0% contamination. By this approach, we estimate the DNA contamination of *Denisova 4* as 65.2–67.0%, and *Denisova 8* as 14.6–15.4% (*SI Appendix, Table S4*). That the nuclear DNA contamination is high, particularly of *Denisova 4*, is compatible with an estimate based on cytosine deamination patterns at the 3'- and 5'- ends of the aligned sequences (*SI Appendix*).

In the third approach, we first determined the sex of the individuals from which the molars derive by counting the number of DNA fragments that map to the X chromosome and autosomes, respectively. To limit the influence of present-day DNA contamination in this part of the analysis, we restricted our analysis to DNA fragments that, at their 5'- and/or 3'-ends, carry thymines (T) at positions where the human reference nuclear genome carries cytosines (C). Such apparent C to T substitutions are frequently caused by deamination of cytosine to uracil toward the ends of ancient DNA fragments (16, 17). We find that both teeth come from males ( $P \sim 0.4$ ), rather than females ( $P < 0.01$ ) (*SI Appendix, Table S6*). We then estimated the amount of female DNA contamination among the aligned sequences

as the fraction of DNA fragments that match the X chromosome in excess of what is expected for a male bone. This yields a female DNA contamination rate of 28.4% (95% CI, 27.3–29.5%) for *Denisova 4* and 8.6% (95% CI, 8.3–8.9%) for *Denisova 8*.

The estimates based on mtDNA and nuclear DNA differ drastically (Table 1), presumably because the ratios of mitochondrial to nuclear DNA differ between the endogenous and the contaminating source or sources of DNA, whereas the two estimates based on nuclear DNA suggest that more males than females are among the contaminating individuals. It is clear that although these methods yield different contamination estimates, they all suggest that the nuclear DNA contamination in both libraries is substantial, particularly in *Denisova 4*, where it is likely to exceed 50%. To reduce the influence of DNA contamination (18, 19), we therefore restrict the analyses of nuclear DNA to fragments that carry thymine residues at the first and/or last two positions at sites where the human reference sequence carries cytosine residues (but remove these C/T sites themselves in the analyses). Using these criteria, a total of 1.0 Mb of nuclear DNA sequences for *Denisova 4* and 24.1 Mb for *Denisova 8* (Table 1 and *SI Appendix, Table S3*) can be analyzed.

**mtDNA Relationships.** A phylogenetic tree relating the mtDNAs from *Denisova 3*, *Denisova 4*, and *Denisova 8*; seven Neandertals from Spain, Croatia, Germany, the Russian Caucasus, and the Altai Mountains (5, 20); and five present-day humans (Fig. 2 *A* and *B*)



**Fig. 2.** Evolutionary relationships of Denisovan mtDNAs. (*A*) Bayesian tree relating the mtDNAs of three Denisovans, seven Neandertals, and five present-day humans. Posterior probabilities are indicated. A chimpanzee mtDNA was used to root the tree. (*B*) Numbers of differences between the two molar mtDNAs and the inferred common mtDNA ancestor of the three Denisovan mtDNAs. (*C*) Pairwise nucleotide differences among the Denisovans and Neandertals (*Left*) and among the Denisovans and 311 present-day human mtDNAs (*Right*).



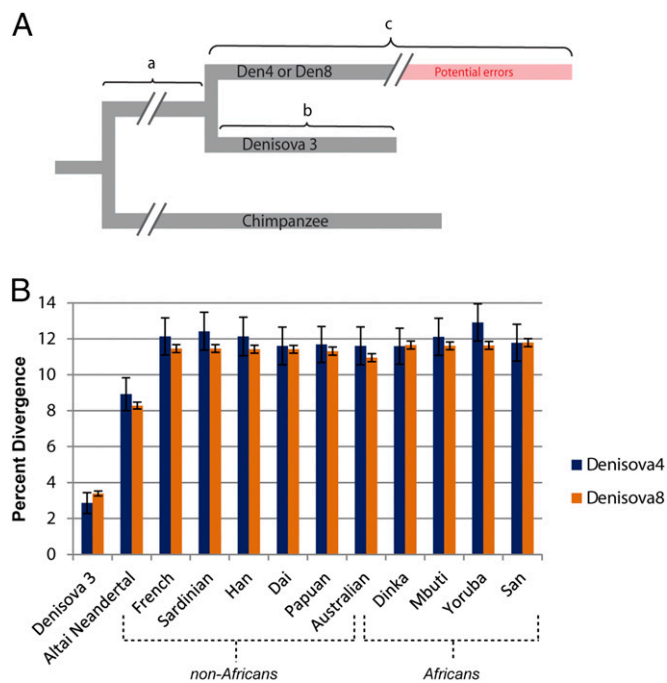
shows that the mtDNAs of the two Denisovan molars form a clade with *Denisova 3* to the exclusion of the Neandertals. The largest number of differences seen among the three Denisovan mtDNAs is 86, whereas the largest number of differences seen among seven Neandertal mtDNAs is 51, and among 311 present-day humans, it is 118 (Fig. 2C). When comparing Watterson's estimator  $\theta_w$ , which to some extent takes the numbers of samples into account, among the populations, the mtDNA diversity of the three Denisovans is  $3.5 \times 10^{-3}$ , that of Neandertals is  $1.8 \times 10^{-3}$ , that of present-day Europeans is  $4.0 \times 10^{-3}$ , and that of present-day humans worldwide is  $16.1 \times 10^{-3}$ . Thus, mtDNA diversity among late Neandertals seems to be low relative to Denisovans, as well as present-day humans.

The number of nucleotide changes inferred to have occurred from the most recent common ancestor of the three Denisovan mtDNAs to the *Denisova 4* molar, the *Denisova 3* phalanx, and the *Denisova 8* molar are 55, 57, and 29 respectively (Fig. 2B and *SI Appendix*, Table S7). The corresponding number of substitutions from the most recent common ancestor of the seven Neandertal mtDNAs to each of the Neandertal mtDNAs varies between 17 and 25 (*SI Appendix*, Table S7). This suggests that the time back to the mtDNA of the most recent common ancestor from the *Denisova 3* and the *Denisova 4* mtDNAs was almost twice as long as that from the *Denisova 8* mtDNA.

**Autosomal Analyses.** To estimate the divergence of the low-coverage DNA sequences retrieved from *Denisova 4* and *Denisova 8* to the high-quality genomes of *Denisova 3* (3), as well as to the Neandertal from Denisova Cave and to 10 present-day humans (5), we first counted nucleotide substitutions inferred to have occurred on the lineages from the human–chimpanzee ancestor to each of the high-coverage genomes (Fig. 3A, *a* and *b*). We then used the low-coverage molar sequences to estimate the fraction of those substitutions that occurred after their divergence from the high-coverage lineages; that is, the fraction of such substitutions not seen in the molars (Fig. 3A, *b*). To the Denisovan high-coverage genome, these fractions are 2.9% (95% CI, 2.28–3.44%) and 3.4% (95% CI, 3.25–3.53%) for *Denisova 4* and *Denisova 8*, respectively. Divergences of *Denisova 4* and *Denisova 8* are 8.9% (95% CI, 8.01–9.83%) and 8.3% (CI, 8.01–8.48%) to the high-coverage Neandertal genome and 10.9–12.9% to 10 present-day humans (Fig. 3B and *SI Appendix*, Tables S7 and S8). These results show that the two teeth come from Denisovans and confirm that Denisovans were a sister group of Neandertals.

The average pairwise divergence among six low-coverage Neandertals to the Altai Neandertal genome is 2.5% (range, 2.5–2.6%) (*SI Appendix*, Table S11). This is slightly lower than the divergence of 2.9% and 3.4% of the two Denisovan molars from the Denisova genome and shows that the individuals from whom the two molars derive are almost as closely related to the *Denisova 3* genome as are the Neandertals to the Altai Neandertal genome. In comparison, the range of divergences among 10 present-day human genomes is 4.2–9.5%, among the four Europeans 6.0–6.4%, and between the two individuals from the South American tribal group Karitiana 4.2%. Thus, nuclear DNA diversity appears low among the archaic individuals, especially the Neandertals.

Using the high-coverage *Denisova 3* genome, it was shown that Denisovans have contributed DNA to present-day people in Oceania (2–5). As expected, we found that *Denisova 8* also shares more derived alleles with Papuans and Australians than with other non-Africans ( $D$ :  $-0.04$  to  $-0.07$ ;  $[Z] = 1.8$ – $3.0$ , excluding CpG sites; *SI Appendix*, Table S13). However, when we subsample, from the high-coverage Denisovan genome, the DNA segments covered by fragments sequenced from *Denisova 4*, we find that there are not enough data to similarly detect gene flow from *Denisova 4* to Oceanians (*SI Appendix*, Table S14). This precludes us from asking whether either *Denisova 4* or *Denisova 8* is more closely related to the introgressing Denisovan than *Denisova 3*. Similarly, there are not enough data to determine whether gene



**Fig. 3.** Nuclear DNA divergence between *Denisova 4* and *Denisova 8* and the Denisovan genome. (A) DNA sequences from *Denisova 4* and *8* were each compared with the genomes of *Denisova 3* (3) and the inferred human–chimpanzee ancestor (25, 26). The differences from the human–chimpanzee ancestor common to the two Denisovans (*a*) as well as differences unique to each Denisovan are shown (*b* and *c*). Errors in the low-coverage Denisova genomes result in artificially long branches (*c*). Divergences of the molar genomes to *Denisova 3* are therefore calculated as the percentage of all differences between *Denisova 3* and the human–chimpanzee ancestor that are not shared with the molar genomes,  $b/(a + b) \times 100$ . (B) Autosomal divergences of *Denisova 4* and *Denisova 8* to the *Denisova 3* genome, the Neandertal genome, and 10 present-day human genomes calculated as in A. All estimates are based on DNA fragments from the two molars that carry putative deamination-induced C to T substitutions. Bars indicate 95% confidence intervals.

flow from Neandertals at the level detected in the high-coverage *Denisova 3* genome (5) is present in *Denisova 4* and *8* (*SI Appendix*, Table S14).

## Discussion

The nuclear DNA sequences retrieved from *Denisova 4* and *Denisova 8* are more closely related to the *Denisova 3* genome used to define the Denisovans as a hominin group than to present-day human or Neandertal genomes. Furthermore, the mtDNAs of the two molars form a clade with *Denisova 3*. Thus, the present work extends the number of Denisovan individuals identified by mitochondrial and nuclear DNA from one to three. Although the number of Denisovan individuals is small and restricted to one locality, and they differ in age, it is nevertheless interesting to note that the nuclear DNA sequence diversity among the three Denisovans is slightly higher than that found among seven Neandertals, although these are widely geographically distributed, but lower than that seen among present-day humans worldwide or in Europe.

Although the three Denisovans come from a single cave, they may differ significantly in age, as indicated by the branch length of the mtDNA of the *Denisova 8* molar, which is shorter than those of *Denisova 4* and the *Denisova 3*, an observation that is congruent with the stratigraphy. If we assume that the mtDNA mutation rate of  $\sim 2.5 \times 10^{-8}$ /site/year (95% CI, 1.8–3.2) that is estimated for modern humans (21) applies also to Denisovan mtDNA, *Denisova 8* is on the order of 60,000 years older than *Denisova 3* and *Denisova 4*. A similar or even larger age difference between

