ARTICLE

The complete genome sequence of a Neanderthal from the Altai Mountains

Kay Prüfer[1], Fernando Racimo[2], Nick Patterson[3], Flora Jay[2], Sriram Sankararaman[3,4], Susanna Sawyer[1], Anja Heinze[1], Gabriel Renaud[1], Peter H. Sudmant[5], Cesare de Filippo[1], Heng Li[3], Swapan Mallick[3,4], Michael Dannemann[1], Qiaomei Fu[1,6], Martin Kircher[1,5], Martin Kuhlwilm[1], Michael Lachmann[1], Matthias Meyer[1], Matthias Ongyerth[1], Michael Siebauer[1], Christoph Theunert[1], Arti Tandon[3,4], Priya Moorjani[4], Joseph Pickrell[4], James C. Mullikin[7], Samuel H. Vohr[8], Richard E. Green[8], Ines Hellmann[9]†, Philip L. F. Johnson[10], Hélène Blanche[11], Howard Cann[11], Jacob O. Kitzman[5], Jay Shendure[5], Evan E. Eichler[5,12], Ed S. Lein[13], Trygve E. Bakken[13], Liubov V. Golovanova[14], Vladimir B. Doronichev[14], Michael V. Shunkov[15], Anatoli P. Derevianko[15], Bence Viola[16], Montgomery Slatkin[2], David Reich[3,4,17], Janet Kelso[1] & Svante Pääbo[1]

We present a high-quality genome sequence of a Neanderthal woman from Siberia. We show that her parents were related at the level of half-siblings and that mating among close relatives was common among her recent ancestors. We also sequenced the genome of a Neanderthal from the Caucasus to low coverage. An analysis of the relationships and population history of available archaic genomes and 25 present-day human genomes shows that several gene flow events occurred among Neanderthals, Denisovans and early modern humans, possibly including gene flow into Denisovans from an unknown archaic group. Thus, interbreeding, albeit of low magnitude, occurred among many hominin groups in the Late Pleistocene. In addition, the high-quality Neanderthal genome allows us to establish a definitive list of substitutions that became fixed in modern humans after their separation from the ancestors of Neanderthals and Denisovans.

In 2008, a hominin finger phalanx was discovered during excavation in the east gallery of Denisova Cave in the Altai Mountains. From this bone, a genome sequence was determined to ~30-fold coverage[1]. Analysis showed that it came from a previously unknown group of archaic humans related to Neanderthals which we named 'Denisovans'[2]. Thus, at least two distinct human groups, Neanderthals and the related Denisovans, inhabited Eurasia when anatomically modern humans emerged from Africa. In 2010, another hominin bone, this time a proximal toe phalanx (Fig. 1a), was recovered in the east gallery of Denisova Cave[3]. Layer 11, where both the finger and the toe phalanx were found, is thought to be at least 50,000 years old. The finger was found in sublayer 11.2, which has an absolute date of 50,300 ± 2,200 years (OxA-V-2359-16), whereas the toe derives from the lowest sublayer 11.4, and may thus be older than the finger (Supplementary Information sections 1 and 2a). The phalanx comes from the fourth or the fifth toe of an adult individual and its morphological traits link it with both Neanderthals and modern humans[3].

Genome sequencing

In initial experiments to determine if DNA was preserved in the toe phalanx, we extracted and sequenced random DNA fragments. This revealed that about 70% of the DNA fragments present in the specimen aligned to the human genome. Initial inspection of the fragments with similarity to the mitochondrial (mt) genome suggested that its mtDNA was closely related to Neanderthal mtDNAs. We therefore assembled the full mitochondrial sequence by aligning DNA fragments to a complete Neanderthal mitochondrial genome[4] (Supplementary Information section 2b). A phylogenetic tree (Fig. 2a) shows that the toe phalanx mtDNA shares a common ancestor with six previously published Neanderthal mtDNAs[5] to the exclusion of present-day humans and the Denisova finger phalanx. Among Neanderthal mtDNAs, the toe mtDNA is most closely related to the mtDNA from infant 1 from Mezmaiskaya Cave in the Caucasus[6].
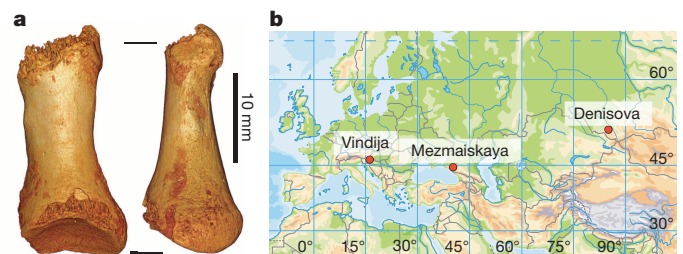

Figure 1 | Toe phalanx and location of Neanderthal samples for which genome-wide data are available. a, The toe phalanx found in the east gallery of Denisova Cave in 2010. Dorsal view (left image), left view (right image). Total length of the bone is 26 mm. **b**, Map of Eurasia showing the location of Vindija Cave, Mezmaiskaya Cave and Denisova Cave, where Neanderthal samples used here were found.

[1]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. [2]Department of Integrative Biology, University of California, Berkeley, California 94720-3140, USA. [3]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [4]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [5]Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA. [6]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China. [7]Genome Technology Branch and NIH Intramural Sequencing Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. [8]Department of Biomolecular Engineering, University of California, Santa Cruz, California 95064, USA. [9]Max F. Perutz Laboratories, Mathematics and Bioscience Group, Campus Vienna Biocenter 5, Vienna 1030, Austria. [10]Department of Biology, Emory University, Atlanta, Georgia 30322, USA. [11]Fondation Jean Dausset, Centre d'Étude du Polymorphisme Humain (CEPH), 75010 Paris, France. [12]Howard Hughes Medical Institute, Seattle, Washington 98195, USA. [13]Allen Institute for Brain Science, Seattle, Washington 98103, USA. [14]ANO Laboratory of Prehistory 14 Linia 3-11, St. Petersburg 1990 34, Russia. [15]Palaeolithic Department, Institute of Archaeology and Ethnography, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia. [16]Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. [17]Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. †Present address: Ludwig-Maximilians-Universität München, Martinsried, 82152 Munich, Germany.
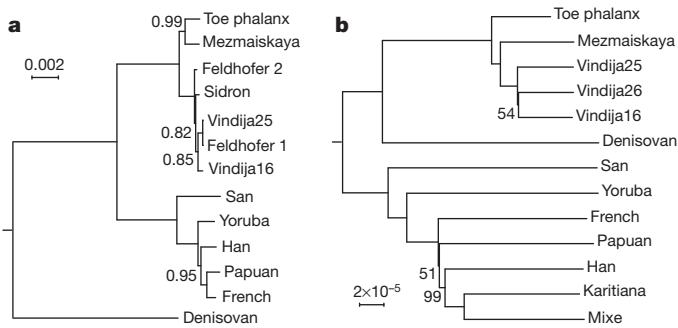
Figure 2 | Phylogenetic relationships of the Altai Neanderthal. a, Bayesian tree of mitochondrial sequences of the toe phalanx, the Denisovan finger phalanx, six Neanderthals and five present-day humans. Posterior probabilities are given for branches whose support is less than one (Supplementary Information section 2b). **b,** Neighbour-joining tree based on autosomal transversion differences among the toe phalanx, four Neanderthals, the Denisova genome and seven present-day human individuals. Bootstrap values are shown for branches supported by less than 100% of 1,000 bootstrap replicates (Supplementary Information section 6).

We generated four DNA libraries using a recently published protocol that is particularly efficient in retrieving DNA from ancient samples[1,7]. These libraries, together with one library prepared using a previous protocol[8], were treated with uracil-DNA glycosylase to remove uracil residues, a common miscoding lesion in ancient DNA that results from the deamination of cytosine[9–11] (Supplementary Information section 5a). In total, these five DNA libraries provided 52-fold sequence coverage of the genome. We estimated present-day human DNA contamination in the libraries with four complementary approaches (Supplementary Information section 5) using mtDNA and nuclear DNA and conclude that present-day human contamination among the DNA fragments sequenced is around 1%. After genotype calling, which is designed to be insensitive to low levels of error, we expect that the inferred genome sequence is largely free from contamination.

Relationship to other hominins

We compared the toe phalanx genome to the Denisovan genome[1], the draft Neanderthal genome of 1.3-fold coverage determined from three individuals from Vindija Cave, Croatia[12], the genome of a Neanderthal infant estimated to be 60,000 to 70,000-years-old[13] from Mezmaiskaya Cave in the Caucasus that we sequenced to 0.5-fold genomic coverage (Supplementary Information section 1; Fig. 1b) as well as 25 genomes of present-day humans: 11 previously sequenced to between 24- and 31-fold coverage[1] (Panel A), and 14 sequenced to between 35- and 42-fold coverage for this study (Panel B). We used pooled fosmid sequencing to resolve the sequences of the two chromosomes carried by 13 of these individuals[14] (Supplementary Information section 4).

A neighbour-joining tree (Fig. 2b) based on transversions, that is, purine–pyrimidine differences, among 7 present-day humans and the low-coverage Mezmaiskaya and Vindija genomes corrected for errors (Supplementary Information section 6a), shows that the toe phalanx nuclear genome forms a clade with the genomes of Neanderthals. The average DNA sequence divergence between the toe phalanx genome and the Mezmaiskaya and Vindija Neanderthal genomes is approximately a third of that between the Neanderthal and Denisova genomes. We conclude that the individual from whom the toe phalanx derives is a Neanderthal. Hereafter we refer to it as the 'Altai Neanderthal'.

Branch shortening

The length of the branches leading from the common ancestor shared with chimpanzee to the high-coverage Altai Neanderthal and Denisovan genomes are 1.02% (range of point estimates: 0.99–1.05%) and 0.81% (range: 0.77–0.84%) shorter, respectively, than the branches to the present-day human genomes (Table 1; Supplementary Information section 6b). This is expected because the archaic genomes ceased accumulating substitutions at the death of the individual tens of thousands of years ago. We previously estimated the shortening of the Denisovan lineage to be 1.16% (range: 1.13–1.27%)[1]. The fact that using present-day human genomes of higher quality and more stringent quality filtering reduces the estimate by about a third shows that at present, estimates of lineage lengths are unstable, probably owing to differences in the error rates among the genomes used. Nevertheless, the fact that the Neanderthal lineage is about 20% shorter than the Denisovan lineage suggests that the Neanderthal toe phalanx is older than the Denisovan finger phalanx, consistent with the stratigraphy of the cave.

Population split times

Figure 2b reflects the average divergence between DNA sequences. The times when the ancestral populations of archaic and modern humans separated are by necessity earlier. We used two approaches to estimate these population split times (Supplementary Information section 12). We caution that for these and other age estimates we rely on dates for the divergence of human and chimpanzee DNA sequences that in turn depend on the human mutation rate, which is currently controversial. In the text we present estimates based on a mutation rate of 0.5×10^{-9} base pairs per year, estimated from comparisons of the genomes of parents and children[15–19]. In Table 1 we also present estimates based on a rate of 1.0×10^{-9} base pairs per year derived from the fossil record which was used in previous studies of archaic genomes[1,2,12]. We also caution that the split times are at the best approximate because the models of population history used are likely to be inaccurate.

We first estimated population split times by extending the pairwise sequentially Markovian coalescent model (PSMC) to estimate the distribution of coalescence times between two single chromosomes that come from different populations[20,21] (Supplementary Information section 12). Using Sub-Saharan African genomes that were experimentally

Table 1 | Dating for branch shortening and population splits

Event	As % of human–chimp divergence	Absolute date calibration number 1 in kyr ($\mu = 1 \times 10^{-9}$ per bp per year)	Absolute date calibration number 2 in kyr ($\mu = 0.5 \times 10^{-9}$ per bp per year)	Supplementary Information section
Altai Neanderthal branch shortening	0.99–1.05	64–68	129–136	6b
Denisova branch shortening	0.77–0.84	50–54	100–109	6b
San-West African split	0.66–1.00	43–65	86–130	12
Introgressing Neanderthal–Altai split	0.58–0.88	38–57	77–114	13
Introgressing Denisovan–Denisovan split	2.12–3.10	138–202	276–403	13
Neanderthal–Denisova split*	2.93–3.64	190–236	381–473	12
Archaic–African split*	4.23–5.89	275–383	550–765	12
Unknown archaic split	7.90–31.12	450–2027	900–4054	16a, b

This table gives date ranges for two calibrations. The first assumes human–chimpanzee divergence of 6.5 million years and 1.30% for human–chimp divergence, or a mutation rate of 1×10^{-9} bp per year[1,2,12]. The second is based on direct measurement of per generation mutation rates[15–17], corresponding to a mutation rate of 0.5×10^{-9} per bp per year or 13 million years ago for human–chimpanzee divergence, and may fit better with some aspects of the fossil record[45,46]. Intervals give the range of values over tested human genomes for branch shortening; lowest and highest estimate for two or three methods for San-West African, Neanderthal–Denisova, Neanderthal–African and Denisova–African split; jackknife confidence interval over introgressed chunks for the Introgressing-Archaic–Archaic splits; and a union of the jackknife confidence interval in Supplementary Information section 16a and the highest posterior mode in Supplementary Information section 16b for the unknown archaic split.
* The indicated values are corrected for branch shortening where relevant as described in the Supplementary Information.

phased (Supplementary Information section 14) and segments of the archaic genomes in which the two chromosomes within an individual are closely related, we estimate the population split time between modern humans on the one hand, and Neanderthals and Denisovans on the other, to between 553,000 and 589,000 years ago, and the split time between Neanderthals and Denisovans to 381,000 years ago (Supplementary Information section 12).

In a second approach we counted how often randomly chosen alleles in an individual from one population are derived (that is, different from the apes) at positions where both the derived and ancestral alleles are seen in an individual from a second population[1,12]. Such derived alleles will be less frequent the older the population separation time is because more derived alleles in the second population will then be because of mutations that occurred after the split. Using this approach and the demographic history inferred from the PSMC (Supplementary Information section 12), we estimate the population split of Neanderthals and Denisovans from modern humans to 550,000–765,000 years ago, and the split time of Neanderthals and Denisovans to 445,000–473,000 years ago.

Inbreeding

We noticed that the Altai Neanderthal genome contains several long runs of homozygosity, indicating that her parents were closely related (Fig. 3a). To estimate the extent of their relatedness, we scanned the genome for 1 Mb regions where most non-overlapping 50-kb windows were devoid of heterozygous sites and merged adjacent regions (Supplementary Information section 10). The Neanderthal genome has 20

Double first cousins | Grandfather and granddaughter* | Half siblings | Uncle and niece*

Figure 3 | Indications of inbreeding in the Altai Neanderthal individual.
a, Time since the most recent common ancestor in log-scale for the two alleles of a French, the Denisovan and the Altai Neanderthal individual (Supplementary Information section 12) along 40 Mb of chromosome 21. **b**, Pedigrees showing four possible scenarios of parental relatedness for the Altai Neanderthal (that is, the child at the bottom of each pedigree). Two additional scenarios can be derived by switching the sex of the parents for the panels marked with an asterisk. **c**, Fraction of the genome in runs of homozygosity between 2.5 and 10 cM in length for Altai Neanderthal, Denisovan and the three present-day human individuals with the largest fractions (grey bars). The fractions for the Altai Neanderthal (bottom four bars) are reduced by the fraction expected from the four inbreeding scenarios in **b**. Error bars represent the full range of values obtained from 700 simulations for each scenario.

such regions longer than 10 cM, whereas the Denisovan genome has one. We performed simulations of inbreeding scenarios that can result in regions of this number and length, and find that the inbreeding coefficient is 1/8, indicating that the parents were as closely related as half-siblings. As the Altai individuals is a female (Supplementary Information section 5) and the X chromosome also has long runs of homozygosity, we can exclude parental relationships in which none or only one of the two X chromosomes was inherited from closely-related common ancestor(s), that is, scenarios that include two successive males in the pedigree. We conclude that the parents of this Neanderthal individual were either half-siblings who had a mother in common, double first cousins, an uncle and a niece, an aunt and a nephew, a grandfather and a granddaughter, or a grandmother and a grandson (Fig. 3b).

To investigate whether mating between closely related individuals may have been typical of the Altai Neanderthal population, we examined the distribution of runs of homozygosity between 2.5 and 10 cM in length. After removing the runs expected from recent inbreeding, the Altai Neanderthal genome still contains more runs than the Denisovan genome ($P < 2.2 \times 10^{-16}$), and both archaic genomes contain more than the Karitiana, a present-day population known to have a small effective size[22] (Fig. 3c; Supplementary Information 10). The sequencing of additional Neanderthal genomes to high quality will address whether breeding among close relatives was common also among Neanderthals in other geographic areas.

Heterozygosity and population size

The Neanderthal autosomal genome carries 1.7–1.8 heterozygous sites per 10,000 bp (Supplementary Information section 9). This is 84% of the number of heterozygous sites in the Denisovan genome, 22–30% of that in present-day non-African genomes, and 16–18% of that in present-day African genomes (Extended Data Fig. 1). When regions of homozygosity longer than 2.5 cM stemming from recent as well as long-term inbreeding in the Neanderthal are removed, 2.1–2.2 sites per 10,000 are heterozygous, similar to what is observed in the Denisovan genome. Thus, heterozygosity in Neanderthals as well as Denisovans appears to have been lower than in present-day humans and is among the lowest measured for any organism[23].

The demographic history of the population can be reconstructed from the distribution of the times since the most recent common ancestor of the two copies of the genome that a single person carries. We use the PSMC[20] to infer changes in the size of the Neanderthal population over time and compare this to inferences from the Denisovan and present-day human genomes (Fig. 4) (Supplementary Information section 12). All genomes analysed show evidence of a reduction in population size that occurred sometime before 1.0 million years ago. Subsequently, the population ancestral to present-day humans increased in size, whereas the Altai and Denisovan ancestral populations decreased further in size. It is thus clear that the demographic histories of both archaic populations differ substantially from that of present-day humans.

Neanderthal gene flow into modern humans

We have previously shown that Neanderthals contributed parts of their genomes to present-day populations outside Africa[12] and that Denisovans contributed to the genomes of present-day populations in Oceania[2,24] (used here to refer to Australia, Melanesia and the Philippines). Using the high-coverage Neanderthal genome in conjunction with the two other Neanderthal genomes, we now estimate that the proportion of Neanderthal-derived DNA in people outside Africa is 1.5–2.1% (Supplementary Information 14; Extended Data Table 1). Second, we find that the Neanderthal-derived DNA in all non-Africans is more closely related to the Mezmaiskaya Neanderthal from the Caucasus than it is to either the Neanderthal from Siberia (Extended Data Table 2; Supplementary Information section 14) (Z-score range: 4.0–6.4) or to the Vindija Neanderthals from Croatia[12] (Z-score range: 1.7–3.9). These results cannot be explained by present-day human contamination in
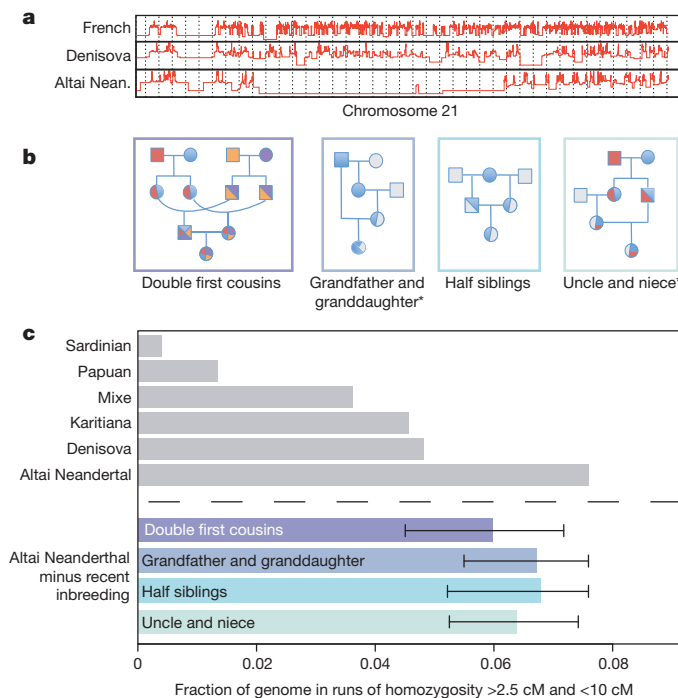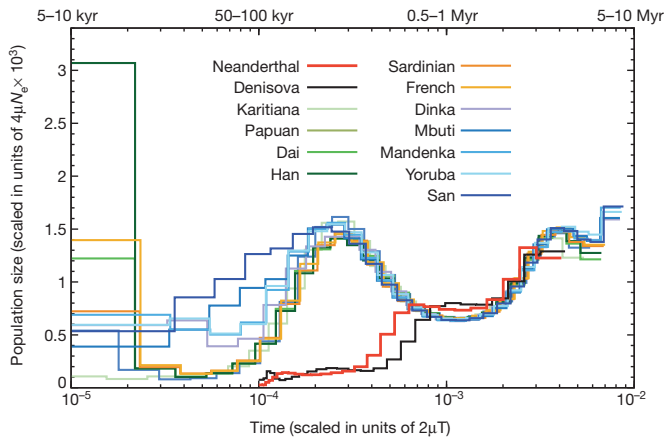
Figure 4 | Inference of population size change over time. The *y* axis specifies a number proportional to the population size N_e. The *x* axis specifies time in units of divergence per base pair (along the top in years for mutation rates of 0.5×10^{-9} to 1.0×10^{-9} per site per year). The analysis assumes that the Neanderthal and Denisova remains are of the same age, whereas archaeological evidence and the branch shortening indicate that the Neanderthal bone is older than the Denisovan bone. However, because the exact difference in ages is not known, it is not possible to determine whether the reduction in population size experienced by both archaic groups (but not by modern humans) coincided in time.

the Mezmaiskaya Neanderthal data, as a contamination level on the order of 2.0–5.4% would be needed to account for the excess relatedness to the Mezmaiskaya Neanderthal whereas the contamination in the Mezmaiskaya data are estimated to be 0–1.1% (Supplementary Information 5a).

Denisovan gene flow in mainland Asia

We used the two high-coverage archaic genomes and a hidden Markov model (HMM) to identify regions of specifically Neanderthal and specifically Denisovan ancestry in 13 experimentally phased present-day human genomes[1,14] (Supplementary Information sections 4 and 13). In the Sardinian and French genomes from Europe we find genomic regions of Neanderthal origin and few or no regions of Denisovan origin. In contrast, in the Han Chinese, the Dai in southern China, and the Karitiana and Mixe in the Americas, we find, in addition to regions of Neanderthal origin, regions that are consistent with being of Denisovan origin (Z-score = 4.3 excess relative to the Europeans) (Supplementary Information section 13), in agreement with previous analysis based on low-coverage archaic genomes[25]. These regions are also more closely related to the Denisova genome than the few regions identified in Europeans (Supplementary Information section 13). We estimate that the Denisovan contribution to mainland Asian and Native American populations is ~0.2% and thus about 25 times smaller than the Denisovan contribution to populations in Papua New Guinea and Australia. The failure to detect any larger Denisovan contribution in the genome of a 40,000-year-old modern human from the Beijing area[26] suggests that any Denisovan contribution to modern humans in mainland Asia was always quantitatively small. In fact, we cannot, at the moment, exclude that the Denisovan contribution to people across mainland Asia is owing to gene flow from ancestors of present-day people in Oceania after they mixed with Denisovans. We also note that in addition to this Denisovan contribution, the genomes of the populations in Asia and America appear to contain more regions of Neanderthal origin than populations in Europe[1,27] (Supplementary Information sections 13 and 14).

Archaic population differentiation

To estimate how closely related the archaic populations that contributed DNA to present-day humans were to the archaic individuals from which high-coverage genomes have been determined, we compared

the regions of Neanderthal and Denisovan ancestry in present-day human genomes identified by an HMM to the sequenced archaic genomes (Supplementary Information section 13). We find that the DNA sequence divergence in the regions that are most similar between the Altai Neanderthal genome and the Neanderthals that contributed DNA to present-day Eurasians is ~1.35% of the human–chimpanzee divergence, whereas the regions with the smallest sequence divergence between the Denisovan genome and the Denisovans that contributed DNA to present-day Papuans and Australians is ~3.18%. Regions of similarly low divergence are also identified by a window-based comparison (Fig. 5).

We estimate the population split time between the introgressing Neanderthal and the Altai Neanderthal genome to 77,000–114,000 years ago, and the split time between the introgressing Denisovan and the Denisovan genome to 276,000–403,000 years ago (Supplementary Information section 13) (Table 1). This is consistent with the Denisovan population being larger, more diverse and/or more subdivided than Neanderthal populations, and with the idea that Denisovans may have populated a wide geographical area. It is also in agreement with the low diversity among Neanderthal nuclear[2] and mitochondrial[5] genomes.

Neanderthal gene flow into Denisovans

If gene flow occurred between Neanderthals and Denisovans, we would expect that regions of the genome where the divergence between Denisovan and Neanderthal haplotypes is low would carry many differences between the two haplotypes of the individual who harbours the introgressed genetic material. This is because this individual carries two haplotypes that have accumulated differences independently in the two populations. In contrast, in the absence of gene flow, regions of low divergence between a Neanderthal and a Denisovan haplotype are not expected to have particularly elevated diversity (Supplementary Information section 15).

We plotted the number of differences between the Neanderthal genome and the closest inferred DNA sequences in the Denisovan genome against Denisovan heterozygosity (Fig. 6). We find that Denisovan heterozygosity is increased in regions where the Neanderthal and one Denisovan allele are close, indicating that gene flow from Neanderthals into Denisovans occurred, and estimate that a minimum of 0.5% of the Denisovan genome was contributed by Neanderthals. The Denisovan genome shares more derived alleles with the Altai Neanderthal genome than with the Croatian or Caucasus Neanderthal genomes (Z-score range: 5.6–10.2) (Extended Data Table 2; Supplementary Information section 15), suggesting that the gene flow into Denisovans came from a
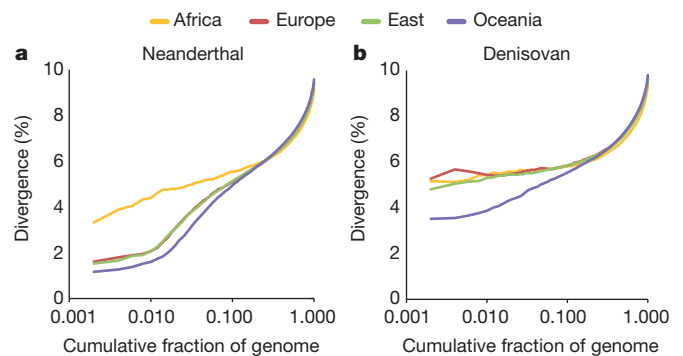

Figure 5 | Relatedness of introgressing archaic and sequenced archaic samples. Divergence of phased present-day human genomes to archaic genomes in windows of size 0.01 cM with a minimum of 25,000 analysed bases. Windows are sorted by sequence divergence measured on the archaic side of the tree (Supplementary Information section 13) and the *y* axis reports the divergence relative to human–chimpanzee divergence for cumulative fractions of the sorted windows over the entire genomes. Regions of low divergence between non-Africans and Neanderthals (**a**) and between Oceanians and Denisovans (**b**) indicate gene flow between these groups and the relative divergences between the introgressing archaic and sequenced archaic samples.
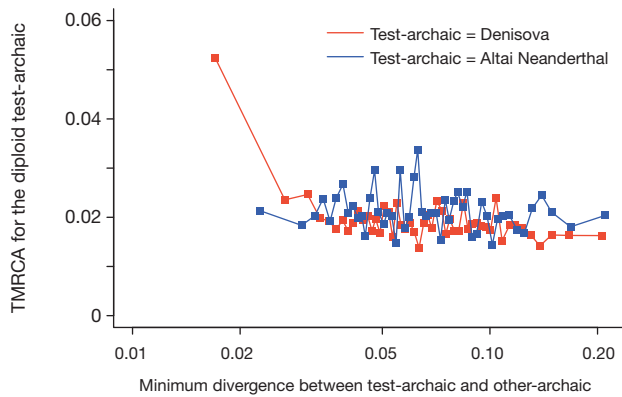
Figure 6 | Neanderthal gene flow into Siberian Denisovans. Divergence in 0.01 cM sized windows with at least 50 kb analysed bases between a 'test'-archaic genome and effectively haploid regions of the other archaic genome plotted against the most recent common ancestor of the two alleles of the 'test'-archaic. The plot shows 50 equally sized bins of windows for the 'test' Denisovan against the effectively haploid Neanderthal (red) and for the 'test'-archaic Altai Neanderthal against the effectively haploid Denisovan (blue). Divergence is given as percentage of human–chimpanzee divergence. Windows that show a close relationship between the effective haploid Altai Neanderthal and the closest inferred Denisovan haplotype show a deep divergence to the second Denisovan haplotype, indicating gene flow from Neanderthal into Denisovan.

Neanderthal population more related to the Altai Neanderthal than to the other two Neanderthals. In the reciprocal analysis, we find no corresponding increase in Neanderthal heterozygosity.

Particularly strong signals of Neanderthal gene flow into Denisovans are found in the human leucocyte antigen (HLA) region and the CRISP gene cluster on chromosome 6 (Extended Data Fig. 2), where we find many segments for which one of the Denisova haplotypes and the Altai Neanderthal share a common ancestor within a few tens of thousands of years before the death of the Altai individual (Supplementary Information section 15). This suggests the possibility that introgressed Neanderthal alleles may have contributed to the Denisovan functional variation at the HLA and the CRISP cluster, which are involved in immunity and sperm function, respectively. This is interesting as it has been suggested that HLA alleles from Neanderthals and Denisovans have been of functional relevance in modern humans[28].

Unknown archaic gene flow into Denisovans

As the ancestors of both Neanderthals and Denisovans left Africa before the emergence of modern humans, one might expect present-day Africans to share equal proportions of derived alleles with these two archaic groups. However, we find that African genomes share about 7% more derived alleles with the Neanderthal genome than with the Denisova genome ($Z = 11.6$ to 13.0; Extended Data Table 2; Supplementary Information section 16a) and that this is particularly the case for derived alleles that are fixed in Africans, of which 13–16% more are shared with the Neanderthal than with the Denisovan genome (Fig. 7).

We tested three non-mutually exclusive scenarios that could explain these observations. First, gene flow from the ancestor of Neanderthals after the split from Denisovans into the ancestors of all present-day humans would result in more sharing of derived alleles between present-day Africans and Neanderthals. However, because gene flow contributes alleles at low frequency, the sharing of derived alleles with Neanderthals would grow weaker with higher African derived allele frequency (Supplementary Information section 16a), whereas we observe the opposite (Fig. 7). Second, gene flow from the ancestors of present-day humans to Neanderthals after their split from Denisovans would also result in more sharing of derived alleles. However, the amount of allele frequency change (genetic drift) that has occurred in present-day Africans since the split from Neanderthals is too small to explain the extent of sharing
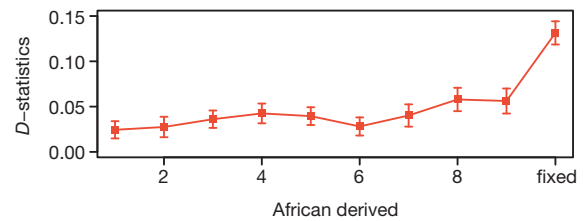

Figure 7 | Altai and Denisovan allele sharing with Africans stratified by African allele frequency. The plot shows the D-statistic of the form D(Neanderthal, Denisova; Africa, chimpanzee) binned by derived allele count in 10 deeply sequenced African genomes. Error bars represent ± 1 standard error. High-frequency and fixed derived alleles in Africa are more often shared with the Neanderthal than with the Denisovan genome.

of derived alleles fixed in Africans (Supplementary Information section 16a). Third, we considered a scenario where Denisovans received gene flow from a hominin whose ancestors diverged deeply from the lineage leading to Neanderthals, Denisovans and present-day humans. We find that this scenario is consistent with the data, as also suggested by others[29], and estimate that 2.7–5.8% (jackknife 95% confidence interval) of the Denisova genome comes from this putative archaic hominin which diverged from the other hominins 0.9–1.4 million years ago (Supplementary Information section 16a). An approximate Bayesian computation[30] again supports the third scenario (Supplementary Information section 16b) and estimates that 0.5–8% of the Denisovan genome comes from an unknown hominin which split from other hominins between 1.1 and 4 million years ago.

We caution that these analyses make several simplifying assumptions. Despite these limitations, we show that the Denisova genome harbours a component that derives from a population that lived before the separation of Neanderthals, Denisovans and modern humans. This component may be present due to gene flow, or to a more complex population history such as ancient population structure maintaining a larger proportion of ancestral alleles in the ancestors of Denisovans over hundreds of thousands of years.

The putative admixture into Denisovans from an unknown archaic group raises the possibility that the apparent Denisovan contribution to the genomes of Papuans and Australians could originate from admixture with the same unknown archaic population instead of with Denisovans. However, we tested this hypothesis and found that the archaic component in the genomes of people in Papua New Guinea and Australia comes from a group related to the Denisovans and not from an unknown archaic hominin (Supplementary Information section 17).

Copy number differences

The high-quality archaic genomes allow us to identify genetic changes that may have been relevant for putative biological traits that set modern humans apart from archaic humans. To identify genomic regions that have changed in copy number during hominin evolution, we used the variation of coverage along the two archaic genomes and 25 present-day human genomes (Supplementary Information section 8). We find three regions that have been duplicated only on the modern human lineage (Extended Data Table 3). One region overlaps *BOLA2,* which occurs as a single copy per haploid genome in the archaic genomes but has two to five copies in all but one of 675 present-day humans analysed, and which is near a microdeletion associated with developmental delay, intellectual disability and autism[31].

Catalogue of modern human changes

We compiled a genome-wide catalogue of sites where all or nearly all of 1,094 present-day humans[32] carry the same nucleotide but differ from the Neanderthal, Denisovan and great ape genomes (Supplementary Information section 18). In the regions of the genome to which short fragments can be mapped, there are 31,389 such single nucleotide
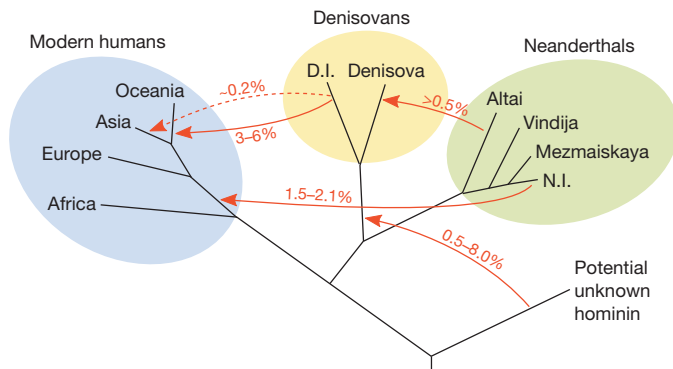
Figure 8 | A possible model of gene flow events in the Late Pleistocene. The direction and estimated magnitude of inferred gene flow events are shown. Branch lengths and timing of gene flows are not drawn to scale. The dashed line indicates that it is uncertain if Denisovan gene flow into modern humans in mainland Asia occurred directly or via Oceania. D.I. denotes the introgressing Denisovan, N.I. the introgressing Neanderthal. Note that the age of the archaic genomes precludes detection of gene-flow from modern humans into the archaic hominins.

substitutions and 4,113 short insertions and deletions (indels) shared by all present-day humans analysed, and a further 105,757 substitutions and 3,900 indels shared by 90% of present-day humans. This list of simple DNA sequence changes that distinguish modern humans from our nearest extinct relatives is thus comparatively small. For example, it contains only 96 fixed amino acid substitutions in a total of 87 proteins and in the order of three thousand fixed changes that potentially influence gene expression in present-day humans (Supplementary Information section 18).

Because the manner in which modern and archaic humans may have differed in aspects of their cognition is particularly interesting, we focused on the expression in the developing human brain of transcripts encoding the 87 proteins with fixed amino acid changes (Supplementary Information section 20). In comparison to a control set of transcripts that carry 108 silent substitutions fixed in present-day humans, genes carrying fixed amino acid changes are more often expressed in the ventricular zone of the developing neocortex ($P = 0.06$, corrected for multiple testing). Out of the five genes which are expressed in the proliferative layers (ventricular and subventricular zones combined) during mid-fetal development (*CASC5*, *KIF18A*, *TKTL1*, *SPAG5*, *VCAM1*), three (*CASC5*, *KIF18A*, *SPAG5*) are associated with the kinetochore of the mitotic spindle. This may be relevant phenotypically as the orientation of the mitotic cleavage plane in neural precursor cells during cortex development is thought to influence the fate of the daughter cells and the number of neurons generated (for example, see ref. 33). Another of these five genes, *VCAM1*, is essential for maintenance of neural stem cells in the adult subventricular zone[34].

Another way to prioritize changes in the catalogue for functional studies is to identify those that show signs of having risen to high frequency rapidly as they may have been affected by positive selection. We implemented an HMM to scan the genome for regions where the Neanderthal and Denisovan genomes fall outside of the variation of present-day humans (Supplementary Information section 19a). We ranked these regions, which cover less than 100 Mb of the genome, according to genetic length, because regions that rose rapidly to fixation are expected to be longer as they have been less affected by recombination events. A set of 63 regions likely to have been affected by positive selection were identified (Supplementary Information Table S19a.3). They contain 2,123 substitutions and 61 indels that are fixed or of high-frequency (>90%) in modern humans (Supplementary Information section 19b). They include, for example, the gene *RB1CC1* (also called *FIP200*) which encodes a transcription factor which, like *VCAM1*, is essential for maintenance of neuronal stem cells in the adult subventricular zone[35]. In present-day humans, but not Neanderthals and Denisovans, *RB1CC1*

carries a substitution inferred to change an amino acid in the encoded protein as well as a substitution that affects a conserved site in a motif that occurs across the genome[36]. Functional investigations will be necessary to clarify whether these and other such changes affect any phenotypes in present-day humans.

Discussion

We present evidence for three to five cases of interbreeding among four distinct hominin populations (Fig. 8). Clearly the real population history is likely to have been even more complex. For example, most cases of gene flow are likely to have occurred intermittently, often in both directions and across a geographic range. Thus, combinations of gene flow among different groups and substructured populations may have yielded the patterns detected rather than the discrete events considered here. Nevertheless, our analyses show that hominin groups met and had offspring on many occasions in the Late Pleistocene, but that the extent of gene flow between the groups was generally low.

We note that the observation that the Neanderthal DNA sequences in non-Africans share more derived alleles with the Neanderthal from the Caucasus than with Neanderthals from either Croatia or the Altai indicates that the archaic gene flow into non-Africans occurred at a time when Neanderthal populations had separated from each other. We also note that the introgressed Neanderthal DNA sequences suggest a population split from the Altai Neanderthal between 77,000 and 114,000 years ago (Supplementary Information section 13), well after ~230,000 years ago when Neanderthal features appear in the fossil record[37]. These and other results[38,39] show that the allele sharing between Neanderthals and non-African populations is owing to recent admixture rather than ancient population subdivision, an alternative which we and others previously considered possible[12,40].

The evidence suggestive of gene flow into Denisovans from an unknown hominin is interesting. The estimated age of 0.9 to 4 million years for the population split of this unknown hominin from the modern human lineage is compatible with a model where this unknown hominin contributed its mtDNA to Denisovans since the Denisovan mtDNA diverged from the mtDNA of the other hominins about 0.7–1.3 million years ago[41]. The estimated population split time is also compatible with the possibility that this unknown hominin was what is known from the fossil record as *Homo erectus*. This group started to spread out of Africa around 1.8 million years ago[42], but Asian and African *H. erectus* populations may have become finally separated only about one million years ago[43]. However, further work is necessary to establish if and how this gene flow event occurred.

METHODS SUMMARY

Sequences were generated on the Illumina HiSeq 2500 and base-calling was carried out using Ibis[44]. For all present-day human samples there is informed consent consistent with their use for whole genome sequencing and dissemination of data. Reads were merged and adaptor trimmed as described[1] and mapped to the human reference genome using BWA (version 0.5.10). Genotyping was carried out using GATK (version 1.3). We restricted analyses to regions of the genome that are non-repetitive (excluding tandem repeats), unique (requiring at least 50%, or all, overlapping 35-mers covering a position to map uniquely, allowing for one mismatch), and fall within the central 95% of the coverage distribution corrected for GC bias (Supplementary Information section 5b). The Supplementary Information describes the details of data processing and other analyses.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
2. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
3. Mednikova, M. B. A proximal pedal phalanx of a paleolithic hominin from Denisova cave, Altai. *Archaeol. Ethnol. Anthropol. Eurasia* **39**, 129–138 (2011).
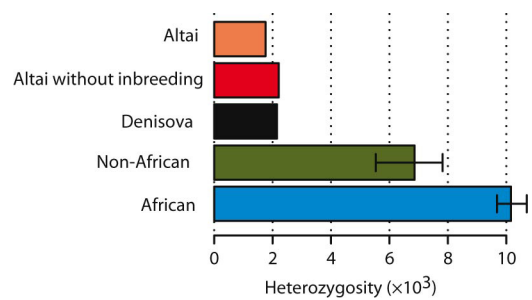
4. Green, R. E. et al. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell **134**, 416–426 (2008).
5. Briggs, A. W. et al. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science **325**, 318–321 (2009).
6. Golovanova, L. V., Hoffecker, J. F., Kharitonov, V. M. & Romanova, G. P. Mezmaiskaya cave: A Neanderthal occupation in the Northern Caucasus. Curr. Anthropol. **40**, 77–86 (1999).
7. Gansauge, M. T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nature Protocols **8**, 737–748 (2013).
8. Kircher, M. Analysis of high-throughput ancient DNA sequencing data. Methods Mol. Biol. **840**, 197–228 (2012).
9. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl Acad. Sci. USA **104**, 14616–14621 (2007).
10. Briggs, A. W. et al. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. **38**, e87 (2010).
11. Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A. & Paabo, S. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. **29**, 4793–4799 (2001).
12. Green, R. E. et al. A draft sequence of the Neandertal genome. Science **328**, 710–722 (2010).
13. Skinner, A. R. et al. ESR dating at Mezmaiskaya Cave, Russia. Appl. Radiat. Isot. **62**, 219–224 (2005).
14. Kitzman, J. O. et al. Haplotype-resolved genome sequencing of a Gujarati Indian individual. Nature Biotechnol. **29**, 59–63 (2011).
15. Abecasis, G. R. et al. A map of human genome variation from population-scale sequencing. Nature **467**, 1061–1073 (2010).
16. Awadalla, P. et al. Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. Am. J. Hum. Genet. **87**, 316–324 (2010).
17. Roach, J. C. et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science **328**, 636–639 (2010).
18. Kong, A. et al. Rate of de novo mutations and the importance of father's age to disease risk. Nature **488**, 471–475 (2012).
19. Campbell, C. D. et al. Estimating the human mutation rate using autozygosity in a founder population. Nature Genet. **44**, 1277–1281 (2012).
20. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. Nature **475**, 493–496 (2011).
21. Prado-Martinez, J. et al. Great ape genetic diversity and population history. Nature **499**, 471–475 (2013).
22. Kirin, M. et al. Genomic runs of homozygosity record population history and consanguinity. PLoS ONE **5**, e13996 (2010).
23. Leffler, E. M. et al. Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol. **10**, e1001388 (2012).
24. Reich, D. et al. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. Am. J. Hum. Genet. **89**, 516–528 (2011).
25. Skoglund, P. & Jakobsson, M. Archaic human ancestry in East Asia. Proc. Natl Acad. Sci. USA **108**, 18301–18306 (2011).
26. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl Acad. Sci. USA **110**, 2223–2227 (2013).
27. Wall, J. D. et al. Higher levels of Neanderthal ancestry in East Asians than in Europeans. Genetics **194**, 199–209 (2013).
28. Abi-Rached, L. et al. The shaping of modern human immune systems by multiregional admixture with archaic humans. Science **334**, 89–94 (2011).
29. Waddell, P. J. & Tan, X. New g%AIC, g%AICc, g%BIC, and power divergence fit statistics expose mating between modern humans, Neanderthals and other archaics. Preprint at http://arxiv.org/abs/1212.6820 (2012).
30. Wegmann, D., Leuenberger, C., Neuenschwander, S. & Excoffier, L. ABCtoolbox: a versatile toolkit for approximate Bayesian computations. BMC Bioinformatics **11**, 116 (2010).
31. Kumar, R. A. et al. Recurrent 16p11.2 microdeletions in autism. Hum. Mol. Genet. **17**, 628–638 (2008).
32. Abecasis, G. R. et al. An integrated map of genetic variation from 1,092 human genomes. Nature **491**, 56–65 (2012).
33. Fietz, S. A. & Huttner, W. B. Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective. Curr. Opin. Neurobiol. **21**, 23–35 (2011).
34. Kokovay, E. et al. VCAM1 is essential to maintain the structure of the SVZ niche and acts as an environmental sensor to regulate SVZ lineage progression. Cell Stem Cell **11**, 220–230 (2012).
35. Wang, C., Liang, C. C., Bian, Z. C., Zhu, Y. & Guan, J. L. FIP200 is required for maintenance and differentiation of postnatal neural stem cells. Nature Neurosci. **16**, 532–542 (2013).
36. Rios, D. et al. A database and API for variation, dense genotyping and resequencing data. BMC Bioinformatics **11**, 238 (2010).
37. Hublin, J. J. Out of Africa: modern human origins special feature: the origin of Neandertals. Proc. Natl Acad. Sci. USA **106**, 16022–16027 (2009).
38. Sankararaman, S., Patterson, N., Li, H., Pääbo, S. & Reich, D. The date of interbreeding between Neandertals and modern humans. PLoS Genet. **8**, e1002947 (2012).
39. Yang, M. A., Malaspinas, A. S., Durand, E. Y. & Slatkin, M. Ancient structure in Africa unlikely to explain Neanderthal and non-African genetic similarity. Mol. Biol. Evol. **29**, 2987–2995 (2012).
40. Eriksson, A. & Manica, A. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. Proc. Natl Acad. Sci. USA **109**, 13956–13960 (2012).
41. Krause, J. et al. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature **464**, 894–897 (2010).
42. Gabunia, L. et al. Dmanisi and dispersal. Evol. Anthropol. **10**, 158–170 (2001).
43. Asfaw, B. et al. Remains of Homo erectus from Bouri, Middle Awash, Ethiopia. Nature **416**, 317–320 (2002).
44. Kircher, M., Stenzel, U. & Kelso, J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. Genome Biol. **10**, R83 (2009).
45. Langergraber, K. E. et al. Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. Proc. Natl Acad. Sci. USA **109**, 15716–15721 (2012).
46. Scally, A. & Durbin, R. Revising the human mutation rate: implications for understanding human evolution. Nature Rev. Genet. **13**, 745–753 (2012).

Extended Data Figure 1 | Heterozygosity estimates for the Altai Neanderthal individual, the Denisovan individual, non-Africans and Africans. The bars for the latter two give the range of heterozygosity observed among 15 non-African and 10 African individuals, respectively (Supplementary Information section 9).
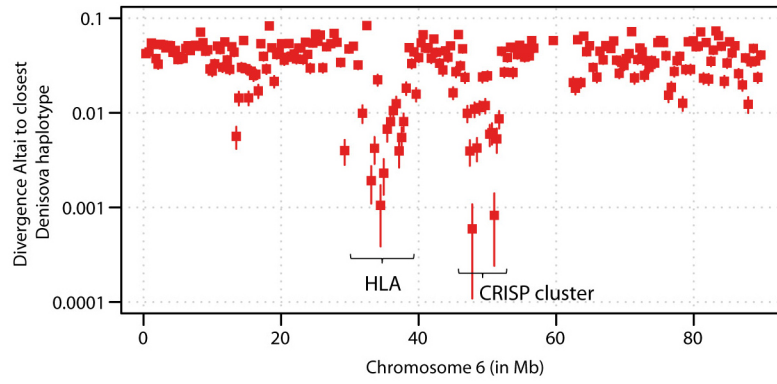
Extended Data Figure 2 | Neanderthal-introgressed loci in Denisova.
Divergence of the Altai Neanderthal to the most closely related Denisovan haplotype in windows of at least 200 kb on chromosome 6. Divergence is given as percentage of human–chimpanzee divergence and bars represent ± 1 standard error.

Extended Data Table 1 | Neanderthal ancestry estimate

| | Other Neanderthal = Mezmaiskaya | | | | Other Neanderthal = Vindija | | | |
| | Panel A | | Panel B | | Panel A | | Panel B | |
	$\hat{\alpha}$	Std. Err.	$\hat{\alpha}$	Std. Err.	$\hat{\alpha}$	Std. Err.	$\hat{\alpha}$	Std. Err.
French	0.020	0.003	0.019	0.003	0.016	0.002	0.017	0.002
Sardinian	0.019	0.002	0.017	0.003	0.018	0.002	0.018	0.002
Han	0.022	0.003	0.018	0.003	0.023	0.002	0.019	0.002
Dai	0.019	0.003	0.016	0.003	0.019	0.002	0.016	0.002
Karitiana	0.020	0.003	0.019	0.003	0.018	0.002	0.019	0.002
Mixe	-	-	0.018	0.003	-	-	0.017	0.002

Note: we estimate ancestry using the equation $\hat{\alpha} = \dfrac{f_4(\text{Denisova, Altai, Africa, X})}{f_4(\text{Denisova, Altai, Africa, Other Neanderthal})}$

Extended Data Table 2 | Selected *D*-statistics supporting inferences about gene flows

Statistic	D	Z	Interpretation
D(French, Dinka; Altai, Chimp)	5.4%	9.2	SI 14: Neanderthals share more derived alleles with non-Africans than with Africans.
D(Han, Dinka; Altai, Chimp)	7.3%	11.4	
D(Han, Papuan; Denisova, Chimp)	-7.0%	-9.5	SI 14: Denisovans share more derived alleles with Oceanian populations than with other non-Africans.
D(Han, Australian; Denisova, Chimp)	-7.7%	-10.7	
D(Altai, Mezmaiskaya; French, Dinka)	-16.4%	-5.8	SI 14: The archaic material in non-Africans falls within late Neanderthal variation: Non-Africans share more alleles with some Neanderthals (Mezmaiskaya/Vindija) than others (Altai).
D(Altai, Vindija; French, Dinka)	-7.0%	-4.3	
D(Altai, Mezmaiskaya; Denisova, Chimp)	13.2%	5.9	SI 15: Gene flow between Altai related Neanderthals and Denisovans (Denisovans share more derived alleles with Altai than with Mezmaiskaya)
D(Altai, Vindija; Denisova, Chimp)	7.9%	5.6	
D(Altai, Denisova; 12 Africans, Chimp)	7.0%	11.6	SI 16: Unknown archaic gene flow into Denisova: Africans share more derived alleles with Altai than with Denisova, a signal that strengthens for fixed derived alleles
D(Altai, Denisova; 12 Africans Fixed, Chimp)	13.4%	10.0	

Extended Data Table 3 | Lineage-specific segmental duplications along each of the terminal branches and genes encompassed

Locus	Length	Lineage	Genes	Genotypes			
				Modern Humans (median)	Denisova	Altai	Mezmai skaya
chr12:122079832-122087495	7663	Altai-Neanderthal	ORAI1	2	2	4	3
chr12:132295389-132391442	96053	Altai-Neanderthal	MMP17,ULK1	2	2	4	2
chr19:9284044-9291195	7151	Altai-Neanderthal		2	2	4	4
chr20:281880-290717	8837	Altai-Neanderthal		2	2	10	9
chr3:12639069-12641393	2324	Altai-Neanderthal	RAF1	2	2	7	3
chr6:95473793-95532866	59073	Altai-Neanderthal		2	2	3	2
chr11:39901956-39909545	7589	Denisova		2	4	2	2
chr1:161272681-161274838	2157	Denisova	MPZ	2	4	2	2
chr12:49894191-49897733	3542	Denisova	SPATS2	2	4	2	2
chr19:55302094-55315197	13103	Denisova	KIR3DP1,KIR2DL4	2	4	2	2
chr2:48781187-48787915	6728	Denisova		2	3	2	2
chr4:68542692-68577288	34596	Denisova	UBA6,LOC550112	2	3	2	2
chr4:68579206-68581585	2379	Denisova	LOC550112	2	3	2	2
chr7:140872574-140879065	6491	Denisova	LOC100131199	2	6	2	2
chr1:108924526-108990191	65665	Modern Human		4	2	2	2
chr16:30200098-30206185	6087	Modern Human	CORO1A,LOC606724,BOLA2	6	2	2	2
chr2:87417089-87420544	3455	Modern Human		4	2	2	2