Supporting Information Appendix
Nuclear and mitochondrial DNA sequences from two Denisovan individuals

Section 1: Morphology of Denisova 8.

Denisova 8 is a fragmentary of molar found in August 2010 in Denisova cave, Square G4, subsquares V or G at the limit of layer 11.4 and Layer 12. It has been reassembled from four fragments, which fit together well, but slight cracks remain between the fragments. On the mesial surface an about 7x4.3x5 mm (bl x md x height) chip of enamel is missing, and the root is broken off just below the cervix.

We identify Denisova 8 as a left upper molar based on the presence of a marked *Crista obliqua* connecting the protocone and metacone both on the enamel surface and on the enamel dentine junction (Suppl. Figure S1A and B). The mesial half of the crown is relatively worn, with most of the relief removed (wear stage 3 of (1)). The paracone retains a small salient buccally, and a small dentine exposure is visible on the protocone. There is no wear on the distal part of the crown.

The lack of wear in the distal half and the lack of a distal interproximal facet leads us to identify Denisova 8 as an M^3. An alternative explanation would be that it is an M^2 from an individual where the M^3 did not erupt yet, but this is in our opinion less likely. Neandertal molars with comparable wear from Krapina (e.g. D165 and 177, both slightly less worn than Denisova 8) show distal interproximal facets indicating that the M^3 was already erupted. Other examples are St. Césaire 1 and Shanidar 2 that show wear comparable to Denisova 8 and erupted M^3s. The emergence of the M^3 happens in general early in Neandertals (2-4), as is well visible in La Chaise BD8 (5) and Le Moustier 1 (6), that both show only very light wear on the M^2s, while the M^3s are in the course of eruption. This early M^3 emergence seems to have been present in *Homo heidelbergensis* as well (7). In the Atapuerca Sima de los Huesos assemblage there are several specimens (AT-46, AT-4326, AT-815) with less wear than Denisova 8 that show clear distal interproximal facets. Identifying Denisova 8 as an upper M^2 of a young individual would thus necessitate wear rates well in excess of those seen in Neandertals and *Homo heidelbergensis*.

A last possibility is that Denisova 8 is the M^2 of an individual with M^3 agenesis. This occurs frequently in recent modern humans, but is rare in pre-mid-Upper Palaeolithic modern

humans and archaic humans. An early *Homo* mandible, Omo 175-14a&b exhibits unilateral M_3 agenesis (8), and it also occurs in *Homo erectus*, such as Dmanisi D2735 (unilateral M_3 agenesis (9)) and Lantian (bilateral M_3 agenesis (10)). Yunxian EV9002 from the Late Middle Pleistocene of Hubei province (China) shows a strongly reduced, peg-like M^3 on the left and a small M^3 on the right that did not erupt despite the fully adult age of the individual (11), while the Penghu 1 mandible from the Taiwan strait exhibits agenesis of the right M_3. Finally, the holotype of *Homo floresiensis*, LB-1 probably had a strongly reduced left M^3, while the right M^3 was congenitally absent. No cases of M^3 agenesis have been described in Neandertals and *Homo heidelbergensis*. In summary, the relatively unusual morphology of Denisova 8, with several accessory cusps and the absence of a distal interproximal facet makes its identification as an M^3 more likely in our view, but we acknowledge the alternative possibility that it is an M^2 with a still unerupted or absent M^3.

The crown is a rounded pentagon in shape, with five major cusps. The largest cusp is the protocone, followed by the metacone, paracone, hypocone (ASUDAS grade 4,(12)) and cusp 5. The lingual surface of the protocone shows no evidence of a Carabelli's cusp. The enamel on the mesial surface and in the area of the mesial marginal ridge is damaged, but at the enamel-dental junction (EDJ), a marked mesial marginal ridge is visible, which was likely also apparent on the enamel surface. A protoconule (accessory cusp on lingual part of the mesial marginal ridge) was probably also present, as there is a slight sulcus just distal of the enamel break line that probably delimited an accessory tubercle, and on the enamel dentine junction a small cuspule is apparent in the mesiolingual corner of the crown. The protocone and metacone are connected by a wide *Crista obliqua*, which on the EDJ is uninterrupted and was likely continuous on the enamel surface, though somewhat obscured by the wear.

A large cusp 5 (ASUDAS grade 5), comparable in size to the hypocone, is situated at the distal end of the crown, connected by a marked distal marginal ridge to the mesiolingual aspect of the hypocone. A relatively large accessory cusp is visible on the distal marginal ridge near the fissure separating the hypocone and cusp 5, delimited bilaterally by marked grooves descending onto the distal surface. The buccal and lingual sides are relatively vertical, while the distobuccal aspect is somewhat bulging.

The previously described Denisovan molar, *Denisova 4*, a right $M^{2/3}$ is characterized by its large size, flaring buccal and lingual sides, strong distal tapering and massive and strongly diverging roots (13). Due to preservation, 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 seen in Denisova 4. The crown of *Denisova 8* also seems lower and

with straighter sides, although this has probably been exaggerated by the stronger wear. *Denisova 8* is somewhat larger than *Denisova 4*, with a mesiodistal length of 14.3 mm and a buccolingual breadth of 14.65 mm. Both teeth from Denisova are much larger than most Neandertal and Upper Paleolithic upper M^2s and M^3s, the length of *Denisova 8* is more than three standard deviations above the Upper Paleolithic modern human and Neandertal means, and in the range of Pliocene hominins (Suppl. Figure S1c,d; Suppl. Table 1). Based on its large size it is likely that the M^3 was the largest of the molars.

Two Late Pleistocene specimens are comparably large in size, the M^3s of the early Upper Paleolithic modern human Oase 2 and the $M^{2/3}$ of Obi-Rakhmat 1 (14, 15). Oase 2 does not show large extra cusps, but instead strong crenulation (16). Obi-Rakhmat shows a large extra cusp, but mesially, not distally (Main text, Figure 1), and a large number of accessory cusps possibly due to gemination (17).

Comparative morphology of Denisova 8 (identified as an M^3)

M^3s are in general very variable, and thus morphologically not very diagnostic. Neandertal M^3s 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* (15, 18). Similarly, in the M^3s of Sima de los Huesos and other Middle Pleistocene European samples we also see a reduction of the hypocone and metacone and lack of a *Crista obliqua*, as well as no expression of a cusp 5 (18).

In South-East Asian *Homo erectus*, M^3s are also in general reduced, with small hypocones and metacones, and frequently interrupted *Cristae obliquae* ((19), own observations). Despite the crown reduction, these specimens frequently have massive and flaring roots (20, 21), similar to those seen in *Denisova 4*.

Early modern humans and recent modern humans show the most morphological variability in the M^3, and here we can find some specimens that show large hypocones, metacones or continuous *Cristae obliquae* (18).

The combination of unreduced metacone and hypocone, continuous *Crista obliqua*, a large cusp 5, and a very large size is something that is not present in any of these samples, and more reminiscent of earlier *Homo*, but Denisova lacks the multiple distal accessory cusps frequently seen in early *Homo* and Australopithecines.

Comparative morphology of Denisova 8 (identified as an M^2)

Neandertal and European Middle Pleistocene M^2s are usually rhomboid or nearly rectangular, with medium sized metacones. The hypocones are usually smaller than in M^1s, while in the Atapuerca SH sample they are frequently reduced (18). The majority of both Neandertal and Middle Pleistocene European M^2s show continuous *cristae obliquae*, and small and medium sized Cusp 5s are frequent (18, 22). Denisova 8 is differentiated from these groups by the presence of a large metacone, hypocone and very large Cusp 5 extending the crown distally.

In East Asian *Homo erectus* and Middle Pleistocene *Homo* the M^2 is frequently trapezoid, with the crown tapering distally, and the hypocone is somewhat reduced when compared with early *Homo* and African *Homo erectus* (19, 21, 23). Similar distal tapering is also present in some African Middle Pleistocene specimens, for example Kabwe 1. The *crista obliqua* in East Asian *Homo erectus* (especially Zhoukoudian) is frequently non-continuous, just like in Dmanisi and earlier hominins, and the Cusp 5, if present, is usually small (19, 21, 24). Denisova 8 lacks distal tapering and the large Cusp 5 is a feature not found in *Homo erectus* or Asian Middle Pleistocene populations.

In recent humans the hypocone is frequently reduced on the M^2, but this reduction is less frequent in fossil modern humans (25). Similarly, a continuous oblique crista is usually absent in recent humans, but often present in Upper Palaeolithic and Early anatomically modern humans (18). Strong expressions of Cusp 5 are rare in all modern humans,

Denisova 4, as described previously ((13), SI. p. 183) is quite similar in its distal tapering to the morphology seen in the M^2s of some Middle Pleistocene *Homo*, but is differentiated from them by its lingually skewed hypocone and metacone, and the large talon basin (features which are more similar to the Neandertal condition), as well as its massive and flaring roots. Denisova 8 on the other hand shows few similarities to any group, mostly due to its unusually large cusp five. Until the recovery of more complete Denisovan material, their morphological affinities remain unclear.

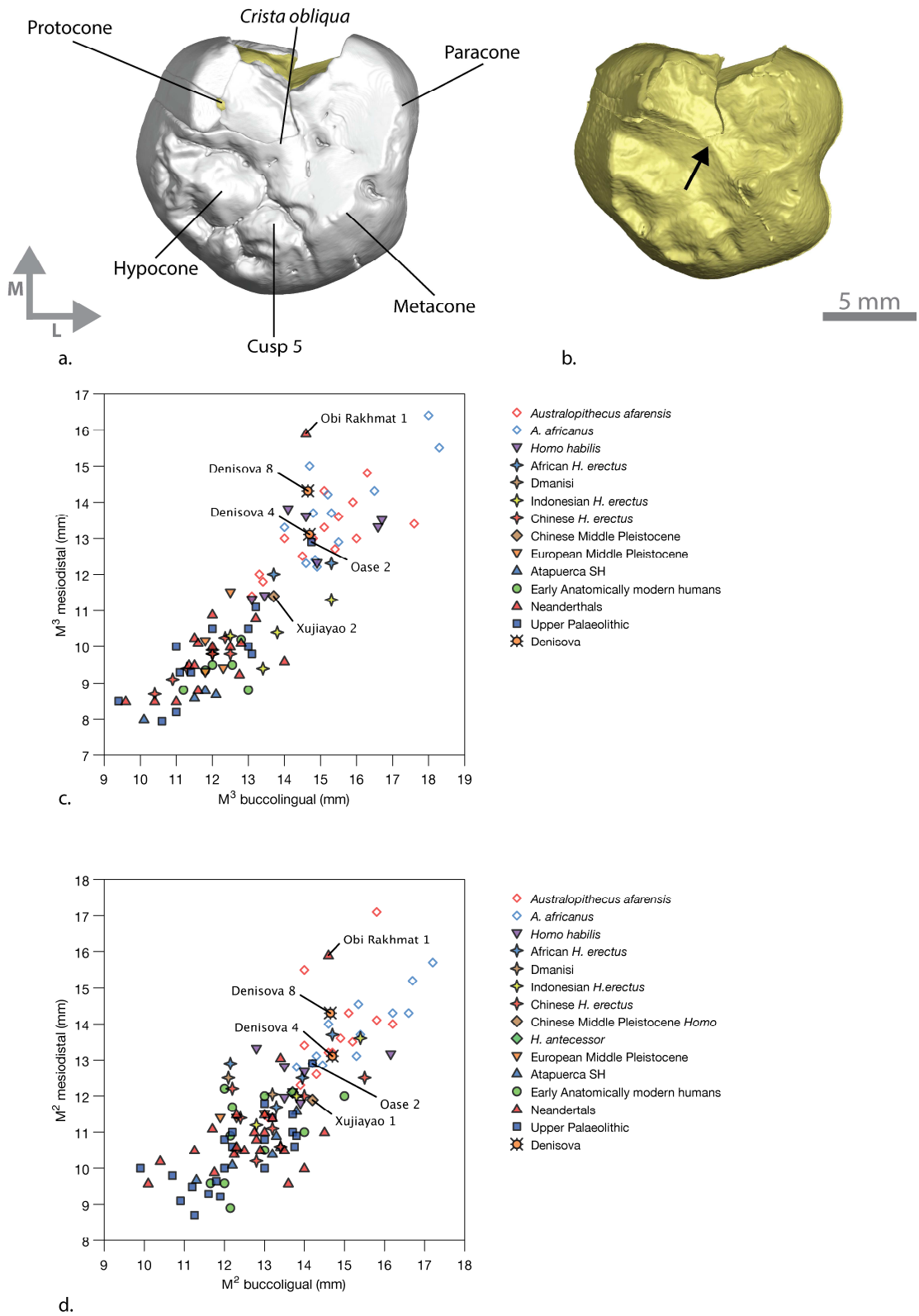

Figure S1. Morphology of *Denisova 8* molar. . **a:** occlusal view (surface model from µCT scan); **b:** enamel dentine junction in occlusal view, The arrow indicates the marked *Crista obliqua* on the enamel-dentine junction; **c**: Biplot of the mesiodistal (md) and labiolingual (bl) diameters of *Denisova 8* and other hominin M^3s. For comparative sample used and sources for data see

Supplementary Table 1. **d:** Biplot of the mesiodistal (md) and labiolingual (bl) diameters of *Denisova 8* and other hominin M^2s. For comparative sample used and sources for data see Supplementary Table 1.

Table S1. Metric comparisons of M^2 and M^3 length and breadth in various fossil hominins and the Denisova remains.

	M^2 md[1]	M^2 bl[2]	M^3 md	M^3 bl
A. afarensis	13.7±1.4 (13)[3]	14.7±0.9 (13)	13.1±1 (14)	15±1.3 (14)
A. africanus	13.9±1 (12)	15.3±1.1 (12)	13.8±1.3 (12)	15.6±1.4 (12)
Homo habilis	12.6±0.6 (6)	14±1.1 (6)	12.7±1.1 (7)	14.8±1.4 (7)
Dmanisi	12.3 (12.05-12.5; 2)[4]	12.7 (12.1-13.2; 2)	9.8 (1)	12 (1)
***H. erectus* (Africa)**	12.7 (11.7-13.7; 4)	13.5 (12.15-14.7; 4)	12.2 (12-12.3; 2)	14.5 (13.7-15.3; 2)
***H. erectus* (Indonesia)**	12.3 (11.2-13.6; 3)	14 (12.8-15.4; 3)	10.4 (9.4-11.3; 4)	13.8 (12.5-15.3; 4)
***H. erectus* (China)**	11.3±0.9 (8)	13.2±1.1 (8)	9.6±0.5 (7)	11.6±0.8 (7)
Atapuerca SH	10.6±0.7 (6)	12.9±0.9 (6)	8.5±0.4 (4)	11.4±0.9 (4)
***H. heidelbergensis* (Europe)**	11.6 (11.4-12.1; 4)	12.7 (11.9-13.7; 4)	10.1 (9.3-11.5; 4)	12.1 (11.8-12.5; 4)
Neandertals	11±1.4 (21)	12.7±1.2 (21)	10.1±1.8 (17)	12±1.3 (17)
Neandertals (w/o Obi-Rakhmat)	10.7±0.8 (20)	12.6±1.1 (20)	9.8±1 (16)	11.8±1.1 (16)
Early AMH	10.8±1.2 (10)	12.7±1.1 (10)	9.4±0.5 (6)	12.2±0.7 (6)
Upper Palaeolithic	10.4±1 (21)	12.3±1.2 (21)	9.8±1.4 (12)	12±1.5 (12)
Denisova 4	13.1	14.7	13.1	14.7
Denisova 8	-	-	14.3	14.65

1. Mesiodistal length measured following the definition of (26)
2. Buccoligual breadth measured following the definition of (26)
3. Mean+-standard deviation (N)
4. Mean (range; N)

Sources of metric data:

A. afarensis: Hadar, Omo (own measurements)
A. africanus: Stekfontein, Makapansgat (27)
Homo habilis: Olduvai (28), East Turkana (27)
Dmanisi (24)
H. erectus (Africa): East Turkana (27), Nariokotome (29), Konso (30), Swartkrans (27)
H. erectus (China): Zhoukoudian (21), Hexian (31)
H. erectus (Indonesia): Trinil (27), Sangiran (own measurements, (19))
Atapuerca SH (18)
H. heidelbergensis (Europe): La Chaise (5), Biache (32), Arago (33), Petralona (5)
Neandertals: Amud (34), Châteauneuf (35), St. Brelade (26), Krapina (2), La Croze de Dua (26), La Quina (26), Le Moustier (26), Obi-Rakhmat (own measurements), Saccopastore (26), Shanidar (36), Spy (26), Tabun (26), Vergisson la Falaise (26)
Early AMH: Skhul (37), Qafzeh (38), Temara (39)
Upper Paleolithic: Brno (26), Changwu (31), Cro-Magnon (26), Dolni Vestonice (40), Grotte des Enfants (26), Kostenki (own measurements), La Rochette (26), Leuca (26), Mladec (26), Oase (16), Predmosti (26), Sungir (own measurements)

Section 2: DNA Extraction, library preparation and sequencing.

Thirty six millligrams (mg) of dentin were removed from the inside of the enamel cusp of *Denisova 8* using a dentistry drill and used to produce 100 microliters (μL) of extract as described (41). From 1/20th of this extract, as well as from 1/10th of a previous 100μL extract made from 40mg of *Denisova 4* (13), we produced Illumina libraries, using a single-stranded library preparation protocol that maximizes the yield of sequences from ancient DNA (42). The libraries were treated with *E. coli* Uracil DNA Glycosylase (UDG) and endonuclease VIII to remove uracils (U) (43). UDG does not effectively excise terminal Us (42). The *Denisova 4* library (L9234, see Suppl. Table S2) had a final volume of 40μL in EBT (10mM Tris-HCl, pH 8.0; 0.05% Tween-20), while *Denisova 8* (B1113) had a final volume of 20μL in EBT.

The concentrations of L9234 and B1113 were measured by qPCR. L9234 from *Denisova 4* was split into two equal parts and used as template for an indexing PCR using two distinct indexing primers per library. The indexing PCR was performed using AccuPrime Pfx DNA polymerase (Life Technologies) and purified with the MinElute purification system as described (42). The purified and indexed libraries were each eluted in 30μL of EB (Qiagin MinElute Kit) to produce L9243 and L9250. An indexing PCR was also performed on B1113 from *Denisova 8* as described above except that all of B1113 was used in one indexing reaction to produce L9108.

To produce larger amounts of amplified library for the mtDNA enrichment, 5μL of L9243 from *Denisova 4* and of L9108 from *Denisova 8* were further amplified with Herculase II Fusion using adapter primers IS5 and IS6 (42, 44), purified with MinElute and eluted into 20μL of EB. DNA concentration was measured on a Nanodrop (ND-1000) and 500ng of the amplified DNA were enriched for human mtDNA via a bead-based protocol where PCR products are sheared, ligated to biotinylated linkers and immobilized on streptavidin-coated beads (45). The enriched libraries were quantified by qPCR and amplified with Herculase II Fusion, taking care not to reach PCR plateau. After measuring DNA concentration on a Bioanalyzer 2100 (Agilent) the *Denisova 4* capture product (L9320) was sequenced on 1/7th of an Illumina MiSeq lane and the *Denisova 8* capture product (L9126) on 1/10th of an Illumina GAII lane.

For shotgun sequencing, the two libraries from *Denisova 4*, L9243 and L9250 (see Table S2), were amplified with Herculase II Fusion. Molecules with insert sizes between 35 and 450 bp were isolated using gel electrophoresis as described to produce L9349 and L9350 (42). L9108 from *Denisova 8* was also size fractionated to isolate molecules of lengths between 40 and 200 bp using gel electrophoresis without prior amplification to produce L9133. This library was amplified and quantified on the Bioanalyzer 2100 (Agilent) along with L9349 and L9350. The two *Denisova 4* libraries (L9349 and L9350) were pooled in equimolar amounts and sequenced on two Illumina HiSeq 2500 High Output flowcells, while the *Denisova 8* library (L9133) was sequenced on one High Output flowcell.

Table S2. Extraction and library IDs. IDs of *Denisova 4* and *8* after each processing step are given. The *Denisova 4* single-stranded (ss) library was split into two aliquots for the indexing amplification.

	Extract ID	(ss)Lib ID	Lib ID after Indexing	Lib ID after mtDNA capture	Lib ID after gel excision for shotgun seq
Denisova 4	E324	L9234	L9243	L9320	L9349
			L9250	-	L9350
Denisova 8	E652	B1113	L9108	L9126	L9133

Section 3: Sequence processing and mapping

Ibis v1.1.6 (46) was used for base calling and sequence processing was carried out as described (47). Briefly, after base-calling, reads were demultiplexed allowing a single mismatch in the indexes; Illumina adapters were identified and removed, and overlapping read-pairs merged when the overlap was at least 11 bp. For all sequences, the following basic filters were applied:

- Sequences with more than 5 bases with base qualities less than 15 (phred score) were removed
- Sequences having a base with a quality less than 10 (phred score) in the index reads were removed
- Sequences shorter than 35 bp were removed
- PCR duplicates were identified based on the same beginning and end coordinates and collapsed

MtDNA sequences were aligned to the mitochondrial sequence of the high coverage *Denisova 3* phalanx (NC_013993.1) using MIA (parameters: -c, -i) ((48), https://github.com/udo-stenzel/mapping-iterative-assembler) which was also used to generate what approximates a 75% consensus sequence.

The shotgun-sequenced fragments were aligned to hg19 (49) using BWA v.0.5.10 (50) with a maximum edit distance (-n option) of 0.01, a maximum of 2 gap openings (-o 2), and without a seed (-l 16500).

Table S3. DNA sequences yields.

	Mg of bone powder for extract[a]	% of extract used for library	% endogenous[b]	Mb aligned to human genome[c]	Mb aligned after duplicate removal	% unique[d]	Mb aligned after deamination filter[e]
Denisova 4	40	20%	0.05%	80.7 Mb	54.6 Mb	67.6%	1.0
Denisova 8	36	10%	0.9%	1,128 Mb	265 Mb	23.5%	24.1

a. Milligrams of bone powder used to make 100uL of extract
b. Percent endogenous is calculated as the Mb aligned to the human genome (after filtering for mapped sequences with a length above 35) divided by the total Mb sequenced (after filtering for a length above 35) times 100.
c. Mb aligned to hg19 after passing the following filters: length > 35, map quality > 37, merging of paired reads with minimum 11 bp overlap, fewer than 5 bases with base quality below 15, index reads with base qualities above 10.
d. Percent unique is Mb aligned with filters to the human genome after duplicate removal divided by aligned Mb before duplicate removal times 100
e. For deamination filter, see the supplemental text.

Section 4: Ancient DNA Authenticity

We used four methods to estimate present-day human contamination in *Denisova 4* and *8*.

(i) MtDNA contamination. We identified 183 and 174 "diagnostic positions" in *Denisova 4* and *Denisova 8, respectively*, where their consensus mtDNA sequences as estimated by MIA (see Section 3) differ from every individual in a panel of 311 present-day humans from around the world.

We then re-aligned all captured sequences from the two molars to the human mtDNA reference sequence (51) using BWA version 0.5.10 (50) with relaxed parameters (-n 0.01, -o 2, -l 16500). This allows modern human mtDNA fragments that differ from the Denisovan mtDNA to be identified. Fragments carrying present-day human variants at the diagnostic sites were counted as contaminants, while fragments carrying consensus variants were counted as endogenous. 95% confidence intervals were calculated using a Wilson score interval. We estimated the mtDNA contamination of *Denisova 4* to 5.2% (95% CI: 4.5-6.0%) of *Denisova 8* to 3.2% (95% CI: 2.9-3.6%).

The shotgun sequences were aligned to the human mtDNA reference sequence as described above, and, using the same diagnostic positions as above, mtDNA contamination estimated for the shotgun data. The shotgun data gave an mtDNA contamination estimate of 4.9% (95% CI: 4.2-5.8%) for *Denisova 4* and 4.0% (95% CI: 3.5-4.6%) for *Denisova 8*.

(ii) Nuclear DNA contamination. To estimate present-day human contamination in the nuclear sequence data, we calculated the divergences of two French individuals to each other as well as two Sardinian individuals to each other (see Figure 3A for explanation of divergence calculation) and used these divergences as a hypothetical contamination of 100% (*c,* Suppl. Figure S2). Similarly, we used the divergence of the *Denisova 3* phalanx sequences to the four Europeans as a proxy for 0% contamination (*a*, Suppl. Figure S2). We then calculated the divergence of *Denisova 4* and *Denisova 8* to the French and Sardinians using sequences that had not been filtered for a terminal C to T change (*b*, Suppl. Figure S2). The percent contamination in the *Denisova 4* and *Denisova 8* sequences were then calculated as (a-b/a-c)x100. For *Denisova 4* this results in a contamination estimate of 65.2 to 67.0% and for *Denisova 8* 14.6 to 15.4% (Suppl. Table S4).

Table S4. Nuclear contamination estimate. An estimate of the nuclear contamination using the method described in Figure S2 applied to fragments without filtering for deamination.

European used to calc div[a]	% Divergence European[b]	% Divergence *Denisova 3*[c]	% Divergence *Denisova 4*	% Divergence *Denisova 8*	Div Den3 – Div human[d]	Div Den3 – Div *Den4*	Div Den3 – Div *Den8*	% contamination Den4[e]	% contamination Den8
French1	6.36 *(to Fr2)*	11.85	8.22	11.02	5.49	3.63	0.83	**66.1**	**15.1**
French2	6.09 *(to Fr1)*	11.62	7.98	10.81	5.53	3.64	0.81	**65.8**	**14.6**
Sardinian1	6.34 *(to Sa2)*	11.86	8.26	11.05	5.52	3.6	0.81	**65.2**	**14.7**
Sardinian2	6.06 *(to Sa1)*	11.64	7.9	10.78	5.58	3.74	0.86	**67.0**	**15.4**

a. The European present-day humans to whom divergence is calculated and whose mutations are used to calculate divergence
b. Divergence calculation using pairs of Europeans. Thus: French2 to French 1, and vice versa, as well as Sardinian2 to Sardinian1 and vice versa. As an example, French2 to French1 uses the mutations on the branch to French1 to calculate the divergence and gives a result of 6.36%.
c. Divergence of Denisova 3 to each of the European present-day humans listed.
d. Differences in divergence, calculated e.g. divergence of Den3 to French1 minus the divergence of French2 to French1 (in this case *a* – *c* in Figure S2).
e. Percent contamination, calculated e.g. (divergence of Den3 to Fr1 – divergence of Den8 to Fr1) / (divergence of Den3 to Fr1 – divergence of Fr2 to Fr1)*100. In this case this would be (a-b)/(a-c)*100 in Figure S2.

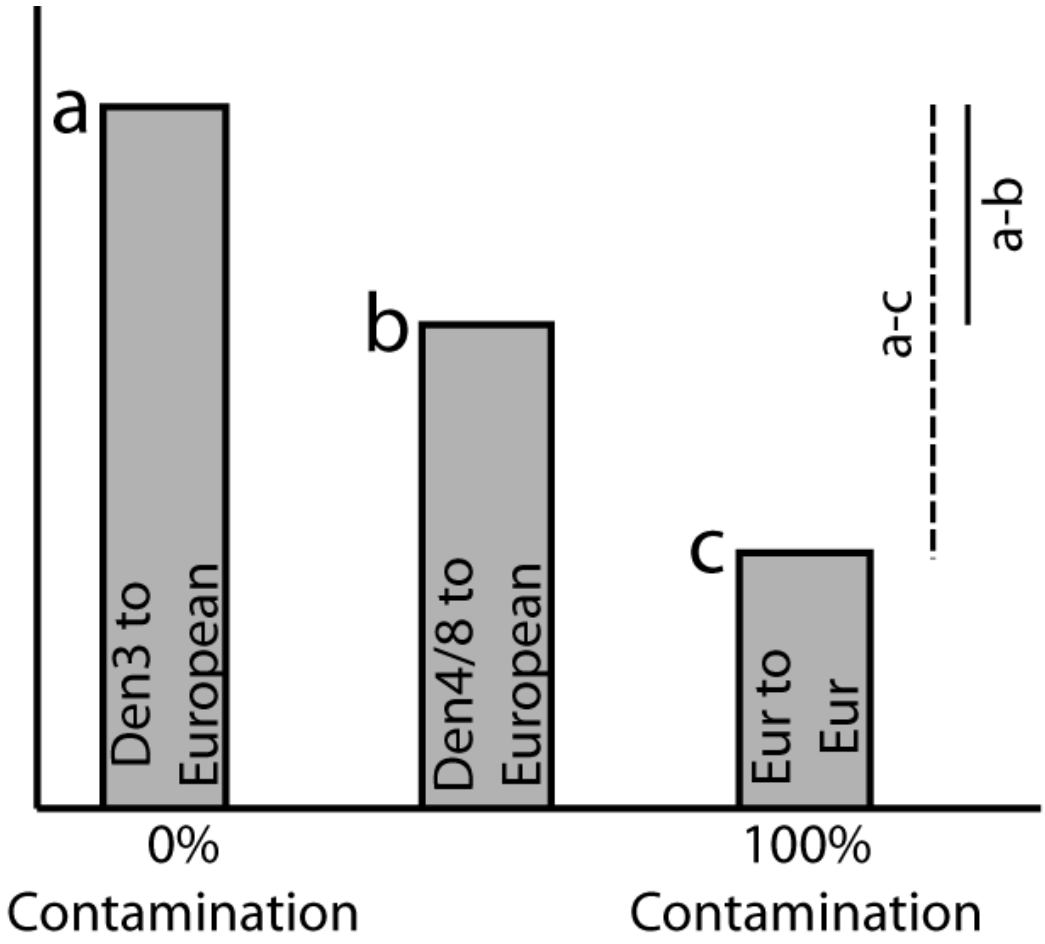

Figure S2. Divergence-based contamination estimates. The divergence of the *Denisovan 3* to two French and two Sardinians (left bar, *a*) is assumed to represent 0 % present-day human contamination. The divergence of French-French and Sardinian-Sardinian (right bar, *c*) is assumed to represent 100 % contamination. The divergence of *Denisova 4* or *8* to the French and Sardinians (middle bar, *b*) is then gauged as the reduction in divergence to the present-day humans as a fraction of the divergence among the present-day humans ((*a*-*b*) / (*a*-*c*)).

(iii) C to T substitutions: To determine whether different populations of molecules that differ in their extent of cytosine deamination-induced C to T substitutions occur in the libraries, we calculated the apparent C to T substitution rate at the 5'- and 3'-ends of DNA fragments. We then calculated the 5' C to T rate of fragments that have a 3' C to T and *vice versa*. Since deamination-induced misincorporations are rare in modern DNA that contaminates ancient DNA preparations (52, 53), it is unlikely that such DNA fragments carry C to T changes on both ends. In contrast, DNA molecules that carry a C to T change at one end are likely to be ancient and the C to T rate at the other end of such molecules can thus be taken to approximate the deamination rate in ancient, endogenous molecules (under the assumption that deamination at the two ends of molecules is independent). By comparing the C to T rates of all sequences to those that carry C to T at one end we can thus gauge if two or more populations of molecules that differ in their rates of deamination occur in the libraries and thus if contamination may exist in a library. 95% CIs were calculated using Wilson score intervals. Although this approach may be affected by factors that we do not fully understand, it yields contamination estimates for *Denisova 4* of 54-69% and 1.3-6.1% for *Denisova 8* (Suppl. Table S5) which are qualitatively compatible with ones based on divergence above. For the mtDNA, the 95% CIs of the C to T rates of the two populations of molecules overlap (Suppl. Table S5).

Table S5. Terminal C to T substitutions nuclear and mtDNA fragments. C to T substitutions relative to the corresponding mtDNA consensus sequences are shown for mtDNA and nuclear DNA fragments sequenced from *Denisova 4* and *Denisova 8*, respectively. "3' filtered" and "5' filtered" refer to fragments that carry C to T substitutions at their 3'- and 5'-ends, respectively. The 95% CI is given in parenthesis.

		5 prime	**3 prime**
***Denisova 4* mtDNA**	***No filter***	11.3 *(9.7-13.0)*	22.4 *(20.9-24.1)*
	3' filtered	17 *(9.7-27.8)*	100
	5' filtered	100	30.5 *(22.2-40.4)*
***Denisova 4* nuclear**	***No filter***	7.2 *(6.9-7.4)*	14.6 *(14.3-14.8)*
	3' filtered	18.9 *(16.0-22.2)*	100
	5' filtered	100	35.7 *(32.6-39.1)*
***Denisova 8* mtDNA**	***No filter***	23.7 *(21.9-25.6)*	46.0 *(44.5-47.6)*
	3' filtered	20.8 *(16.2-26.3)*	100
	5' filtered	100	46.9 *(39.9-54.2)*
***Denisova 8* nuclear**	***No filter***	31.4 *(31.2-31.6)*	49.8 *(49.7-49.9)*
	3' filtered	32.5 *(32.0-33.2)*	100
	5' filtered	100	52.3 *(51.8-52.8)*

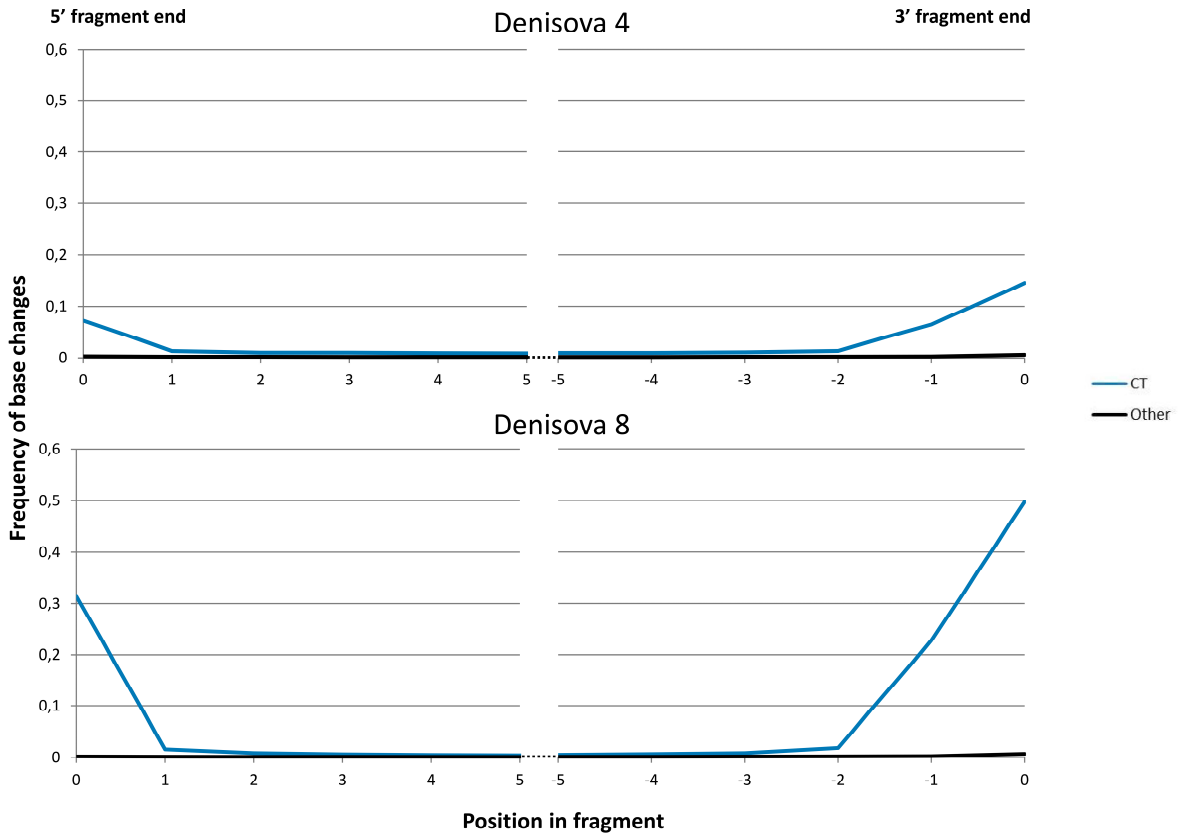

Figure S3. Nucleotide differences to the human reference genome as a function of distance from fragment ends. Differences are given as percent of a base in the reference genome that occurs as a different base in the sequenced DNA fragments. C to T differences are largely due to deamination of cytosine residues in ancient DNA fragments. Libraries were treated with *E.coli* uracil DNA glycosylase, which is not efficient at the first, the last and second to last bases.

(iv) Sexing and female DNA contamination: For sex determination, we used sequences that passed the filters described in Section 3 have a minimum map quality of 37 (phred scale).

We identified regions on the sex chromosomes that are ≥500 bps long and pass the mappability filter. The mappability filter removes positions where at least one overlapping window of 35bp length maps to a different position in the genome with up to one mismatch (54). On the Y-chromosome, we in addition excluded positions that overlap with sequences from four females from the 1000 Genomes Project (NA12878, NA12892, NA19240, NA19238) (54). This left us with 627,426 bp on the Y chromosome and 40,661,238 bp on the X chromosome.

The number of sequenced fragments expected to fall in these regions if the individuals were male is: (Number of fragments aligned to the whole genome) × (the number positions in the X or Y-chromosome) / (genome size), where genome size is: 2 × (autosomal positions) + (X-chromosomal positions) + (Y-chromosomal positions).

We then determined the number of fragments that actually fall within these regions using either (*i*) all fragments or (*ii*) only those that carry putative deamination-induced C to T substitutions. We determined if the observed and expected numbers are significantly different from the male expectation using a Chi-square test (chisq.test) in the R package 3.1.0 (55). For the X-chromosomal fragments carrying C to T substitutions, we also determined if there is a significant difference under the female expectation. Both *Denisova 4* and *8* are more likely to come from males than from females. See Suppl. Table S6.

Because the molars come from male individuals, we can estimate the fraction of fragments due to female contamination using the number of "extra" fragments mapped to the X-chromosome relative to the expected number if the individual is male and all Y-chromosome fragments are assumed to be endogenous. The contamination rate is then the difference between the number of fragments mapped to the X chromosome and the number expected if the individual is male divided by number expected if the individual is male. A Wilson score interval was used to calculate 95% CIs. We find that *Denisova 4* has a female contamination of 28.4% (95% CI: 27.3-29.5%) and *Denisova 8* 8.6% (95% CI: 8.3-8.9%) (Table S6).

Table S6. Sex determination and female contamination. The number of X- and Y-chromosomal sequences mapped and expected to map if the molars are from males. DNA sequences carrying terminal C to T substations as well as all sequences were analyzed.

		Y-chromosome			X-chromosome			
Denisova	**Analysis/ Sequences**	# of sequences mapped	# sequences expected to map if male	*χ^2-test p-value*	# of sequences mapped	# sequences expected to map if male	*χ^2-test p-value*	Percent female contamination
4	**Sex determination (Terminal C->T seqs)**	8	3	-	231	222	*0.42 (5.9e-14 if female)*	-
8		94	86	*0.26*	5,535	5,576	*0.43 (<2.2e-16 if female)*	-
4	**Contamination estimate (all seqs)**	75	93	*0.006*	7,764	6,048	*<2.2e-16*	28.4% *(27.3-29.5)*
8		617	599	*0.32*	42,175	38,829	*<2.2e-16*	8.6% *(8.3-8.9)*

Section 5: MtDNA Analyses

MtDNA relationships among Denisovans. The mtDNA sequences of the three Denisovan individuals, seven Neandertals (Altai – KC879692, Mezmaiskaya 1 – FM865411.1, Feldhofer 1 – FM865407.1, Feldhofer 2 – FM865408.1, Vindija 33.16 – AM948965, Vindija 33.25 – FM865410.1 and Sidron 1253 – FM865409.1) (48, 54), five present-day humans (San – AF347008, Yoruba – AF347014, Han Chinese – AF346972, French – AF346981 and Papuan – AF347004) (56) and the chimpanzee (X93335.1) (57) were aligned using the software MAFFT v6.708b (58, 59). Pairwise mtDNA differences among the seven Neandertals and three Denisovans were calculated using MEGA 6.06 (60) (Suppl. Table S7). In addition, the three Denisovan mtDNAs were aligned with 311 modern human mtDNAs and the pairwise differences among these individuals were calculated.

To estimate phylogenetic relationships, Modeltest 3.7 (61) was used to identify an appropriate substitution model (GTR+G+I) and MrBayes 3.2 (62, 63) was run with default MCMC parameters for 5,000,000 generations, sampling every 1,000 generations, using a burn-in of 1,000,000 generations. The 4,000 resulting trees were combined to a consensus using TreeAnnotator v1.6.2 from the BEAST package (64) (Figure 2A).

A tree including the partial mtDNA sequence of a hominid from Sima de los Huesos, Spain (KF683087.1) (65) was estimated as above (Suppl. Figure S5).

Table S7. Number of differences to mtDNA MRCAs. The number of differences between each Denisovan mtDNA and their inferred MRCA as well as between each Neandertal mtDNA and their inferred MRCA.

Denisovan	Number of diffs to MRCA of Denisovans	Neandertal	Number of diffs to MRCA of Neandertals
Denisova 3	57	Mezmaiskaya 1	25
Denisova 4	55	Altai	24
Denisova 8	29	Feldhofer 1	21
		Feldhofer 2	17
		Sidron 1253	19
		Vi33.16	23
		Vi33.25	21

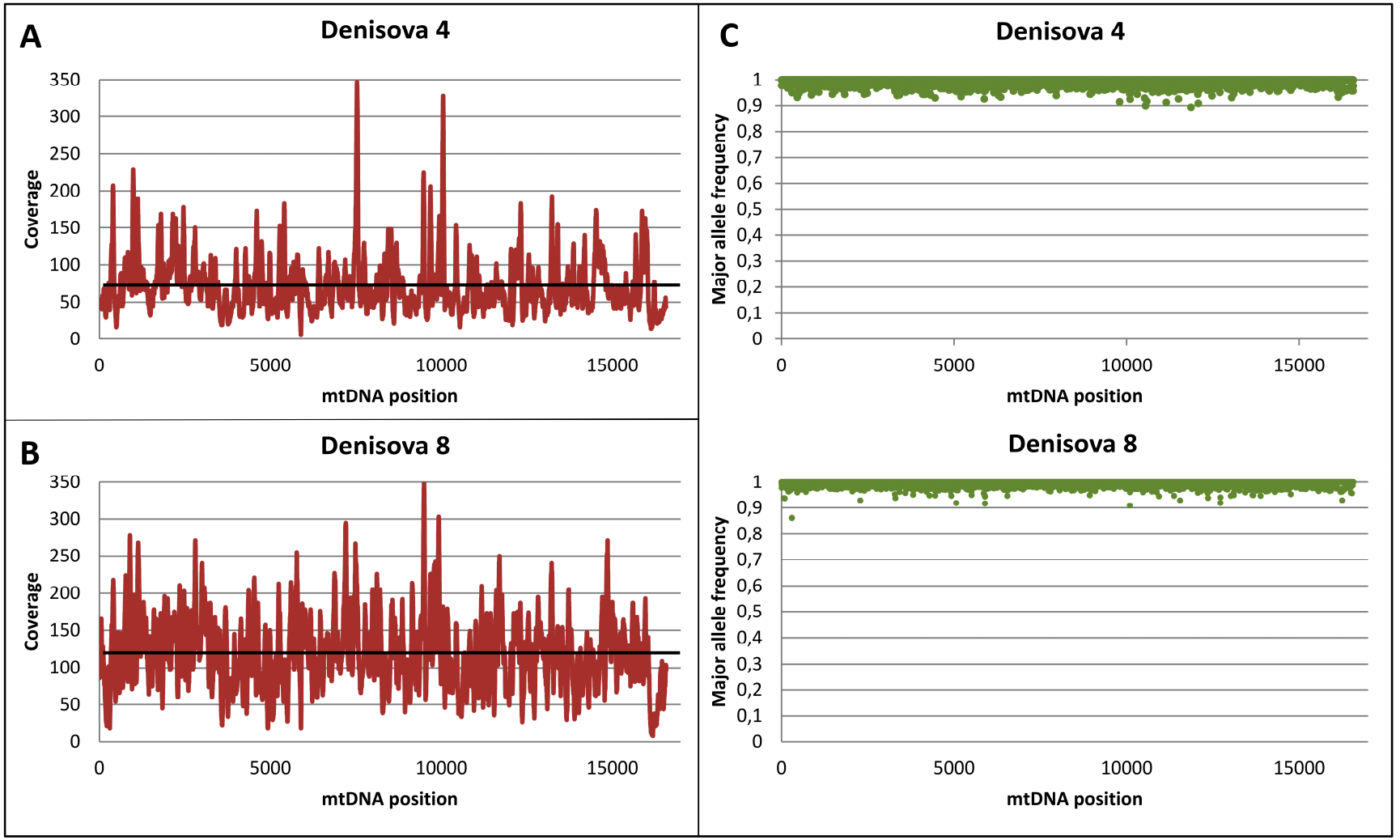

Figure S4. Quality of mtDNA sequences from *Denisova 4* and *8*. A, B: Coverage across the mitochondrial genomes. Black lines denote the average coverage. **C, D**: Consensus support across the genomes.

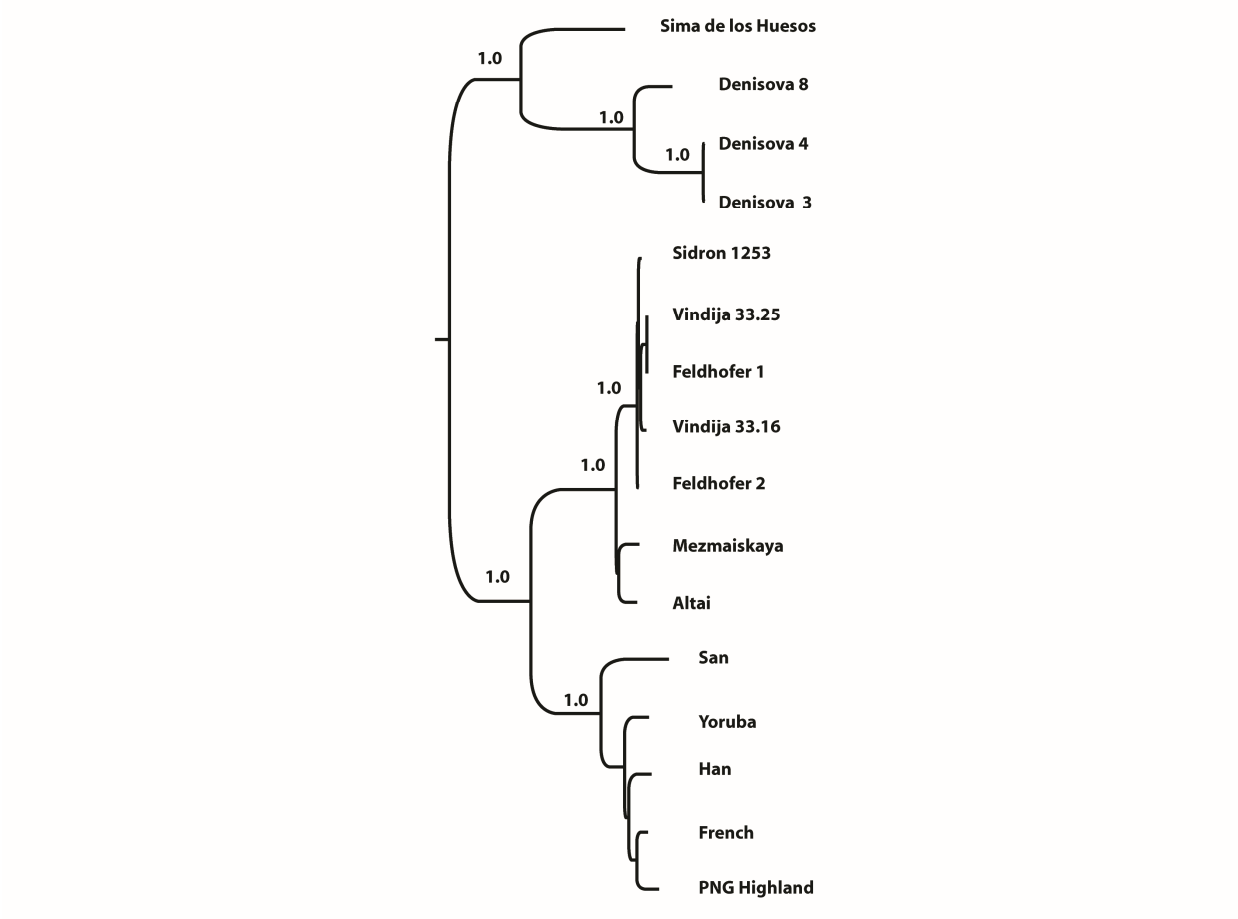

Figure S5. MtDNA tree of three Denisovans, seven Neandertals, a hominin from Sima de los Huesos (65), and five present-day humans. The Bayesian tree was computed using 16,286 mtDNA positions and a chimpanzee mtDNA (X93335.1) as outgroup (not shown). Important posterior probabilities are shown.

Branch Shortening. The most recent common ancestor (MRCA) of the three Denisovans was estimated using parsimony and a Yoruba mtDNA (AF347014). There were two positions where the MRCA was not resolvable. The MRCA of the seven Neandertals was calculated in the same way, with five unresolvable positions. The pairwise differences between the MRCAs and each individual were then calculated (Table S7).

Watterson's estimator θw. θw was calculated for the three Denisovan individuals and the seven Neandertal, 31 Europeans (Italians, Germans, Spanish, Saami, English, Dutch, Finnish and French) and 311 present-day humans (including the Europeans) (Table S8). θw was calculated as follows: $K/a_n/16{,}595$, where K is the number of segregating sites, and a_n is $\sum_{i=1}^{n-1} \frac{1}{i}$. The numbers of segregating sites were ascertained using DNA Sequence Polymorphism (DnaSP) version 5.10.01 (66).

Table S8. Watterson's estimator *(θw)* for mtDNA.

Population	# segregating sites	n (# indv)	*θw*
Denisovans	**86**	**3**	**3.46E-03**
Neandertals	**73**	**7**	**1.80E-03**
Present-day humans	**1,689**	**311**	**16.1E-03**
Present-day Europeans	**262**	**31**	**3.96E-03**

Bayesian dating. We estimated the age of the two molars and the divergence times between the three Denisovans, five radiocarbon-dated Neandertals (18), ten radiocarbon-dated ancient modern humans (67) and the five present-day humans used for tree estimations (Fig. 2) using BEAST v1.6.2. The age of *Denisova 3* date was set to either 50,000 years or 100,000 years as in ref. (54). A strict as well as a relaxed uncorrelated lognormal molecular clock was used with a normally distributed substitution rate prior of 2.67×10^{-8} per site per year (67) (standard deviation 1.0×10^{-8}), a Bayesian skyline coalescent tree prior with a uniform population size prior of 1,000 to 1,000,000 individuals, and a TN93 substitution model (68) . MCMC runs were carried out for 100,000,000 generations, sampling every 10,000 generations, with a burn-in of 10,000,000 generations. As expected, the relaxed clock is a better fit to the data and was used for the estimates presented in Table S9.

Table S9. Age estimates of the two molars and mtDNA lineages divergences based on mtDNA. Estimates using dates of 50,000 years as well as 100,000 years for *Denisova 3* and 95% upper and lower highest posterior densities (HPD) are given in thousand years (kyr).

	Age of *Denisova 3* set to 50,000 years BP			**Age *of Denisova 3* set to 100,000 years BP**		
Mitochondrial lineage	Estimate	95% HPD lower	95% HDP upper	Estimate	95% HPD lower	95% HDP upper
***Denisova 8* age**	177 kyr	97 kyr	265 kyr	226 kyr	143 kyr	313 kyr
***Denisova 4* age**	56 kyr	45 kyr	69 kyr	106 kyr	094 kyr	121 kyr
Denisova-Human/Neandertal	808 kyr	622 kyr	1,016 kyr	846 kyr	652 kyr	1056 kyr
Den8 – Den4*/*Den3	262 kyr	187 kyr	343 kyr	314 kyr	238 kyr	393 kyr
Human-Neandertal	405 kyr	312 kyr	511 kyr	413 kyr	318 kyr	522 kyr
San-rest of humans	173 kyr	128 kyr	223 kyr	176 kyr	128 kyr	225 kyr
Mezmaiskaya 1-rest of Neandertals	128 kyr	101 kyr	155 kyr	129 kyr	103 kyr	157 kyr

Section 6: Autosomal Analyses

Data Filtering. The following filters were implemented for the *Denisova 4* and *Denisova 8* autosomal analyses:

- Filters outlined in Section 3
- A minimum map quality of 37 (PHRED scale)
- Base quality set to 2 (phred scale) for Ts at the first or last two positions of fragments (to avoid errors induced by cytosine deamination)
- A minimum base quality of 30 (PHRED scale) (results in removal thymines with low base quality from step above)
- mappability filter that retains all positions where all possible overlapping 35-mers do not have match elsewhere in the genome allowing for one mismatch (54)
- Removal of triallelic sites
- Removal of CpG sites if the CpG occurs in either human, chimpanzee, gorilla or orangutan
- Removal of sites with a coverage higher than 2-fold
- When estimating nucleotide misincorporations due to cytosine deamination positions where the human reference (hg19) carries a C but one or more present-day human from the 1000 Genomes carries a T were excluded.

For high-coverage genomes, the following filters were used:

- mappability filter that retains all positions where all possible overlapping 35-mers do not have match elsewhere in the genome allowing for one mismatch (54)
- Root mean square of the map quality >= 30
- Coverage cut-off of 2.5% on each side of the coverage distribution; corrected for GC content for the *Denisova 3* and the Altai Neandertal (54)

Divergence Estimates. We estimate the divergence for *Denisova 4* and *Denisova 8* to ten present-day humans (French - HGDP00521, Sardinian - HGDP00665, Han - HGDP00778, Dai - HGDP01307, Papuan - HGDP00542, Australian - SS6004477, Dinka - DNK02, Mbuti - HGDP0456, Yoruba - HGDP00927, San - HGDP01029) (42, 54)), the high-coverage *Denisova 3* genome (42) and the high-coverage Altai Neandertal genome (54). The variant call format (VCF) files for the present-day humans as well as the *Denisova 3* and the Altai Neandertal were filtered as stated above.

Divergences between low-coverage and high-coverage genomes are estimated as the percentages of substitutions from the human-chimp ancestor to high-coverage genomes that occurred after the split of the low-coverage genomes from high-coverage genomes (see Figure 3A). Ancestral states for the human-chimpanzee ancestor was taken from the 6-way primate EPO alignments from

Ensembl version 69 (genome-wide alignments of human, chimpanzee, gorilla, orangutan, macaque, marmoset) (69, 70) and substitutions were parsimoniously assigned to one of the three lineages. Random alleles were picked at heterozygous sites in the high-coverage genomes while for the low-coverage Denisovan molars a random fragment was picket to represent each site analyzed. Standard errors for the divergence estimates (Suppl. Table S10-13) were estimated by running 5,000 jackknife replicates of the divergences in 5 Mb windows. Standard errors were multiplied by 1.96 to generate 95% CIs.

We similarly estimated divergences to the high-coverage Altai Neandertal genome (54) for low-coverage data from Vindija Cave, Croatia (Vindija 33.16, Vindija 33.25, Vindija 33.26), from El Sidron Cave, Spain (Sidron 1253), from Feldhofer Cave, Germany (Feldhofer 1) (all available from ERP000119, (71)), and from Mezmaiskaya Cave, Russia (Mezmaiskaya 1) (54). We excluded regions with a coverage higher than 2-fold for Feldhofer 1, 3-fold for the Vindija Neandertals and 4-fold for the Mezmaiskaya 1 Neandertal. We removed putative deamination-induced C to T substitutions at first and last two positions of the fragments from the Mezmaiskaya 1 Neandertal, as a double-stranded library preparation method and *E. coli* UDG was used, which does not remove uracils efficiently at these positions. For the other low-coverage Neandertals, which were not UDG treated, we removed putative deamination-induced C to T substitutions at the first and last five bases. We calculated the divergence of these six low-coverage Neandertals to the Altai Neandertal along with a 95% CI as above (Suppl. Table S13).

Table S10. Divergences for *Denisova 4*. Divergences for the deaminated sequences, not deaminated sequences as well as all sequences combined. Divergence is the percent divergence of *Denisova 4* along the branch to the human-chimpanzee ancestor from the high-coverage genomes (first column). 95% CI are given.

	Deaminated fragments				**Not deaminated fragments**				**All fragments**			
High-coverage genomes	Shared[1]	Genome[2]	*Den4*[3]	%	Shared	Genome	*Den4*	%	Shared	Genome	*Den4*	%
Denisova 3	3,699	109	3,767	**2.86** *2.28-3.44*	121,663	11,775	77,551	**8.82** *8.66-8.99*	126,716	11,990	81,920	**8.64** *8.48-8.81*
Altai Neandertal	3,471	340	4,029	**8.92** *8.01-9.83*	120,142	13,796	79,546	**10.30** *10.11-10.48*	124,952	14,290	84,303	**10.26** *10.08-10.44*
French	3,482	481	4,164	**12.14** *11.10-13.17*	126,237	11,133	76,306	**8.10** *7.94-8.27*	131,123	11,749	80,963	**8.22** *8.05-8.39*
Sardinian	3,448	489	4,095	**12.42** *11.37-13.47*	124,622	11,049	75,208	**8.14** *7.97-8.31*	129,262	11,634	80,055	**8.26** *8.09-8.42*
Han	3,455	477	4,111	**12.13** *11.06-13.2*	125,153	11,464	76,061	**8.39** *8.21-8.57*	129,955	11,919	80,724	**8.40** *8.23-8.57*
Dai	3,442	452	4,120	**11.61** *10.56-12.66*	124,793	11,519	75,590	**8.45** *8.28-8.62*	129,623	11,993	80,407	**8.47** *8.31-8.63*
Papuan	3,445	456	4,087	**11.69** *10.69-12.69*	124,182	11,617	75,444	**8.55** *8.38-8.73*	129,005	12,275	80,101	**8.69** *8.53-8.85*
Australian	3,418	449	4,098	**11.61** *10.56-12.67*	124,613	11,252	75,620	**8.28** *8.11-8.45*	129,368	11,845	80,360	**8.39** *8.23-8.55*
Dinka	3,418	448	4,159	**11.59** *10.58-12.59*	123,200	12,939	77,631	**9.50** *9.32-9.69*	127,989	13,397	82,318	**9.48** *9.3-9.66*
Mbuti	3,433	473	4,129	**12.11** *11.08-13.14*	122,769	13,726	78,122	**10.06** *9.87-10.24*	127,615	14,241	82,765	**10.04** *9.86-10.22*
Yoruba	3,473	515	4,146	**12.91** *11.88-13.95*	123,623	13,188	78,107	**9.64** *9.46-9.82*	128,425	13,890	82,882	**9.76** *9.57-9.95*
San	3,407	455	4,095	**11.78** *10.76-12.81*	121,951	13,989	77,901	**10.29** *10.10-10.48*	126,739	14,558	82,650	**10.30** *10.11-10.49*

1. The number of allelic states shared by the genome and *Densiova 4* but not shared with the human-chimpanzee ancestor.
2. Allelic states specific to the genome analyzed.
3. Allelic states specific to *Denisova 4*.

Table S11. **Divergences for *Denisova 8***. See Table S10 for explanations.

	Denisova8 deaminated				Denisova8 not deaminated				Denisova8 all			
Individual#1	Shared	Genome	*Den8*	%	Shared	Genome	*Den8*	%	Shared	Genome	*Den8*	%
Denisova 3	88,315	3102	33,574	**3.39** *3.25-3.53*	507,405	26,224	210,931	**4.91** *4.83-5*	637,505	31,657	261,670	**4.73** *4.64-4.82*
Altai Neandertal	84,101	7598	38,370	**8.29** *8.09-8.48*	486,591	47,274	234,493	**8.86** *8.73-8.97*	611,034	58,838	292,030	**8.78** *8.66-8.9*
French	82,999	10741	40,898	**11.46** *11.23-11.69*	486,909	60,026	243,442	**10.97** *10.86-11.09*	609,735	75,858	303,855	**11.02** *10.95-11.17*
Sardinian	82,188	10641	40,463	**11.46** *11.24-11.68*	481,113	59,575	240,320	**11.02** *10.9-11.13*	602,610	74,671	299,982	**11.05** *10.92-11.13*
Han	82,694	10661	40,505	**11.42** *11.2-11.64*	483,764	60,157	242,418	**11.06** *10.95-11.17*	606,187	75,989	302,355	**11.13** *11.03-11.24*
Dai	82,488	10633	40,676	**11.42** *11.2-11.64*	482,321	59,659	242,036	**11.01** *10.89-11.12*	604,506	75,249	302,505	**11.10** *10.97-11.17*
Papuan	82,423	10515	40,375	**11.31** *11.1-11.54*	481,045	59,090	240,568	**10.94** *10.83-11.05*	602,992	74,518	300,472	**11.00** *10.89-11.11*
Australian	82,513	10150	40,374	**10.95** *10.73-11.18*	482,594	57,825	240,792	**10.70** *10.59-10.81*	604,910	72,637	300,738	**10.76** *10.61-10.83*
Dinka	82,250	10846	40,385	**11.65** *11.43-11.87*	480,376	61,308	243,261	**11.32** *11.21-11.43*	601,643	76,990	303,706	**11.31** *11.24-11.45*
Mbuti	82,646	10858	40,571	**11.61** *11.4-11.82*	480,838	62,446	244,989	**11.49** *11.37-11.61*	603,063	78,469	305,286	**11.51** *11.4-11.63*
Yoruba	82,598	10875	40,745	**11.63** *11.42-11.85*	482,785	62,201	244,267	**11.41** *11.29-11.53*	604,950	77,960	304,739	**11.41** *11.31-11.52*
San	82,173	10985	40,645	**11.79** *11.57-12.01*	478,377	62,644	243,639	**11.58** *11.46-11.69*	599,764	79,290	304,396	**11.65** *11.57-11.78*

Table S12. Divergences for *Denisova 3*. See Table S10 for explanations.

	***Denisova 3* deaminated**				***Denisova 3* all**			
Individual#1	Shared	Genome	*Den3*	%	Shared	Genome	*Den3*	%
Denisova 3	-	-	-	-	-	-	-	-
Altai Neandertal	4531663	418624	1180396	**8.46** *8.37-8.54*	6040420	560355	1424811	**8.49** *8.4-8.57*
French	4439597	591694	1303961	**11.76** *11.68-11.84*	5908484	793950	1585350	**11.85** *11.76-11.93*
Sardinian	4391458	584609	1288694	**11.75** *11.67-11.82*	5842629	786441	1568908	**11.86** *11.78-11.94*
Han	4421887	587375	1295063	**11.73** *11.64-11.81*	5882753	788594	1576327	**11.82** *11.73-11.9*
Dai	4431008	587058	1299791	**11.70** *11.62-11.78*	5893395	788793	1581617	**11.80** *11.72-11.89*
Papuan	4410486	577448	1283146	**11.58** *11.49-11.66*	5867117	774918	1559910	**11.67** *11.58-11.75*
Australian	4433269	565793	1285982	**11.32** *11.23-11.4*	5899300	759006	1564065	**11.40** *11.31-11.49*
Dinka				*11.69-11.87*				*11.92-12.27*
Mbuti	4427808	593721	1301891	**11.82** *11.74-11.9*	5889250	795352	1585013	**11.90** *11.82-11.98*
Yoruba	4422950	592266	1297910	**11.81** *11.72-11.89*	5884572	794419	1581895	**11.89** *11.81-11.98*
San	4413422	595874	1297860	**11.90** *11.81-11.98*	5870882	798906	1580382	**11.98** *11.89-12.06*

Table S13. Divergences for Neandertals to the high coverage Altai Neandertal genome. See Table S10 for explanations of labels. All Mezmaiskaya 1 fragments were used for this analysis, because UDG treatment left C to T substitutions at only 4% of fragment ends.

	Neandertal deaminated				Neandertal all			
Neandertal	Shared	AltaiNea	Neandertal	%	Shared	AltaiNea	Neandertal	%
Feldhofer 1	447	6	576	1.32 *0.28-2.37*	2,581	67	3,446	2.53 *1.96-3.1*
Sidron 1253	893	29	1026	3.15 *2.00-4.29*	2,716	73	3,158	2.62 *1.97-3.26*
Vindija33.16	569,284	14,610	750,801	2.50 *2.44-2.57*	1,611,437	42,324	1,991,958	2.56 *2.5-2.61*
Vindija33.25	500,325	12,729	560,651	2.48 *2.41-2.55*	1,730,545	43,780	1,918,680	2.47 *2.41-2.52*
Vindija33.26	477,869	12,296	585,208	2.51 *2.44-2.58*	1,591,266	40,910	1,829,657	2.51 *2.45-2.56*
Mezmaiskaya1	-	-	-	-	2,331,784	59,473	772,431	2.49 *2.43-2.54*

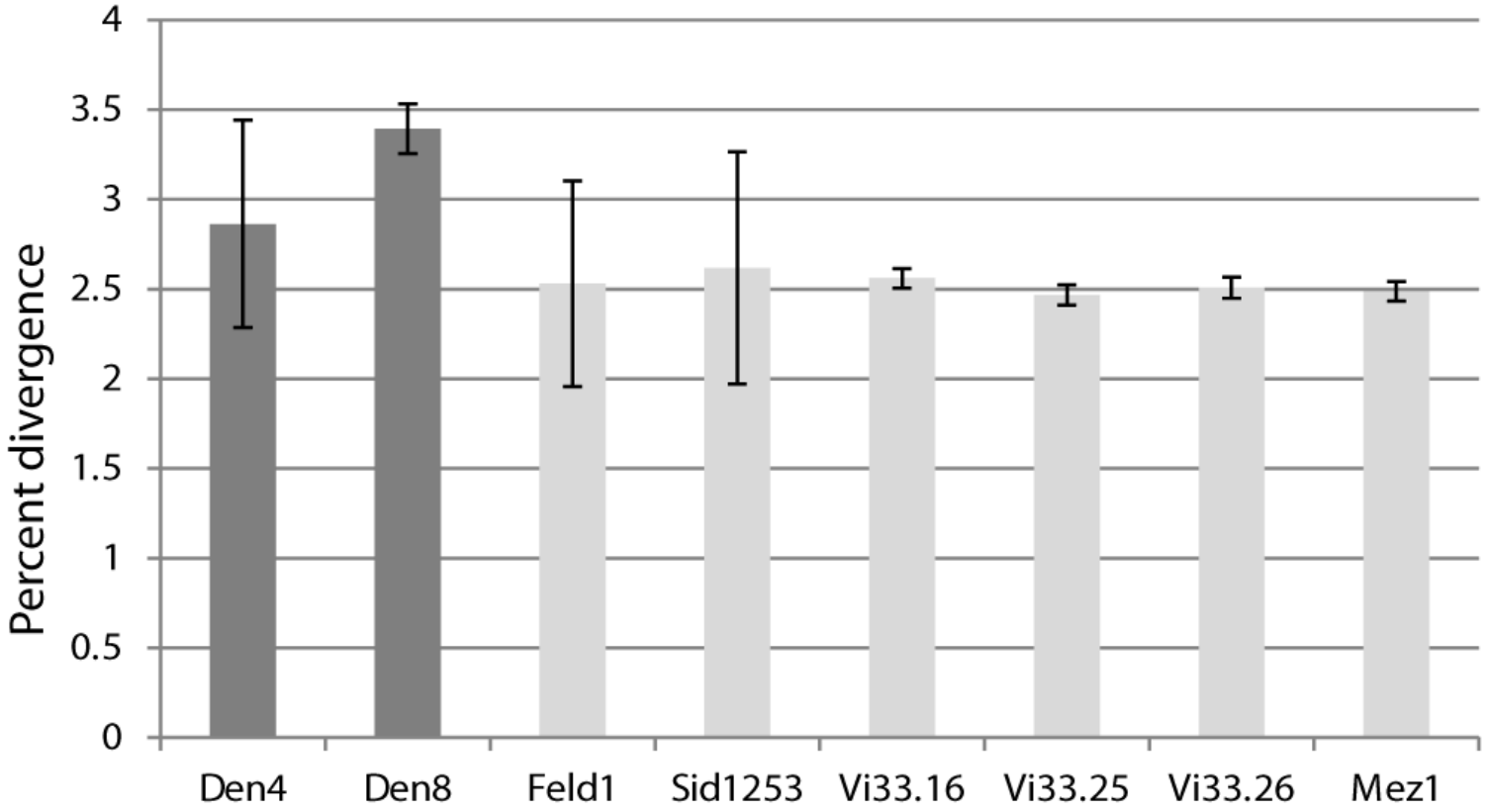

Figure S6. Divergences to *Denisova 3* and Altai Neandertal reference genomes. The percent divergence of the *Denisova 4* and *8* genomes to the *Denisova 3* genome (dark gray) and of six low-coverage Neandertal genomes to the Altai Neandertal genome (light gray) estimated as in main text Fig. 3A. Error bars indicate 95% CIs.

D-statistics. *D*-statistics (72) were calculated from genotype calls for high-coverage genomes, picking random alleles at heterozygous positions, or from random fragments for low-coverage genomes. Ancestral states were from the EPO alignment (69, 70) (Ensembl v69).

When the low-coverage Mezmaiskaya 1 genome was analyzed together with the high-coverage Altai Neandertal genome, random DNA sequences were picked from both genomes to avoid problems resulting from the difference in sequence quality between the two genomes.

Errors in the low coverage genome sequences contribute apparently derived alleles. To test if derived alleles in DNA sequences determined from *Denisova 8* tend match derived allele in one present-day person more than another, we used *Denisova 8* fragments and asked if derived alleles in *Denisova 8* match derived alleles in one or the other of two individuals from different African populations. This is not the case (D=0.01, Z=0.73).

Suppl. Table S14 shows that *Denisova 8* tends to share more derived alleles with the Papuan or Australian genomes using all sites (*D*:-0.03 to -0.08, Z-score: -1.9 to -4.3). However, the amount of data limits the power, as can be seen for similar comparisons using the whole high-coverage *Denisova 3* genome (*D*:-0.05 to -0.07, Z-score: -4.2 to -10.1).

To see if the amount of data determined from *Denisova 8* is enough to detect the excess sharing of derived alleles with the Altai relative to the Mezmaiskaya 1 previously described (42), we restrict the analysis to positions in the *Denisova 3* genome covered by the *Denisova 8* fragments and failed to detect the extra sharing (Suppl. Table S15). As expected from this, we fail to detect any excess sharing of derived alleles between *Denisova 8* and the Altai genome *(*Suppl. Table S15) when we restricted the analysis to transversions in order to avoid aberrant results due to errors in the low-coverage Mezmaiskaya 1 genome (not shown).

Table S14. Sharing of derived alleles between *Denisova 8* and Eurasian populations. Only *Denisova 8* fragments carrying a C to T substitutions at the first or last two positions are used.

	Type of sites	$AADA_a$	ADDA	DADA	DDDA	(ADDA-DADA)/ (ADDA+ADDA)	Z_b
Papuan, French, *Den8*, Chimp	all sites	43,502	1,311	1,473	205,735	-0.06	-3.03
	no cpg sites	36,640	906	1,022	179,687	-0.06	**-2.55**
	only cpg sites	6,862	405	451	26,048	-0.05	-1.57
	transitions	25,093	913	1,004	136,322	-0.05	-2.03
	transversions	18,409	398	469	69,413	-0.08	-2.33
Papuan, Sardinian, *Den8*, Chimp	all sites	43,387	1,358	1,454	205,685	-0.03	-1.90
	no cpg sites	36,519	930	1,023	179,680	-0.05	**-2.24**
	only cpg sites	6,868	428	431	26,005	0.00	-0.10
	transitions	25,031	944	1,010	136,224	-0.03	-1.56
	transversions	18,356	414	444	69,461	-0.03	-1.02
Papuan, Han, *Den8*, Chimp	all sites	43,255	1,232	1,352	204,023	-0.05	-2.32
	no cpg sites	36,435	832	951	178,188	-0.07	**-2.84**
	only cpg sites	6,820	400	401	25,835	0.00	-0.03
	transitions	24,989	855	913	135,233	-0.03	-1.36
	transversions	18,266	377	439	68,790	-0.08	-2.09
Papuan, Dai, *Den8*, Chimp	all sites	43,215	1,199	1,356	204,110	-0.06	-3.31
	no cpg sites	36,360	833	956	178,239	-0.07	**-3.01**
	only cpg sites	6,855	366	400	25,871	-0.04	-1.32

	transitions	24,981	816	927	135,238	-0.06	-2.81
	transversions	18,234	383	429	68,872	-0.06	-1.66
Australian, French, *Den8*, Chimp	all sites	43,027	1,224	1,451	204,126	-0.08	-4.32
	no cpg sites	36,314	861	966	178,308	-0.06	**-2.43**
	only cpg sites	6,713	363	485	25,818	-0.14	-4.19
	transitions	24,847	865	979	135,092	-0.06	-2.66
	transversions	18,180	359	472	69,034	-0.14	-3.82
Australian, Sardinian, *Den8*, Chimp	all sites	43,118	1,313	1,482	204,409	-0.06	-3.19
	no cpg sites	36,335	915	1,023	178,581	-0.06	**-2.40**
	only cpg sites	6,783	398	459	25,828	-0.07	-2.13
	transitions	24,892	896	1,009	135,205	-0.06	-2.60
	transversions	18,226	417	473	69,204	-0.06	-1.91
Australian, Han, *Den8*, Chimp	all sites	43,016	1,228	1,389	202,806	-0.06	-3.06
	no cpg sites	36,281	844	944	177,174	-0.06	**-2.34**
	only cpg sites	6,735	384	445	25,632	-0.07	-2.07
	transitions	24,852	875	927	134,265	-0.03	-1.20
	transversions	18,164	353	462	68,541	-0.13	-3.92
Australian, Dai, *Den8*, Chimp	all sites	42,767	1,243	1,391	202,727	-0.06	-2.98
	no cpg sites	36,047	894	969	177,058	-0.04	**-1.80**

	only cpg sites	6,720	349	422	25,669	-0.09	-2.75
	transitions	24,757	848	929	134,146	-0.05	-2.04
	transversions	18,010	395	462	68,581	-0.08	-2.20
Papuan, Han, *Den3*, Chimp	all sites	71,720	8,606	9,909	1,397,467	-0.07	-9.5
	no cpg sites	60,186	6,052	7,040	1,225,758	-0.08	**-8.60**
	only cpg sites	11,534	2,554	2,869	171,709	-0.06	-4.211
	transitions	48,439	5,944	6,801	927,866	-0.07	-7.52
	transversions	23,281	2,662	3,108	469,601	-0.08	-5.87
Papuan, French, *Den3*, Chimp	all sites	71,440	8,886	10,258	1,397,118	-0.07	-10.0
	no cpg sites	59,920	6,284	7,224	1,225,378	-0.07	**-8.30**
	only cpg sites	11,520	2,602	3,034	171,740	-0.08	-5.69
	transitions	48,215	6,168	7,094	927,573	-0.07	-8.08
	transversions	23,225	2,718	3,164	469,545	-0.08	-6.00
Papuan, French, *Den3*, Chimp	all sites	10111	1290	1480	206318	-0.07	-3.57
(sites covered by *Den8*)	no cpg sites	8382	894	1033	179798	-0.07	-3.10
	only cpg sites	1729	396	447	26520	-0.06	-1.73
	transitions	6883	884	1004	136864	-0.06	-2.73
	transversions	3228	406	476	69454	-0.08	-2.26
Papuan, French, *Den3*, Chimp	all sites	399	42	59	8872	-0.17	-1.75
(sites covered by *Den4*)	no cpg sites	340	29	41	7889	-0.17	-1.52

	only cpg sites	59	13	18	983	-0.16	-0.85
	transitions	273	27	40	5777	-0.19	-1.54
	transversions	126	15	19	3095	-0.12	-0.72

a. 'A' refers to an ancestral state and 'D' refers to a derived state. Thus, this column shows the number of sites where populations 1 and 2 share the ancestral allele with population 4 (Ancestral), and population 3 (Derived) has a derived state.
b. The Z-score is the difference between the *D*-statistics using all data and the mean of the same statistics for bootstrap replicates divided by the standard deviation for those replicates.

Table S15. Sharing of derived alleles between *Denisova 8* and Neandertals. *Denisova 8* fragments carrying a C to T substitutions at the first or last two positions (*Den8*_deaminated) as well as all fragments (*Den8*_all) are used. Only estimates based on transversions can be used due to errors in the low coverage Mezmaiskaya 1 genome.

	Type of sites	AADA	ADDA	DADA	DDDA	(ADDA-DADA)/(ADDA+ADDA)	Z
Mez, AltaiNea, *Den8*_deaminated, Chimp	all sites	15,245	511	376	77,110	0.15	4.49
	no cpg sites	12,142	179	139	64,649	0.13	2.27
	only cpg sites	3,103	332	237	12,461	0.17	4.00
	transitions	8,898	431	313	52,358	0.16	4.34
	transversions	**6,347**	**80**	**63**	**24,752**	**0.12**	**1.44**
Mez, AltaiNea, *Den8*_all, Chimp	all sites	104,707	3,586	2,532	521,739	0.17	13.88
	no cpg sites	87,986	1,382	1,138	441,125	0.10	4.92
	only cpg sites	16,721	2,204	1,394	80,614	0.23	14.12
	transitions	56,272	3,063	2,041	354,226	0.20	14.75
	transversions	**48,435**	**523**	**491**	**167,513**	**0.03**	**1.02**
Mez, AltaiNea, *Den3*, Chimp	all sites	3,392	296	498	77,271	-0.25	-7.15
(sites covered by *Den8*_deaminated)	no cpg sites	2,655	121	177	64,648	-0.19	-3.21
	only cpg sites	737	175	321	12,623	-0.29	-6.75
	transitions	2,371	234	420	52,530	-0.28	-7.44
	transversions	**1,021**	**62**	**78**	**24,741**	**-0.11**	**-1.32**
Mez, AltaiNea, *Den3*, Chimp	all sites	23,573	3,463	2,024	523,579	0.26	20.23
(sites covered by *Den8*_all)	no cpg sites	18,757	1,348	957	441,914	0.17	8.38
	only cpg sites	4,816	2,115	1,067	81,665	0.33	20.29

	transitions	16,333	2,977	1,599	355,784	0.30	21.29
	transversions	**7,240**	**486**	**425**	**167,795**	**0.07**	**2.04**
Mez, AltaiNea, *Den3*, Chimp	all sites	295,159	42,000	24,746	6,550,020	0.26	70.25
(all *Den3* sites, not conditioned on	no cpg sites	232,239	16,149	11,699	5,547,171	0.16	27.62
***Den8*)**	only cpg sites	62,920	25,851	13,047	1,002,849	0.33	67.79
	transitions	205,685	36,111	19,517	4,490,038	0.30	76.13
	transversions	**89,474**	**5,889**	**5,229**	**2,059,982**	**0.06**	**6.35**

References

1. Molnar S (1971) Human tooth wear, tooth function and cultural variability. *American Journal of Physical Anthropology* 34(2):175-189.
2. Wolpoff MH (1979) The Krapina dental remains. *American Journal of Physical Anthropology* 50(1):67-113.
3. Stringer C, Dean M, & Martin R (1990) A comparative study of cranial and dental development in a recent British population and Neandertals. *De Rousseau CJ, ed. Primate life history and evolution* New York: Wiley Liss, Inc:115-152.
4. Tompkins RL (1996) Relative dental development of Upper Pleistocene hominids compared to human population variation. *American Journal of Physical Anthropology* 99(1):103-118.
5. Condemi S (2001) Les Néanderthaliens de La Chaise (Abri Bourgeois-Delaunay). *Éditions du Comité des Travaux Historiques et Scientifiques, Paris*.
6. Bilsborough A & Thompson JL (2005) The dentition of the Le Moustier 1 Neandertal. *H. Ullrich (Ed.), The Neandertal adolescent Le Moustier 1 - New aspects, new results* Berlin: Staatliche Museen zu Berlin - Preußischer Kulturbesitz:157-186.
7. Bermudez De Castro JM & Rosas A (2001) Pattern of dental development in Hominid XVIII from the Middle Pleistocene Atapuerca-Sima de los Huesos site (Spain). *American Journal of Physical Anthropology* 114(4):325-330.
8. Wallace JA (1977) Gingival eruption sequences of permanent teeth in early hominids.5. *American Journal of Physical Anthropology* 46(3):483-493.
9. Margvelashvili A, Zollikofer, C. P., Lordkipanidze, D., Peltomäki, T., & Ponce de León, M. S. (2013) Tooth wear and dentoalveolar remodeling are key factors of morphological variation in the Dmanisi mandibles. . *Proceedings of the National Academy of Sciences of the United States of America* 110(43):17278-17283.
10. Woo J (1964) Mandible of sinanthropus lantianensis. *Curr Anthropol* 5(2):98-101.
11. Li T & Etler DA (1992) New Middle Pleistocene hominid crania from Yunxian in China. *Nature* 357(6377):404-407.
12. Turner CG, Nichol, C.R. & Scott, G.R. (1991) Scoring procedures for key morphological traits of the permanent dentition: the Arizona State University Dental Anthropology System. *Advances in Dental Anthropology, Wiley Liss, New York*:13-31.
13. Reich D*, et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468(7327):1053-1060.
14. Glantz M, Viola, B., Wrinn, P., Chikisheva, T., Derevianko, A., Krivoshapkin, A., Islamov, U., Suleimanov, R. & Ritzman, T. (2008) New hominin remains from Uzbekistan. *Journal of Human Evolution* 55(2):223-237.
15. Bailey S, Glantz, M., Weaver, T.D. & Viola, B (2008) The affinity of the dental remains from Obi-Rakhmat Grotto, Uzbekistan. *Journal of Human Evolution* 55(2):238-248.
16. Trinkaus E (2010) Denisova Cave, Peştera cu Oase, and Human Divergence in the Late Pleistocene. *PaleoAnthropology* 2010:196-200.
17. Smith TM, Reid, D. J., Olejniczak, A. J., Bailey, S. E., Glantz, M., Viola, B (2011) Dental development and age at death of a Middle Paleolithic juvenile hominin from Obi-Rakhmat Grotto, Uzbekistan. In S. Condemi & G.C. Weniger (eds). *Continuity and discontinuity in the peopling of Europe - 150 Years of Neanderthals Discoveries*:155-164.
18. Martinón-Torres M, Bermúdez de Castro, J.M., Gómez-Robles, A., Prado-Simón, L. & Arsuaga, J.L. (2012) Morphological description and comparison of the dental remains from Atapuerca-Sima de los Huesos site (Spain). *Journal of Human Evolution* 62(1):7-58.
19. Grine FE & Franzen JL (1994) Fossil hominid teeth from the Sangiran dome (Java, Indonesia). *Courier Forsch Senckenberg* 171:75-103.

20. Kaifu Y (2006) Advanced dental reduction in Javanese Homo erectus. *Anthropol Sci* 114:35-43.
21. Weidenreich F (1937) The dentition of Sinanthropus pekinensis: a comparative odontography of the hominids. Paleaontologia Sinica new Series D. *Chinese Academy of Sciences, Beijing*.
22. Compton T, & Stringer, C (2012) The human remains. Neanderthals in Wales - Pontnewydd and the Elwy valley caves *Oxford: Oxbow Books. In S. Aldhouse-Green, R. Peterson, & E. A. Walker (Eds.)*:118-230.
23. Xing S, Martinon-Torres, M., Bermudez de Castro, J.M., Wu, X., Liu, W (2015) Hominin Teeth From the Early Late Pleistocene Site of Xujiayao, Northern China. *Journal of Physical Anthropology* 156(2):224-240.
24. Martinón-Torres M, Bermúdez de Castro, J.M., Gómez-Robles, A., Margvelashvili, A., Prado, L., Lordkipanidze, D. & Vekua, A. (2008) Dental remains from Dmanisi (Republic of Georgia): Morphological analysis and comparative study. *Journal of Human Evolution* 55(2):249-273.
25. Bailey S (2006) Beyond shovel-shaped incisors: Neandertal dental morphology in a comparative context. . *Periodicum Biologorum* 108(3):253-267.
26. Frayer D (1976) Evolutionary Dental Changes in Upper Palaeolithic and Mesolithic Human Populations *PhD Dissertation, University of Michigan, Ann Arbor*.
27. Wood BA (1991) Hominid cranial remains. Koobi Fora Research Project, Vol. 4. *Clarendon Press, Oxford*.
28. Tobias PV (1991) Olduvai Gorge (The skulls, endocasts and teeth of Homo habilis, Vol. 4). *Cambridge: Cambridge University Press*.
29. Brown B & Walker A (1993) The dentition. *The Nariokotome Homo erectus Skeleton*:161-192.
30. Suwa G, Asfaw, B., Haile-Selassie, Y., White, T., Katoh, S., WoldeGabriel, G., Hart, W.K., Nakaya, H. & Beyene, Y. (2007) Early Pleistocene Homo erectus fossils from Konso, southern Ethiopia. *Anthropological Science* 115(2):133-151.
31. Wu X & Poirier FE (1995) Human Evolution in China. *Oxford University Press, Oxford*.
32. Rougier H (2003) Étude descriptive et comparative de Biache-Saint-Vaast 1 (Biache-Saint-Vaast, Pas-de-Calais, France). *PhD Thesis, University of Bordeaux I*.
33. de Lumley H, & de Lumley, M.A (1982) L'Homo erectus et sa place de l'homme de Tautavel parmi les hominidés fossiles. *Nice, Palais des expositions*.
34. Suzuki H & Takai F (1970) The Amud Man and his Cave Site. *Academic Press of Japan, Tokyo.*
35. Tillier A (1979) La dentition de l'enfant moustérien Châteauneuf 2, découvert á l'Abri de Hauteroche (Charente). *L'Anthropologie* 83(3):417-438.
36. Trinkaus E (1983) The Shanidar Neanderthals. *Academic Press, New York*.
37. McCown T & Keith A (1939) The Stone Age of Mount Carmel II: The human remains from the Levalloiso-Mousterien. *Oxford: Clarendon Press*.
38. Vandermeersch B (1981) Les hommes fossiles de Qafzeh (Israel). *Paris: Editions du CNRS* 319.
39. Hublin J-J, Verna, C., Bailey, S., Smith, T., Olejniczak, A., Sbihi-Alaoui, F. Z., & Zouak, M. (2012) Dental evidence from the Aterian human populations of Morocco. In J.- J. Hublin & S. McPherron (Eds.). *Vertebrate Paleobiology and Paleaanthropology: Modern origins - A North African perspective. Dordrecht: Springer*:189-204.
40. Sladek V, Trinkaus, E., HIllson, S.W. & Holliday, T.W. (2000) The People of the Pavlovian: Skeletal Catalogue and Osteometrics of the Gravettian Fossil Hominids from Dolní Vestonice and Pavlov. *Inst. of Archaeology, Academy of Sciences of the Czech Republic, Brno.*
41. Rohland N & Hofreiter M (2007) Comparison and optimization of ancient DNA extraction. *Biotechniques* 42(3):343-352.
42. Meyer M*, et al.* (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338(6104):222-226.
43. Briggs AW*, et al.* (2010) Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res* 38(6):e87.

44. Dabney J & Meyer M (2012) Length and GC-biases during sequencing library amplification: A comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques* 52(2):87-94.
45. Maricic T, Whitten M, & Paabo S (2010) Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 5(11):e14004.
46. Kircher M, Stenzel U, & Kelso J (2009) Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol* 10(8):R83.
47. Kircher M (2012) Analysis of high-throughput ancient DNA sequencing data. *Methods in molecular biology* 840:197-228.
48. Briggs AW*, et al.* (2009) Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science* 325(5938):318-321.
49. Genomes Project Consortium*, et al.* (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56-65.
50. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754-1760.
51. Andrews RM*, et al.* (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23(2):147.
52. Krause J*, et al.* (2010) A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr Biol* 20(3):231-236.
53. Sawyer S, Krause J, Guschanski K, Savolainen V, & Paabo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7(3):e34131.
54. Prufer K*, et al.* (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505(7481):43-49.
55. R Core Team (2014) R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
56. Ingman M, Kaessmann H, Paabo S, & Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408(6813):708-713.
57. Arnason U, Xu X, & Gullberg A (1996) Comparison between the complete mitochondrial DNA sequences of Homo and the common chimpanzee based on nonchimeric sequences. *J Mol Evol* 42(2):145-152.
58. Katoh K, Misawa K, Kuma K, & Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30(14):3059-3066.
59. Katoh K & Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* 30(4):772-780.
60. Koichiro Tamura GS, Daniel Peterson, Alan Filipski, and Sudhir Kumar (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. . *Molecular Biology and Evolution* 30:2725-2729.
61. Posada D & Crandall K (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14(9):817-818.
62. Huelsenbeck JP, Ronquist F, Nielsen R, & Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294(5550):2310-2314.
63. Ronquist F & Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572-1574.
64. Drummond AJ, Suchard MA, Xie D, & Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969-1973.
65. Meyer M*, et al.* (2014) A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* 505(7483):403-406.
66. Librado P & Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
67. Fu Q*, et al.* (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* 23(7):553-559.

68. Kass R & Raftery A (1995) Bayes Factors. *Journal of the American Statistical Association* 90(430):773-795.
69. Paten B, Herrero J, Beal K, Fitzgerald S, & Birney E (2008) Enredo and Pecan: genome-wide mammalian consistency-based multiple alignment with paralogs. *Genome Res* 18(11):1814-1828.
70. Paten B*, et al.* (2008) Genome-wide nucleotide-level mammalian ancestor reconstruction. *Genome Res* 18(11):1829-1843.
71. Green RE*, et al.* (2010) A draft sequence of the Neandertal genome. *Science* 328(5979):710-722.
72. Patterson N*, et al.* (2012) Ancient admixture in human history. *Genetics* 192(3):1065-1093.