

# The genetic history of Ice Age Europe

Qiaomei Fu<sup>1,2,3</sup>, Cosimo Posth<sup>4,5\*</sup>, Mateja Hajdinjak<sup>3\*</sup>, Martin Petr<sup>3</sup>, Swapan Mallick<sup>2,6,7</sup>, Daniel Fernandes<sup>8,9</sup>, Anja Furtwängler<sup>4</sup>, Wolfgang Haak<sup>5,10</sup>, Matthias Meyer<sup>3</sup>, Alissa Mittnik<sup>4,5</sup>, Birgit Nickel<sup>3</sup>, Alexander Peltzer<sup>4</sup>, Nadin Rohland<sup>2</sup>, Viviane Slon<sup>3</sup>, Sahra Talamo<sup>11</sup>, Iosif Lazaridis<sup>2</sup>, Mark Lipson<sup>2</sup>, Iain Mathieson<sup>2</sup>, Stephan Schiffels<sup>5</sup>, Pontus Skoglund<sup>2</sup>, Anatoly P. Derevianko<sup>12,13</sup>, Nikolai Drozdov<sup>12</sup>, Vyacheslav Slavinsky<sup>12</sup>, Alexander Tsybankov<sup>12</sup>, Renata Grifoni Cremonesi<sup>14</sup>, Francesco Mallegni<sup>15</sup>, Bernard Gély<sup>16</sup>, Eligio Vacca<sup>17</sup>, Manuel R. González Morales<sup>18</sup>, Lawrence G. Straus<sup>18,19</sup>, Christine Neugebauer-Maresch<sup>20</sup>, Maria Teschler-Nicola<sup>21,22</sup>, Silviu Constantin<sup>23</sup>, Oana Teodora Moldovan<sup>24</sup>, Stefano Benazzi<sup>11,25</sup>, Marco Peresani<sup>26</sup>, Donato Coppola<sup>27,28</sup>, Martina Lari<sup>29</sup>, Stefano Ricci<sup>30</sup>, Annamaria Ronchitelli<sup>30</sup>, Frédérique Valentin<sup>31</sup>, Corinne Thevenet<sup>32</sup>, Kurt Wehrberger<sup>33</sup>, Dan Grigorescu<sup>34</sup>, Hélène Rougier<sup>35</sup>, Isabelle Crevecoeur<sup>36</sup>, Damien Flas<sup>37</sup>, Patrick Semal<sup>38</sup>, Marcello A. Mannino<sup>11,39</sup>, Christophe Cupillard<sup>40,41</sup>, Hervé Bocherens<sup>42,43</sup>, Nicholas J. Conard<sup>43,44</sup>, Katerina Harvati<sup>43,45</sup>, Vyacheslav Moiseyev<sup>46</sup>, Dorothee G. Drucker<sup>42</sup>, Jiří Svoboda<sup>47,48</sup>, Michael P. Richards<sup>11,49</sup>, David Caramelli<sup>29</sup>, Ron Pinhasi<sup>8</sup>, Janet Kelso<sup>3</sup>, Nick Patterson<sup>6</sup>, Johannes Krause<sup>4,5,43§</sup>, Svante Pääbo<sup>3§</sup> & David Reich<sup>2,6,7§</sup>

**Modern humans arrived in Europe ~45,000 years ago, but little is known about their genetic composition before the start of farming ~8,500 years ago. Here we analyse genome-wide data from 51 Eurasians from ~45,000–7,000 years ago. Over this time, the proportion of Neanderthal DNA decreased from 3–6% to around 2%, consistent with natural selection against Neanderthal variants in modern humans. Whereas there is no evidence of the earliest modern humans in Europe contributing to the genetic composition of present-day Europeans, all individuals between ~37,000 and ~14,000 years ago descended from a single founder population which forms part of the ancestry of present-day Europeans. An ~35,000-year-old individual from northwest Europe represents an early branch of this founder population which was then displaced across a broad region, before reappearing in southwest Europe at the height of the last Ice Age ~19,000 years ago. During the major warming period after ~14,000 years ago, a genetic component related to present-day Near Easterners became widespread in Europe. These results document how population turnover and migration have been recurring themes of European prehistory.**

Modern humans arrived in Europe around 45,000 years ago and have lived there ever since, even during the Last Glacial Maximum 25,000–19,000 years ago when large parts of Europe were covered in ice<sup>1</sup>. A major question is how climatic fluctuations influenced the population history of Europe and to what extent changes in material cultures documented by archaeology corresponded to movements of people. To date, it has been difficult to address this question because genome-wide ancient DNA has been retrieved from just four Upper Palaeolithic

individuals from Europe<sup>2–4</sup>. Here we assemble and analyse genome-wide data from 51 modern humans dating from 45,000 to 7,000 years ago (Extended Data Table 1; Supplementary Information section 1).

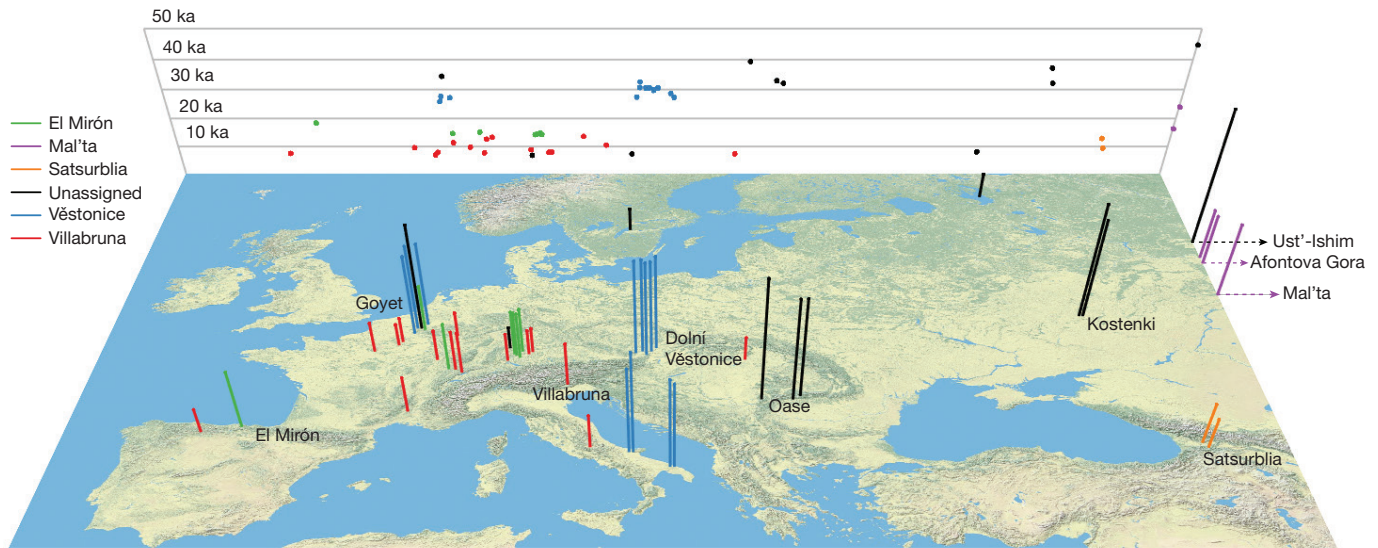
## Ancient DNA retrieval

We extracted DNA from human remains in dedicated clean rooms<sup>5</sup>, and transformed the extracts into Illumina sequencing libraries<sup>6–8</sup>. A major challenge in ancient DNA research is that the vast majority

<sup>1</sup>Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, IVPP, CAS, Beijing 100044, China. <sup>2</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>3</sup>Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>4</sup>Institute for Archaeological Sciences, Archaeology and Palaeogenetics, University of Tübingen, 72070 Tübingen, Germany. <sup>5</sup>Department of Archaeogenetics, Max Planck Institute for the Science of Human History, 07745 Jena, Germany. <sup>6</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. <sup>7</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>8</sup>School of Archaeology and Earth Institute, University College Dublin, Belfield, Dublin 4, Ireland. <sup>9</sup>CIAS, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. <sup>10</sup>Australian Centre for Ancient DNA, School of Biological Sciences, The University of Adelaide, SA-5005 Adelaide, Australia. <sup>11</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. <sup>12</sup>Institute of Archaeology and Ethnography, Russian Academy of Sciences, Siberian Branch, 17 Novosibirsk, RU-630090, Russia. <sup>13</sup>Altai State University, Barnaul, RU-656049, Russia. <sup>14</sup>Dipartimento di Civiltà e Forme del Sapere, Università di Pisa, 56126 Pisa, Italy. <sup>15</sup>Department of Biology, University of Pisa, 56126 Pisa, Italy. <sup>16</sup>Direction régionale des affaires culturelles Rhône-Alpes, 69283 Lyon, Cedex 01, France. <sup>17</sup>Dipartimento di Biologia, Università degli Studi di Bari 'Aldo Moro', 70125 Bari, Italy. <sup>18</sup>Instituto Internacional de Investigaciones Prehistóricas, Universidad de Cantabria, 39005 Santander, Spain. <sup>19</sup>Department of Anthropology, MSC01 1040, University of New Mexico, Albuquerque, New Mexico 87131-0001, USA. <sup>20</sup>Quaternary Archaeology, Institute for Oriental and European Archaeology, Austrian Academy of Sciences, 1010 Vienna, Austria. <sup>21</sup>Department of Anthropology, Natural History Museum Vienna, 1010 Vienna, Austria. <sup>22</sup>Department of Anthropology, University of Vienna, 1090 Vienna, Austria. <sup>23</sup>“Emil Racoviță” Institute of Speleology, 010986 Bucharest 12, Romania. <sup>24</sup>“Emil Racoviță” Institute of Speleology, Cluj Branch, 400006 Cluj, Romania. <sup>25</sup>Department of Cultural Heritage, University of Bologna, 48121 Ravenna, Italy. <sup>26</sup>Sezione di Scienze Preistoriche e Antropologiche, Dipartimento di Studi Umanistici, Università di Ferrara, 44100 Ferrara, Italy. <sup>27</sup>Università degli Studi di Bari 'Aldo Moro', 70125 Bari, Italy. <sup>28</sup>Museo di “Civiltà preclassiche della Murgia meridionale”, 72017 Ostuni, Italy. <sup>29</sup>Dipartimento di Biologia, Università di Firenze, 50122 Florence, Italy. <sup>30</sup>Dipartimento di Scienze Fisiche, della Terra e dell'Ambiente, U.R. Preistoria e Antropologia, Università degli Studi di Siena, 53100 Siena, Italy. <sup>31</sup>CNRS/UMR 7041 ArScAn MAE, 92023 Nanterre, France. <sup>32</sup>INRAP/UMR 8215 Trajectoires 21, 92023 Nanterre, France. <sup>33</sup>Ulmer Museum, 89073 Ulm, Germany. <sup>34</sup>University of Bucharest, Faculty of Geology and Geophysics, Department of Geology, 01041 Bucharest, Romania. <sup>35</sup>Department of Anthropology, California State University Northridge, Northridge, California 91330-8244, USA. <sup>36</sup>Université de Bordeaux, CNRS, UMR 5199-PACEA, 33615 Pessac Cedex, France. <sup>37</sup>TRACES – UMR 5608, Université Toulouse Jean Jaurès, Maison de la Recherche, 31058 Toulouse Cedex 9, France. <sup>38</sup>Royal Belgian Institute of Natural Sciences, 1000 Brussels, Belgium. <sup>39</sup>Department of Archaeology, School of Culture and Society, Aarhus University, 8270 Højbjerg, Denmark. <sup>40</sup>Service Régional d'Archéologie de Franche-Comté, 25043 Besançon Cedex, France. <sup>41</sup>Laboratoire Chronoenvironnement, UMR 6249 du CNRS, UFR des Sciences et Techniques, 25030 Besançon Cedex, France. <sup>42</sup>Department of Geosciences, Biogeology, University of Tübingen, 72074 Tübingen, Germany. <sup>43</sup>Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, 72072 Tübingen, Germany. <sup>44</sup>Department of Early Prehistory and Quaternary Ecology, University of Tübingen, 72070 Tübingen, Germany. <sup>45</sup>Institute for Archaeological Sciences, Paleoanthropology, University of Tübingen, 72070 Tübingen, Germany. <sup>46</sup>Museum of Anthropology and Ethnography, Saint Petersburg 34, Russia. <sup>47</sup>Department of Anthropology, Faculty of Science, Masaryk University, 611 37 Brno, Czech Republic. <sup>48</sup>Institute of Archaeology at Brno, Academy of Science of the Czech Republic, 69129 Dolní Věstonice, Czech Republic. <sup>49</sup>Department of Archaeology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada.

\*These authors contributed equally to this work.

§These authors jointly supervised this work.



**Figure 1 | Location and age of the 51 ancient modern humans.** Each bar corresponds to an individual, the colour code designates the genetically defined cluster of individuals, and the height is proportional to age (the background grid shows a projection of longitude against age). To help in

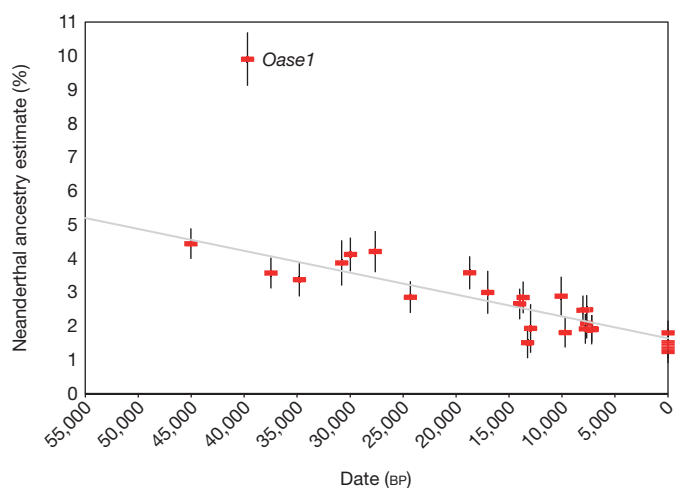
visualization, we add jitter for sites with multiple individuals from nearby locations. Four individuals from Siberia are plotted at the far eastern edge of the map. ka, thousand years ago.

of the DNA extracted from most specimens is of microbial origin, making random shotgun sequencing prohibitively expensive. We addressed this problem by enriching the libraries for between 390,000 and 3.7 million single nucleotide polymorphisms (SNPs) in the nuclear genome via hybridizing to pools of previously synthesized 52-base-pair oligonucleotide probes targeting these positions. This makes it possible to generate genome-wide data from samples with high percentages of microbial DNA that are not practical to study by shotgun sequencing<sup>3,9</sup>. We sequenced the isolated DNA fragments from both ends, and mapped the consensus sequences to the human genome (hg19), retaining fragments that overlapped the targeted SNPs. After removing fragments with identical start and end positions to eliminate duplicates produced during library amplification, we chose one fragment at random to represent each individual at each SNP.

Contamination from present-day human DNA is a danger in ancient DNA research. To address this, we took advantage of three characteristic features of ancient DNA (Supplementary Information section 2). First, for an uncontaminated specimen, we expect only a single mitochondrial DNA sequence to be present, allowing us to detect contamination as a mixture of mitochondrial sequences. Second, because males carry a single X chromosome, we can detect contamination in male specimens as polymorphisms on chromosome X<sup>10</sup>. Third, cytosines at the ends of genuine ancient DNA molecules are often deaminated, resulting in apparent cytosine to thymine substitutions<sup>11</sup>, and thus we can filter out contaminating molecules by restricting analysis to those with evidence of such deamination<sup>12</sup>. For libraries from males with evidence of mitochondrial DNA contamination or X chromosomal contamination estimates  $>2.5\%$ —as well as for all libraries from females—we restricted the analyses to sequences with evidence of cytosine deamination (Supplementary Information section 2). After merging libraries from the same individual and limiting to individuals with  $>4,000$  targeted SNPs covered at least once, 38 individuals remained, which we merged with newly generated shotgun sequencing data from the *Karelia* individual<sup>9</sup> (2.0-fold coverage), and published data from ancient<sup>2–4,7,13–19</sup> and present-day humans<sup>20</sup>. The final data set includes 51 ancient modern humans, of which 16 had at least 790,000 SNPs covered (Fig. 1; Extended Data Table 1).

### Natural selection reduced Neanderthal ancestry over time

We used two previously published statistics<sup>3,7,21</sup> to test if the proportion of Neanderthal ancestry in Eurasians changed over the last 45,000 years. Whereas on the order of 2% of present-day Eurasian DNA is of Neanderthal origin (Extended Data Table 2), the ancient modern human genomes carry significantly more Neanderthal DNA (Fig. 2) ( $P \ll 10^{-12}$ ). Using one statistic, we estimate a decline from 4.3–5.7% from a time shortly after introgression to 1.1–2.2% in Eurasians today (Fig. 2). Using the other statistic, we estimate a decline from 3.2–4.2% to 1.8–2.3% (Extended Data Fig. 1 and Extended Data Table 3). Because all of the European individuals we analysed dating to between 37,000 and 14,000 years ago are consistent with descent from a single founding



**Figure 2 | Decrease of Neanderthal ancestry over time.** Plot of radiocarbon date against Neanderthal ancestry for individuals with at least 200,000 SNPs covered, along with present-day Eurasians (standard errors are from a block jackknife). The least squares fit (grey) excludes the data from *Oase1* (an outlier with recent Neanderthal ancestry) and three present-day European populations (known to have less Neanderthal ancestry than east Asians). The slope is significantly negative for all eleven subsets of individuals we analysed ( $10^{-29} < P < 10^{-11}$  based on a block jackknife) (Extended Data Table 3). BP, before present.

population, admixture with populations with lower Neanderthal ancestry cannot explain the steady decrease in Neanderthal-derived DNA that we detect during this period, showing that natural selection against Neanderthal DNA must have driven this phenomenon (Fig. 2). We also obtained an independent line of evidence for selection from our observation that the decrease in Neanderthal-derived alleles is more marked near genes than in less constrained regions of the genome ( $P=0.010$ ) (Extended Data Table 3; Supplementary Information section 3)<sup>22–25</sup>.

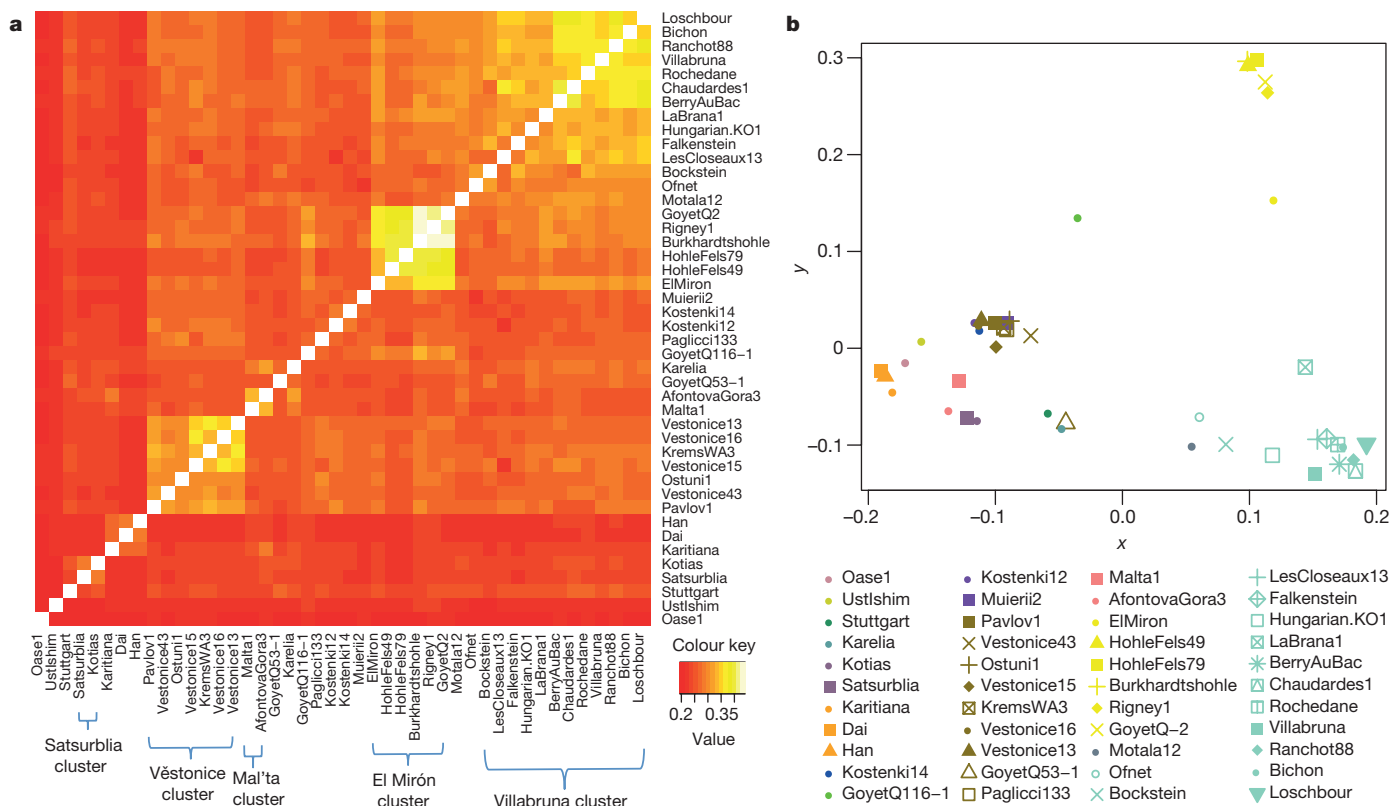
**Chromosome Y, mtDNA, and significant mutations**

We used the proportion of sequences mapping to the Y chromosome to infer sex (Extended Data Table 4; Supplementary Information section 4), and determined Y chromosome haplogroups for the males. We were surprised to find haplogroup R1b in the ~14,000-year-old *Villabruna* individual from Italy. While the predominance of R1b in western Europe today owes its origin to Bronze Age migrations from the eastern European steppe<sup>9</sup>, its presence in *Villabruna* and in a ~7,000-year-old farmer from Iberia<sup>9</sup> documents a deeper history of this haplotype in more western parts of Europe. Additional evidence of an early link between West and East comes from the *HERC2* locus, where a derived allele that is the primary driver of light eye colour in Europeans appears nearly simultaneously in specimens from Italy and the Caucasus ~14,000–13,000 years ago. Extended Data Table 5 presents results for additional alleles of biological importance. When analysing the mitochondrial genomes we noted the presence of haplogroup M in a ~27,000-year-old individual from southern Italy (*Ostuni1*) in agreement with the observation that this haplogroup, which today occurs in Asia and is absent in Europe, was

present in pre-Last Glacial Maximum Europe and was subsequently lost<sup>26</sup>. We also find that the ~33,000-year-old *Muierii2* from Romania carries a basal version of haplogroup U6, in agreement with the hypothesis that the presence of derived versions of this haplogroup in North Africans today is due to back-migration from western Eurasia<sup>27</sup>.

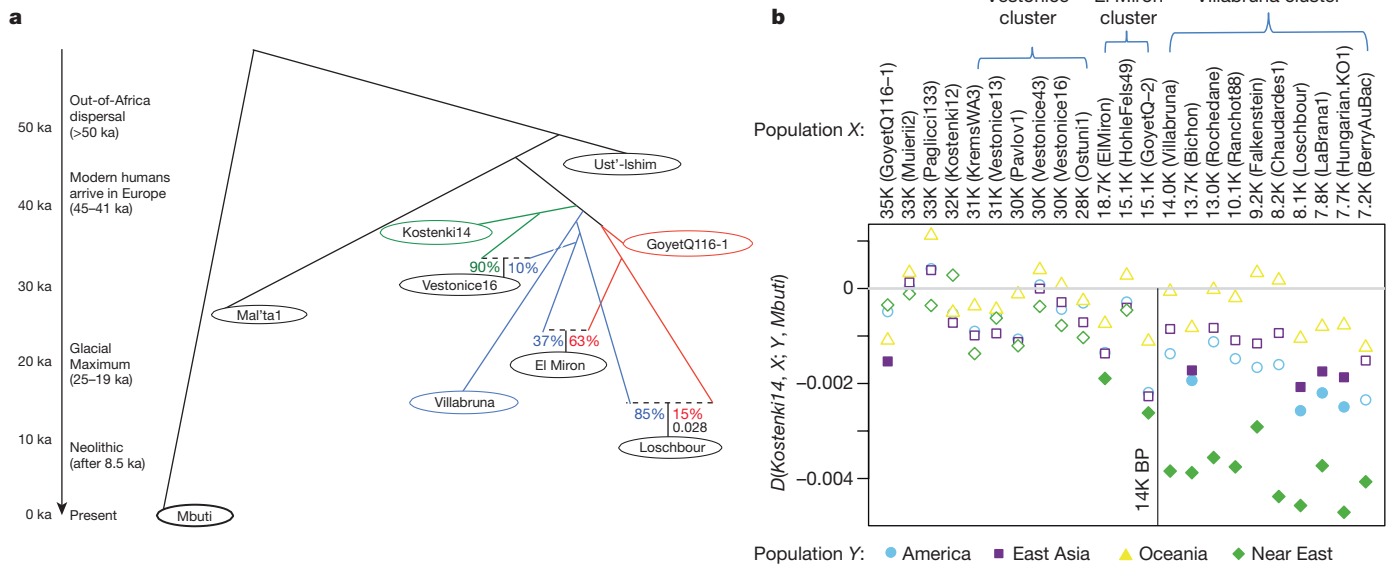
**Genetic clustering of the ancient specimens**

This data set provides an unprecedented opportunity to study the population history of Upper Palaeolithic Europe over more than 30,000 years. In order not to prejudice any association between genetic and archaeological groupings among the individuals studied, we first allowed the genetic data alone to drive the groupings of the specimens, and only afterward examined their associations with archaeological cultural complexes. We began by computing  $f_3$ -statistics<sup>14</sup> of the form  $f_3(X, Y; Mbuti)$ , which measure shared genetic drift between a pair of ancient individuals after divergence from an outgroup (here *Mbuti* from sub-Saharan Africa) (Fig. 3a and Extended Data Fig. 2). Through multi-dimensional scaling (MDS) analysis of this matrix (Fig. 3b), as well as through *D*-statistic analyses<sup>28</sup> (Supplementary Information section 5), we identify five clusters of individuals who share substantial amounts of genetic drift. We name these clusters after the oldest individual in each cluster with >1.0-fold coverage (Supplementary Information section 5; Extended Data Table 1). In contrast, we were not able to identify clear structure among the individuals studied based on model-based clustering<sup>29,30</sup>, which may reflect the fact that many of the individuals are so ancient that present-day human variation is not very relevant to understanding their patterns of genetic differentiation<sup>4,13</sup>. The ‘Vestonice Cluster’ is composed of 14 pre-Last



**Figure 3 | Genetic clustering of the ancient modern humans.** a, Shared genetic drift measured by  $f_3(X, Y; Mbuti)$  among individuals with at least 30,000 SNPs covered (for *AfontovaGora3*, *ElMiron*, *Falkenstein*, *GoyetQ-2*, *GoyetQ53-1*, *HohleFels49*, *HohleFels79*, *LesCloseaux13*, *Ofnet*, *Ranchot88* and *Rigney1*, we use all sequences for higher resolution). Lighter colours indicate more shared drift. b, Multi-dimensional scaling (MDS) analysis,

computed using the R software *cmdscale* package, highlights the main genetic groupings analysed in this study: Vestonice Cluster (brown), Mal'ta Cluster (pink), El Mirón Cluster (yellow), Villabruna Cluster (light green), and Satsurbliya Cluster (dark purple). The affinity of *GoyetQ116-1* (dark green) to the El Mirón Cluster is evident in both views of the data.



**Figure 4 | Population history inferences.** **a**, Admixture graph relating selected high coverage individuals. Dashed lines show inferred admixture events; the estimated mixture proportions fitted using the ADMIXTUREGRAPH software are labelled<sup>28</sup> (the estimated genetic drift on each branch is given in a version of this graph shown in Supplementary Information section 6). The individuals are positioned vertically based on their radiocarbon date, but we caution that the population split times are not accurately known. Colour is used to highlight important early branches of the European founder population: the *Kostenki14* lineage is modelled as the predominant contributor to the Věstonice Cluster

Glacial Maximum individuals from 34,000–26,000 years ago, who are all associated with the archaeologically defined Gravettian culture. The ‘Mal’ta Cluster’ is composed of three individuals who lived between 24,000–17,000 years ago from the Lake Baikal region of Siberia. The ‘El Mirón Cluster’ is composed of seven post-Last Glacial Maximum individuals from 19,000–14,000 years ago, who are all associated with the Magdalenian culture. The ‘Villabruna Cluster’ is composed of 15 post-Last Glacial Maximum individuals from 14,000–7,000 years ago, associated with the Azilian, Epipaleolithic and Mesolithic cultures. The ‘Satsurblia Cluster’ is composed of two individuals from 13,000–10,000 years ago from the southern Caucasus<sup>2</sup>. Ten individuals were not assigned to any cluster, either because they represented distinct early lineages (*Ust’-Ishim*, *Oase1*, *Kostenki14*, *GoyetQ116-1*, *Muierii2*, *Cioclovina1* and *Kostenki12*), because they were admixed between clusters (*Karelia* or *Motala12*), or because they were of very different ancestry (*Stuttgart*). To classify the ancestry of additional low coverage individuals, we built an admixture graph that fits the allele frequency correlation patterns among high-coverage individuals<sup>28</sup> (Fig. 4a; Supplementary Information section 6). We fit each low-coverage individual into the graph in turn, using all DNA fragments from these individuals, rather than just fragments with evidence of cytosine deamination, and account for contamination by modelling (Supplementary Information section 7).

### A founding population for Europeans 37–14 ka

A previous genetic analysis of early modern humans in Europe using data from the ~37,000-year-old *Kostenki14* suggested that the population to which *Kostenki14* belonged harboured within it the three major lineages that exist in mixed form in Europe today<sup>4,15</sup>: (1) a lineage related to all later pre-Neolithic Europeans, (2) a ‘Basal Eurasian’ lineage that split from the ancestors of Europeans and east Asians before they separated from each other; and (3) a lineage related to the ~24,000-year-old *Mal’ta1* from Siberia. With our more extensive sampling of Ice Age Europe, we find no support for this. When we test whether the ~45,000-year-old *Ust’-Ishim*—an early Eurasian without

(green); the *GoyetQ116-1* lineage as the predominant contributor to the El Mirón Cluster (red); and the *Villabruna* lineage as broadly represented across many clusters. **b**, Drawing together of European and Near Eastern populations ~14,000 years ago. Plot of affinity of each pre-Neolithic European population *X* to non-Africans outside Europe *Y* moving forward in time, comparing to *Kostenki14* as a baseline; values  $Z < -3$  standard errors below zero are indicated with filled symbols (we restricted to individuals with >50,000 SNPs). We observe an affinity to Near Easterners beginning with the Villabruna Cluster, and another to east Asians that affects a subset of the Villabruna Cluster.

any evidence of Basal Eurasian ancestry—shares more alleles with one test individual or another by computing statistics of the form  $D(Test_1, Test_2; Ust’-Ishim, Mbuti)$ , we find that the statistic is consistent with zero when the *Test* populations are any pre-Neolithic Europeans or present-day east Asians<sup>3,13</sup>. This would not be expected if some of the pre-Neolithic Europeans, including *Kostenki14*, had Basal Eurasian ancestry (Supplementary Information section 8). We also find no evidence for the suggestion that the *Mal’ta1* lineage contributed to Upper Palaeolithic Europeans<sup>4</sup>, because when we compute the statistic  $D(Test_1, Test_2; Mal’ta1, Mbuti)$ , we find that the statistic is indistinguishable from zero when the *Test* populations are any pre-Neolithic Europeans beginning with *Kostenki14*, consistent with descent from a single founder population since separation from the lineage leading to *Mal’ta1* (Supplementary Information section 9). A corollary of this finding is that the widespread presence of *Mal’ta1*-related ancestry in present-day Europeans<sup>15</sup> is probably explained by migrations from the Eurasian steppe in the Neolithic and Bronze Age periods<sup>9</sup>.

### Resurfacing of a European lineage in the Glacial Maximum

Among the newly reported individuals, *GoyetQ116-1* from present-day Belgium is the oldest at ~35,000 years ago. This individual is similar to the ~37,000-year-old *Kostenki14* and all later individuals in that it shares more alleles with present-day Europeans (for example, *French*) than with east Asians (for example, *Han*). In contrast, *Ust’-Ishim* and *Oase1*, which predate *GoyetQ116-1* and *Kostenki14*, do not show any distinctive affinity to later Europeans (Extended Data Table 6). Thus, from about 37,000 years ago, populations in Europe shared at least some ancestry with present Europeans. However, *GoyetQ116-1* differs from *Kostenki14* and from all individuals of the succeeding Věstonice Cluster in that both  $f_3$ -statistics (Fig. 3; Extended Data Fig. 2) and  $D$ -statistics show that it shares more alleles with members of the El Mirón Cluster who lived 19,000–14,000 years ago than with other pre-Neolithic Europeans (Supplementary Information section 10). Thus, *GoyetQ116-1* has an affinity to individuals who lived more than 15,000 years later. While at least half of the ancestry of all

El Mirón Cluster individuals comes from the lineage represented by *GoyetQ116-1*, this proportion varies among individuals with the largest amount found outside Iberia ( $Z = -4.8$ ) (Supplementary Information section 10).

### Europe and the Near East drew together around 14 ka

Beginning around 14,000 years ago with the Villabruna Cluster, the strong affinity to *GoyetQ116-1* seen in El Mirón Cluster individuals who belong to the Late Glacial Magdalenian culture becomes greatly attenuated (Supplementary Information section 10). To test if this change might reflect gene flow from populations that did not descend from the >37,000-year-old European founder population, we computed statistics of the form  $D(\text{Early European}, \text{Later European}; Y, \text{Mbuti})$  where  $Y$  are various present-day non-Africans. If no gene flow from exogenous populations occurred, this statistic is expected to be zero. Figure 4b shows that it is consistent with zero ( $|Z| < 3$ ) for nearly all individuals dating to between about 37,000 and 14,000 years ago. However, beginning with the Villabruna Cluster, it becomes highly significantly negative in comparisons where the non-European population ( $Y$ ) is Near Easterners (Fig. 4b; Extended Data Fig. 3; Supplementary Information section 11). This must reflect a contribution to the Villabruna Cluster from a lineage also found in present-day Near Easterners (Fig. 4b).

The Satsurblia Cluster individuals from the Caucasus dating to ~13,000–10,000 years ago<sup>2</sup> share more alleles with the Villabruna Cluster individuals than they do with earlier Europeans, indicating that they are related to the population that contributed new alleles to people in the Villabruna Cluster, although they cannot be the direct source of the gene flow. One reason for this is that the Satsurblia Cluster carries large amounts of Basal Eurasian ancestry while Villabruna Cluster individuals do not<sup>2</sup> (Supplementary Information section 12; Extended Data Fig. 4). One possible explanation for the sudden drawing together of the ancestry of Europe and the Near East at this time is long-distance migrations from the Near East into Europe. However, a plausible alternative is population structure, whereby Upper Palaeolithic Europe harboured multiple groups that differed in their relationship to the Near East, with the balance shifting among groups as a result of demographic changes after the Glacial Maximum.

The Villabruna Cluster is represented by the largest number of individuals in this study. This allows us to study heterogeneity within this cluster (Supplementary Information section 13). First, we detect differences in the degree of allele sharing with members of the El Mirón Cluster, as revealed by significant statistics of the form  $D(\text{Test}_1, \text{Test}_2; \text{El Mirón Cluster}, \text{Mbuti})$ . Second, we detect an excess of allele sharing with east Asians in a subset of Villabruna Cluster individuals—beginning with an ~13,000-year-old individual from Switzerland—as revealed by significant statistics of the form  $D(\text{Test}_1, \text{Test}_2; \text{Han}, \text{Mbuti})$  (Fig. 4b and Extended Data Fig. 3). For example, *Han* Chinese share more alleles with two Villabruna Cluster individuals (*Loschbour* and *LaBranca1*) than they do with *Kostenki14*, as reflected in significantly negative statistics of the form  $D(\text{Kostenki14}, \text{Loschbour/LaBranca1}; \text{Han}, \text{Mbuti})$ <sup>4</sup>. This statistic was originally interpreted as evidence of Basal Eurasian ancestry in *Kostenki14*. However, because this statistic is consistent with zero when *Han* is replaced with *Ust'-Ishim*, these findings cannot be driven by Basal Eurasian ancestry (as we discuss earlier), and must instead be driven by gene flow between populations related to east Asians and the ancestors of some Europeans (Supplementary Information section 8).

### Conclusions

We show that the population history of pre-Neolithic Europe was complex in several respects. First, at least some of the initial modern humans to appear in Eurasia, exemplified by *Ust'-Ishim* and *Oase1*, failed to contribute appreciably to the current European gene pool<sup>3,13</sup>. Only from around 37,000 years ago do all the European individuals analysed share ancestry with present-day Europeans. Second, from

the time of *Kostenki14* about 37,000 years ago until the time of the Villabruna Cluster about 14,000 years ago, all individuals seem to derive from a single ancestral population with no evidence of substantial genetic influx from elsewhere. It is interesting that during this time, the Mal'ta Cluster is not represented in any of the individuals we sampled from Europe. Thus, while individuals assigned to the Gravettian cultural complex in Europe are associated with the Věstonice Cluster, there is no genetic connection between them and the *Mal'ta1* individual in Siberia, despite the fact that Venus figurines are associated with both. This suggests that if this similarity is not a coincidence<sup>31</sup>, it reflects diffusion of ideas rather than movements of people. Third, we find that *GoyetQ116-1* derives from a different deep branch of the European founder population than the Věstonice Cluster which became predominant in many places in Europe between 34,000 and 26,000 years ago including at Goyet. *GoyetQ116-1* is chronologically associated with the Aurignacian cultural complex. Thus, the subsequent spread of the Věstonice Cluster shows that the diffusion of the Gravettian cultural complex was mediated at least in part by population movements. Fourth, the population represented by *GoyetQ116-1* did not disappear, as its descendants became widespread again after ~19,000 years ago in the El Mirón Cluster when we detect them in Iberia. The El Mirón Cluster is associated with the Magdalenian culture and may represent a post-Glacial Maximum expansion from southwestern European refugia<sup>32</sup>. Fifth, beginning with the Villabruna Cluster at least ~14,000 years ago, all European individuals analysed show an affinity to the Near East. This correlates in time to the Bølling-Allerød interstadial, the first significant warming period after the Glacial Maximum<sup>33</sup>. Archaeologically, it correlates with cultural transitions within the Epigravettian in southern Europe<sup>34</sup> and the Magdalenian-to-Azilian transition in western Europe<sup>35</sup>. Thus, the appearance of the Villabruna Cluster may reflect migrations or population shifts within Europe at the end of the Ice Age, an observation that is also consistent with the evidence of mitochondrial DNA turnover<sup>26,36</sup>. One scenario that could explain these patterns is a population expansion from southeastern European or west Asian refugia after the Glacial Maximum, drawing together the genetic ancestry of Europe and the Near East. Sixth, within the Villabruna Cluster, some, but not all, individuals have an affinity to east Asians. An important direction for future work will be to generate similar ancient DNA data from southeastern Europe and the Near East to arrive at a more complete picture of the Upper Palaeolithic population history of western Eurasia.

**Online Content** Methods, along with any additional 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.

**Received 18 December 2015; accepted 12 April 2016.**

**Published online 2 May 2016.**

- Gamble, C., Davies, W., Pettitt, P. & Richards, M. Climate change and evolving human diversity in Europe during the last glacial. *Philosoph. Trans. Royal Soc. B* **359**, 243–253 (2004).
- Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Commun.* **6**, 8912 (2015).
- Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
- Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
- Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
- Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols* **2010**, pdb.prot5448 (2010).
- Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370**, 20130624 (2015).
- Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
- Korneliusson, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).

11. Krause, J. *et al.* A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr. Biol.* **20**, 231–236 (2010).
12. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
13. Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
14. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
15. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
16. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
17. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nature Commun.* **5**, 5257 (2014).
18. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
19. Prüfer, K. *et al.* The complete genome sequence of a Neandertal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
20. Mallick, S. *et al.* The landscape of human genome diversity. *Nature* (in the press).
21. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
22. Sankararaman, S. *et al.* The genomic landscape of Neandertal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).
23. Vernet, B. & Akey, J. M. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* **343**, 1017–1021 (2014).
24. Harris, K. & Nielsen, R. The genetic cost of Neandertal introgression. *Genetics* <http://dx.doi.org/10.1534/genetics.116.186890> (2016).
25. Juric, I., Aeschbacher, S. & Coop, G. The strength of selection against Neandertal introgression. Preprint at <http://dx.doi.org/10.1101/030148> (2015).
26. Posth, C. *et al.* Pleistocene mitochondrial genomes suggest a single major dispersal of Non-Africans and a Late Glacial population turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).
27. Olivieri, A. *et al.* The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science* **314**, 1767–1770 (2006).
28. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
29. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
30. Skotte, L., Korneliussen, T. S. & Albrechtsen, A. Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693–702 (2013).
31. Conard, N. J. A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany. *Nature* **459**, 248–252 (2009).
32. Straus, L. G. After the deep freeze: confronting “Magdalenian” realities in Cantabrian Spain and beyond. *J. Archaeol. Method Theory* **20**, 236–255 (2013).
33. Weaver, A. J., Saenko, O. A., Clark, P. U. & Mitrovica, J. X. Meltwater pulse 1A from Antarctica as a trigger of the Bölling-Allerød warm interval. *Science* **299**, 1709–1713 (2003).
34. Montoya, C. & Peresani, M. Nouveaux éléments de diachronie dans l'Épigravettien récent des Préalpes de la Vénétie in *D'un monde à l'autre. Les systèmes lithiques pendant le Tardiglaciaire autour de la Méditerranée nord-occidentale* (eds Bracco, J.-P. & Montoya, C.) 123–138 (Mémoire Société Préhistorique Française, 2005).
35. Valentin, B. *Jalons pour une paléohistoire des derniers chasseurs (XIVe-VIe millénaires av. J.-C.)* (Cahiers archéologiques de Paris 1 – 1, Publications de la Sorbonne, 2008).
36. Pala, M. *et al.* Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am. J. Hum. Genet.* **90**, 915–924 (2012).

**Supplementary Information** is available in the online version of the paper.

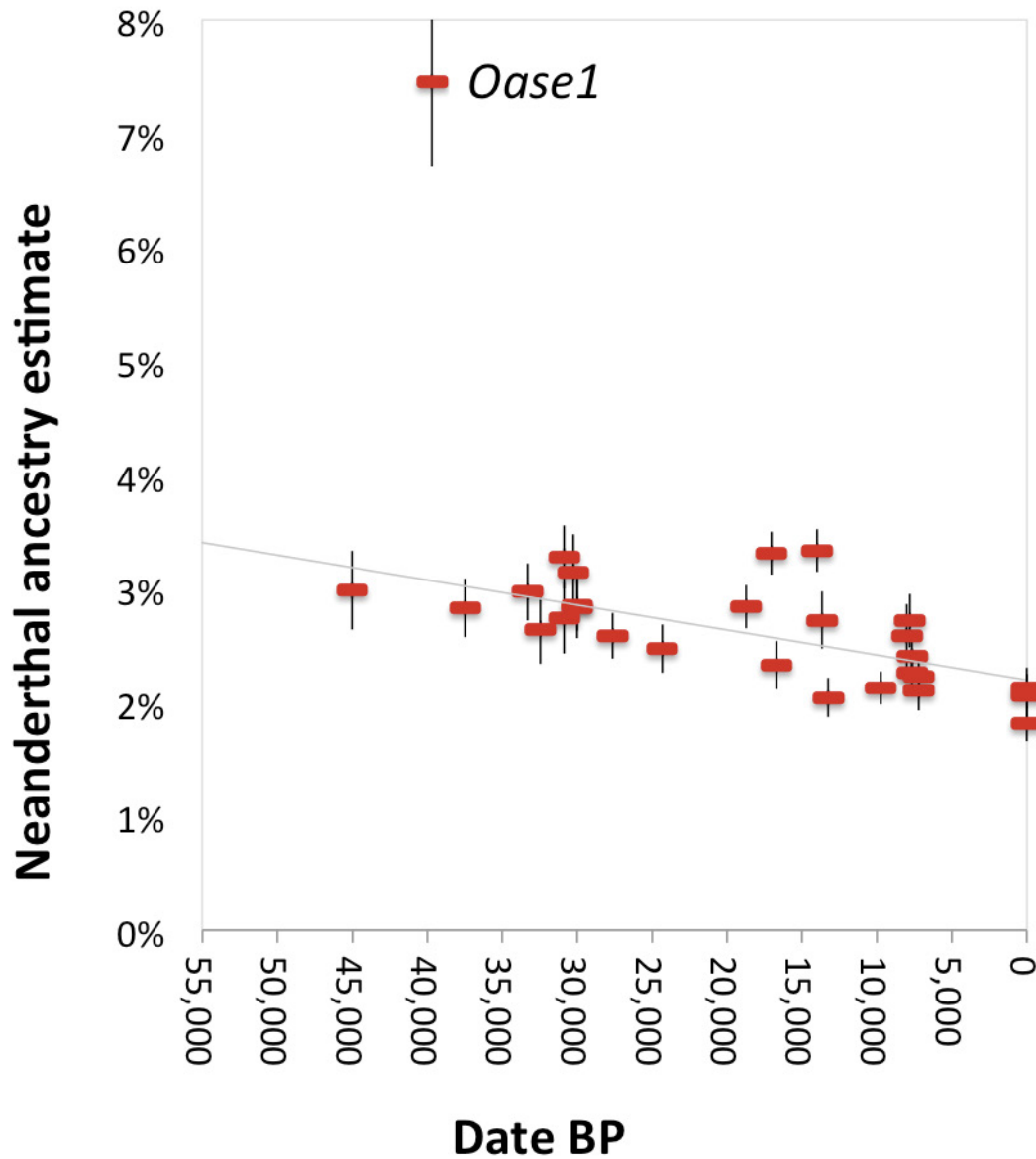
**Acknowledgements** We thank B. Alex, D. Meltzer, P. Moorjani, I. Olalde, S. Sankararaman and B. Viola for comments, K. Stewardson and E. Harney for

sample screening, and F. Hallgren for sharing a radiocarbon date for *Motala12*. The Fig. 1 map is plotted using data available under the Open Database License © OpenStreetMap (<http://www.openstreetmap.org/copyright>). The Goyet project led by H.R. was funded by the Wenner-Gren Foundation (Gr. 7837), the College of Social and Behavioral Sciences of CSUN, the CSUN Competition for Research, Scholarship and Creative Activity Awards, and the RBINS. The excavation of the El Mirón Cave burial, led by L.G.S. and M.R.G.M., was supported by the Gobierno de Cantabria, the L.S.B. Leakey Foundation, the University of New Mexico, the Stone Age Research Fund (J. and R. Auel, principal donors), the town of Ramales de la Victoria and the Universidad de Cantabria. Excavations at Grotta Paglicci were performed by A. Palma di Cesnola in collaboration with the Soprintendenza Archeologia della Puglia (founded by MIUR and local Institutions). Research at Riparo Villabruna was supported by MIBACT and the Veneto Region. Q.F. was funded by the Special Foundation of the President of the Chinese Academy of Sciences (2015–2016), the Bureau of International Cooperation of the Chinese Academy of Sciences, the Chinese Academy of Sciences (XDA05130202), the National Natural Science Foundation of China (L1524016) and the Chinese Academy of Sciences Discipline Development Strategy Project (2015-DX-C-03). D.Fe was supported by an Irish Research Council grant (GOIPG/2013/36). I.M. was supported by a long-term fellowship from the Human Frontier Science Program LT001095/2014-L. P.Sk was supported by the Swedish Research Council (VR 2014-453). S.T. and M.P.R. were funded by the Max Planck Society. C.N.-M. was funded by FWF P-17258, P-19347, P-21660 and P-23612. S.C. and O.T.M. were funded by a ‘Karstives’ Grant PCCE 31/2010 (CNCS-UEFISCDI, Romania). A.P.D., N.D., V.Sla and N.D. were funded by the Russian Science Foundation (project No.14-50-00036). M.A.M. was funded by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Programme (grant number PIEF-GA-2008-219965). M.La and D.C. were funded by grants PRIN 2010-11 and 2010EL8TXP\_003. C.C. and the research about the French Jura sites of Rochedane, Rigney and Ranchot was funded by the Collective Research Program (PCR) (2005-2008). K.H. was supported by the European Research Council (ERC StG 283503) and the Deutsche Forschungsgemeinschaft (DFG INST37/706-1FUGG, DFG FOR2237). D.G.D. was funded by the European Social Fund and Ministry of Science, Research and Arts of Baden-Württemberg. R.P. was funded by ERC starting grant ADNABIOARC (263441). J.Ke was funded by a grant from the Deutsche Forschungsgemeinschaft (SFB1052, project A02). J.Kr was funded by DFG grant KR 4015/1-1, the Baden Württemberg Foundation, and the Max Planck Society. S.P. were funded by the Max Planck Society and the Krekeler Foundation. D.R. was funded by NSF HOMINID grant BCS-1032255, NIH (NIGMS) grant GM100233, and the Howard Hughes Medical Institute.

**Author Contributions** J.Kr, S.P. and D.R. conceived the idea for the study. Q.F., C.P., M.H., W.H., M.Me, V.Slo, R.G.C., A.P.D., N.D., V.Sla, A.T., F.M., B.G., E.V., M.R.G.M., L.G.S., C.N.-M., M.T.-N., S.C., O.T.M., S.B., M.Per, D.Co, M.La, S.R., A.R., F.V., C.T., K.W., D.G., H.R., I.C., D.FI, P.Se, M.A.M., C.C., H.B., N.J.C., K.H., V.M., D.G.D., J.S., D.Ca, R.P., J.Kr, S.P. and D.R. assembled archaeological material. Q.F., C.P., M.H., D.Fe, A.F., W.H., M.Me, A.M., B.N., N.R., V.Slo, S.T., H.B., D.G.D., M.P.R., R.P., J.Kr, S.P. and D.R. performed or supervised wet laboratory work. Q.F., C.P., M.H., M.Pet, S.M., A.P., I.L., M.Li, I.M., S.S., P.Sk, J.Ke, N.P. and D.R. analysed data. Q.F., C.P., M.H., M.Pet, J.Ke, S.P. and D.R. wrote the manuscript and supplements.

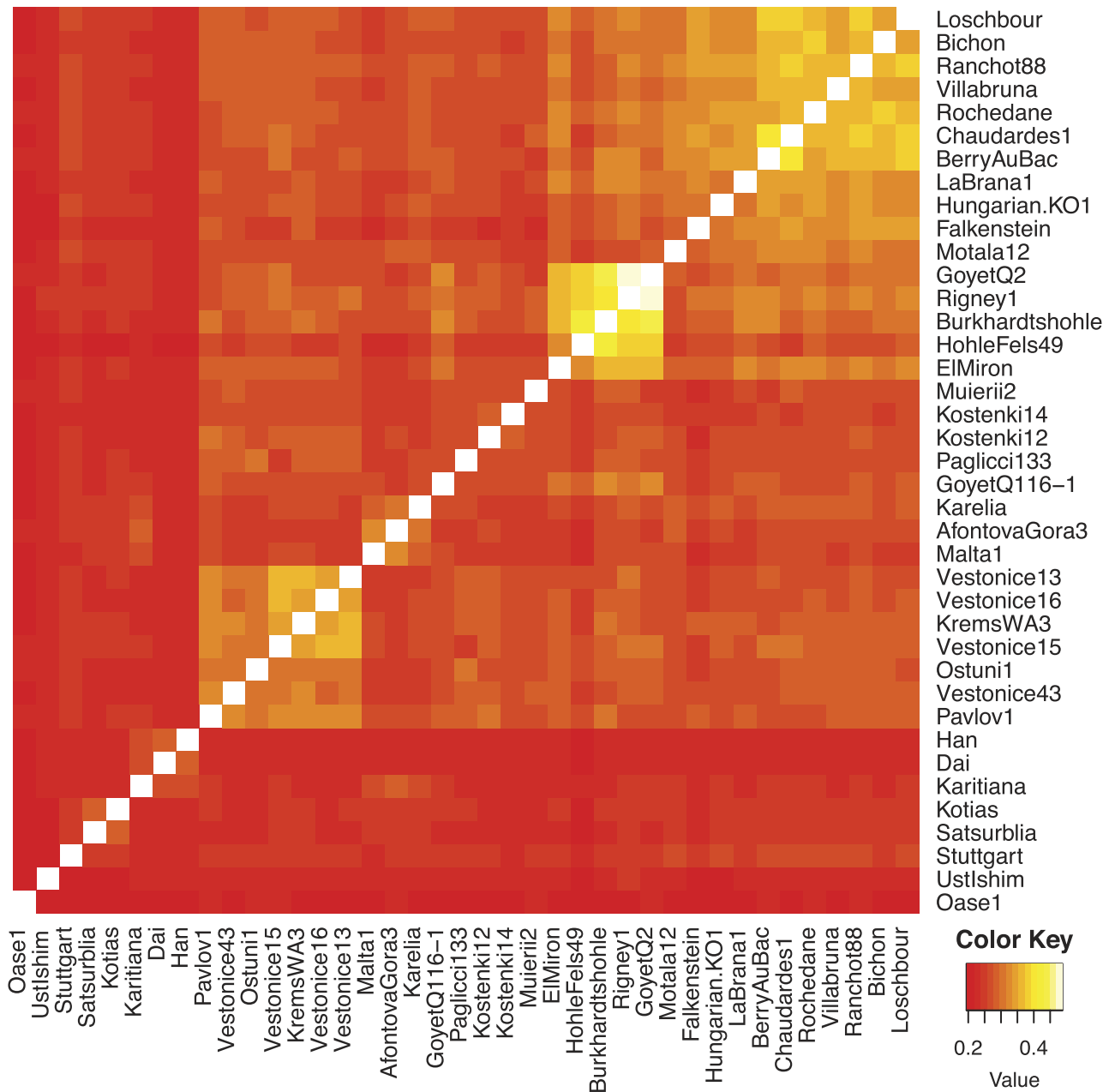
**Author Information** The aligned sequences are available through the European Nucleotide Archive under accession number PRJEB13123. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R. ([reich@genetics.med.harvard.edu](mailto:reich@genetics.med.harvard.edu)).

**Reviewer Information** *Nature* thanks C. Lalueza-Fox and the other anonymous reviewer(s) for their contribution to the peer review of this work.



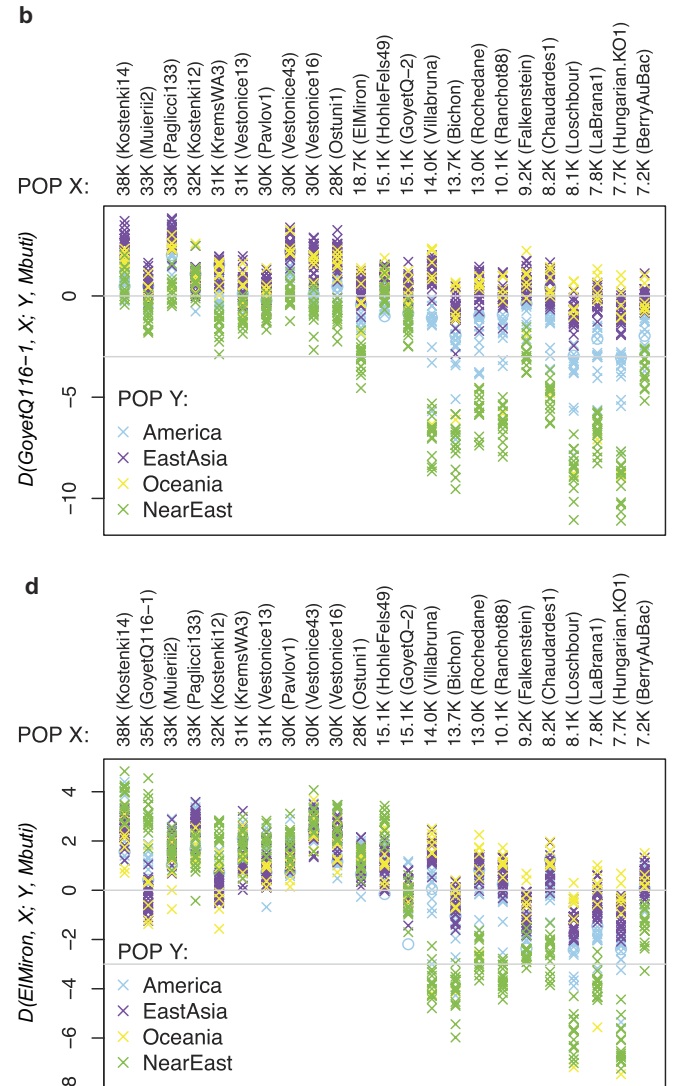
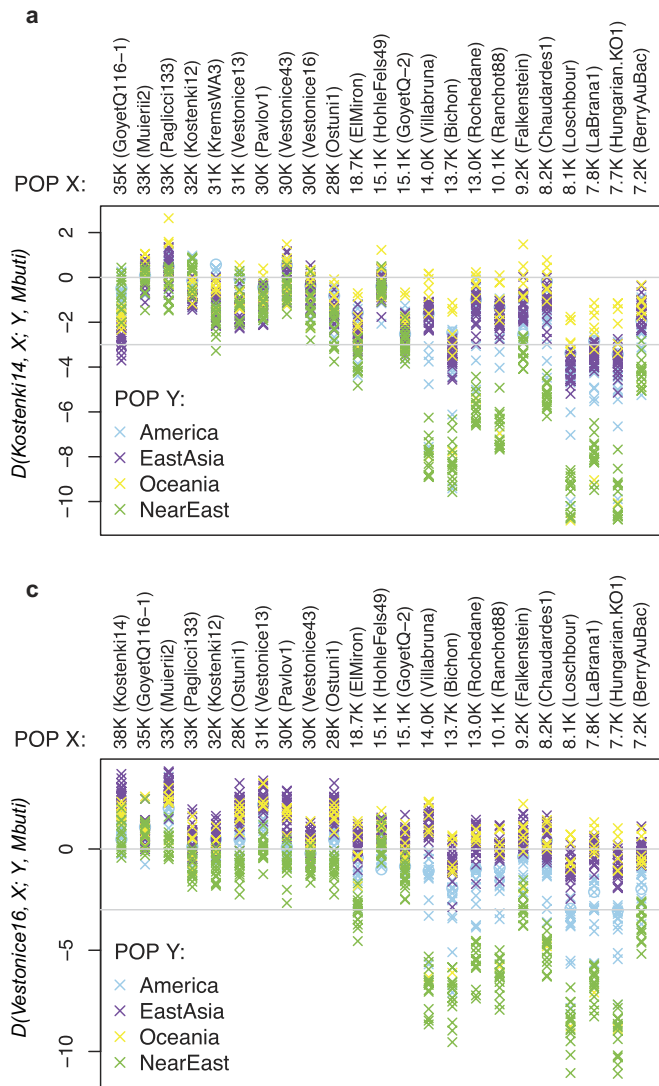
**Extended Data Figure 1 | A decrease in Neanderthal ancestry in the last 45,000 years.** This is similar to Fig. 2, except we use ancestry estimates from rates of alleles matching to Neanderthal rather than  $f_4$ -ratios, as described in Supplementary Information section 3. The least-squares

fit excludes *Oase1* (as an outlier with recent Neanderthal ancestry) and Europeans (known to have reduced Neanderthal ancestry). The regression slope is significantly negative ( $P = 0.00004$ , Extended Data Table 3).



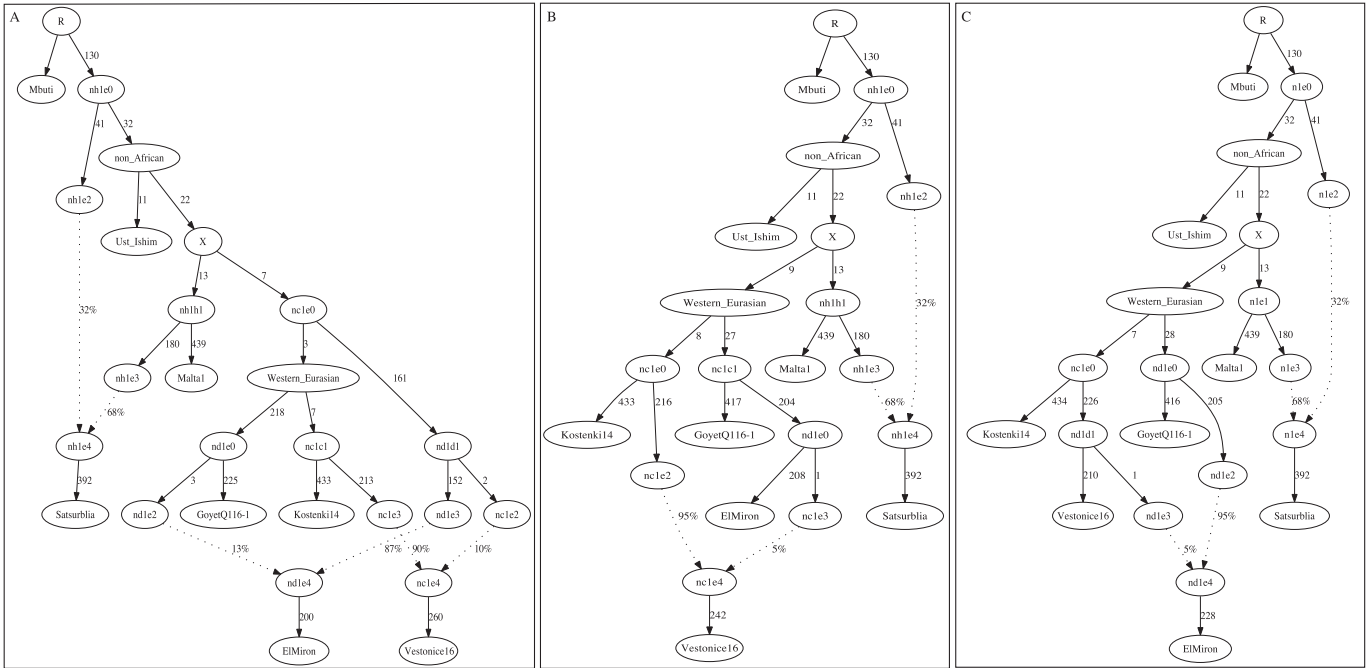
Extended Data Figure 2 | Heat matrix of pairwise  $f_3(X, Y; Mbuti)$  for selected ancient individuals. Only individuals with at least 30,000 SNPs covered at least once are analysed.





**Extended Data Figure 3 | Studying how the relatedness of non-European populations to pairs of European hunter-gatherers changes over time.** Statistics were examined of the form  $D(W, X; Y, \text{Mbuti})$ , with the Z-score given on the y axis, where  $W$  is an early European hunter-gatherer,  $X$  is another European hunter-gatherer (in chronological order on the x axis), and  $Y$  is a non-European population (see legend).

**a,  $W = \text{Kostenki14}$ . b,  $W = \text{GoyetQ116-1}$ . c,  $W = \text{Vestonice16}$ . d,  $W = \text{ElMiron}$ .**  $|Z| > 3$  scores are considered statistically significant (horizontal line). The similar Fig. 4b gives absolute  $D$ -statistic values rather than Z-scores (for  $W = \text{Kostenki14}$ ) and uses pooled regions rather than individual populations  $Y$ .



**Extended Data Figure 4 | Three admixture graph models that fit the data for Satsurbliia, an Upper Palaeolithic individual from the Caucasus.** These models use 127,057 SNPs covered in all populations. Estimated genetic drifts are given along the solid lines in units of  $f_2$ -distance (parts per thousand), and estimated mixture proportions are given along the dotted lines. All three models provide a fit to the

allele frequency correlation data among *Mbuti*, *Ust-Ishim*, *Kostenki14*, *Vestonice16*, *Malta1*, *ElMiron* and *Satsurbliia* to within the limits of our resolution, in the sense that all empirical  $f_2$ -,  $f_3$ - and  $f_4$ -statistics relating the individuals are within three standard errors of the expectation of the model. Models in which *Satsurbliia* is treated as unadmixed cannot be fit.



Extended Data Table 2 | Estimated proportion of Neanderthal ancestry

Sample Code	Age BP	$f_2$ -ratios				Archaic Ancestry Informative SNPs							
		SNPs	Est.	95% CI		SNPs	Est.	95% CI		Increase in Neanderthal ancestry with B	S.E.		
UstIshim	45,020	2,137,615	4.4%	3.6%	-	5.3%	778,774	3.0%	2.3%	-	3.7%	-0.9%	1.3%
Oase1	39,610	285,076	9.9%	8.4%	-	11.4%	59,854	7.5%	6.0%	-	8.9%	2.5%	1.8%
Kostenki14	37,470	1,774,156	3.6%	2.7%	-	4.4%	632,748	2.8%	2.3%	-	3.3%	-1.0%	1.0%
GoyetQ116-1	34,795	846,983	3.4%	2.4%	-	4.3%							
Muierii2	33,300	98,618	5.2%	3.0%	-	7.4%	22,189	3.0%	2.5%	-	3.5%	0.6%	1.1%
Paglicci133	32,895	82,330	4.1%	2.1%	-	6.0%							
Cioclovina1	32,435	12,784	4.1%	-1.1%	-	9.3%							
Kostenki12	32,415	61,228	1.9%	-0.7%	-	4.4%	13,385	2.6%	2.1%	-	3.2%	1.7%	1.5%
KremsWA3	30,970	203,986	3.9%	2.6%	-	5.2%							
Vestonice13	30,870	139,568	4.6%	2.6%	-	6.5%	35,983	3.3%	2.7%	-	3.8%	0.3%	1.3%
Vestonice15	30,870	30,900	4.3%	0.6%	-	7.9%	5,855	2.7%	2.1%	-	3.4%	-1.5%	1.3%
Vestonice14	30,870	5,677	2.6%	-5.9%	-	11.0%							
Pavlov1	30,260	57,005	4.4%	1.6%	-	7.1%	9,327	3.1%	2.5%	-	3.8%	0.7%	1.2%
Vestonice43	30,010	163,946	6.9%	5.2%	-	8.5%	38,749	2.9%	2.4%	-	3.3%	0.9%	0.9%
Vestonice16	30,010	945,292	4.1%	3.1%	-	5.1%	268,157	2.8%	2.3%	-	3.3%	-0.1%	1.0%
Ostuni2	28,975	17,017	1.6%	-3.2%	-	6.3%	2,746	2.3%	1.4%	-	3.1%	1.3%	1.6%
GoyetQ53-1	27,975	12,567	4.8%	-0.7%	-	10.3%							
Paglicci108	27,750	4,330	3.4%	-6.0%	-	12.7%							
Ostuni1	27,620	369,313	4.2%	3.0%	-	5.4%	88,449	2.6%	2.2%	-	3.0%	0.1%	0.9%
GoyetQ376-19	27,515	25,400	6.5%	2.7%	-	10.2%							
GoyetQ56-16	26,320	9,988	3.6%	-1.9%	-	9.1%							
Malta1	24,305	1,439,501	2.9%	1.9%	-	3.8%	437,187	2.5%	2.1%	-	2.9%	1.0%	0.8%
ElMiron	18,720	797,714	3.6%	2.6%	-	4.5%	250,071	2.8%	2.5%	-	3.2%	0.6%	0.9%
AfontovaGora3	16,710	286,355	3.0%	1.8%	-	4.2%	96,237	3.3%	2.9%	-	3.7%	-1.5%	1.0%
AfontovaGora2	16,710	143,751	2.2%	0.4%	-	4.0%	37,280	2.3%	1.9%	-	2.7%	-0.3%	0.9%
Rigney1	15,465	35,600	0.8%	-2.6%	-	4.2%							
HohleFels49	15,130	63,151	2.3%	-0.6%	-	5.2%							
GoyetQ-2	15,005	72,263	1.7%	-0.6%	-	4.0%							
Brillenhohle	14,780	13,459	2.5%	-3.0%	-	8.1%							
HohleFels79	14,670	11,211	1.7%	-5.1%	-	8.5%							
Burkhardtshohle	14,615	38,376	1.7%	-1.6%	-	5.0%							
Villabruna	13,980	1,215,433	2.7%	1.8%	-	3.5%	425,148	3.3%	3.0%	-	3.7%	1.1%	0.9%
Bichon	13,665	2,116,782	2.9%	1.9%	-	3.8%	769,422	2.7%	2.2%	-	3.2%	0.7%	1.3%
Satsurbli	13,255	1,460,368	1.5%	0.6%	-	2.4%	542,561	2.0%	1.7%	-	2.4%	0.9%	0.6%
Rochedane	12,960	237,390	1.9%	0.5%	-	3.3%							
Ibousseries39	11,725	9,659	6.4%	-0.8%	-	13.7%							
Continenza	10,855	11,717	4.1%	-1.4%	-	9.6%	1,733	2.9%	1.8%	-	4.0%	-10.6%	4.4%
Ranchot88	10,085	414,863	2.9%	1.8%	-	4.0%							
LesClosaux13	9,900	8,635	-	-9.7%	-	3.8%							
Kotias	9,720	2,133,968	1.8%	1.0%	-	2.7%	779,146	2.1%	1.8%	-	2.4%	0.7%	0.5%
Falkenstein	9,200	64,428	4.8%	1.7%	-	7.8%							
Karelia	8,375	1,754,410	1.9%	1.1%	-	2.7%	582,444	2.2%	1.9%	-	2.6%	-0.2%	0.7%
Bockstein	8,265	21,977	5.7%	1.0%	-	10.5%							
Ofnet	8,245	6,263	9.8%	1.4%	-	18.1%							
Chaudardes1	8,205	92,657	1.9%	-0.2%	-	3.9%							
Loschbour	8,050	2,091,584	2.5%	1.6%	-	3.3%	774,139	2.6%	2.0%	-	3.1%	2.7%	1.7%
LaBranal	7,815	1,884,745	1.9%	1.1%	-	2.8%	642,231	2.7%	2.3%	-	3.2%	0.4%	0.8%
Hungarian.KO1	7,660	1,410,303	2.1%	1.2%	-	3.0%	439,408	2.4%	2.0%	-	2.8%	-0.1%	1.2%
Motala12	7,625	1,874,519	2.5%	1.6%	-	3.3%	655,685	2.3%	1.9%	-	2.7%	-0.1%	0.7%
BerryAuBac	7,245	54,690	2.5%	-0.2%	-	5.1%							
Stuttgart	7,140	2,078,724	1.9%	1.1%	-	2.7%	767,813	2.1%	1.8%	-	2.5%	0.0%	0.7%
Dai	0	2,144,502	1.4%	0.7%	-	2.1%	782,066	1.8%	1.5%	-	2.1%	1.4%	0.4%
Han	0	2,144,502	1.8%	1.1%	-	2.5%	782,164	2.1%	1.8%	-	2.5%	1.9%	0.7%
English	0	2,144,502	1.5%	0.8%	-	2.2%							
French	0	2,144,502	1.5%	0.9%	-	2.1%	782,386	1.7%	1.4%	-	1.9%	1.4%	0.6%
Sardinian	0	2,144,502	1.2%	0.6%	-	1.9%	782,351	1.7%	1.4%	-	2.0%	0.7%	0.5%
Karitiana	0						782,037	2.1%	1.7%	-	2.4%	1.5%	1.0%

Extended Data Table 3 | Significant correlation of Neanderthal ancestry estimate with specimen age

Subset of samples	N	P-value for date correlation	Decrease in ancestry per 10,000 years	Estimate of Neanderthal ancestry at different time points			
				0 years ago (present)	50,000 years ago	55,000 years ago	60,000 years ago
<b><i>f<sub>i</sub></i>-ratio estimates</b>							
Core Set 1 (all ancient samples (except <i>Oase1</i> ) + <i>Han</i> + <i>Dai</i> )	57	$5 \times 10^{-22}$	0.48-0.73%	1.1-2.2%	4.0-5.4%	4.3-5.7%	4.5-6.0%
Subset of Core Set 1 (<32kya)	50	$2 \times 10^{-15}$	0.59-0.98%	0.9-2.1%	4.5-6.4%	4.8-6.9%	5.1-7.4%
Subset of Core Set 1 (>32kya or <25kya)	44	$4 \times 10^{-18}$	0.44-0.69%	1.0-2.2%	3.7-5.2%	4.0-5.5%	4.2-5.8%
Subset of Core Set 1 (>25kya or <14kya)	47	$5 \times 10^{-21}$	0.48-0.73%	1.0-2.2%	3.9-5.3%	4.2-5.7%	4.5-6.0%
Subset of Core Set 1 (>14kya or present day)	37	$2 \times 10^{-18}$	0.47-0.74%	1.1-2.4%	4.1-5.5%	4.3-5.8%	4.6-6.2%
Subset of Core Set 1 (only ancient samples)	50	$4 \times 10^{-15}$	0.46-0.76%	1.0-2.3%	4.0-5.4%	4.3-5.8%	4.5-6.1%
Subset of Core Set 1 (individuals with >200,000 SNPs)	28	$4 \times 10^{-19}$	0.46-0.71%	1.1-2.3%	3.9-5.3%	4.2-5.7%	4.4-6.0%
Modification of Core Set 1 (replace East Asians with Europeans)	58	$2 \times 10^{-23}$	0.49-0.73%	1.1-2.3%	4.0-5.4%	4.3-5.8%	4.6-6.1%
All ancient samples including <i>Oase1</i> + <i>Han</i> + <i>Dai</i>	58	$8 \times 10^{-29}$	0.57-0.81%	1.0-2.2%	4.3-5.7%	4.7-6.1%	5.0-6.5%
All ancient samples	51	$1 \times 10^{-20}$	0.57-0.86%	0.9-2.2%	4.4-5.8%	4.7-6.2%	5.0-6.6%
All ancient samples except <i>Oase1</i> or <i>UstIshim</i>	49	$8 \times 10^{-12}$	0.45-0.81%	1.0-2.3%	4.0-5.6%	4.2-6.0%	4.5-6.4%
<b>Ancestry informative SNPs</b>							
Core Set 2 (all ancient samples (except <i>Oase1</i> ) + <i>Han</i> + <i>Dai</i> + <i>Karitiana</i> )	29	$4 \times 10^{-11}$	0.21-0.39%	1.8-2.3%	3.1-4.0%	3.2-4.2%	3.3-4.3%
Subset of Core Set 2 (no <i>Han</i> , <i>Dai</i> , <i>Karitiana</i> , <i>Stuttgart</i> )	25	$1 \times 10^{-4}$	0.11-0.36%	1.8-2.5%	2.9-3.8%	3.0-4.0%	3.0-4.1%
Subset of Core Set 2 (no <i>Han</i> , <i>Dai</i> , <i>Karitiana</i> , <i>Stuttgart</i> , <i>UstIshim</i> )	24	$2 \times 10^{-4}$	0.11-0.37%	1.8-2.5%	2.9-3.8%	2.9-4.0%	3.0-4.2%

'Core set 1' used for the *f<sub>i</sub>*-ratio analyses, refers to 50 ancient individuals (removing *Oase1* as an outlier) along with 7 east Asians (*Dai* and *Han*). 'Core set 2' used for the analyses of Neanderthal ancestry informative SNPs, refers to 26 ancient individuals (removing *Oase1*, *Han*, *Dai* and *Karitiana*).

Extended Data Table 4 | Sex determination for newly reported individuals

Sample	Target	Type	N <sub>auto</sub>	N <sub>X</sub>	N <sub>Y</sub>	N <sub>X</sub> /N <sub>auto</sub>	N <sub>Y</sub> /N <sub>auto</sub>	X-rate	Y-rate	Sex
	1240k or 2.2M		1151240	49711	32681	0.0432	0.0284			
	390k		388745	1819	2242	0.0047	0.0058			
Kostenki14	2.2M	all	29633405	395534	262846	0.0133	0.0089	0.309	0.312	M
GoyetQ116-1	1240k	all	2122620	36391	22256	0.0171	0.0105	0.397	0.369	M
Cioclovina1	1240k	Damage	11521	184	125	0.0160	0.0108	0.370	0.382	M
Kostenki12	2.2M	Subset	63908	856	504	0.0134	0.0079	0.310	0.278	M
Muierii2	2.2M	Damage	81165	2177	8	0.0268	0.0001	0.621	0.003	F
Vestonice13	2.2M	Damage	119094	1578	1059	0.0133	0.0089	0.307	0.313	M
Vestonice15	2.2M	Damage	28762	338	227	0.0118	0.0079	0.272	0.278	M
Vestonice14	390k	Damage	4846	8	11	0.0017	0.0023	0.353	0.394	M
Vestonice43	2.2M	Damage	136933	1826	1204	0.0133	0.0088	0.309	0.310	M
Pavlov1	2.2M	Damage	54429	631	404	0.0116	0.0074	0.268	0.261	M
Vestonice16	2.2M	Subset	2433741	30463	20976	0.0125	0.0086	0.290	0.304	M
KremsWA3	1240k	all	235069	4119	2661	0.0175	0.0113	0.406	0.399	M
Ostuni2	2.2M	Damage	15749	138	1	0.0088	0.0001	0.203	0.002	F
Ostuni1	2.2M	Damage	427199	10868	47	0.0254	0.0001	0.589	0.004	F
Paglicci108	1240k	Damage	3883	124	2	0.0319	0.0005	0.740	0.018	F
GoyetQ53-1	1240k	Damage	10771	311	4	0.0289	0.0004	0.669	0.013	F
GoyetQ376-19	1240k	Damage	20052	680	10	0.0339	0.0005	0.785	0.018	F
GoyetQ56-16	1240k	Damage	8702	304	7	0.0349	0.0008	0.809	0.028	F
Paglicci133	1240k	Subset	81092	1641	983	0.0202	0.0121	0.469	0.427	M
ElMiron	2.2M	Damage	1765696	40647	196	0.0230	0.0001	0.533	0.004	F
HohleFels79	390k	Damage	10188	28	22	0.0027	0.0022	0.587	0.374	M
AfontovaGora3	2.2M	Damage	291798	8705	37	0.0298	0.0001	0.691	0.004	F
HohleFels49	390k	Damage	61051	113	111	0.0019	0.0018	0.396	0.315	M
Rigney1	1240k	Damage	32797	1131	9	0.0345	0.0003	0.799	0.010	F
GoyetQ-2	1240k	Damage	65563	1123	706	0.0171	0.0108	0.397	0.379	M
Brillenhohle	390k	Damage	12603	22	22	0.0017	0.0017	0.373	0.303	M
Burkhardtshohle	1240k	Damage	34207	563	407	0.0165	0.0119	0.381	0.419	M
Villabruna	2.2M	Subset	5505838	72055	52110	0.0131	0.0095	0.303	0.333	M
Rochedane	1240k	Subset	256325	4780	2830	0.0186	0.0110	0.432	0.389	M
Continenza	2.2M	Damage	10647	208	2	0.0195	0.0002	0.452	0.007	F
Ibousieres39	390k	Damage	8246	12	22	0.0015	0.0027	0.311	0.463	M
Ranchot88	1240k	Damage	594962	18520	119	0.0311	0.0002	0.721	0.007	F
LesCloseaux13	1240k	Damage	7326	275	2	0.0375	0.0003	0.869	0.010	F
Falkenstein	390k	Damage	58970	113	102	0.0019	0.0017	0.410	0.300	M
Bockstein	390k	Damage	20214	62	0	0.0031	0.0000	0.655	0.000	F
Ofnet	390k	Damage	5294	13	1	0.0025	0.0002	0.525	0.033	F
Chaudardes1	1240k	Damage	84052	1429	865	0.0170	0.0103	0.394	0.363	M
BerryAuBac	1240k	All	49670	902	554	0.0182	0.0112	0.421	0.393	M

\*We restrict analysis to the 1240k target set for study of the 2.2M capture datasets.

Y-rate is the ratio of  $N_Y/N_{\text{auto}}$  divided by the same quantity for the genome-wide target set. Female sex (F) is inferred as Y-rate < 0.05 and male sex (M) as Y-rate > 0.

Extended Data Table 5 | Allele counts at SNPs affected by selection in individuals with &gt;1-fold coverage

SNP		<i>LCT</i> rs4988235	<i>SLC45A2</i> rs16891982	<i>SLC24A5</i> rs1426654	<i>EDAR</i> rs3827760	<i>HERC2</i> rs12913832
Ancestral		G	C	G	A	A
Derived		A	G	A	G	G
UstIshim	Coverage	31	46	52	42	50
	Derived allele frequency	0%	0%	2%	0%	0%
Kostenki14	Coverage	140	113	6	45	52
	Derived allele frequency	0%	2%	17%	0%	0%
GoyetQ116-1	Coverage	8	6	0	9	1
	Derived allele frequency	0%	0%	n/a	0%	0%
Vestonice16	Coverage	13	18	0	4	5
	Derived allele frequency	0%	6%		0%	0%
Malta1	Coverage	1	0	2	2	2
	Derived allele frequency	0%		0%	0%	0%
ElMiron	Coverage	2	10	0	7	5
	Derived allele frequency	0%	0%		0%	0%
Villabruna	Coverage	17	52	5	19	10
	Derived allele frequency	0%	0%	0%	0%	100%
Bichon	Coverage	11	4	25	16	9
	Derived allele frequency	0%	0%	0%	0%	33%
Satsurblia	Coverage	1	2	4	1	4
	Derived allele frequency	0%	0%	100%	0%	50%
Kotias	Coverage	16	22	13	20	15
	Derived allele frequency	0%	0%	100%	0%	0%
Loschbour	Coverage	19	18	20	17	21
	Derived allele frequency	0%	0%	0%	0%	100%
LaBranca1	Coverage	8	6	2	11	3
	Derived allele frequency	12%	0%	0%	0%	100%
Hungarian.KO1	Coverage	1	2	2	1	2
	Derived allele frequency	0%	0%	50%	0%	100%
Motala12	Coverage	2	0	3	3	1
	Derived allele frequency	0%		0%	33%	100%
Karelia	Coverage	1	9	4	0	1
	Derived allele frequency	0%	67%	0%		0%
Stuttgart	Coverage	25	21	15	29	21
	Derived allele frequency	0%	0%	100%	0%	0%

rs4988235 is responsible for lactase persistence in Europe<sup>60,61</sup>. The SNPs at *SLC24A5* and *SLC45A2* are responsible for light skin pigmentation. The SNP at *EDAR*<sup>62,63</sup> affects tooth morphology and hair thickness. The SNP at *HERC2* (refs 64, 65) is the primary determinant of light eye colour in present-day Europeans. We present the fraction of fragments overlapping each SNP that are derived; the observation of a low rate of derived alleles does not prove that the individual carried the allele, and instead may reflect sequencing error or ancient DNA damage. Sites highlighted in light grey were judged (based on the derived allele count) likely to be heterozygous for the derived allele, and dark grey sites are likely to be homozygous.

60. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nature Genet.* **30**, 233–237 (2002).
61. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).
62. Fujimoto, A. *et al.* A scan for genetic determinants of human hair morphology: *EDAR* is associated with Asian hair thickness. *Hum. Mol. Genet.* **17**, 835–843 (2008).
63. Kimura, R. *et al.* A common variation in *EDAR* is a genetic determinant of shovel-shaped incisors. *Am. J. Hum. Genet.* **85**, 528–535 (2009).
64. Sturm, R. A. *et al.* A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *Am. J. Hum. Genet.* **82**, 424–431 (2008).
65. Eiberg, H. *et al.* Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the *HERC2* gene inhibiting *OCA2* expression. *Hum. Genet.* **123**, 177–187 (2008).

**Extended Data Table 6 | All European hunter-gatherers beginning with *Kostenki14* share genetic drift with present-day Europeans**

Test	SNPs used	D-value	Z score
Ust'-Ishim	2,050,358	0.003	6.6
Oase1	278,785	0.005	10.6
Kostenki14	1,676,253	-0.002	-5.5
Muierii2	95,787	-0.004	-6.3
GoyetQ116-1	811,756	-0.004	-8.0
Kostenki12	59,850	-0.004	-5.1
Paglicci133	79,624	-0.004	-5.5
Vestonice13	136,598	-0.004	-7.1
Vestonice15	30,252	-0.006	-6.4
Vestonice16	914,141	-0.004	-9.1
Pavlov1	55,835	-0.005	-6.3
Vestonice43	160,463	-0.004	-6.9
KremsWA3	229,187	-0.005	-10.2
Ostuni1	360,347	-0.004	-8.6
Malta1	1,401,718	-0.005	-11.3
ElMiron	777,654	-0.007	-14.7
AfontovaGora2	141,073	-0.007	-13.6
AfontovaGora3	707,617	-0.006	-13.6
HohleFels49	62,816	-0.004	-5.2
Rigney1	34,445	-0.006	-6.1
GoyetQ-2	70,210	-0.006	-8.8
Burkhardtshohle	37,234	-0.006	-6.2
Villabruna	1,170,777	-0.010	-24.7
Bichon	2,034,069	-0.009	-23.6
Satsurbliia	1,419,824	-0.005	-13.1
Rochedane	229,806	-0.011	-20.8
Ranchot88	402,274	-0.010	-21.8
Kotias	2,047,856	-0.006	-15.8
Falkenstein	64,043	-0.008	-11.6
Chaudardes1	90,047	-0.011	-16.0
Loschbour	2,037,082	-0.011	-25.4
LaBrana1	1,824,307	-0.009	-23.0
Motala12	1,816,201	-0.009	-23.8
Hungarian.KO1	1,372,801	-0.010	-26.5
Karelia	1,701,664	-0.009	-21.9
Stuttgart	2,023,939	-0.009	-23.9
BerryAuBac	53,028	-0.011	-14.0

The statistic  $D(Han, Test; French, Mbuti)$  was computed measuring whether present-day French share more alleles with *Han* or with a *Test* population (restricting to ancient individuals with at least 30,000 SNPs covered at least once). Present-day Europeans share significantly more genetic drift with European hunter-gatherers from *Kostenki14* onward than they do with *Han*. Thus, by the date of *Kostenki14*, there was already west Eurasian-specific genetic drift.