doi:10.1038/nature17993

Table of Contents 1

Section 1
Archaeological information

Radiocarbon date calibrations used in this study
We recalibrated all radiocarbon dates using IntCal13[1] in OxCal4.2[2]. We give 95.4% confidence intervals for calibrated dates in years before present (where the present is defined as 1950 CE). We round calibrated dates to the nearest decade.

Previously published genome-wide ancient DNA data
We do not provide archaeological context summaries for genome-wide ancient DNA data sets that have previously been published. However, we list the dates here:

- *UstIshim* at 47,480-42,560 cal BP (OxA-25516: 41,400 ±1300 ^{14}C; OxA-30190: 41,400 ±1400 ^{14}C)[3]
  (direct date, using collagen ultrafiltration)

- *Oase1* at 41,640-37,580 cal BP (GrA-22810: 34,290, +970, -870 ^{14}C; OxA-11711: >35,200 ^{14}C; Combined ^{14}C Age: OxA-GrA: 34,950 +900 -890 ^{14}C)[4].
  (direct date, using collagen ultrafiltration)

- *Malta1* at 24,520-24,090 cal BP (UCIAMS-79666: 20,240 ± 60 ^{14}C)[5]
  (direct date, using collagen ultrafiltration)

- *AfontovaGora2* at 16,940-16,480 cal BP (UCIAMS-79661: 13,810 ± 35 ^{14}C)[5]
  (direct date, using collagen ultrafiltration)

- *Bichon* at 13,770-13,560 cal BP (OxA-27763: 11,855 ± 50 ^{14}C)[6]
  (direct date, using collagen ultrafiltration)

- *Satsurblia* at 13,380-13,130 cal BP (OxA-34632: 11,415 ± 50 ^{14}C)[6]
  (direct date, using collagen ultrafiltration)

- *Kotias* at 9,890-9,550 cal BP (RTT-5246: 8,665 ± 65 ^{14}C; OxA-28256: 8,745 ± 40 ^{14}C)[6]
  (direct date, using collagen ultrafiltration)

- *Loschbour* at 8,160-7,940 cal BP (OxA-7738: 7,205 ± 50 ^{14}C)[7]
  (direct date, using collagen without ultrafiltration)

- *LaBrana1* at 7,940-7,690 cal BP (Beta-226472: 6,980 ± 50 ^{14}C)[8]
  (direct date, using collagen ultrafiltration)

- *Hungarian.KO1* at 7,730-7,590 cal BP (OxA-23757: 6835 ± 34 ^{14}C)[9]
  (direct date, using collagen ultrafiltration)

- *Motala12* at 7,670-7,580 cal BP (Ua-51723-16.8: 6,773 ± 30 ^{14}C)[7]
  (direct date, using collagen ultrafiltration – the present study is the first report of this date, courtesy of Fredrik Hallgren)

- *Karelia* at 8,800-7,950 cal BP (OxA-1665, OxA-2266, OxA-1667, OxA1668, OxA-1669, OxA-2124, OxA-2125, OxA-1973 ^{14}C)[10]
  We genetically analysed individual UZOO-74/I0061 (MAE RAS collection number 5773–74, grave number 142), collected from the Mesolithic site Yuzhnyy Oleni' Ostrov (*Oleneostrovski' mogilnik* or *Deer Island cemetery*), Onega Lake, Karelia, Russian Federation (61°30'N 35°45'E). Table S1.1 gives the dates we used.
  (layer date, not from collagen ultrafiltration, based on other individuals in the graveyard)[11]

Table S1.1. Layer dates used for the *Karelia* sample

Lab number	burial no.	^{14}C years BP	Cal BP 95.4% probability
OxA-1665	57	7,280 ± 80	8,220-7,950
OxA-2266	57	7,350 ± 90	8,360-8,000
OxA-1667	80	7,330 ± 90	8,340-7,980
OxA-1668	80	7,560 ± 90	8,540-8,190
OxA-1669	80	7,560 ± 90	8,540-8,190
OxA-2124	89	7,280 ± 90	8,310-7,950
OxA-2125	85	7,510 ± 90	8,480-8,160
OxA-1973	108	7,750 ± 110	8,800-8,360

- *Stuttgart* at 7,260-7,020 cal BP (MAMS-24635: 6,246 ± 30 ^{14}C)[7]
  (direct date, using collagen ultrafiltration – the present study is the first report of this date)

Peştera Cioclovina Uscată (South Transylvania, Romania)
We analysed an occipital bone fragment from the cranium, which gave a date between the late phases of the Aurignacian and the beginning of the Gravettian. The remains were not associated with artifacts, so we consider its archaeological context to be "Unassigned":

- *Cioclovina1* at 33,090-31,780 cal BP (OxA-15527: 28,510 ± 170 ^{14}C)[12]
  (direct date, using collagen ultrafiltration)

La Rochette (Dordogne, France)
The site of La Rochette, in the proximity of Saint-Léon-sur-Vézère, was excavated beginning in 1910 by Otto Hauser. There is evidence of human occupation of the site spanning from the Middle Palaeolithic (Mousterian) until the Upper Palaeolithic (Magdalenian). A revision of the site stratigraphy divided the layer with human remains into Gravettian, Aurignacian and Châtelperronian periods[13]. The right ulna was redated in 2011, which confirmed a previous date from 2002 associating it with the Gravettian horizon[14]. The ancient DNA analysis was performed on an ulna, yielding data from fewer than 4,000 SNPs covered at least once after quality control. Thus, we did not use the data from this individual in our main analysis.

- *LaRochette* at 27,780-27,400 cal BP (OxA-23413: 23,400 ± 110 ^{14}C)[15]
  (direct date, using collagen ultrafiltration)

Dolní Věstonice and Pavlov (South Moravia, Czech Republic)
Three Gravettian burial sites were discovered beginning in 1924 on the northern slope of the Pavlovské Hills in South Moravia: Dolní Věstonice I, Dolní Věstonice II and Pavlov I. From Dolní Věstonice I we analysed the triple burial comprised of *Vestonice13* (a femur), *Vestonice14* (a femur) and *Vestonice15* (a femur). Associated charcoal gives a date of:

- *Vestonice13, Vestonice14, and Vestonice15* at 31,070-30,670 cal BP (GrN-14831: 26,640 ± 110 ^{14}C)[16]
  (layer date, based on associated charcoal)

The *Vestonice16* (a femur), *Vestonice42* (a femur) and *Vestonice43* (a femur) samples were retrieved close to Dolní Věstonice I, suggesting similar depositional age (*Vestonice42* provided fewer than 4,000 SNPs and we did not use it for whole genome analysis). We obtained two dates on charcoal associated with these specimens:

- *Vestonice16, Vestonice42,* and *Vestonice43* at 30,710-29,310 cal BP (GrN-15277: 25,740 ± 210 ^{14}C and GrN-15276: 25,570 ± 280 ^{14}C)[16]
  (layer dates, based on associated charcoal)

Charcoal from the Pavlov I site from which we analysed *Pavlov1* (a femur) gave a date of:

- *Pavlov1* at 31,110-29,410 cal BP (GrN-20391: 26,170 ± 450 ^{14}C)[16]
  (layer date, based on associated charcoal)

Grotta Paglicci (Rignano Garganico - Foggia, Italy)
Paglicci Cave is situated in the Apulia region in southeast Italy. Excavation began in 1961 led by F. Zorzi of the Natural History Museum of Verona, followed, in 1971, by A. Palma di Cesnola and then by A. Ronchitelli of the University of Siena in collaboration with the Soprintendenza Archeologia della Puglia. Human occupation at this site is attested during the Early Middle and throughout the entirety of the Upper Palaeolithic. Paglicci is also important for the presence of the only Palaeolithic wall paintings discovered so far in Italy and for the most ancient evidence of flour production[17]. The main stratigraphic sequence is 12 meters thick. The Upper Palaeolithic layers yielded several human remains spanning from the Early Gravettian to the Final Epigravettian[18]. The three human remains used in ancient DNA analyses came from three distinct stratigraphic units.

We performed genetic analysis on *Palgicci71*, a right patella discovered in layer 8 whose ^{14}C date confirms the archaeological attribution to the Evolved Epigravettian, although this sample was not used in our main analysis as it yielded fewer than 4,000 SNPs:

- *Paglicci71* at 19,250-18,210 cal BP (F-66: 15,460 ± 220 ^{14}C)[19]
  (layer date, based on associated charcoal)

We also performed genetic analysis on *Paglicci108*, a phalanx discovered in layer 21B. A date from charcoal is consistent with the Evolved Gravettian material found in this layer:

- *Paglicci108* at 28,430-27,070 cal BP (F-52: 23,470 ± 370 ^{14}C)[20]
  (layer date, based on associated charcoal)

We finally performed genetic analysis on *Paglicci133*, a tooth (M_3 dx) found in layer 23C2, which is not directly dated but which can be associated on the basis of its stratigraphic position and the associated artifacts to the Early Gravettian culture:

- *Paglicci133* was found in a layer (23C2) whose chronology can be derived from the occurrence of tephra Codola elsewhere dated to around 33,000 years ago[21]. Moreover this layer is located between layer 23A (Early Gravettian) with a date of 33,110-31,210 cal BP (UTC-1415: 28,100 ± 400 ^{14}C) and layer 24A1 (Aurignacian) with a date of 34,580-31,860 cal BP (UTC-1789: 29,300 ± 600 ^{14}C)[18].
  (both dates are based on charcoal in the layers above and below)

Troisième Caverne (Goyet, Belgium)

The Troisième Caverne of Goyet in Belgium was excavated between the second half of the 19th century and the beginning of the 20th century, and more recently in the last decade of the 20th century. Edouard Dupont led the main excavations in 1868 and revealed Palaeolithic industries[22]. Follow-up studies found evidence of human occupation represented by Mousterian material, Lincombian-Ranisian-Jerzmanowician material, and Upper Palaeolithic complexes, including Aurignacian, Gravettian and Magdalenian material[23]. In 2008, a study was initiated with the goal of revisiting the human and faunal collections, and resulted in the identification of numerous new human remains. However, due to the lack of detailed excavation data, it has been challenging to assign the human remains to specific stratigraphic horizons. Taphonomic and morphometric characteristics were therefore first evaluated to sort human remains in different groups. All the specimens we analysed were directly radiocarbon dated in order to confirm the temporal attribution[24]. Those results, in combination with isotopic and genetic analyses[25-27], allowed specimen assignment either to late Neanderthal or to modern human origin. For ancient DNA analysis we analysed two individuals dated to the Aurignacian period[24]: *GoyetQ116-1* (a humerus), and *GoyetQ376-3* (a humerus):

- *GoyetQ116-1* at 35,160-34,430 cal BP (GrA-46175: 30,880 + 170 -160 ^{14}C)[24] (direct date, without ultrafiltration)

- *GoyetQ376-3* at 33,940-33,140 cal BP (GrA-60034: 29,370 + 180 -170 ^{14}C)[24] (direct date, without ultrafiltration)

We also analysed five individuals dated to the Gravettian period[24]: *GoyetQ376-19* (a humerus), *GoyetQ53-1* (a fibula), *GoyetQ55-2* (a fibula), *GoyetQ56-16* (a fibula) and *Goyet2878-21* (a clavicle):

- *GoyetQ376-19* at 27,720-27,310 cal BP (GrA-54026: 23,260 + 110 -100 ^{14}C)[24] (direct date, without ultrafiltration)

- *GoyetQ53-1* at 28,230-27,720 cal BP (GrA-46169: 23,920 ± 100 ^{14}C)[24] (direct date, without ultrafiltration)

- *GoyetQ55-2* at 27,730-27,310 cal BP (GrA-54031: 23,270 + 120 -110 ^{14}C)[24] (direct date, without ultrafiltration)

- *GoyetQ56-16* at 26,600-26,040 cal BP (GrA-59991: 22,100 ± 100 ^{14}C)[24] (direct date, without ultrafiltration)

- *Goyet2878-21* at 27,060-26,270 cal BP (GrA-62455: 22,360 ± 110 ^{14}C)[24] (direct date, without ultrafiltration)

We finally analysed an individual dated to the Magdalenian period[24]: *GoyetQ-2* (a humerus).

- *GoyetQ-2* at 15,230-14,780 cal BP (GrA-46168: 12,650 ± 50 ^{14}C)[24] (direct date, without ultrafiltration)

Three samples (*GoyetQ376-3*, *GoyetQ55-2* and *Goyet2878-21*) were not used for genome-wide analyses because fewer than 4,000 SNPs were covered after quality control.

Swabian Jura sites (Baden-Württemberg, Germany)

We carried out ancient DNA analysis on seven specimens from six caves in the Swabian Jura region, Southwest Germany. The geology of this area was formed during the Jurassic period and is mainly comprised of limestone that favors the formation of caves. The Ach and Lone rivers form two valleys of the same name before they merge into the Danube.

(i) Hohle Fels (Ach valley): This site is famous for the discoveries of Upper Palaeolithic ivory figurines and flutes associated with the early Aurignacian[28,29]. We performed genetic analysis on two remains, *HohleFels49* (a femur) and *HohleFels79* (a cranial fragment) both dated to the Magdalenian period:

- *HohleFels49* at 16,000-14,260 cal BP (H5312-4907: 12,770 ± 220 ^{14}C and H5119-4601: 13,085 ± 95 ^{14}C)[30]
  (layers date, based on a direct dates on bones from collagen without ultrafiltration)

- *HohleFels79* at 15,070-14,270 cal BP (MAMS-25564: 12,490 ± 70 ^{14}C)[24]
  (direct date, using collagen ultrafiltration)

(ii) Brillenhöhle (Ach valley): An excavation led by Gustav Riek beginning in 1956 discovered at least four individuals, all assigned to the Magdalenian culture based on associated artifacts[31]. For ancient DNA analysis we analysed a parietal bone:

- *Brillenhohle* at 15,120-14,440 cal BP (OxA23414: 12,535 ± 50 ^{14}C)[15]
  (direct date, using collagen ultrafiltration)

(iii) Burkhardtshöhle (Westerheim, Württemberg): An excavation beginning in 1933-1934 and led by Gustav Riek found five cranial fragments, probably from the same individual. These specimens were associated with the Magdalenian culture and gave a date of:

- *Burkhardtshohle* at 15,080-14,150 cal BP (ETH-7613: 12,450 ± 110 ^{14}C)[32]
  (direct date, without collagen ultrafiltration)

(iv) Bockstein-Höhle (Lone valley): This cave yielded a double burial consisting of an infant and an adult. The adult was morphologically identified as female and the sex is confirmed here through genetic analysis of an incisor. Two skeletal elements of the infant were dated and linked the burial to the Late Mesolithic period:

- *Bockstein* at 8,370-8,160 cal BP (UtC-7887: 7,350 ± 70 ^{14}C and UtC-6796: 7,460 ± 60 ^{14}C)[33]
  (layer date, based on a direct date of a second human skeleton in the same layer, without collagen ultrafiltration)

(v) Falkensteiner Höhle: This cave is situated in a karstic area in the vicinity of Bad Urach. The excavations of the site started in 1933 and in later analyses, around 40 human remains were identified[31]. The same site yielded typical Mesolithic stone tools as well as the human fibula that we investigated genetically. The remains gave a date:

- *Falkenstein* at 9,410-8,990 cal BP (ETH-7615: 8,185 ± 80 ^{14}C)[34]
  (direct date, using collagen without ultrafiltration)

(vi) Hohlenstein-Stadel (Lone valley): Signs of human occupation at this site spanned from the Middle Palaeolithic to the Neolithic. A triple burial only represented by human skulls and cervical vertebrae was found in the cave and was dated to a late phase of the Mesolithic. We did not include this sample in our main analysis because fewer than 4,000 SNPs were covered at least once after quality control.

- *HohlensteinStadel* at 8,980-8,440 cal BP (ETH-5732: 7,835 ± 80 ^{14}C)[35]
  (direct date, using collagen without ultrafiltration)

Felsdach Inzigkofen (Upper Danube valley, Germany)
This site is located in the Upper Danube valley, in close proximity of Beuron village. Wolfgang Taute excavated the site where a 3rd molar was discovered in a layer associated to the early period of the Late Mesolithic, also called the Early Atlantic. We did not include this sample in our main analysis because it gave fewer than 4,000 SNPs after quality control.

- *Felsdach* at 8,980-8,380 cal BP (B-933: 7,770 ± 120 ^{14}C)[31,36]
  (layer date)

Große Ofnet Höhle (Franconian Jura, Germany)
Große Ofnet cave is found close to Nördlingen, in Bavaria. Within the cave, two sites yielded a total of 34 crania[35], all facing west and some presenting cut-marks or sign of injuries[37]. For ancient DNA analysis, we analysed a molar tooth that was excavated by Oscar Fraas in 1875-1876 and is part of the human osteological collection of the University of Tübingen (OSUT 4043). Here we quote a radiocarbon date that associates the remains to the Late Mesolithic:

- *Ofnet* at 8,430-8,060 cal BP (OxA1574: 7,480 ± 80 ^{14}C)[35]
  (direct date, using collagen without ultrafiltration)

French Jura sites (Franche-Comté, France)
This mountainous region is located between Eastern France and Western Switzerland. The entire area is composed of Mesozoic limestone, where a karstic landscape leads to cave formation[38]. We analysed samples from three sites.

(i) Cabônes rockshelter (Ranchot, Jura department): This site was discovered in the 1950s, and excavation occurred between 1978 and 1990. At least five individuals were assigned to the Mesolithic[39]. The specimen we analysed, *Ranchot88*, is a right parietal fragment:

- *Ranchot88* at 10,240-9,930 cal BP (GrA-38019: 8,985 ± 40 ^{14}C)[24]
  (direct date, using collagen without ultrafiltration)

(ii) Rigney 1 cave (Doubs department): We performed ancient DNA analysis on a mandible fragment recovered at Rigney 1 cave in 1986-87 during a rescue excavation. This site is linked to the Magdalenian material culture[40].

- *Rigney1* at 15,690-15,240 cal BP (Ly-6515(OxA): 12,930 ± 55 ^{14}C)[38]
  (direct date, using collagen without ultrafiltration)

(iii) Rochedane rockshelter (Villars-sous-Dampjoux, Doubs department): A. Thévenin excavated this site between 1968 and 1976 after its discovery at the end of the 19th century. The mandible we analysed was assigned to the EpiPalaeolithic culture.

- *Rochedane* at 13,090-12,830 cal BP (GrA-41739: 11,120 ± 50 ^{14}C)[38]
  (direct date, using collagen without ultrafiltration)

Paris Basin (France)
In this region northeast of Paris, several Mesolithic human remain were discovered[41]. The local geology is characterized by a plateau of sedimentary rocks. We analysed three sites.

(i) Les Closeaux at Rueil-Malmaison: A single human skeleton was buried in squatting position inside a circular pit along the river Seine.

- *LesCloseaux13* at 10,240-9,560 cal BP (OxA-7109 (Ly-612): 8,870±130 ^{14}C)[42]
  (direct date, using collagen without ultrafiltration)

(ii) Les Fontinettes at Cuiry-lès-Chaudardes: This Mesolithic burial, which we call *Chaudardes1*, was found at the side of a Neolithic settlement with a single individual in seated position accompanied by three flints and a necklace[43].

- *Chaudardes1* at 8,360-8,050 cal BP (GrA-28268: 7,400 ± 60 ^{14}C)[24]
  (direct date, using collagen without ultrafiltration)

(iii) Le Vieux Tordoir at Berry-au-Bac: This site revealed human occupation in several periods including a fortified Neolithic site. For ancient DNA analysis we studied a radius from a skeleton buried with a bone tool and covered with ochre[44].

- *BerryAuBac* at 7,320-7,170 cal BP (SacA-5455: 6,325 ± 35 ^{14}C)[24]
  (direct date, using collagen without ultrafiltration)

Aven des Iboussières à Malataverne (Rhône-Alpes, France)
In this site located in South France, at least eight human individuals were discovered and associated to the EpiPalaeolithic culture[45]. We studied the left femur of a juvenile.

- *Iboussieres39* at 12,040-11,410 cal BP (GrA-43700: 10,140 ± 50 ^{14}C)[24]
  (direct date, using collagen without ultrafiltration)

Peştera Muierilor (Romania)
Excavations at Muierilor Cave in the 1950s yielded human remains from up to three individuals. The *Muierii2* temporal bone that we analysed genetically gave a date of:

- *Muierii2* at 33,760-32,840 cal BP (OxA-16252: 29,110 ± 190 ^{14}C yr BP)[46]
  (direct date, using collagen ultrafiltration)

El Mirón Cave (Cantabria, Spain)
This large cave is located in the Cantabrian Cordillera, around 20 km from the present shore and equidistant between Bilbao and Santander. Since 1996, it has been excavated under the direction of L.G. Straus and M.R. González Morales, uncovering a nearly complete sequence of cultural occupations from the late Mousterian through the early Bronze Age, with 86 radiocarbon dates ranging from >46,000 years to 3,200 years uncal. BP[47]. Among the richest levels are those of the Cantabrian Lower Magdalenian with classic lithic and osseous artifacts

and engraved red deer scapulae, as well as rock art dated *terminus ante* and *post quem* to this period and intimately associated with the human burial—the first human interment of Magdalenian age to be discovered in the Iberian Peninsula. The "Red Lady of El Mirón" skeleton, which consists of about 100 individual elements (including the mandible, but missing the cranium and most of the large long bones), was recovered between 2010-2013[48]. The human bones, the layer in which they were buried, and the face of the large, engraved block against which the individual's back seems to have rested, were all stained or impregnated with specially prepared, non-local, red ochre (specular hematite). The genetic analysis was performed on a toe bone, and the radiocarbon date came from a fibula of the individual, who was a female, 35-40 years old at the time she died, relatively tall and robust, and in apparently in good health, with a mixed diet of terrestrial game, fish and plants.

- *ElMiron* at 18,830-18,610 cal BP (MAMS-14585: 15,460 ± 40 ^{14}C)[49]
  (direct date, using collagen ultrafiltration)

Villabruna (Sovramonte – Belluno, Italy)
The burial of Riparo Villabruna was discovered in 1988 by A. Broglio in the small rockshelter named Riparo Villabruna A in the Veneto Dolomites. It contains a partial skeleton with lower limbs severed at the distal femoral shafts associated with burial goods of the Epigravettian culture[50]. The date quoted here comes from the skull[51], whereas the genetic analysis is of a left femur. This individual bears the earlier known example of treatment of dental caries[52].

- *Villabruna* at 14,180-13,780 cal BP (KIA-27004: 12,140±70^{14}C)[51]
  (direct date, using collagen ultrafiltration)

Kostenki 14 – Markina Gora (Voronezh region, Russia)
The *Kostenki14* skeleton was found in 1954 in a crouched position inside an oval burial below the third cultural layer. The layer lacked diagnostic stone tools and thus its cultural assignment is unclear. The bones were partially covered with ochre and associated with a small number of flints as well as isolated faunal bones[53,54]. We performed ancient DNA analysis on a tibia from the skeleton.

- *Kostenki14* at 38,680-36,260 cal BP (OxA-X-2395-15: 33,250 ± 500 ^{14}C)[55]
  (HPLC-separated hydroxyproline fraction in collagen)

Kostenki 12 – Volkovskaya (Voronezh region, Russia)
The *Kostenki12* cranial bone belongs to a perinatal child, maybe deliberately buried, found in layer 1 of the Kostenki 12 – Volkovskaya site. This layer yielded an assemblage that has been attributed to the Gorodtsovian, an Early Upper Palaeolithic culture that is characteristic of the region, and that overlaps in time with the Early Gravettian (the Early Gravettian is documented at Kostenki 8 in layer 2). The association between the dated bone and the Gordotsovian material is not secure because of reworking of the deposit[56]. *Kostenki12* was not assigned to a genetic cluster:

- *Kostenki12* at 32,990-31,840 cal BP (GrA-5552: 28,500 ± 140 ^{14}C)[56]
  (layer date, based on associated charcoal)

Ostuni (Apulia, Italy)
The Grotta di Santa Maria di Agnano is a cave near the village of Ostuni, located at about 170 meters above sea level in a limestone formation (*Calcare di Altamura*) on the south-

eastern slopes of the Murge tablelands in central Apulia (southeast Italy). Excavations at this site led by one of the authors (D. Coppola) have demonstrated that the cave was occupied for tens of thousands of years from the Middle Palaeolithic until modern times, including Middle Palaeolithic (Mousterian) and Upper Palaeolithic (Aurignacian, Gravettian and Epigravettian) industries[57]. The excavation discovered two burials, which on stratigraphic and typological grounds are attributable to the Gravettian culture. The first of these burials (*Ostuni1*) is of a pregnant female of around 20 years of age, about to give birth[57]. The skeletons of both the mother and the fetus are remarkably well preserved and, to the best of our knowledge, constitute the oldest reported burial of a pregnant female in the world. The second burial (*Ostuni2*) contains the remains of an adult whose gender has not yet been established based on morphology (due to the partial excavation of this inhumation) but who genetically is female. Around their crania, both *Ostuni1* and *Ostuni2* had numerous ornaments made of perforated shells of the marine gastropod *Cyclope neritea*, similar to other Gravettian burials from Italy. We extracted collagen from the ribs of both individuals using the protocol of Talamo and Richards[58]. The radiocarbon dates for both samples are based on ultrafiltration, and give dates that confirm the assignment of both samples to the Gravettian:

- *Ostuni1* at 27,810-27,430 cal BP (MAMS-11449: 23,446 ± 107 ^{14}C)
  (direct date, using collagen ultrafiltration; the present study is the first report of this date)

- *Ostuni2* at 29,310-28,640 cal BP (MAMS-11450: 24,910 ± 125 ^{14}C)
  (direct date, using collagen ultrafiltration; the present study is the first report of this date)

Afontova Gora (Russia)

Afontova Gora 3 was found in 2014 during salvage excavations connected to bridge construction in the Yenisei River basin. Excavation area No. 24 cut into the third river terrasse and yielded a first cervical vertebra, a mandible, and five teeth from an *in situ* discovery in occupation layer 2. Preliminary analyses suggest that the mandible and five teeth may be associated with a 14-15 year old girl; the sex is confirmed in this study genetically. The mandible demonstrates a gracile structure: its longitudinal and transverse measurements fall into the category of "small" or "very small". The gracile mandible is quite different from the children of similar age recovered of Sungir-2 and Sungir-3, which might indicate different affinities. We performed genetic analysis on a left molar (M_2). We assume that the *AfontovaGora3* specimen is approximately contemporaneous with the *AfontovaGora2* sample found at the site in the 1920s, and so use that sample to provide a layer date:

- *AfontovaGora3* at 16,930-16,490 cal BP (UCIAMS-79661: 13,810 ± 35 ^{14}C)[5]
  (layer date, based on ultrafiltration of remains from the *AfontovaGora2* individual)

Grotta Continenza (Abruzzo, Italy)

Grotta Continenza is a cave at an elevation of 700 meters above sea level overlooking the Fucino plain in an area of Abruzzo, known as Marsica, in central Italy. It contains an outer and an inner chamber, occupied from the late Upper Palaeolithic to the Bronze Age, as well as during the Roman period[59,60]. The prehistoric deposits are stratified in a 9 meter sequence which from the most recent to the oldest is: Neolithic (spits 4 to 23), Castelnovian (spit 24), Sauveterrian (spits 25 to 29) and Late Epigravettian (spits 30 to 48) layers. Human remains are numerous throughout the Mesolithic deposits (spits 24 to 29). A cranial fragment recovered within the Mesolithic layers in stratigraphic spit 24 was selected for ancient DNA analysis. We obtained three radiocarbon dates on a bone from the top of the stratigraphic unit in which the cranial element analysed for DNA was recovered, as well as two bones from

immediately below. The method of bone collagen extraction that we adopted is that proposed by Talamo and Richards[58]. The dates suggest that the deposit from which the cranial fragment was recovered dates to 11,210-10,510 cal years BP. It can therefore be concluded that the specimen for which we obtained ancient DNA is Mesolithic. Despite being recovered in a stratigraphic spit associated with Castelnovian lithic industries, the radiocarbon dates suggest that a tentative attribution to the Sauveterrian is more feasible.

- *Continenza* at 11,200-10,510 calBP, based on a range of three dates from stratigraphic contexts immediately above and below:
  10,870-10,700 cal BP (MAMS-11444: 9,521 ± 31 ^{14}C)
  10,690-10,510 cal BP (MAMS-11445: 9,379 ± 30 ^{14}C)
  11,200-11,080 cal BP (MAMS-11448: 9,680 ± 32 ^{14}C)
  (layer date, based on collagen ultrafiltration of three bones – the present study is the first report of these dates)

Krems-Wachtberg (Lower Austria)

Between 2005-2015, the Austrian Academy of Sciences carried out an interdisciplinary research project at Wachtberg in Krems, Lower Austria. This site is found on the southern slope of a promontory between the Danube and Krems River. The profiles are up to 8 meters high and represent a significant record of the time between 40,000-20,000 years ago. At a depth of 5 meters, about 150 square meters of a well-preserved cultural layer with associated features including hearths and two burials were recovered. The living floor dates to 27.0 ^{14}C ka BP. Burial 1 is a unique double burial of newborns (individuals 1 and 2), while burial 2 (individual 3) is a single burial of an immature individual (its estimated age at death is around 3 months). All three individuals were buried in flexed position and embedded in red ochre. The sample for the present study was taken from the parietal bone of individual 3 (burial 2). The date is from charcoal on the living floor[61].

- *Krems-Wachtberg* at 31,250-30,690 cal BP (VERA-3941: 26,870 ± 220^{14}C)[61,62]
  (layer date, based on associated charcoal)

References

1 Reimer, P. J. *et al.* Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0-50,000 Years cal BP. *Radiocarbon* **55**, 1869-1887 (2013).
2 Ramsey, C. B. & Lee, S. Recent and planned developments of the program OxCal. *Radiocarbon* **55**, 720-730 (2013).
3 Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449 (2014).
4 Rougier, H. *et al.* Peştera cu Oase 2 and the cranial morphology of early modern Europeans. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1165-1170 (2007).
5 Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91 (2014).
6 Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* **6**, 8912 (2015).
7 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413 (2014).
8 Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225-228 (2014).

9 Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications* **5**, 5257 (2014).

10 Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522,** 207-211 (2015).

11 Price, T. D. & Jacobs, K. Olenii Ostrov - 1st radiocarbon dates from a major Mesolithic cemetery in Karelia, USSR. *Antiquity* **64**, 849-853 (1990).

12 Soficaru, A., Dobo**ş**, A. & Trinkaus, E. Early modern humans from the Pe**ş**tera Muierii, Baia de Fier, Romania. Proc. Natl Acad. Sci. USA **103,** 17196–17201 (2006).

13 Delporte, H. Le gisement paléolithique de la Rochette, commune de Saint-Léon-sur-Vézère, Dordogne. *Gallia Préhistoire* **V**, 1-22 (1962).

14 Orschiedt, J. Datation d'un vestige humain provenant de la Rochette (Saint Léon-sur-Vézère, Dordogne) par la méthode du carbone 14 en spectrométrie de masse. *Paléo* **14**, 239-240 (2002).

15 Benazzi, S. *et al.* Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature* **479**, 525-528 (2011).

16 Trinkaus, E. & Svoboda, J. *Early Modern Human Evolution in Central Europe: The People of Dolní Věstonice and Pavlov*. Vol. 12 (Oxford University Press, 2006).

17 Mariotti Lippi, M., Foggi, B., Aranguren, B., Ronchitelli, A. & Revedin, A. Multistep food plant processing at Grotta Paglicci (Southern Italy) around 32,600 cal B.P. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 12075-12080 (2015).

18 Palma di Cesnola, A. *Paglicci. L'Aurignaziano e il Gravettiano antico* (Claudio Grenzi Ed., 2004).

19 Azzi, C. M., Bigliocca, L. & Piovan, F. Florence Radiocarbon Dates III. *Radiocarbon* **19**, 165-169 (1977).

20 Azzi, C. M., Bigliocca, L. & Piovan, F. Florence Radiocarbon Dates II. *Radiocarbon* **16**, 10-14 (1974).

21 Giaccio, B. *et al.* The Campanian Ignimbrite and Codola tephra layers: Two temporal/stratigraphic markers for the Early Upper Palaeolithic in southern Italy and eastern Europe. *Journal of Volcanology and Geothermal Research* **177**, 208-226 (2008).

22 Dupont, E. *L'Homme pendant les âges de la pierre dans les environs de Dinant-sur-Meuse, 2ème édition* (C. Muquardt Ed., 1872).

23 Flas, D. *La transition du Paléolithique moyen au supérieur dans la plaine septentrionale de l'Europe. Les problématiques du Lincombien-Ranisien-Jerzmanowicien*. PhD dissertation (Université de Liège, 2006).

24 Posth, C. *et al.* Pleistocene mitochondrial genomes suggest a single major dispersal of Non-Africans and a Late Glacial population turnover in Europe. *Current Biology* **26,** 827-833 (2016).

25 Rougier, H. *et al.* First evidence of Neandertal cannibalism in Northern Europe. (subm.).

26 Wißing, C. *et al.* Isotopic evidence for dietary ecology of Late Neandertals in North-Western Europe. *Quaternary International*, doi:10.1016/j.quaint.2015.09.091 (2015).

27 Rougier, H., *et al.* The first Upper Paleolithic human remains from Belgium: Aurignacian, Gravettian and Magdalenian fossils at the "Troisième caverne" of Goyet. *PaleoAnthropology* **A33** (2013).

28 Conard, N. J., Malina, M. & Munzel, S. C. New flutes document the earliest musical tradition in southwestern Germany. *Nature* **460**, 737-740 (2009).

29 Conard, N. J. A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany. *Nature* **459**, 248-252 (2009).

30 Housley, R. A., Gamble, C. S., Street, M. & Pettitt, P. B. Radiocarbon evidence for the Lateglacial human recolonisation of Norther Europe. *Proc. Prehist. Soc.* **63**, 25-54 (1997).

31 Haas-Campen, S. *Die menschlichen Skelettreste des Spätpleistozäns und Frühholozäns in Baden-Württemberg*, Magisterarbeit (1993).

32 Simon, U. *Die Burkhardtshöhle - eine Magdalénienstation am Nordrand der Schwäbischen Alb*, Magisterarbeit (1993).

33 Wehrberger, K. "Der Streit ward definitiv beendet..." Eine mesolithische Bestattung aus der Bocksteinhöhle im Lonetal, Alb-Donau-Kreis. Zur Erinnerung an Ludwig Bürger (1844-1898). *Archäologisches Korrespondenzblatt* **30**, 15-31 (2000).

34 Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326**, 137-140 (2009).

35 Orschiedt, J. Manipulationen an menschlichen Skelettresten. Taphonomische Prozesse, Sekundärbestattungen oder Kannibalismus. Dissertation. *Urgeschichtliche Materialhefte* **13** (1999).

36 Holdermann, C.-S. & Kind, C.-J. Zeitwechsel in Schichten. Bedeutende Fundstellen an der oberen Donau. *Eiszeit. Kunst und Kultur. Begleitband zur Großen Landesausstellung Eiszeit - Kunst und Kultur im Kunstgebäude Stuttgart*, 332–335 (2009).

37 Mollison, T. Zeichen gewaltsamer Verletzungen an den Ofnet-Schädeln. *Anthrop. An.* **13**, 79-88 (1936).

38 Cupillard, C. *et al.* Changes in ecosystems, climate and societies in the Jura Mountains between 40 and 8 ka cal BP. *Quaternary International* **378**, 40-72 (2015).

39 Valentin, F. Les restes humains de l'abri des Cabônes à Ranchot (Jura). In : Les derniers chasseurs-cueilleurs (13000-5500 av. JC) dans le massif du Jura et ses marges. *Centre Jurassien du Patrimoine*, 185-186 (1998).

40 Cupillard, C., Malgarini, R. & Fornage-Bontemps, S. Le Paléolithique supérieur ancien dan le quart nord-est de la France : l'exemple de la Franche-Comté. Environnement, chronologie et faciès culturels. *Mémoire LVI de la Société préhistorique française*, 351-363 (2013).

41 Bosset, G. & Valentin, F. Mesolithic burial practices in the northern half of France: Isolated burials and their spatial organisation. *Paris: Société préhistorique française*, 207-216 (2013).

42 Valentin, F. et al. Découvertes récentes d'inhumations et d'une incinération datées du Mésolithique en Île-de-France. Revue Archéologique d'Île-de-France **1,** 21–42 (2008).

43 Ilett, M. Cuiry-lès-Chaudardes "les Fontinettes" (Ministère de la Culture, Université de Paris 1-Panthéon Sorbonne, 1998).

44 Auxiette, G. & Hachem, L. Berry-au-Bac, "Le Chemin de la Pêcherie ouest", "le Vieux-Tordoir", "la culée" (Ministère de la Culture, Fouille Protohistorique de la Vallée de l'Aisne, 1989).

45 Gely, B. & Morand, P. Les sépultures épipaléolithiques de l'aven des Iboussières à Malataverne (Drôme, France): Premiers résultats. *Ardèche Archéologie* **15**, 13-18 (1998).

46 Soficaru, A., Dobo**ş**, A. & Trinkaus, E. Early modern humans from the Peştera Muierii, Baia de Fier, Romania. Proc. Natl Acad. Sci. USA **103,** 17196–17201 (2006).

47 Straus, L. G., González Morales, M.R. (eds.) El Mirón Cave, Cantabrian Spain. *University of New Mexico Press* (2012).

48 Straus, L. G., Gonzalez Morales, M.R., Carretero, J.M. (eds.) The Red Lady of El Miron Cave': Lower Magdalenian Human Burial in Cantabrian Spain. *J. Archaeol. Sci.* **60**, 1-137 (2015).

49 Straus, L. G., Gonzalez Morales, M. R., Higham, T., Richards, M. & Talamo, S. Radiocarbon dating the late Upper Paleolithic of Cantabrian Spain: El Miron Cave Date List IV. *Radiocarbon* **57**, 183-188 (2015).

50 Aimar, A., et al. . Les Abris Villabruna dans la Vallèe du Cismon. *Preistoria Alpina* **28**, 227- 254 (1992).

51 Vercellotti, G., Alciati, G., Richards, M. P. & Formicola, V. The Late Upper Paleolithic skeleton Villabruna 1 (Italy): a source of data on biology and behavior of a 14.000 year-old hunter. *Journal of Anthropological Sciences* **86**, 143-163 (2008).

52 Oxilia, G. *et al.* Earliest evidence of dental caries manipulation in the Late Upper Palaeolithic. *Scientific Reports* **5**, 12150 (2015).

53 Sinitsyn, A. A. Kostenki 14 (markina gora): Data, problems, and perspectives. *Prehistoire Europeene*, 273-313 (1996).

54 Krause, J. *et al.* A complete mtDNA genome of an early modern human from Kostenki, Russia. *Current Biology* **20**, 231-236 (2010).

55 Marom, A., McCullagh, J. S., Higham, T. F., Sinitsyn, A. A. & Hedges, R. E. Single amino acid radiocarbon dating of Upper Paleolithic modern humans. *Proc Natl Acad Sci U S A* **109**, 6878-6881 (2012).

56 Sinitsyn, A. A. Les sépultures de Kostenki : chronologie, attribution culturelle, rite funéraire. *In OTTE M. (dir.), La spiritualité, Proceedings of UISPP conference, Liege, ERAUL* **106**, 237-244 (2004).

57 Coppola, D. Il riparo di Agnano nel paleolitico superiore. La sepoltura Ostuni ed i suoi simboli, vol.1. . *Università di Roma Tor Vergata, Roma.* (2012).

58 Talamo, S. & Richards, M. A comparison of bone pretreatment methods for AMS dating of samples >30,000 BP. *Radiocarbon* **53**, 443-449 (2011).

59 Serradimigni, M. Grotta Continenza (Trasacco, Prov. de l'Aquila). *Notiziario di Preistoria e Protostoria* **2014 -1.I, 12-14** (2014).

60 Grifoni Cremonesi, R., Serradimigni, M. & Usala, M. . in Il Fucino e le aree limitrofe nell'antichità. *Atti del III Convegno di Archeologia in ricordo di Walter Cianciusi, Avezzano* (2011).

61 Haendel, M., Einwoegerer, T., Simon, U. & Neugebauer-Maresch, C. Krems-Wachtberg excavations 2005-12: Main profiles, sampling, stratigraphy, and site formation. *Quaternary International* **351**, 38-49 (2014).

62 Einwogerer, T. *et al.* Upper Palaeolithic infant burials. *Nature* **444**, 285 (2006).

Section 2
Ancient DNA processing and quality control

DNA extraction and library preparation

All ancient DNA extracts were prepared using the method described in ref. [1], starting from between 28 and 350 mg of bone or tooth powder.

A total of 44 libraries generated at the Max Planck Institute for Evolutionary Anthropology (MPI-EVA) in Leipzig were prepared using a single-stranded protocol[2,3]. All but three were treated with uracil-DNA-glycosylase (UDG) from *Escherichia coli* and endonuclease (Endo VIII)[4]. These libraries retain characteristic ancient DNA damage at the terminal 5' nucleotide as well as the two terminal 3' nucleotides[2], and we denote them as "ss UDG". We denote the three MPI-EVA libraries prepared without UDG-treatment as "ss noUDG" (Table S2.1).

Table S2.1: Libraries prepared at MPI-EVA in Leipzig

Sample	Library	Extract	% endogenous	SNPs	mt cov	Match consens. Est.	Match consens. 95% CI	mt. Hg	Library type	Damage	Chr. X contam.	Library decision*	Analyzed?
Vestonice13	A5280	E1500	1.94	2.2M	93	1	0.91-1	U8c	ss UDG	0.37	5.5%	Damage	Yes
Vestonice14	A5281	E1501	n/a	390k	114	1	0.9-1	H7	ss UDG	0.2	n/a	Damage	Yes
Vestonice14	A5282	E1798	0.37	390k	81	0.99	0.76-1	H7a1	ss UDG	0.34	n/a	Damage	Yes
Vestonice15	A5271	E1502	0.47	2.2M	88	1	0.91-1	U5	ss UDG	0.36	16.0%	Damage	Yes
Vestonice16	A5188	E1799	0.1	2.2M	97	1	0.91-1	U	ss UDG	0.22	27.4%	Fail	Yes
Vestonice16	A5272	E1503	1.57	2.2M	115	1	0.92-1	U	ss UDG	0.27	26.3%	Fail	Yes
Vestonice16	A5306	E1789	0.76	2.2M	108	1	0.92-1	U5	ss UDG	0.3	1.9%	Pass	Yes
Vestonice16	A5307	E1799	1.28	390k	121	0.99	0.92-1	U	ss UDG	0.22	23.6%	Fail	Yes
Vestonice42	A5284	E1504	0.28	390k	107	0.99	0.87-1	H	ss UDG	0.19	34.1%	Damage	No
Vestonice43	A5287	E1505	0.18	390k	85	0.99	0.84-1	U	ss UDG	0.3	33.3%	Fail	Yes
Vestonice43	A5308	E1800	0.89	2.2M	102	1	0.92-1	U5	ss UDG	0.33	9.0%	Damage	Yes
ElMiron	A5268	E1796	2.58	2.2M	104	1	0.92-1	U5b	ss UDG	0.33	n/a	Pass	Yes
ElMiron	A5279	E1796	2.73	2.2M	105	0.98	0.88-1	U5b	ss UDG	0.39	n/a	Pass	Yes
ElMiron	A5301	E1796	2.69	2.2M	133	0.99	0.9-1	U5b	ss UDG	0.34	n/a	Pass	Yes
Continenza	A5189	E1788	n/a	2.2M	33	0.99	0.66-1	U5b1	ss UDG	0.24	n/a	Damage	Yes
Continenza	A5206	E1788	0.03	2.2M	43	0.99	0.83-1	U5b1	ss UDG	0.58	n/a	Damage	Yes
Continenza	A5207	E1788	0.03	2.2M	43	0.99	0.8-1	U5b1	ss UDG	0.32	n/a	Damage	Yes
Continenza	A5213	E1788	0.03	390k	21	0.96	0.47-0.99	U5b1	ss UDG	0.26	n/a	Fail	Yes
Continenza	A5214	E1788	0.02	390k	26	0.84	0.23-0.95	U5b1	ss UDG	0.33	n/a	Fail	Yes
Kostenki12	A5212	E1794	1.81	390k	11	0.92	0.03-0.98	U	ss UDG	0.5	n/a	Fail	Yes
Kostenki12	A5273	E1785	0.07	2.2M	20	0.76	0.16-0.96	U2	ss UDG	0.32	11.8%	Fail	Yes
Kostenki12	A5275	E1785	0.14	390k	6	n/a	n/a	R	ss UDG	0.28	n/a	Fail	Yes
Kostenki12	A5276	E1648	0.66	2.2M	43	0.99	0.72-1	U2	ss UDG	0.32	n/a	Pass	Yes
Kostenki12	A5288	E1844	0.07	2.2M	40	0.98	0.6-1	U2	ss UDG	0.27	15.0%	Fail	Yes
Kostenki14	A5187	E1782	n/a	2.2M	89	1	0.88-1	U2	ss UDG	0.22	2.1%	Pass	Yes
Kostenki14	A5195	E1781	8.6	390k	120	0.97	0.86-1	U2	ss UDG	0.48	1.2%	Pass	Yes
Kostenki14	A5196	E1781	9.65	2.2M	125	1	0.93-1	U2	ss UDG	0.21	1.5%	Pass	Yes
Kostenki14	A5197	E1781	8.66	2.2M	112	0.99	0.88-1	U2	ss UDG	0.2	1.5%	Pass	Yes
Kostenki14	A5198	E1781	8.48	2.2M	122	1	0.89-1	U2	ss UDG	0.2	2.0%	Pass	Yes
Kostenki14	A5278	E1781	12.15	390k	110	0.97	0.84-1	U2	ss UDG	0.33	1.2%	Fail	Yes
Kostenki14	A5305	E1782	4.72	2.2M	113	1	0.91-1	U2	ss UDG	0.23	1.8%	Pass	Yes
Muierii2	A5270	E1407	0.67	390k	27	0.57	0.39-0.73	U	ss UDG	0.23	n/a	Fail	Yes
Muierii2	A5289	E1850	1.99	390k	25	0.66	0.44-0.79	U	ss UDG	0.23	n/a	Fail	Yes
Muierii2	A5299	E1851	1.09	2.2M	109	0.93	0.88-0.97	U6	ss UDG	0.23	n/a	Damage	Yes
Oberkassel	A5302	E1783	0.09	390k	62	0.99	0.75-1	U5b1	ss UDG	0.24	n/a	Damage	No
Oberkassel	A5303	E1783	0.08	390k	60	0.86	0.34-0.99	U5b1	ss UDG	0.25	n/a	Damage	No
Ostuni1	A5180	E1750	1.04	2.2M	32	0.69	0.52-0.84	M	ss UDG	0.41	n/a	Damage	Yes
Ostuni1	A5181	E1750	0.97	2.2M	44	0.88	0.74-0.94	M	ss UDG	0.41	n/a	Damage	Yes
Ostuni1	A5182	E1750	0.91	2.2M	61	0.87	0.77-0.93	M	ss UDG	0.4	n/a	Damage	Yes
Ostuni1	A5201	E1751	0.76	2.2M	93	0.89	0.8-0.93	M	ss UDG	0.33	n/a	Damage	Yes
Ostuni1	A5265	E1750	1.06	2.2M	48	0.77	0.65-0.86	M	ss UDG	0.4	n/a	Damage	Yes
Ostuni2	A5203	E1752	0.08	2.2M	67	0.83	0.56-0.94	U2	ss UDG	0.46	19.5%	Damage	Yes
Pavlov1	A5277	E1506	0.17	2.2M	83	0.99	0.89-1	U	ss UDG	0.34	25.3%	Damage	Yes
Pavlov1	A5304	E1801	0.09	2.2M	80	0.99	0.86-1	U5	ss UDG	0.35	11.8%	Damage	Yes
Villabruna	A5290	E1845	n/a	2.2M	92	1	0.91-1	U5b2b	ss noUDG	0.45	1.60%	Pass	Yes
Villabruna	A5294	E1849	n/a	2.2M	104	0.95	0.93-0.96	U5b2b	ss noUDG	0.37	5.90%	Fail	Yes
AfontovaGora3	L5121	E2670	1.36	2.2M	2504	0.99	0.97-0.99	R1b	ss noUDG	0.37	n/a	Pass	Yes

Note: Cells in red have mitochondrial contamination estimates that fail our criteria: <95% for the point estimate, or a ninety five percent confidence interval that includes values <85%. The final column indicates whether the sample was among the 51 used in the genome-wide analysis (some samples were not used because there were <4,000 SNPs after quality control).
* Some samples do not have an X chromosome contamination estimate because of likely female sex or because fewer than 200 SNPs were covered at least twice.
+ Libraries marked "Fail" are not used in analysis. For libraries marked "Damage," analyses are restricted to fragments showing deamination.

A total of 46 libraries generated at the University of Tübingen were prepared by both single-stranded and double-stranded protocols[5,6]. The single-stranded libraries were prepared using the same protocols as the libraries at MPI-EVA in Leipzig. The double-stranded libraries were prepared either without UDG-treatment ("ds noUDG"), or with a protocol that also removes deaminated cytosines at the final nucleotide ("ds UDG") (Table S2.2).

Table S2.2: Libraries prepared at the University of Tübingen

Sample	Library	Extract	Endogenous%	SNPs	mt cov	Match consens. Est.	Match consens. 95% CI	mt. Hg	Library type	Dam-age	*Chr. X contam.	Library decision+	Analyzed?
Iboussieres39	GA162	GX45	n/a	390k	31	0.92	0.85-0.96	U5b2b	ds UDG	0.04	39.7%	Fail	Yes
	GA77	GX45	0.5	390k					ds noUDG	0.43	n/a	Damage	Yes
Paglicci108	GA264	B1	4.23	1240k	20	0.93	0.86-0.98	U2'3'4'7'	ds noUDG	0.12	n/a	Damage	Yes
BerryAuBac	GA261	GX81	0.2	1240k	107	0.97	0.94-0.99	U5b1a	ds noUDG	0.46	-1.6%	Pass	Yes
Bockstein	GA165	GX37	n/a	390k	267	0.97	0.95-0.99	U5b1d1	ds UDG	0.03	n/a	Fail	Yes
	GA89	GX37	2.67	390k					ds noUDG	0.25	n/a	Pass	Yes
Brillenhohle	GA163	GX52	n/a	390k	19	0.9	0.82-0.95	U8a	ds UDG	0.01	26.2%	Fail	Yes
	GA79	GX52	1.08	390k					ds noUDG	0.16	14.3%	Damage	Yes
Burkhardtshohle	GA260	GX53	0.35	1240k	45	0.95	0.89-0.98	U8a	ds noUDG	0.26	15.7%	Damage	Yes
Paglicci133	GA252	C2	3.67	1240k	28	0.83	0.74-0.89	U8c	ds noUDG	0.41	-0.7%	Pass	Yes
	MA160	C2	n/a	1240k					ss noUDG	0.08	n/a	Fail	Yes
Cioclovina1	GA259	GX51	0.21	1240k	19	0.88	0.76-0.95	U	ds noUDG	0.20	15.9%	Damage	Yes
Chaudardes1	CRC	CRC	1.65	1240k	17	0.93	0.91-0.95	U5b1b	ds noUDG	0.35	18.3%	Damage	Yes
	MA169	CRC	n/a	1240k					ss UDG	0.28	n/a	Damage	Yes
Paglicci71	GA265	FA	0.27	1240k	12	0.83	0.68-0.92	U5b2b	ds noUDG	0.21	44.4%	Damage	No
Falkenstein	FLA	FL	5.45	390k	600	0.97	0.97-0.98	U5a2c	ds noUDG	0.23	7.7%	Damage	Yes
Falkenstein	GA53	FL	n/a	390k					ds UDG	0.02	9.6%	Fail	Yes
Falkenstein	GA54	FL	n/a	390k					ds UDG	0.02	6.3%	Fail	Yes
Felsdach	GA258	GX49	0.5	1240k	55	0.76	0.67-0.80	U5b2a	ds noUDG	0.04	47.2%	Damage	No
HohleFels49	GA164	GX55	n/a	390k	364	0.99	0.98-0.99	U8a	ds UDG	0.03	5.5%	Fail	Yes
	GA82	GX55	1.2	390k					ds noUDG	0.31	4.9%	Damage	Yes
HohleFels79	GA166	GX47	n/a	390k	42	0.98	0.95-1.00	U8a	ds UDG	0.02	4.2%	Fail	Yes
	GA90	GX47	1.52	390k					ds noUDG	0.19	n/a	Damage	Yes
HohlensteinStadel	MA162	VE	n/a	1240k	34	0.77	0.69-0.84	U5b2c1	ss noUDG	0.14	n/a	Damage	No
	MA171	VE	n/a	1240k					ss noUDG	0.15	n/a	Damage	No
LaRochette	GA253	LA	0.28	1240k	40	0.80	0.67-0.88	M	ds noUDG	0.09	7.0%	Damage	No
LesCloseaux13	GA256	GX43	0.17	1240k	19	0.98	0.91-1.00	U5a2	ds noUDG	0.27	n/a	Pass	Yes
Ofnet	GA167	GX50	n/a	390k	1859	1	1.00-1.00	U5b1d1	ds UDG	0.02	n/a	Fail	Yes
	GA93	GX50	1.25	390k					ds noUDG	0.21	n/a	Pass	Yes
GoyetQ116-1	GA63	GX58	4.88	1240k	56	0.99	0.97-1.00	M	ds noUDG	0.22	1.0%	Pass	Yes
	MA167	GX58	n/a	1240k					ss UDG	0.25	0.9%	Pass	Yes
GoyetQ-2	GA231	GX176	8.6	1240k	406	1	0.99-1.00	U8a	ds noUDG	0.30	5.5%	Damage	No
	MA166	GX176	n/a	1240k					ss UDG	0.28	3.9%	Damage	No
Goyet2878-21	GA248	GX177	1.76	1240k	21	0.99	0.95-1.00	U5	ds noUDG	0.16	n/a	Damage	No
	MA168	GX177	n/a	1240k					ss UDG	0.11	8.0%	Damage	No
GoyetQ376-19	GA250	GX60	0.9	1240k	43	0.92	0.86-0.95	U2	ds noUDG	0.13	n/a	Damage	Yes
	MA158	GX60	n/a	1240k					ss noUDG	0.26	27.6%	Damage	Yes
GoyetQ376-3	MA161	GX59	0.27	1240k	46	0.9	0.82-0.94	M	ss noUDG	0.17	n/a	Damage	No
	MA170	GX59	n/a	1240k					ss noUDG	0.24	-7.1%	Damage	No
GoyetQ53-1	GA251	GX64	0.62	1240k	48	0.82	0.73-0.88	U2	ds noUDG	0.13	32.4%	Damage	Yes
	MA159	GX64	n/a	1240k					ss noUDG	0.16	6.9%	Damage	Yes
GoyetQ55-2	GA254	GX62	0.1	1240k	17	0.87	0.78-0.92	U2	ds noUDG	0.16	n/a	Damage	No
GoyetQ56-16	GA255	GX63	0.19	1240k	45	0.84	0.77-0.89	U2	ds noUDG	0.18	24.6%	Damage	Yes
Ranchot88	GA262	GX83	20.65	1240k	86	0.99	0.96-1.00	U5b1	ds noUDG	0.39	n/a	Pass	Yes
Rigney1	GA263	GX89	0.53	1240k	41	0.9	0.86-0.93	U2'3'4'7'	ds noUDG	0.22	n/a	Damage	Yes
Rochedane	GA127	GX96	4.25	1240k	104	0.98	0.96-0.99	U5b2b	ds noUDG	0.39	5.6%	Damage	Yes
	MA165	GX96	n/a	1240k					ss UDG	0.38	1.6%	Pass	Yes

Note: Cells in red have mitochondrial contamination estimates that fail our criteria: <95% for the point estimate, or a ninety five percent confidence interval that includes values <85%.
The final column indicates whether the sample was among the 51 used in the genome-wide analysis (some samples were not used because there were <4,000 SNPs after quality control).
* Some samples do not have an X chromosome contamination estimate because of likely female sex or because fewer than 200 SNPs were covered at least twice.
+ Libraries marked "Fail" are not used in analysis. For libraries marked "Damage," analyses are restricted to fragments showing deamination.

Libraries generated at Harvard Medical School were prepared using double-stranded protocols and were UDG-treated in a way that retains some characteristic ancient DNA damage at the last nucleotide[7]. We denote these libraries as "ds partial UDG" in Table S2.3.

Table S2.3: Libraries prepared at Harvard Medical School

Sample	Library	Extract	Endogenous%	SNPs	mt cov	Match consens. Est.	Match consens. 95% CI	mt. Hg	Library type	Dam-age	Chr. X contam.	Library decision	Analyzed?
KremsWA3	S1577.E1.L2	S1577.E1	0.50	1240k	85	1.00	1.00-1.00	U5	ds partial UDG	0.084	Pass	n/a	Yes
	S1577.E1.L3	S1577.E1	0.39	1240k	43	1.00	1.00-1.00	U5	ds partial UDG	0.077	Pass	1.8%	Yes

In solution capture of mitochondrial DNA

We hybridized the libraries to oligonucleotide probes overlapping the mitochondrial DNA genome (mtDNA). We used the method of ref. [8] for the libraries from MPI-EVA and of ref. [9] for the libraries from Tübingen. We sequenced the enriched libraries on the Illumina MiSeq or HiSeq2500 platforms using a double index configuration (2×75bp or 2x100bp reads).

To analyse the mtDNA capture data for the MPI-EVA libraries, we demultiplexed the reads according to the expected index pairs, allowing one mismatch for each. We merged paired-end reads into a single fragment by requiring an overlap of at least 11 bp (with one mismatch allowed), using a modified version of SeqPrep[10] in which the base and quality score is determined by the read that has higher quality. After stripping adapters, we mapped merged fragments which we required to be at least 30bp in length to the mtDNA revised Cambridge Reference Sequence (rCRS) with BWA (v0.6.1) using the *samse* command. We identified duplicated fragments based on having the same orientation, start and end positions, and kept the highest quality fragment. We excluded fragments with mapping quality <30.

To analyse the mtDNA capture data for the University of Tübingen libraries, we clipped adapters and merged paired reads that overlapped by at least 10 bp using the program Clip&Merge[11]. We restricted to merged fragments, and filtered out ones less than 30 bp in length. We mapped fragments to the mtDNA revised Cambridge Reference Sequence (rCRS) and removed duplicates[11]. We excluded fragments with mapping quality below 30.

In solution capture of nuclear DNA for 38 samples

We hybridized libraries in solution to oligonucleotide probes synthesized by Agilent Technologies[8]. For 8 samples, we enriched for a targeted set of 394,577 SNPs (SNP Panel 1, "390k"), using the probe sequences specified in Supplementary Data 2 of ref. [12] (http://www.nature.com/nature/journal/v522/n7555/abs/nature14317.html#supplementary-information). For 16 samples, we enriched for ~1.24 million SNPs (SNP Panels 1 and 2, "1240k"), and for 14 samples we enriched for ~3.7 million SNPs (SNP Panels 1, 2, 3 and 4, "3.7M"), using the additional probe sequences specified in Supplementary Data 1, 2 and 3 of [13] (http://www.nature.com/nature/journal/v524/n7564/full/nature14558.html#supplementary-information). Extended Data Table 1, as well as Table S2.1, Table S2.2 and Table S2.3, specify which samples were enriched for which targeted set of SNPs.

Sequencing and alignment to the nuclear genome

We generated 2×75bp reads on Illumina HiSeq2500 or NextSeq500 instruments. We processed the fragments as for mtDNA, except we aligned to the human reference genome, *hg19*, and required an overlap of at least 15 bp. We mapped with the command bwa –n 0.01 and –l 16500.

Four tests for contamination

(1) We required all analysed libraries to have a damage profile consistent with ancient DNA

All single stranded libraries ("ss UDG" and "ss noUDG"), as well as all non-UDG-treated double stranded libraries ("ds noUDG") retain damage in the last nucleotide. We restricted analyses of such libraries to ones in which ≥10% of terminal nucleotides that are cytosines in the reference genome are read as thymines, as expected for authentic ancient DNA[14]. For the "ds partial UDG" protocol, we restricted analysis to libraries that had a rate of ≥3% cytosine-to-thymine substitutions in the terminal nucleotide[7]. To be conservative, for our whole

genome analyses, we did not use any data from double stranded UDG-treated libraries ("ds UDG"), as this protocol does not retain damage at the terminal nucleotide.

(2) We tested for contamination based on the match rate to the mtDNA consensus
We used mtDNA data to flag samples as possibly contaminated (in red in Table S2.1 and Table S2.2). We declared a library possibly contaminated if, after running the contamination estimator ContamMix[8], either of two criteria were met:
(i) The fraction of fragments matching the reconstructed consensus better than any of 311 worldwide mtDNA sequences used for comparison is <95%.
(ii) The ninety-five percent confident lower bound of the fraction of fragments matching the consensus better than any of 311 worldwide mtDNA sequences is <85%.

(3) We tested for contamination based on consistency of damaged and undamaged fragments
For each sample, we restricted to fragments with characteristic ancient DNA damage, and checked whether the population genetic affinities inferred from damaged fragments match those from the consensus of all fragments. In the case of mtDNA data, only *Vestonice14* showed a change in haplogroup comparing damaged fragments to all fragments, suggesting mtDNA contamination. Further evidence of contamination in *Vestonice14* comes from the fact that when we determine sex based on the proportion of Y chromosome fragments (Supplementary Information section 3), the sex for this individual switched from female when all fragments are analysed to male for damaged fragments only. The genetically determined sex of *Ostuni2*, *Paglicci108*, *GoyetQ53-1*, *GoyetQ376-19*, *GoyetQ56-16*, *GoyetQ55-2* and *Goyet2878-21* changes from male when all fragments are analysed to female for damaged fragments only.

(4) We tested for contamination based on the X chromosome polymorphism rate in males
Males have only one X chromosome and thus are not expected to be polymorphic in this part of their genome. This can be used to obtain a conservative estimate of contamination in males given sufficient X chromosome coverage[15,16]. We used the ANGSD software to run this test on all males where it gave good resolution (we only used X chromosome estimates for males for whom at least 200 SNPs covered at least twice). We considered libraries as effectively uncontaminated if their X chromosome contamination estimates were less than 2.5%.

Selecting libraries for analysis and restricting to damaged fragments
For nine male individuals, we identified libraries with no evidence of contamination by the four criteria above, and with high enough coverage to perform a chromosome X contamination estimate and to determine that the contamination was less than 2.5%. We used all fragments from these libraries. For the remaining individuals, we restricted to fragments carrying characteristic ancient DNA damage at their terminal ends, which is known to reduce contamination albeit at the cost of losing data[17,18] (Box S2.1). Specifically:

(i) We restricted to damaged fragments for all libraries with evidence of contamination according to criteria 1-4, and that we did not fail outright (Table S2.1 and Table S2.2). We failed libraries outright if (a) they had evidence of contamination and were from samples that had multiple libraries some of which had no evidence of contamination, or (b) if they were prepared with a double-stranded UDG-treated protocol ("ds UDG"), as damage is not retained for this class of libraries even in the last nucleotide.

(ii) We restricted to damaged fragments for libraries with evidence of contamination based on criteria 1-3, and for which we could not perform an X contamination estimate. While this is a severe step—for example, it means that we restricted to damaged fragments for all

female samples—we decided that it was the only safe thing to do in cases where we could not perform an X chromosome contamination assessment. In particular, we found that a substantial fraction of male samples with no evidence of contamination by criteria 1-3 gave evidence of substantial contamination by the X chromosome method.

After damage restriction, we merged fragments from all libraries from the same sample. At each SNP covered at least once, we used a randomly sampled fragment to determine an allele. Thus, we do not attempt to determine diploid genotypes, and none of the SNPs are assigned a heterozygous genotype. The final dataset is given in Extended Data Table 1, after restricting to samples with at least 4,000 SNPs hit at least once.

Box S2.1. Strategy used to retain damaged fragments for contaminated libraries

ss UDG-treated libraries: Restrict to fragments with C→T substitutions in the first position at the 5'-end and the last two positions at the 3'-end.

ss noUDG-treated libraries: Restrict to fragments with C→T substitutions in the first three positions at the 5'-end and the last three positions bases at the 3'-end.

ds UDG- partial treated libraries: Restrict to fragments with C→T substitutions in the first position at the 5'-end and G→A substitutions in the last position at the 3'-end.

ds noUDG-treated libraries: Restrict to fragments with C→T substitutions in the first three positions at the 5'-end, and G→A substitutions in the last three positions on the 3'-end.

ds UDG-treated libraries: Cannot restrict to damaged fragments so do not use.

References

1 Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* **110**, 15758-15763, doi:10.1073/pnas.1314445110 (2013).

2 Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222-226, doi:10.1126/science.1224344 (2012).

3 Gansauge, M. T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nature protocols* **8**, 737-748, doi:10.1038/nprot.2013.038 (2013).

4 Briggs, A. W. *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14616-14621, doi:10.1073/pnas.0704665104 (2007).

5 Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic acids research* **40**, e3, doi:10.1093/nar/gkr771 (2012).

6 Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor protocols* **2010**, pdb prot5448, doi:10.1101/pdb.prot5448 (2010).

7 Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370**, 20130624, doi:10.1098/rstb.2013.0624 (2015).

8 Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2223-2227, doi:10.1073/pnas.1221359110 (2013).

9 Maricic, T., Whitten, M. & Paabo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PloS one* **5**, e14004, doi:10.1371/journal.pone.0014004 (2010).
10 John, J. S. *SeqPrep*, <https://github.com/jstjohn/SeqPrep> (2011).
11 Peltzer, A. *et al.* EAGER: efficient ancient genome reconstruction. *Genome biology* **17**, 60, doi:10.1186/s13059-016-0918-z (2016).
12 Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, doi:10.1038/nature14317 (2015).
13 Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, doi:10.1038/nature14558 (2015).
14 Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Paabo, S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PloS one* **7**, e34131, doi:10.1371/journal.pone.0034131 (2012).
15 Rasmussen, M. *et al.* An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334**, 94-98, doi:10.1126/science.1211177 (2011).
16 Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC bioinformatics* **15**, 356, doi:10.1186/s12859-014-0356-4 (2014).
17 Skoglund, P. *et al.* Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 2229-2234, doi:10.1073/pnas.1318934111 (2014).
18 Meyer, M. *et al.* A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**, 403-406, doi:10.1038/nature12788 (2014).

Section 3
Evidence for a decrease in Neanderthal ancestry over time

Change in Neanderthal ancestry over time as inferred from f_4-ratio statistics

We used f_4-ratio statistics[1] to estimate the proportion of Neanderthal ancestry in the 51 ancient Eurasian samples analysed in this study as well as in a handful of present-day Europeans and East Asians for comparison. We used the dataset of approximately 2.2 million SNPs (the SNPs captured in Panel 1, Panel 2 and Panel 3) (Supplementary Information section 2).

To estimate Neanderthal ancestry proportion in each individual, we used the previously reported f_4-ratio statistic shown in Equation S3.1[2-4]. Under the assumption that West and Central African populations are outgroups to East African *Dinka* and to the modern human ancestry in non-Africans today, this quantity provides an unbiased estimate of the proportion of modern human ancestry in a present-day non-African. We use 1 minus this quantity to estimate the Neanderthal ancestry. We compute a Block Jackknife on non-overlapping 5 centimorgan blocks to determine empirical standard errors[5]. There are a number of f_4-ratio statistics that have been proposed for estimating Neanderthal ancestry, and while each has different merits, some have large standard errors[4]. We use Equation S3.1 as we find empirically that it has relatively small standard errors.

$$Q(Test) = 1 - \frac{f_4(West\_and\_Central\_Africans, Chimp; Test, Archaic)}{f_4(West\_and\_Central\_Africans, Chimp; Dinka, Archaic)} \qquad \text{(Equation S3.1)}$$

We compute allele frequencies by pooling data from each of the following sets of samples:

West_and_Central_Africans: a pool of 9 samples from the Mbuti, Yoruba and Mende populations (S_Mbuti-1, S_Mbuti-2, S_Mbuti-3, B_Mbuti-4, S_Yoruba-1, S_Yoruba-2, S_Yoruba-3, S_Mende-1, S_Mende-2)

Dinka: a pool of 3 samples (S_Dinka-1, S_Dinka-2, B_Dinka-3)

Archaic: a pool of 2 samples (Altai Neanderthal and the Siberian Denisovan)

For *Test*, we analyse 51 ancient individuals along with 8 Europeans (S_French-1, S_French-2, B_French-3, S_English-1, S_English-2, S_Sardinian-1, S_Sardinian-2, B_Sardinian-3) and 7 East Asians (S_Dai-1, S_Dai-2, S_Dai-3, B_Dai-4, S_Han-1, S_Han-2, B_Han-3).

Extended Data Table 2 presents the results for all samples, while Figure 2 plots the results for the subset of individuals with at least 200,000 SNPs covered (these gave low enough standard errors to produce a visually clear plot). Figure 2 reveals a clear decrease in Neanderthal ancestry proportion over time. There is one outlier—*Oase1*—which has previously been shown to have a Neanderthal ancestor 4-6 generations back in its family tree[4], and which we estimate here had 9.9 ± 0.8% Neanderthal ancestry. Such an ancestor is recent enough that *Oase1* is not expected to be representative of the population in which he lived. Here we are interested in how Neanderthal ancestry proportion changed over time in populations that reached equilibrium; that is, in populations in which all individuals had approximately the same Neanderthal ancestry proportion. Thus, we remove *Oase1* for most of the analyses that

follow. This is conservative for testing whether there has been a decrease in the proportion of Neanderthal ancestry over time.

To test whether the slope of the fitted line is significantly negative, we needed to compute a standard error on the slope taking into account the fact that we had information from multiple individuals with different amounts of data and different extents of correlated history. To do this, we used a Weighted Block Jackknife. As this is a novel context in which to use a Weighted Block Jackknife for population history analysis, we describe this in more detail.

Let n_i be our estimate of the mean Neanderthal ancestry in sample i using the f_4-ratio. We estimate the covariance matrix $\boldsymbol{V}$ of the errors using a Weighted Block Jackknife[5,6] (with 5 centimorgan blocks and weights equal to the number of SNPs used). Set $\boldsymbol{Q} = \boldsymbol{V}^{\boldsymbol{1}}$. If a is the ancestry proportion now (time=0) and s is the "slope" in units of ancestry per year, then we can model n_i as:

$$n_i = a + sd_i + e_i \qquad \text{(Equation S3.2)}$$

Here, d_i is the date in years (BP) of sample i and the error e_i has covariance $\boldsymbol{V}$. It is now natural to estimate a and s by minimizing:

$$L(a,s) = \sum_{i,j} Q_{ij}\,(n_i - a - sd_i)\big(n_j - a - sd_j\big) \qquad \text{(Equation S3.3)}$$

This is a generalized least squares problem. If we set $\boldsymbol{L_2}=max(\boldsymbol{L(a,\ s)})$ and $\boldsymbol{L_1}=max(\boldsymbol{L(a,\ 0)})$, then $(L_1\text{-}L_0)$ is approximately χ^2 with 1 degree of freedom under the null (s=0).

Extended Data Table 3 shows the results of the regression analysis for different subsets of samples. Our "Core Set 1" of individuals consists of 50 ancient modern humans (removing *Oase1* as an outlier) along with 7 East Asians that we use to represent present-day individuals. We use East Asians rather than Europeans to represent present-day individuals, since present-day East Asians are known to harbor more Neanderthal ancestry than Europeans[3,7,8]. Using East Asians to represent present-day people is conservative for a test that searches for evidence of a decrease in Neanderthal ancestry proportion over time.

For our Core Set 1 of samples, we observe a highly significant $P=5\times10^{-22}$) correlation of Neanderthal ancestry with sample date. We also observe highly statistical significant signals for a 11 alternate sample sets (always in the range ($10^{-29}<P<10^{-11}$):

- 5×10^{-22} for Core Set 1
- 2×10^{-15} remove the oldest samples (restrict to <32 kya)
- 4×10^{-18} remove most Věstonice Cluster samples (restrict to >32 kya or <25 kya)
- 5×10^{-21} remove the El Mirón and Mal'ta Cluster samples (restrict to >25 kya or <14 kya)
- 2×10^{-18} remove Villabruna Cluster, Neolithic samples (restrict to >14 kya or present-day)
- 4×10^{-15} remove present-day samples (restrict to ancient samples)
- 4×10^{-19} restrict to samples with >200,000 SNPs (same set of samples as in Figure 2)
- 2×10^{-23} replace East Asians with Europeans
- 8×10^{-29} add *Oase1*
- 1×10^{-20} restrict to ancient samples only (including *Oase1*)
- 8×10^{-12} restrict to ancient samples only (but excluding both *Oase1* and *UstIshim*)

Having established a significant decrease in Neanderthal over time, we used the parameters of the least squares fit to estimate the rate of decrease in Neanderthal ancestry. There is no theoretical reason to expect that the proportion of Neanderthal ancestry in modern humans decreased at a linear rate immediately after introgression. Indeed, reasonable models of selection against Neanderthal ancestry suggest that there might have been a much faster decrease in the proportion of Neanderthal ancestry under the pressure of natural selection for the first couple of dozens of generations after introgression[9]. However, the line is useful in providing a meaningful measurement of the rate of decrease in Neanderthal ancestry over most of the period since Neanderthals and modern humans interbred. By extrapolating to the time when introgression occurred, we can estimate what the Neanderthal ancestry proportion was some time after introgression, and can likely obtain a minimum estimate.

The results of the least squares fitting are summarizing in Extended Data Table 3. For the Core Set 1 of samples, we estimate a 0.48-0.73% decrease in Neanderthal ancestry per 10,000 years (95% confidence interval). We obtain similar intervals for other sample subsets.

To obtain an estimate of the proportion of Neanderthal DNA in the ancestral population some time after introgression occurred, we need to make an assumption about when Neanderthal introgression occurred. If we use a date of 55,000 years, the best estimate from the analyses of the *UstIshim*[10] and *Kostenki14*[11] genomes, we infer a 95% confidence interval of 4.3-5.7%. We also considered how uncertainty in the date of Neanderthal admixture—estimated to be 50,000-60,000 years ago in the analysis of the *UstIshim* genome—affects this estimate[10]. Assuming 50,000 years, we estimate 4.0-5.4%, and assuming 60,000 years, we estimate 4.5-6.0%). Thus, a conservative range is 4.0-6.0%.

The inferred 4.0-6.0% Neanderthal ancestry some time after introgression is substantially higher than the proportion of Neanderthal ancestry in Eurasian populations today, which in the same analyses we estimate to be 1.1-2.2%. This is higher than the values of ~3-3.5% considered by Juric et al.[12] and Harris and Nielsen[9]. An important direction for future work is to explore what distributions of selection coefficients and demographic scenarios might produce this inferred 2-4-fold magnitude of reduction of Neanderthal ancestry in present-day compared to ancient individuals.

Neanderthal ancestry estimates from ancestry informative SNPs

The f_4-ratio statistic analysis depends on assumptions about the deep historical relationships of the populations used in the statistic. To obtain an estimate of Neanderthal ancestry that is not strongly dependent on such assumptions, we analysed SNPs where Neanderthals carry an allele that differs from the great majority of present-day sub-Saharan Africans. Such alleles have a high chance of deriving from Neanderthal ancestors, and by counting these in each individual, it is possible to obtain a number that is linearly related to and highly sensitive to an individual's proportion of Neanderthal ancestry[3,4].

We used in-solution hybridization capture[13] to enrich the libraries from 15 individuals (the individuals listed in Table S2.1 in Supplementary Information section 2) for 1,749,385 SNPs where all Yoruba individuals in the 1000 Genomes Project are fixed or nearly fixed for one allele, and an archaic genome (a Denisovan or a Neanderthal) carries another allele[4]. For the analyses that follow, we focused on a subset of 783,747 specifically Neanderthal informative SNPs where a randomly selected allele from the high coverage Altai Neanderthal genome differs from the great majority of Yoruba in the 1000 Genomes project, and also from a randomly selected allele in the deeply sequenced Yoruba individual from "Panel B" of ref. [8].

We also further required that the analysed SNPs had an assignment of a B-statistic measuring the intensity of linked selection at the position[14].

We sequenced the enriched libraries on the Illumina HiSeq 2500 platform using a double index configuration (2 x 76bp), and spiked in an indexed ΦX174 control library in each sequencing run. We performed base calling using the machine-learning algorithm implemented in freeIBIS[15], merged overlapping pair-end reads[3], and mapped them to the human reference genome (*hg19*) using the Burrows-Wheeler Aligner (BWA)[16]. We adjusted BWA parameters to allow for more mismatches and indels and disabled seeding ("-n 0.01 –o 2 –l 16500"), as appropriate for analysis of error-prone ancient DNA sequences[3].

We restricted our analyses to merged fragments that had perfect matches to the expected index combinations. After removing fragments identified as being likely duplicates (having the same mapped orientation and start and stop positions) (https://github.com/udo-stenzel/biohazard), we retained for analysis fragments that were longer than 35bp and that had a mapping quality of at least 37. For some individuals, we further restricted to fragments with evidence of deamination (Extended Data Table 1). To do this for UDG-treated libraries, we retained fragments that showed C→T substitutions in the first base at the 5' and/or at the last two bases at the 3'-end relative to *hg19*. For the non-UDG treated libraries (from *AfontovaGora3*), we retained fragments with a C→T substitution at the first three and/or the last three bases. We merged data from all libraries from the same individual, and in Extended Data Table 2 report the number of fragments overlapping positions of interest and the number of SNPs covered at least once. We reprocessed the previously reported *Oase1*[2] data using the same workflow for consistency.

We co-analysed these data with shotgun sequence data from 12 ancient modern humans and a pool of three Vindija Neanderthals to represent the rate of matching to Altai for an individual with 100% Neanderthal ancestry (we obtained qualitatively similar results when replacing the Vindija Neanderthals with the Mezmaiskaya Neanderthal). We represented present-day humans using 7 individuals drawn from "Panel B" of ref.[8] (6 non-Africans and a Dinka East African to represent an individual with 0% Neanderthal ancestry).

To compute the percentage of alleles matching Neanderthal, we represented each individual by a single DNA fragment at each targeted SNP. For the high coverage individuals (*UstIshim* and the present-day humans), we randomly sampled one of the two alleles from the genotype file. For the low coverage individuals, we randomly sampled an allele using the same procedure described in Supplementary Information section 2.

We converted the fraction of alleles matching *Altai* for each *Test* individual into an estimate of Neanderthal ancestry by assuming that the proportion of Neanderthal ancestry can be estimated by how far a *Test* individual's rate of matching E_{Test} is along the path from an individual assumed to have 0% Neanderthal ancestry (E_{Dinka}, the value seen in a *Dinka* individual from East Africa), and 100% Neanderthal ancestry ($E_{Vindija}$, the value seen in then pool of three *Vindija* Neanderthals from Croatia[17]). Concretely, we used the equation:

$$R(Test) = \frac{E_{Test} - E_{Dinka}}{E_{Vindija} - E_{Dinka}} \qquad \text{(Equation S3.4)}$$

We computed a standard error on *R(Test)* using a Block Jackknife with 100 equal blocks[5].

Extended Data Table 2 and Extended Data Figure 1 reveal that *Oase1* has an elevated rate of alleles matching to Neanderthal. This replicates a previous finding[4] and is also qualitatively consistent with the findings from the f_4-ratio statistic analysis. We also replicate previous findings in estimating that *UstIshim* has an elevated rate of alleles matching Neanderthal compared to present-day humans[4].

We performed a least squares fit to the scatterplot of sample age versus estimated Neanderthal ancestry proportion. We repeated this procedure leaving out each of 100 equally sized contiguous chunks of SNPs in turn, and used jackknife statistics to compute a standard error on the two regression parameters (slope and intercept). We tested whether the slope is non-zero by evaluating the number of standard errors it is from zero. We then used this as a significance test for whether Neanderthal ancestry proportion has changed over time. We determined the intercept at 50,000, 55,000, and 60,000 years ago to obtain estimates of the proportion of Neanderthal ancestry shortly after introgression.

The results are shown in Extended Data Table 3. We observe a significant decrease in Neanderthal ancestry over time, whatever subset of samples we analyse, just as for the f_4 ratio analysis. For the "Core Set 2" of 29 samples (all the ancient samples for which we had data with the exception of *Oase1*, and adding in *Han*, *Dai*, and *Karitiana* to represent present-day Eurasians), we observe a significant ($P=4.0\times10^{-11}$) correlation of Neanderthal ancestry with sample date. We also observe significant signals for alternate sample sets:

- 4.0×10^{-11} for Core Set 2
- 1.1×10^{-4} remove *Han, Dai*, *Karitiana* and *Stuttgart*
- 1.6×10^{-4} remove *Han, Dai*, *Karitiana, Stuttgart, and UstIshim*

For the Core Set 2 group of samples, the least squares fit corresponds to an estimated 0.21-0.39% decrease in Neanderthal ancestry every 10,000 years (95% confidence interval based on a standard error from the Block Jackknife), and an extrapolated 3.2-4.2% Neanderthal ancestry 55,000 years ago. These results support the conclusion that the proportion of Neanderthal ancestry has decreased over time, consistent with the findings of the f_4-ratio analysis in the previous section, although the magnitude of the estimated reduction is smaller.

The reduction in Neanderthal ancestry has been faster closer to genes

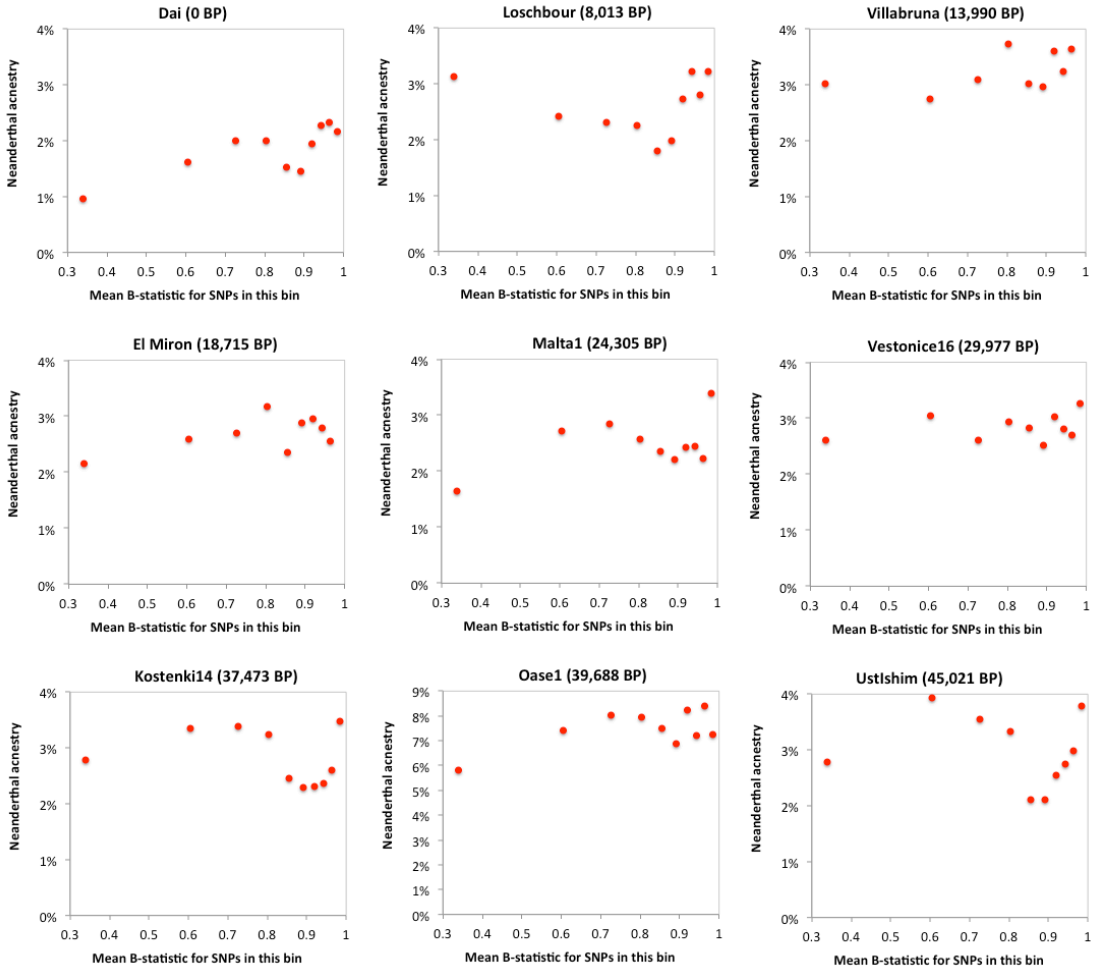
We took advantage of the power of ancestry informative SNPs to study whether there is evidence for a more rapid decrease in Neanderthal ancestry closer to genes. This pattern is predicted based on the reduced Neanderthal ancestry close to genes that is observed today[18,19].

To test whether the reduction in Neanderthal ancestry over time has been more rapid closer to genes, we stratified the genome based on a previously published *B* statistic, which provides a measure of the strength of loss of diversity in a region due to proximity to functionally important regions[14]. We divided the SNPs sites that we studied as being informative about ancestry into 10 bins of equal size based on B, ranging from most strongly constrained by selection (bin 1) to least constrained (bin 10). We calculated the proportion of Neanderthal ancestry in each bin, and converted it into an estimate of Neanderthal ancestry.

Figure S3.1 shows that if there is a trend in any sample toward an increase in Neanderthal ancestry with increasing B-statistic, it is very noisy. We fit a line to the individual-specific plots of Neanderthal ancestry against B-statistic and computed a standard error using Block Jackknife. The regression coefficients and standard errors are presented in Extended Data

Table 3. As expected from the noisy by-individual plots (Figure S3.1), the evidence for positive slopes is weak. The strongest evidence for a positive slope (as measured by a Z-score formed from the ratio of the estimate to the standard error), is Z=3.3 for *Dai*, followed by Z=2.7 for *Han*, Z=2.1 for *French*, and Z=1.6 for *Loschbour*. The strongest evidence for a negative slope is non-significant after multiple-hypothesis testing (Z=-2.4 for *Continenza*).

Figure S3.1. Empirical correlation of Neanderthal ancestry estimates to B-statistic bin. We show results for a selected set of relatively high coverage samples.



Despite the noisy individual estimates, we observed that the fitted slopes for the regressions against B-statistics tended to be positive (Extended Data Table 2), and that this was especially true for more recent samples. We were also motivated to study these patterns based on previous evidence for a significant depletion of Neanderthal ancestry in functional regions of present-day Europeans and East Asians based on Neanderthal introgression maps[18,19].

To test formally whether the dependence of Neanderthal ancestry on B-statistic is stronger for more recent periods, we carried out a least squares fit to the slopes of the "Main set" of samples in Extended Data Table 2, weighting the information from each sample by their empirical standard errors (which can also be read off of Extended Data Table 2). We used a Block Jackknife over the whole procedure to determine the standard error ("slope of slopes"). We observe a weakly significant reduction in the dependence on B-statistic with increasing sample date (Z=-2.32; P=0.010). These results are consistent with the theory that selection against Neanderthal alleles over time has been most intense closest to genes.

References

1 Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature* **461**, 489-494, doi:10.1038/nature08365 (2009).

2 Reich, D. *et al*. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053-1060, doi:nature09710 [pii]

10.1038/nature09710 (2010).

3 Meyer, M. *et al*. A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science*, doi:10.1126/science.1224344 (2012).

4 Fu, Q. *et al*. An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, doi:10.1038/nature14558 (2015).

5 Kunsch, H. R. THE JACKKNIFE AND THE BOOTSTRAP FOR GENERAL STATIONARY OBSERVATIONS. *Annals of Statistics* **17**, 1217-1241, doi:10.1214/aos/1176347265 (1989).

6 Busing, F., Meijer, E. & Van Der Leeden, R. Delete-m jackknife for unequal m. *Statistics and Computing* **9**, 3-8, doi:10.1023/a:1008800423698 (1999).

7 Wall, J. D. *et al*. Higher levels of neanderthal ancestry in East Asians than in Europeans. *Genetics* **194**, 199-209, doi:10.1534/genetics.112.148213 (2013).

8 Prufer, K. *et al*. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43-49, doi:10.1038/nature12886 (2014).

9 Harris, K., Nielsen, R. The Genetic Cost of Neanderthal Introgression. *biorxiv.org* http://dx.doi.org/10.1101/030387 (2015).

10 Fu, Q. *et al*. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449, doi:10.1038/nature13810 (2014).

11 Seguin-Orlando, A. *et al*. Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).

12 Juric, I., Aeschbacher, S., Coop, C. The Strength of Selection Against Neanderthal Introgression. *biorxiv.org* http://dx.doi.org/10.1101/030148 (2015).

13 Fu, Q. *et al*. DNA analysis of an early modern human from Tianyuan Cave, China. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2223-2227, doi:10.1073/pnas.1221359110 (2013).

14 McVicker, G., Gordon, D., Davis, C. & Green, P. Widespread genomic signatures of natural selection in hominid evolution. *PLoS genetics* **5**, e1000471, doi:10.1371/journal.pgen.1000471 (2009).

15 Renaud, G., Kircher, M., Stenzel, U. & Kelso, J. freeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics* **29**, 1208-1209, doi:10.1093/bioinformatics/btt117 (2013).

16 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

17 Green, R. E. *et al*. A draft sequence of the Neandertal genome. *Science* **328**, 710-722, doi:10.1126/science.1188021 (2010).

18 Sankararaman, S. *et al*. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354-357, doi:10.1038/nature12961 (2014).

19 Vernot, B. & Akey, J. M. Resurrecting surviving Neandertal lineages from modern human genomes. *Science* **343**, 1017-1021, doi:10.1126/science.1245938 (2014).

Section 4
Sex determination and Y chromosome analyses

Sex determination
For each individual, we extracted merged sequences of high mapping quality (MAPQ ≥ 37) using samtools 0.1.18. We counted the number of SNPs on the autosomal, chromosome X and chromosome Y targets covered by at least once sequence. We restricted this analysis to the smaller 1240k SNP target set for the samples captured for the 3.7M SNP target set (Extended Data Table 4).

It is tempting to interpret the ratios N_X/N_{auto} and N_Y/N_{auto} for each sample in Extended Data Table 4 as directly informative about sex. However, since the ratio of Autosome:X:Y targets is very different for the 1240k compared to the 390k SNP target set, we need to make an adjustment for the target set. We define "X-rate" as the $N_X/N_{autosomes}$ rate for a sample, divided by the expected value of this quantity based on the number of SNPs in the relevant target set (similarly for the "Y-rate"). For the 1240k capture, these normalizing quantities 0.0432 (X target set) and 0.0284 (Y target set). For the 390k capture set they are 0.0047 (X target set) and 0.0058 (Y target set).

We empirically observe two clusters in the X-rate to Y-rate scatterplot, corresponding to males and females (Figure S4.1). The X-rate histogram does not show a clear separation between males and females but the Y-rate histogram does. We ascribe the greater usefulness of the Y chromosome information to the fact that for true females, the expectation of the number of Y chromosome sequences is extremely low (because females do not have a Y chromosome), and has an extremely low standard deviation, and this makes its distribution simple to distinguish from the non-zero male expectation. In contrast, distinguishing between the expectations for the two sexes for the X chromosome sequences is more difficult because of the high empirical standard deviation for the male and female distributions, which means that the distributions are overlapping even though the means are different by a factor of two. We suspect that the reason why there are large empirical standard deviations when the true number of sequences in the library is not zero (Y chromosomes in females) is bias in the capture experiment in terms of which SNPs were effectively targeted.

Guided by the marginal histogram on the right of Figure S4.1, we determine genetic sex by the rule that Y-rate < 0.05 for a female and Y-rate > 0.2 for a male.

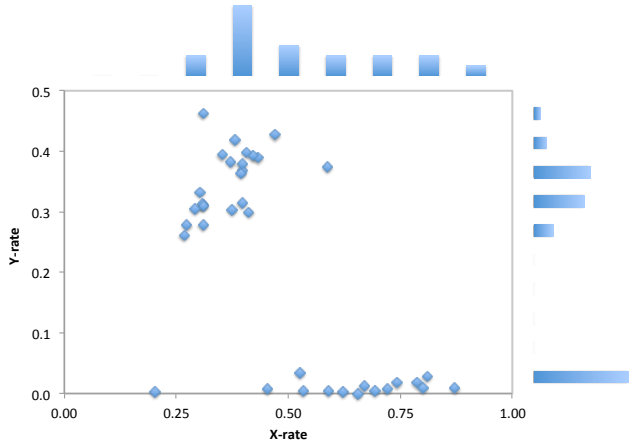


Figure S4.1 X-rate vs. Y-rate plot. There is better separation of the two sexes based on the Y-rate (see the marginal histogram).

Y chromosome haplogroup determination

For the male samples, we determined Y chromosome haplogroups by identifying the most derived allele upstream and the most ancestral allele downstream in the phylogenetic tree in the ISOGG database version 10.01 (http://www.isogg.org/tree). If the most derived Y chromosome SNP upstream was a C→T or G→A substitution (susceptible to ancient DNA damage), we required as least two derived SNPs to assign it to the haplogroup (otherwise, we assigned it to the upstream haplogroup). The results are shown in Table S4.1 and Table S4.2, and are shown in a summary column in Extended Data Table 1:

- We assign *Kostenki14* to haplogroup C1b, as previously described[1] (Table S4.1).
- We assign *GoyetQ116-1* to C1a.
- We assign *Vestonice16* to C1a2. Although our data suggests it carries the derived allele at an A>G SNP that is characteristic of C1b1a1, we find that it carries the ancestral allele at many SNPs that are characteristic of haplogroups upstream of C1b1a1 (i.e. C1b, C1b1, C1b1a) (Table S4.2). Thus, this site may be affected by a sequencing or database error and we ignore the information from it.
- We were surprised to assign *Villabruna* to R1b1 (Table S4.2). When we restrict to damaged sequences, we still assign it to R1b.

Table S4.1. Details of Y haplogroup SNPs for pre-Villabruna Cluster samples.
The most "upstream" ancestral allele observed for a sample in a group of related Y chromosomes is indicated in light red.

Haplogroup	SNP	Ancestral	Derived	Ypos37	Observed	Sequence Depth
Kostenki14	**C1b**					
C	M130	C	T	2734854	T	5
C	V199	C	A	2772928	A	2
C	V232	T	C	7629098	C	23
C	P255	G	A	8685038	A	10
C	V183	G	A	14263271	A	9
C	Page85	G	T	14924643	T	3
C	M216	C	T	15437564	T	3
C	P260	A	C	17286006	C	31
C1	F3393	C	A	23023974	A	2
C1b	F1370	G	C	8643365	C	12
GoyetQ116-1	**C1a**					
C	M130	C	T	2734854	T	1
C	V232	T	C	7629098	C	1
C1	F3393	C	A	23023974	A	2
C1a	CTS11043	G	T	22914979	T	3
Cioclovina1	**CT**					
CT	M5756	T	C	18948988	C	1
Kostenki12	**CT**					
CT	M5609	T	G	7738840	G	1
CT	M5611	C	T	7778691	T	1
CT	M5692	A	C	16325663	C	1
CT	M5712	A	C	17104433	C	1
CT	M5737	C	T	17897543	T	1
CT	CTS9948	G	C	19167672	C	1
Vestonice13	**CT**	**Not IJK**				
CT	CTS109	C	A	2733618	A	1
CT	CTS5318	G	T	16203547	T	1
CT	CTS6327	A	G	16822011	G	1
CT	CTS8243	C	T	17894575	T	1
CT	CTS9556	C	A	18961874	A	1
CT	Z17718	T	C	22263161	C	1
CT	Y1571	G	A	23234852	A	1
CT	M5831	A	C	28685341	C	1
IJ	P126	C	G	21225770	C	1
IJK	L16	G	A	7173143	G	1

Vestonice15	**BT**					
BT	PF1178	G	T	22460746	T	1
Vestonice43	**F**					
F	P145	G	A	8424089	A	1
F	P158	C	T	17493513	T	1
I	PF3641	T	C	7688470	T	1
I	FI4	G	T	8873160	G	1
I	CTS2193	G	T	14214481	G	1
I	CTS4848	C	T	15862842	C	1
I	CTS8963	C	T	18582617	C	1
I	CTS11540	C	T	23156725	C	1
Pavlov1	**I**					
IJK	L16	G	A	7173143	A	1
I	CTS4517	T	G	14986989	G	1
I	FGC2414	C	T	21155653	C	1
Vestonice16	**C1a2**					
C1a	CTS11043	G	T	22914979	T	2
C1a2	V20	G	A	6845955	A	3
C1a2	V86	G	A	6909957	A	1
C1b	F1370	G	C	8643365	G	2
C1b1	M356	C	G	2888203	C	1
C1b1a	K43	G	A	2889366	G	1
C1b1a	F930	C	A	7202706	C	3
C1b1a	K108	C	T	7281157	C	1
C1b1a	Z12437	G	A	7747597	G	4
C1b1a	Z12438	G	T	7821105	G	2
C1b1a	K129	G	A	8292050	G	2
C1b1a	Z12441	A	G	8373844	A	1
C1b1a	Z12442	G	A	8410393	G	1
C1b1a	K141	C	G	8583426	C	2
C1b1a	Z12443	G	A	8635324	G	4
C1b1a	Z12447	G	T	13592515	G	2
C1b1a	K187	G	T	14069571	G	1
C1b1a	Z12450	C	A	14653473	C	2
C1b1a	Z12459	T	C	16549378	T	1
C1b1a	K280	T	C	17237260	T	1
C1b1a	Z12463	C	T	17631240	C	2
C1b1a	Z12464	G	A	18381850	G	2
C1b1a	K319	T	C	18602855	T	2
C1b1a	K396	A	T	22475806	A	1
C1b1a	K414	G	T	23131625	G	3
C1b1a	K415	A	G	23150146	A	2
C1b1a	K417	A	C	23156792	A	1
C1b1a	K426	G	A	23273888	G	2
C1b1a	K435	G	A	23630857	G	3
C1b1a1	K231	A	G	15545270	G	1
Paglicci133	**I**					
I	CTS674	C	T	6943522	T	1
I	CTS9269	C	T	18789763	T	1
I	FGC2416	G	T	7642823	G	1
I	CTS8300	T	A	17924382	T	1
I	PF3815	G	T	21841289	G	2
I1	L80	A	G	14640715	A	1
I1	M253	C	T	15022707	C	1
I1	L81	A	C	22513726	A	1
I1	M307.2	G	A	22750951	G	1
I	FGC2416	G	T	7642823	G	1
I	CTS8300	T	A	17924382	T	1
I	PF3815	G	T	21841289	G	2
HohleFels49	**I**					
I	CTS674	C	T	6943522	T	1
I	CTS9269	C	T	18789763	T	1
GovetQ2	**HIJK**					
HIJK	F929	C	T	7202703	T	1
I	PF3837	G	A	22573702	A	1
Burkhardtshohle	**I**					
I	CTS5650	A	G	16415916	G	1

Table S4.2. Details of Y haplogroup SNPs in Villabruna Cluster samples.

Haplogroup	SNP	Ancestral	Derived	Ypos37	Observed	Read Depth
Villabruna	**R1b1**	**All fragments (subset)**				
R	P224	C	T	17285993	T	7
R	M734	C	T	18066156	T	5
R	P285	C	A	19267344	A	2
R	P227	G	C	21409706	C	7
R1	P294	G	C	7570822	C	2
R1	P242	G	A	7647357	A	2
R1	P238	G	A	7771131	A	6
R1	P245	T	C	8633545	C	2
R1	M173	A	C	15026424	C	1
R1	P286	C	T	17716251	T	6
R1	P236	C	G	17782178	G	10
R1	M306	C	A	22750583	A	5
R1b	M343	C	A	2887824	A	10
R1b1	M415	C	A	9170545	A	2
R1b1	L278	C	T	18914441	T	1
Villabruna	**R1b**	**Damaged fragments**				
R	P224	C	T	17285993	T	4
R	M734	C	T	18066156	T	3
R	P285	C	A	19267344	A	1
R	P280	C	G	21843090	G	1
R1	P294	G	C	7570822	C	2
R1	P238	G	A	7771131	A	2
R1	P286	C	T	17716251	T	1
R1	P236	C	G	17782178	G	1
R1b	M343	C	A	2887824	A	3
Rochedane	**I**					
IJK	L16	G	A	7173143	A	1
I	PF3641	T	C	7688470	C	1
I	CTS4088	T	C	15389836	C	1
I	CTS7593	G	A	17548890	A	1
I	CTS8420	C	A	18018313	A	1
I	CTS10058	A	G	19233673	G	1
I	PF3800	A	G	21402723	G	1
I	PF3803	A	G	21452125	G	1
I2	L68	C	T	18700150	T	1
Falkenstein	**F**					
CF	P143	G	A	14197867	A	1
F	P187	G	T	9108252	T	1
CuiryLesChaudardes1	**I**					
I	PF3640	T	A	7681156	A	1
I	CTS5764	A	G	16471254	G	1
I	Z16987	A	G	22243817	G	1
BerryAuBac	**I**					
HIJK	F929	C	T	7202703	T	1
I	CTS2514	T	C	14337364	C	1
I	PF3836	T	G	22525421	G	1

References

1 Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).

Section 5
Genetic clustering of ancient samples

Overview
We assembled ancient DNA data on a total of 51 ancient samples (Extended Data Table 1). To prepare these data, we combined previously published data from 14 samples (*Malta1*[1], *AfontovaGora2*[1], *UstIshim*[2], *Oase1*[3], *Loschbour*[4], *LaBrana1*[5], *Hungarian.KO1*[6], *Motala12*[4], *Karelia*[7], *Stuttgart*[4], *Satsurblia*[8], *Kotias*[8], *Bichon*[8], and *Kostenki14*[9]), with our new capture data for 38 samples. One sample overlapped between the two collections, *Kostenki14*. For most of our analyses, we use the 16.1-fold coverage data that we newly report in this study, rather than the 2.8-fold coverage published data[9] (we confirmed that the key scientific results were consistent between the two independently collected *Kostenki14* datasets).

Figure 1 shows the geographical distribution of the samples. Below we list the 51 samples by time period, with samples having more than >0.1x coverage highlighted in bold, and samples with fewer than 10,000 SNPs covered underlined.

Early Upper Palaeolithic (>33,000 BP): ***UstIshim***, ***Oase1***, ***Kostenki14***, ***GoyetQ116-1***, *Cioclovina1*, *Kostenki12*, *Muierii2*

Middle Upper Palaeolithic (33,000-24,000 BP): *Vestonice13*, *Vestonice15*, *Vestonice14*, *Vestonice43*, *Pavlov1*, ***Vestonice16***, *Palgicci133*, ***KremsWA3***, *Ostuni2*, ***Ostuni1***, *Paglicci108*, *GoyetQ53-1*, *GoyetQ376-19*, *GoyetQ56-16*, *Malta1*

Late Upper Palaeolithic and Mesolithic (19,000-6,000 BP): ***ElMiron***, *AfontovaGora2*, *AfontovaGora3*, *HohleFels79*, *HohleFels49*, *Rigney1*, *GoyetQ-2*, *Brillenhohle*, *Burkhardtshohle*, ***Villabruna***, ***Bichon***, ***Satsurblia***, ***Rochedane***, *Continenza*, *Iboussieres39*, ***Ranchot88***, *LesCloseaux13*, ***Kotias***, *Falkenstein*, *Bockstein*, *Ofnet*, *Chaudardes1*, ***Loschbour***, ***LaBrana1***, ***Hungarian.KO1***, *BerryAuBac*, ***Motala12***, ***Karelia***

Early Neolithic (7,000 BP): ***Stuttgart***

Outgroup f_3-statistics
We computed statistics of the form $f_3(X, Y; Mbuti)$, which measure the shared genetic drift between populations X and Y since their separation from an outgroup (*Mbuti*)[1]. For the version of this analysis shown in Figure 3A, we did not restrict to damaged sequences for *AfontovaGora3*, *ElMiron*, *Falkenstein*, *GoyetQ-2*, *GoyetQ53-1*, *HohleFels79*, *HohleFels49*, *Les Closeaux13*, *Ofnet*, *Ranchot88* and *Rigney1*. While this allows some present-day human contamination into the data, the increased size of the dataset produces a clearer view of the shared ancestry. A version of this analysis that restricts to damaged sequences only, and to samples with at least 30,000 SNPs, is shown in Extended Data Figure 2.

The shared genetic drift analysis based on these f_3-statistics reveals several apparent clusters (note that this is not the full list of samples we assign to these clusters in Extended Data Table 1, as in the analyses that follow we are also able to assign additional individuals):

- "Věstonice Cluster": *Vestonice13*, *Vestonice15*, *Vestonice14*, *Vestonice43*, *Vestonice16*, *KremsWA3*, *Ostuni1*, *Pavlov*
- "El Mirón Cluster": *Burkhardtshohle*, *ElMiron*, *GoyetQ-2*, *HohleFels79*, *HohleFels49* and *Rigney1*

- "Villabruna Cluster": *BerryAuBac, Bichon, Bockstein, Chaudardes1, Falkenstein, Hungarian.KO1 LaBrana1, LesCloseaux13, Loschbour, Ofnet, Ranchot88, Rochedane* and *Villabruna*.
- "Mal'ta Cluster": *AfontovaGora3* and *Malta1*.
- "Satsurblia Cluster": *Satsurblia, Kotias*

"Věstonice Cluster" (these individuals all lived 34,000-26,000 BP)

We used *D*-statistics to test formally if various pairs of pre-Neolithic samples are consistent with being clades with respect to the other samples and outgroups. We first identified a large group of samples that were consistent with being a clade with samples from the site of Dolní Věstonice. Specifically, *Vestonice13, Vestonice15, Vestonice43, KremsWA3, Ostuni1*, and *Pavlov1* share more alleles with *Vestonice16* (our highest coverage sample that we use to represent it for many analyses) than with other pre-Neolithic samples as revealed by statistics of the form *D(X, Y; Vestonice_Cluster, Mbuti)*. This is consistent with the patterns in Figure 3A and Extended Data Fig. 2 (Table S5.1). We also identified three subgroups within the "Věstonice Cluster":

(a) "Věstonice Central European Cluster" (Vestonice_CE_C): Vestonice, KremsWA3, Pavlov

Statistics of the form $D(Vestonice\_Cluster_1, Vestonice\_Cluster_2; Vestonice\_Cluster_3, Mbuti)$, where $Vestonice\_Cluster_3$ is *Vestonice13, Vestonice15, Vestonice16, Vestonice43, KremsWA3, Ostuni1* or *Pavlov1* in turn, suggest that (*Vestonice16, Vestonice13, Vestonice15, KremsWA3*) form a subgroup (Table S5.2). These statistics also suggest that *Vestonice43* and *Pavlov1* are more distantly related, although one statistic, *D(Vestonice16, KremsWA3; Vestonice43, Mbuti),* is significantly less than 0, providing some evidence against this simple model (Z=-3.6).

(b) "Věstonice Italian Cluster" (Vestonice_I_C) – Ostuni1, Ostuni2 and Palgicci133

The Italian high coverage sample *Ostuni1*, while in the "Věstonice Cluster", is a definite outgroup compared to all samples in the "Věstonice Central European Cluster" (Table S5.2). There is also *D*-statistic evidence that the three Italian samples sub-cluster (Table S5.3).

(c) "Vestonice Goyet Cluster"(Vestonice_Goyet_C) – Multiple samples from Goyet cave

Samples from Goyet cave in Belgium dating from 28,000-26,000 BP are part of the broader Věstonice Cluster*: GoyetQ376-19, Goyet53-1* and *Goyet56-16*. Notably, not all Goyet cave samples are from this cluster: the earlier ~35,000 BP *GoyetQ116-1*, and the later ~15,000 BP *GoyetQ-2*, have very different genetic affinities.

Table S5.1 Z-score of *D(X, Y; Věstonice Cluster, Mbuti)*

D(X, Y; Vestonice13, Mbuti) Vestonice13: 139,568 SNPs											
X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.6	3.3	-12.9	-11.4	-25.5	-17.1	-12.4	-14.4	-13.1	-12.6
UstIshim	-0.6	NA	1.9	-11.5	-10.3	-23.9	-15.6	-11.4	-12.7	-11.4	-11.4
Oase1	-3.3	-1.9	NA	-8.9	-7.8	-18.8	-9.8	-8.7	-8.3	-8.5	-9.1
Kostenki14	12.9	11.5	8.9	NA	0.6	-13.7	-5.7	-0.1	-0.3	0.6	0.5
GoyetQ116-1	11.4	10.3	7.8	-0.6	NA	-13.5	-6.2	-0.8	-0.9	-0.4	-0.2
Vestonice16	25.5	23.9	18.8	13.7	13.5	NA	5.9	13.2	13.8	14.3	13.7
Ostuni1	17.1	15.6	9.8	5.7	6.2	-5.9	NA	4.6	5.5	6.7	6.2
ElMiron	12.4	11.4	8.7	0.1	0.8	-13.2	-4.6	NA	-0.3	0.8	0.8
Villabruna	14.4	12.7	8.3	0.3	0.9	-13.8	-5.5	0.3	NA	1	0.9
Loschbour	13.1	11.4	8.5	-0.6	0.4	-14.3	-6.7	-0.8	-1	NA	-0.2
LaBrana1	12.6	11.4	9.1	-0.5	0.2	-13.7	-6.2	-0.8	-0.9	0.2	NA
Malta1	7.9	7.2	4.9	-4	-3.4	-15.6	-7.9	-3.9	-4.4	-3.1	-3.6
D(X, Y; Vestonice15, Mbuti) Vestonice15: 30,900 SNPs											
X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.5	0.8	-7.5	-7.2	-19.5	-10.5	-7.4	-10.1	-11.5	-7.2
UstIshim	-0.5	NA	0.4	-7.3	-7	-19	-10.2	-7.5	-9	-9.8	-6.6
Oase1	-0.8	-0.4	NA	-4.6	-3.9	-10.7	-6.3	-4.3	-5.9	-6.6	-4.2

Kostenki14	7.5	7.3	4.6	NA	0.1	-11.1	-3.2	0	-1.6	-1.8	0.7
GoyetQ116-1	7.2	7	3.9	-0.1	NA	-11.4	-5	-0.2	-1.9	-2.1	1
Vestonice16	19.5	19	10.7	11.1	11.4	NA	5.4	11	10.2	10.7	12.2
Ostuni1	10.5	10.2	6.3	3.2	5	-5.4	NA	3.9	3	2.6	4.6
ElMiron	7.4	7.5	4.3	0	0.2	-11	-3.9	NA	-1.4	-1.9	1.7
Villabruna	10.1	9	5.9	1.6	1.9	-10.2	-3	1.4	NA	0.2	2.9
Loschbour	11.5	9.8	6.6	1.8	2.1	-10.7	-2.6	1.9	-0.2	NA	3.2
LaBrana1	7.2	6.6	4.2	-0.7	-1	-12.2	-4.6	-1.7	-2.9	-3.2	NA
Malta1	5.9	5.1	3.1	-0.7	-0.7	-10.7	-3.5	-0.8	-2.8	-2.7	-0.2

D(X, Y; Vestonice43, Mbuti) Vestonice43: 163,946 SNPs

X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.3	2.2	-14.2	-12.1	-19.6	-19.4	-15.6	-17.7	-18.3	-14.7
UstIshim	-0.3	NA	2.1	-13.3	-10.9	-17.4	-16.7	-14.1	-15.6	-16.1	-12.7
Oase1	-2.2	-2.1	NA	-9.8	-9.1	-13.3	-12.4	-10	-12.6	-12	-8.8
Kostenki14	14.2	13.3	9.8	NA	1.7	-5.1	-5.7	-2.5	-2.8	-2.8	0.7
GoyetQ116-1	12.1	10.9	9.1	-1.7	NA	-6.1	-7.4	-3	-3.5	-3.9	-0.6
Vestonice16	19.6	17.4	13.3	5.1	6.1	NA	-0.9	3.9	3.2	3.3	6.2
Ostuni1	19.4	16.7	12.4	5.7	7.4	0.9	NA	4.6	3.8	4.3	6.7
ElMiron	15.6	14.1	10	2.5	3	-3.9	-4.6	NA	-0.9	-0.3	2.8
Villabruna	17.7	15.6	12.6	2.8	3.5	-3.2	-3.8	0.9	NA	0	3.7
Loschbour	18.3	16.1	12	2.8	3.9	-3.3	-4.3	0.3	0	NA	4.2
LaBrana1	14.7	12.7	8.8	-0.7	0.6	-6.2	-6.7	-2.8	-3.7	-4.2	NA
Malta1	8	7.2	6.7	-5.2	-3.8	-9.9	-9.3	-6	-7	-7.7	-4.3

D(X, Y; Vestonice16, Mbuti) Vestonice16: 945,292 SNPs

X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	-1.2	3.9	-16.3	-14.4	NA	-24.6	-18.6	-19	-18.5	-16
UstIshim	1.2	NA	4.4	-12.5	-11	NA	-20.2	-14.8	-14.7	-14.5	-12.1
Oase1	-3.9	-4.4	NA	-15	-13.3	NA	-18.5	-15.9	-14.7	-15.9	-13.5
Kostenki14	16.3	12.5	15	NA	0.9	NA	-8.7	-2	-1.5	-1.7	0.6
GoyetQ116-1	14.4	11	13.3	-0.9	NA	NA	-9.5	-2.8	-2.5	-3.2	-0.5
Ostuni1	24.6	20.2	18.5	8.7	9.5	NA	NA	7.4	8.1	7.8	9.7
ElMiron	18.6	14.8	15.9	2	2.8	NA	-7.4	NA	0.9	0.3	3.4
Villabruna	19	14.7	14.7	1.5	2.5	NA	-8.1	-0.9	NA	-0.6	2.9
Loschbour	18.5	14.5	15.9	1.7	3.2	NA	-7.8	-0.3	0.6	NA	3.2
LaBrana1	16	12.1	13.5	-0.6	0.5	NA	-9.7	-3.4	-2.9	-3.2	NA
Malta1	11.8	8.4	10.4	-4.6	-2.6	NA	-11.9	-6.3	-5.6	-6.1	-3.8

D(X, Y; Pavlov1, Mbuti) Pavlov1: 57,005 SNPs

X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.3	0.5	-8.6	-9.5	-16.7	-13.2	-10.7	-12.2	-14	-11.3
UstIshim	-0.3	NA	0.2	-8.1	-9.2	-16.7	-12.6	-10.4	-11.6	-12.3	-10.4
Oase1	-0.5	-0.2	NA	-3.1	-3.4	-8.6	-7	-4.2	-6.1	-6.3	-5.6
Kostenki14	8.6	8.1	3.1	NA	-0.7	-7.6	-5.7	-1.9	-2.6	-2.6	-1.7
GoyetQ116-1	9.5	9.2	3.4	0.7	NA	-7.3	-4.7	-0.8	-2	-2.3	-0.9
Vestonice16	16.7	16.7	8.6	7.6	7.3	NA	0.7	5.2	4.9	5.7	6.2
Ostuni1	13.2	12.6	7	5.7	4.7	-0.7	NA	4.5	3.7	4.2	5
ElMiron	10.7	10.4	4.2	1.9	0.8	-5.2	-4.5	NA	-0.8	-0.8	0.3
Villabruna	12.2	11.6	6.1	2.6	2	-4.9	-3.7	0.8	NA	0.2	1.4
Loschbour	14	12.3	6.3	2.6	2.3	-5.7	-4.2	0.8	-0.2	NA	1.1
LaBrana1	11.3	10.4	5.6	1.7	0.9	-6.2	-5	-0.3	-1.4	-1.1	NA
Malta1	7.3	6.8	2.7	-0.3	-1.7	-8.1	-6.3	-2.7	-4.5	-4	-2.7

D(X, Y; KremsWA3, Mbuti) KremsWA3: 203,986 SNPs

X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	-0.3	2.4	-14.3	-13.5	-26.5	-20.1	-16.2	-18.4	-19.6	-15.9
UstIshim	0.3	NA	3	-11.9	-11.1	-24.8	-16.5	-14.4	-16.3	-15.9	-13.8
Oase1	-2.4	-3	NA	-8.1	-9.6	-18	-10.4	-11.2	-10.7	-11.6	-9.9
Kostenki14	14.3	11.9	8.1	NA	-0.5	-13.9	-7	-2.7	-3.6	-3.2	-1.6
GoyetQ116-1	13.5	11.1	9.6	0.5	NA	-13.1	-6.7	-2.3	-2.8	-2.8	-1
Vestonice16	26.5	24.8	18	13.9	13.1	NA	6.3	11.6	12.1	12.7	13.3
Ostuni1	20.1	16.5	10.4	7	6.7	-6.3	NA	4.1	3.5	5.1	5.9
ElMiron	16.2	14.4	11.2	2.7	2.3	-11.6	-4.1	NA	-0.7	-0.2	1.2
Villabruna	18.4	16.3	10.7	3.6	2.8	-12.1	-3.5	0.7	NA	0.5	2.4
Loschbour	19.6	15.9	11.6	3.2	2.8	-12.7	-5.1	0.2	-0.5	NA	2.1
LaBrana1	15.9	13.8	9.9	1.6	1	-13.3	-5.9	-1.2	-2.4	-2.1	NA
Malta1	9.5	7.5	5.6	-4.5	-3.1	-17.9	-7.8	-5.4	-7.3	-7.4	-5.7

D(X, Y; Ostuni1, Mbuti) Ostuni1: 369,313 SNPs

X/Y	Han	Ust	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.9	3.4	-13.4	-13	-23.3	NA	-17.5	-18.5	-17.3	-16.1
UstIshim	-0.9	NA	1.8	-12.2	-12.3	-21.9	NA	-15.4	-16.4	-14.9	-14.2
Oase1	-3.4	-1.8	NA	-9.8	-10.5	-17.4	NA	-13	-13.3	-13	-12
Kostenki14	13.4	12.2	9.8	NA	-0.4	-10	NA	-3.5	-3.9	-2.4	-1.9
GoyetQ116-1	13	12.3	10.5	0.4	NA	-8.8	NA	-2.7	-3	-2.3	-1.2
Vestonice16	23.3	21.9	17.4	10	8.8	NA	NA	6.7	6.1	7.9	9
Ostuni1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ElMiron	17.5	15.4	13	3.5	2.7	-6.7	NA	NA	-0.5	0.7	1.7
Villabruna	18.5	16.4	13.3	3.9	3	-6.1	NA	0.5	NA	1.7	2.6
Loschbour	17.3	14.9	13	2.4	2.3	-7.9	NA	-0.7	-1.7	NA	1
LaBrana1	16.1	14.2	12	1.9	1.2	-9	NA	-1.7	-2.6	-1	NA
Malta1	9	8	6.4	-3.9	-3.3	-13.8	NA	-6.6	-7.3	-6.7	-6.3

Table S5.2 Z-score *D(Vestonice_Cluster$_1$, Vestonice_Cluster$_2$; Vestonice_Cluster$_3$, Mbuti)*

D(X, Y; Vestonice13, Mbuti) Vestonice13: 139,568 SNPs						
X/Y	Vestonice15	Vestonice43	Pavlov1	Vestonice16	Ostuni1	KremsWA3
Vestonice15	NA	1.8	-0.4	0.7	0.4	0.4
Vestonice43	-1.8	NA	-1.3	-6.8	0.3	-2.8
Pavlov1	0.4	1.3	NA	-2.5	0.8	-2.8
Vestonice16	-0.7	6.8	2.5	NA	5.9	-1.3
Ostuni1	-0.4	-0.3	-0.8	-5.9	NA	-5.3
KremsWA3	-0.4	2.8	2.8	1.3	5.3	NA
D(X, Y; Vestonice15, Mbuti) Vestonice15: 30,900 SNPs						
X/Y	Vestonice13	Vestonice43	Pavlov1	Vestonice16	Ostuni1	KremsWA3
Vestonice13	NA	2.1	1.4	-1	1.9	0.1
Vestonice43	-2.1	NA	-1	-5.4	-1.6	-2.5
Pavlov1	-1.4	1	NA	-2.7	-1.9	-0.1
Vestonice16	1	5.4	2.7	NA	5.4	1.3
Ostuni1	-1.9	1.6	1.9	-5.4	NA	-1
KremsWA3	-0.1	2.5	0.1	-1.3	1	NA
D(X, Y; Vestonice43, Mbuti) Vestonice43: 163,946 SNPs						
X/Y	Vestonice13	Vestonice15	Pavlov1	Vestonice16	Ostuni1	KremsWA3
Vestonice13	NA	0.3	-1	0.6	-0.3	0.6
Vestonice15	-0.3	NA	1	-0.4	-0.2	0.4
Pavlov1	1	-1	NA	3.2	1.6	-0.7
Vestonice16	-0.6	0.4	-3.2	NA	-0.9	-3.6
Ostuni1	0.3	0.2	-1.6	0.9	NA	-2.2
KremsWA3	-0.6	-0.4	0.7	3.6	2.2	NA
D(X, Y; Pavlov1, Mbuti) Pavlov1: 57,005 SNPs						
X/Y	Vestonice13	Vestonice15	Vestonice43	Vestonice16	Ostuni1	KremsWA3
Vestonice13	NA	1.8	0.2	1.2	2.2	-0.2
Vestonice15	-1.8	NA	2	-0.6	-0.7	-0.9
Vestonice43	-0.2	-2	NA	1.8	1.2	-0.7
Vestonice16	-1.2	0.6	-1.8	NA	0.7	-1.5
Ostuni1	-2.2	0.7	-1.2	-0.7	NA	-2
KremsWA3	0.2	0.9	0.7	1.5	2	NA
D(X, Y; Vestonice16, Mbuti) Vestonice16: 945,292 SNPs						
X/Y	Vestonice13	Vestonice15	Vestonice43	Pavlov1	Ostuni1	KremsWA3
Vestonice13	NA	-1.8	7.4	3.8	6.7	2.2
Vestonice15	1.8	NA	5.1	2	7.9	3.4
Vestonice43	-7.4	-5.1	NA	-1.4	-0.5	-6.9
Pavlov1	-3.8	-2	1.4	NA	1.7	-3.7
Ostuni1	-6.7	-7.9	0.5	-1.7	NA	-5.5
KremsWA3	-2.2	-3.4	6.9	3.7	5.5	NA
D(X, Y; Ostuni1, Mbuti) Ostuni1: 369,313 SNPs						
X/Y	Vestonice13	Vestonice15	Vestonice43	Pavlov1	Vestonice16	KremsWA3
Vestonice13	NA	1.7	-0.5	1.4	0.7	1.5
Vestonice15	-1.7	NA	1.4	1.2	2.6	1.3
Vestonice43	0.5	-1.4	NA	-0.4	0.4	-0.3
Pavlov1	-1.4	-1.2	0.4	NA	1	-0.7
Vestonice16	-0.7	-2.6	-0.4	-1	NA	0.8
KremsWA3	-1.5	-1.3	0.3	0.7	-0.8	NA
D(X, Y; KremsWA3, Mbuti) KremsWA3: 203,986 SNPs						
X/Y	Vestonice13	Vestonice15	Vestonice43	Pavlov1	Vestonice16	Ostuni1
Vestonice13	NA	-0.3	3.3	2.6	3.5	6.7
Vestonice15	0.3	NA	2.9	-0.7	2.1	2.3
Vestonice43	-3.3	-2.9	NA	0	-3	1.9
Pavlov1	-2.6	0.7	0	NA	-2	1.4
Vestonice16	-3.5	-2.1	3	2	NA	6.3
Ostuni1	-6.7	-2.3	-1.9	-1.4	-6.3	NA

"El Mirón Cluster" (ElMiron_C) (these individuals all lived 19,000-14,000 BP)

Brillenhohle, *Burkhardtshohle*, *GoyetQ-2*, *HohleFels79*, *HohleFels49* and *Rigney1* share more alleles with *ElMiron* than with other pre-Neolithic Europeans as revealed by statistics like *D(X, Y; El_Miron_Cluster, Mbuti)* (Table S5.4), consistent with outgroup f_3-statistics (Figure 3A and Extended Data Figure 2). The *D*-statistics also suggest that *ElMiron*—the earliest and most southern of the samples—is an outgroup to the later set of samples, as revealed by statistics like *D(El_Miron_Cluster$_1$, El_Miron_Cluster$_2$; El_Miron_Cluster$_3$, Mbuti)* (Table S5.5). We designate the entire group as the "El Mirón Cluster" and designate the subgroup excluding *ElMiron* as the "El Mirón Non-Iberian Cluster" (ElMiron_NI_C).

Table S5.3 Z-score of *D(X, Y; Ostuni2/Paglicci133, Mbuti)*

D(X, Y; Ostuni2, Mbuti) Ostuni2: 17,017 SNPs

X/Y	Han	Ust	Oase1	Kost14	Q116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	0.1	0.4	-6.2	-6.2	-8.5	-11.9	-7.4	-5.6	-6.5	-5.2
UstIshim	-0.1	NA	0.5	-6	-5.8	-8.3	-10.2	-6.5	-5.3	-5.7	-4.8
Oase1	-0.4	-0.5	NA	-3.6	-2.8	-4	-4.2	-3.9	-3.2	-3.4	-3.8
Kostenki14	6.2	6	3.6	NA	-0.1	-2.3	-5.4	-0.9	-0.2	1.1	1.4
GoyetQ116-1	6.2	5.8	2.8	0.1	NA	-2.3	-4.4	-0.4	0.8	1.4	1.4
Vestonice16	8.5	8.3	4	2.3	2.3	NA	-2.9	1.2	2.6	3.6	3.6
Ostuni1	11.9	10.2	4.2	5.4	4.4	2.9	NA	4.4	5.8	6.5	4.9
ElMiron	7.4	6.5	3.9	0.9	0.4	-1.2	-4.4	NA	1.8	2	2.4
Villabruna	5.6	5.3	3.2	0.2	-0.8	-2.6	-5.8	-1.8	NA	0.5	0.8
Loschbour	6.5	5.7	3.4	-1.1	-1.4	-3.6	-6.5	-2	-0.5	NA	0.3
LaBrana1	5.2	4.8	3.8	-1.4	-1.4	-3.6	-4.9	-2.4	-0.8	-0.3	NA
Malta1	3.6	3.3	2.2	-2.3	-2.1	-3.9	-4.9	-2.5	-1.4	-1.1	-1

D(X, Y; Paglicci133, Mbuti) Paglicci133: 82,330 SNPs

X/Y	Han	Ust	Oase1	Kost14	Q116-1	Vestonice16	Ostuni1	ElMiron	Villabruna	Loschbour	LaBrana1
Han	NA	1.5	1.8	-9.2	-9.5	-13.3	-12.7	-9.4	-11.4	-11.9	-10.5
UstIshim	-1.5	NA	1.1	-9.2	-9.5	-13.3	-11.7	-8.9	-11.5	-11.8	-9.9
Oase1	-1.8	-1.1	NA	-5.1	-6.5	-7	-7.2	-4.3	-7.5	-6.9	-5.2
Kostenki14	9.2	9.2	5.1	NA	-0.6	-4.6	-5.6	-0.7	-1.8	-1.4	-0.7
GoyetQ116-1	9.5	9.5	6.5	0.6	NA	-3.3	-4.9	0.9	-0.7	-0.5	0.7
Vestonice16	13.3	13.3	7	4.6	3.3	NA	-1.8	3.5	2.7	3	4.2
Ostuni1	12.7	11.7	7.2	5.6	4.9	1.8	NA	5	3.5	4.5	4.6
ElMiron	9.4	8.9	4.3	0.7	-0.9	-3.5	-5	NA	-1.7	-1.6	-0.4
Villabruna	11.4	11.5	7.5	1.8	0.7	-2.7	-3.5	1.7	NA	0.7	1
Loschbour	11.9	11.8	6.9	1.4	0.5	-3	-4.5	1.6	-0.7	NA	0.2
LaBrana1	10.5	9.9	5.2	0.7	-0.7	-4.2	-4.6	0.4	-1	-0.2	NA
Malta1	6	6.1	3.5	-2.9	-2.6	-6.4	-7.4	-3.4	-4.7	-5	-4.6

Table S5.4 Z-score of *D(X, Y; El_Miron_Cluster, Mbuti)*

D(X, Y; ElMiron, Mbuti) HohleFels79: 797,714 SNPs

X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	0.8	3.8	-12.2	-22.5	-17	-17.4	-21.9	-25.9	-30.5	-33.5
UstIshim	-0.8	NA	3.2	-11.2	-19.7	-15.7	-15.3	-20.4	-23	-26.3	-27.8
Oase1	-3.8	-3.2	NA	-10.6	-17.9	-16.2	-13.8	-13.2	-20.5	-24.6	-23.6
Kostenki14	12.2	11.2	10.6	NA	-11	-5.2	-5	-13.4	-13.5	-16.4	-18.6
GoyetQ116-1	22.5	19.7	17.9	11	NA	6.1	4.5	-8.4	-1.8	-5.2	-5.3
Vestonice16	17	15.7	16.2	5.2	-6.1	NA	-0.4	-12.3	-8.3	-12.1	-12.5
Ostuni1	17.4	15.3	13.8	5	-4.5	0.4	NA	-9.3	-8.1	-11.3	-11.3
Villabruna	25.9	23	20.5	13.5	1.8	8.3	8.1	-7.1	NA	-4.1	-4.1
Loschbour	30.5	26.3	24.6	16.4	5.2	12.1	11.3	-5.5	4.1	NA	-0.3
LaBrana1	33.5	27.8	23.6	18.6	5.3	12.5	11.3	-6.8	4.1	0.3	NA
Malta1	9.1	7.8	8.9	-2.1	-11.9	-7.5	-7	-12.5	-15	-18.5	-19.9

D(X, Y; HohleFels79, Mbuti) HohleFels79: 11,211 SNPs

X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	0.1	0.5	-3.3	-8.7	-3.7	-2.5	-12.2	-4.4	-8.7	-7.7
UstIshim	-0.1	NA	0.8	-3.4	-7.9	-3.7	-2.3	-11	-3.9	-7.9	-6.8
Oase1	-0.5	-0.8	NA	-2.5	-3.3	-2	-1.2	-5	-1.2	-3.9	-3.7
Kostenki14	3.3	3.4	2.5	NA	-3.6	-0.6	0	-6.9	-0.7	-3.9	-3
GoyetQ116-1	8.7	7.9	3.3	3.6	NA	4.4	2.5	-3.1	3.9	1.4	0.9
Vestonice16	3.7	3.7	2	0.6	-4.4	NA	0.7	-7	-0.3	-3.6	-2.9
Ostuni1	2.5	2.3	1.2	0	-2.5	-0.7	NA	-5.4	-1	-3.4	-2.5
ElMiron	12.2	11	5	6.9	3.1	7	5.4	NA	6.6	4.5	3.8
Villabruna	4.4	3.9	1.2	0.7	-3.9	0.3	1	-6.6	NA	-3.5	-2.8
Loschbour	8.7	7.9	3.9	3.9	-1.4	3.6	3.4	-4.5	3.5	NA	-0.1
LaBrana1	7.7	6.8	3.7	3	-0.9	2.9	2.5	-3.8	2.8	0.1	NA
Malta1	3.2	2.4	2.4	-0.1	-3.2	0.3	-0.6	-5.6	-0.2	-3	-2.5

D(X, Y; HohleFels49, Mbuti) HohleFels49: 63,151 SNPs

X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	0.2	1.6	-6.2	-16	-8	-7.1	-23.7	-11.1	-17	-15.5
UstIshim	-0.2	NA	1.9	-5.7	-14.1	-7.4	-6.7	-21.7	-9.9	-14.6	-13.4
Oase1	-1.6	-1.9	NA	-4.2	-9.9	-5.7	-4	-12.3	-6.3	-9.2	-8.6
Kostenki14	6.2	5.7	4.2	NA	-8.7	-0.8	-2.8	-14.4	-4	-7.3	-7
GoyetQ116-1	16	14.1	9.9	8.7	NA	6.6	5.2	-6.9	4.8	1.8	2.1
Vestonice16	8	7.4	5.7	0.8	-6.6	NA	-0.2	-14.4	-2.6	-6	-5
Ostuni1	7.1	6.7	4	2.8	-5.2	0.2	NA	-10	-0.5	-3.9	-3.2
ElMiron	23.7	21.7	12.3	14.4	6.9	14.4	10	NA	10.9	9.1	8.9
Villabruna	11.1	9.9	6.3	4	-4.8	2.6	0.5	-10.9	NA	-3.8	-3
Loschbour	17	14.6	9.2	7.3	-1.8	6	3.9	-9.1	3.8	NA	-0.2
LaBrana1	15.5	13.4	8.6	7	-2.1	5	3.2	-8.9	3	0.2	NA
Malta1	4.4	4.1	2.7	-1.3	-8.1	-1.9	-2.4	-14	-3.6	-7.4	-6.9

D(X, Y; Rigney1, Mbuti) Rigney1: 35,600 SNPs

X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	-0.5	2.2	-5.1	-11.1	-6.5	-6.2	-18.7	-10.9	-14.9	-14.7
UstIshim	0.5	NA	2.2	-4.3	-10.1	-5.9	-5.3	-17	-9.9	-12.8	-12.8
Oase1	-2.2	-2.2	NA	-3.6	-7.3	-4.1	-3.6	-9	-6.8	-7.6	-7.8
Kostenki14	5.1	4.3	3.6	NA	-5.2	-1.9	-1.9	-11.5	-4.8	-7.3	-7.7
GoyetQ116-1	11.1	10.1	7.3	5.2	NA	4.3	2.1	-6.3	0.5	-1.2	-1.5
Vestonice16	6.5	5.9	4.1	1.9	-4.3	NA	-1.6	-10.8	-3.3	-5.1	-5.3
Ostuni1	6.2	5.3	3.6	1.9	-2.1	1.6	NA	-7.7	-1.7	-2.9	-3.3

ElMiron	18.7	17	9	11.5	6.3	10.8	7.7	NA	7	6.4	5.5
Villabruna	10.9	9.9	6.8	4.8	-0.5	3.3	1.7	-7	NA	-2.5	-2.9
Loschbour	14.9	12.8	7.6	7.3	1.2	5.1	2.9	-6.4	2.5	NA	-1
LaBrana1	14.7	12.8	7.8	7.7	1.5	5.3	3.3	-5.5	2.9	1	NA
Malta1	4	3.2	2.6	-0.4	-6.7	-2.5	-2.2	-9.9	-5.4	-7.8	-7.4
D(X, Y; GoyetQ-2, Mbuti) GoyetQ-2: 72,263											
X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	0.8	0.4	-6.3	-18.7	-9	-6.6	-25.4	-12.7	-19.2	-19.2
UstIshim	-0.8	NA	-0.2	-6.1	-16.9	-8.9	-6.8	-23.2	-12	-16.4	-16.9
Oase1	-0.4	0.2	NA	-3.5	-8.1	-5.3	-5.1	-12.9	-7.2	-9	-8.5
Kostenki14	6.3	6.1	3.5	NA	-10.9	-1.8	-2.5	-15.2	-4.9	-8.5	-10.2
GoyetQ116-1	18.7	16.9	8.1	10.9	NA	7.3	4.9	-6.8	5.2	3.8	1.5
Vestonice16	9	8.9	5.3	1.8	-7.3	NA	-0.8	-13.7	-3.1	-6.1	-6.9
Ostuni1	6.6	6.8	5.1	2.5	-4.9	0.8	NA	-9.9	-1.6	-3.3	-5.7
ElMiron	25.4	23.2	12.9	15.2	6.8	13.7	9.9	NA	11.7	10.8	7.6
Villabruna	12.7	12	7.2	4.9	-5.2	3.1	1.6	-11.7	NA	-3.1	-4.9
Loschbour	19.2	16.4	9	8.5	-3.8	6.1	3.3	-10.8	3.1	NA	-2.7
LaBrana1	19.2	16.9	8.5	10.2	-1.5	6.9	5.7	-7.6	4.9	2.7	NA
Malta1	7.3	6.6	3.7	0.9	-8	-1.1	-1.1	-12.8	-3.7	-6.6	-7.9
D(X, Y; Burkhardtshohle, Mbuti) Burkhardtshohle: 38,376 SNPs											
X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	-0.2	2.1	-4.9	-15.2	-8	-5.9	-21.2	-9.4	-14.5	-14.3
UstIshim	0.2	NA	2.5	-3.9	-13.5	-7	-5.5	-18.9	-8.4	-12.2	-12.3
Oase1	-2.1	-2.5	NA	-3.8	-8.5	-4.6	-3.8	-10.4	-6	-8.1	-7.6
Kostenki14	4.9	3.9	3.8	NA	-8.5	-3	-1.2	-13.2	-3.3	-6.6	-7.4
GoyetQ116-1	15.2	13.5	8.5	8.5	NA	5.6	3.9	-5.6	5.2	2.7	1.3
Vestonice16	8	7	4.6	3	-5.6	NA	0.6	-11.4	-0.9	-4.3	-4.6
Ostuni1	5.9	5.5	3.8	1.2	-3.9	-0.6	NA	-8.8	-0.6	-2.5	-3.1
ElMiron	21.2	18.9	10.4	13.2	5.6	11.4	8.8	NA	11.1	9.6	7.4
Villabruna	9.4	8.4	6	3.3	-5.2	0.9	0.6	-11.1	NA	-3.5	-4.2
Loschbour	14.5	12.2	8.1	6.6	-2.7	4.3	2.5	-9.6	3.5	NA	-1.1
LaBrana1	14.3	12.3	7.6	7.4	-1.3	4.6	3.1	-7.4	4.2	1.1	NA
Malta1	4.5	3.8	4.3	0	-7.5	-2.5	-1.4	-11.3	-3.3	-6.9	-7.3
D(X, Y; Brillenhohle, Mbuti) Brillenhohle: 13,459 SNPs											
X/Y	Han	Ust	Oase1	Kost14	GoyetQ116-	Vestonice1	Ostuni	ElMiro	Villabruna	Loschbour	LaBrana
Han	NA	1	1.7	-2	-7.6	-3.5	-4.2	-11.2	-6.3	-9.7	-8.8
UstIshim	-1	NA	0.6	-3	-7.6	-3.8	-4.2	-11.4	-6.6	-8.9	-8.8
Oase1	-1.7	-0.6	NA	-2.2	-5.1	-1.8	-3.2	-6.7	-4	-6.1	-4.9
Kostenki14	2	3	2.2	NA	-4.4	-1.2	-2	-6.7	-3.7	-5.1	-5.9
GoyetQ116-1	7.6	7.6	5.1	4.4	NA	3.5	1.8	-2.8	1.9	0.1	-0.3
Vestonice16	3.5	3.8	1.8	1.2	-3.5	NA	-0.7	-6.7	-1.7	-4	-4.2
Ostuni1	4.2	4.2	3.2	2	-1.8	0.7	NA	-4.2	-0.6	-2.8	-2.4
ElMiron	11.2	11.4	6.7	6.7	2.8	6.7	4.2	NA	5.2	3.6	3.1
Villabruna	6.3	6.6	4	3.7	-1.9	1.7	0.6	-5.2	NA	-2.3	-2.1
Loschbour	9.7	8.9	6.1	5.1	-0.1	4	2.8	-3.6	2.3	NA	-0.5
LaBrana1	8.8	8.8	4.9	5.9	0.3	4.2	2.4	-3.1	2.1	0.5	NA
Malta1	1.7	1.9	1.8	-0.5	-3.9	-1.3	-1.5	-7	-3.1	-4.3	-4.4

Table S5.5 Z-score *D(El_Miron_Cluster$_1$, El_Miron_Cluster$_2$; El_Miron_Cluster$_3$, Mbuti)*

D(X, Y; ElMiron, Mbuti) HohleFels79: 797,714 SNPs						
X/Y	HohleFels79	HohleFels49	Rigney1	GoyetQ-2	Brillenhohle	Burkhardtshohle
HohleFels79	NA	0.5	0.8	1.6	0.7	1.2
HohleFels49	-0.5	NA	-1.2	-0.9	0.6	0.9
Rigney1	-0.8	1.2	NA	0.4	-0.9	0
GoyetQ-2	-1.6	0.9	-0.4	NA	0.3	-0.1
Brillenhohle	-0.7	-0.6	0.9	-0.3	NA	0.8
Burkhardtshohle	-1.2	-0.9	0	0.1	-0.8	NA
D(X, Y; HohleFels79, Mbuti) HohleFels79: 11,211 SNPs						
X/Y	ElMiron	HohleFels49	Rigney1	GoyetQ-2	Brillenhohle	Burkhardtshohle
ElMiron	NA	-2.9	-0.4	0.2	0.5	0.3
HohleFels49	2.9	NA	-0.4	0.8	1.5	0.2
Rigney1	0.4	0.4	NA	0	1.3	0
GoyetQ-2	-0.2	-0.8	0	NA	0.8	-1.1
Brillenhohle	-0.5	-1.5	-1.3	-0.8	NA	1.5
Burkhardtshohle	-0.3	-0.2	0	1.1	-1.5	NA
D(X, Y; HohleFels49, Mbuti) HohleFels49: 63,151 SNPs						
X/Y	ElMiron	HohleFels79	Rigney1	GoyetQ-2	Brillenhohle	Burkhardtshohle
ElMiron	NA	-3.3	-4.2	-3.4	-3	-2.5
HohleFels79	3.3	NA	-0.5	1.3	-0.4	0.4
Rigney1	4.2	0.5	NA	1.8	0.6	0.1
GoyetQ-2	3.4	-1.3	-1.8	NA	-0.1	0.4
Brillenhohle	3	0.4	-0.6	0.1	NA	0.7
Burkhardtshohle	2.5	-0.4	-0.1	-0.4	-0.7	NA
D(X, Y; Rigney1, Mbuti) Rigney1: 35,600 SNPs						
X/Y	ElMiron	HohleFels79	HohleFels49	GoyetQ-2	Brillenhohle	Burkhardtshohle

ElMiron	NA	-1.1	-3.1	-5.1	-1.9	-2.1
HohleFels79	1.1	NA	-0.1	-1	1.2	-0.5
HohleFels49	3.1	0.1	NA	0.2	-0.2	0.3
GoyetQ-2	5.1	1	-0.2	NA	0	-0.3
Brillenhohle	1.9	-1.2	0.2	0	NA	0
Burkhardtshohle	2.1	0.5	-0.3	0.3	0	NA
D(X, Y; GoyetQ-2, Mbuti) GoyetQ-2: 72,263 SNPs						
X/Y	ElMiron	HohleFels79	HohleFels49	Rigney1	Brillenhohle	Burkhardtshohle
ElMiron	NA	-1.5	-2.4	-5.3	-2.5	-2.9
HohleFels79	1.5	NA	0.5	-1.3	1.3	-0.6
HohleFels49	2.4	-0.5	NA	-1.7	0.2	1
Rigney1	5.3	1.3	1.7	NA	0	-0.4
Brillenhohle	2.5	-1.3	-0.2	0	NA	-0.6
Burkhardtshohle	2.9	0.6	-1	0.4	0.6	NA
D(X, Y; Brillenhohle, Mbuti) Brillenhohle: 13,459 SNPs						
X/Y	ElMiron	HohleFels79	HohleFels49	Rigney1	GoyetQ-2	Burkhardtshohle
ElMiron	NA	-0.2	-3.5	-1.1	-2.8	0.3
HohleFels79	0.2	NA	-1.8	0	0	2.1
HohleFels49	3.5	1.8	NA	-0.7	0.3	0
Rigney1	1.1	0	0.7	NA	0	0
GoyetQ-2	2.8	0	-0.3	0	NA	-1
Burkhardtshohle	-0.3	-2.1	0	0	1	NA
D(X, Y; Burkhardtshohle, Mbuti) Burkhardtshohle: 38,376 SNPs						
X/Y	ElMiron	HohleFels79	HohleFels49	Rigney1	GoyetQ-2	Brillenhohle
ElMiron	NA	-0.9	-3.5	-2	-2.6	-0.6
HohleFels79	0.9	NA	0.2	0	0.5	0.8
HohleFels49	3.5	-0.2	NA	0.1	0.6	-0.8
Rigney1	2	0	-0.1	NA	0	0
GoyetQ-2	2.6	-0.5	-0.6	0	NA	-0.5
Brillenhohle	0.6	-0.8	0.8	0	0.5	NA

"Villabruna Cluster" (Villabruna_C) (these individuals all lived 14,000-7,000 BP)

This genetic grouping includes *BerryAuBac*, *Bichon*, *Bockstein*, *Chaudardes1*, *Falkenstein*, *Hungarian.KO1*, *LaBrana1*, *LesCloseaux13*, *Loschbour*, *Ranchot88*, *Rochedane*, *Ofnet*, and *Villabruna*, all of whom lived after around 14,000 BP. This is the largest single grouping of samples in our dataset. Because of the fact that there are many samples, and that a number of them have low coverage, we cannot co-analyse them at the set of SNPs that are covered at least once in all samples, as this would leave us with too few SNPs. In Table S5.6, we therefore analyse each sample in turn along with a selected set of high coverage samples:

- *Kostenki14*
- *GoyetQ116-1*
- *Vestonice16* (as a high coverage representative of the Věstonice Cluster)
- *ElMiron* (as a high coverage representative of the El Mirón Cluster)
- *Malta1* (as a high coverage representative of the Mal'ta Cluster)
- *Loschbour*
- *LaBrana1*
- *Hungarian.KO1*
- *Motala12*

Based on analysis of statistic like *D(X, Y; Villabruna Cluster, Mbuti)*, we find that *BerryAuBac*, *Bichon*, *Bockstein*, *Chaudardes1*, *Falkenstein*, *Ranchot88*, *Rochedane*, and *Villabruna* all show a high degree of allele sharing with Mesolithic Western Europeans including *Loschbour* and *LaBrana1*, which are sometimes also called "Western Hunter Gatherers"[4] (Table S5.6). We view all these samples as closely related, along with *Hungarian.KO1* which clusters with them despite being from an Early Neolithic context[6].

When we include the lower coverage *Continenza* and *Bockstein* (each with 10,000-30,000 SNPs), they too show a genetic affinity to "Western Hunter Gatherers" (end of Table S5.6). We do not include *LesCloseaux13* and *Ofnet* in some analyses involving the Villabruna Cluster, since after restricting to damaged sequences, both have fewer than 10,000 SNPs.

We designate this combined set of samples as a broad "Villabruna Cluster" (Table S5.6). Within this cluster, *LaBrana1* and *Hungarian.KO1* are possible outgroups to the main cluster as revealed by outgroup-f_3 analysis (Figure S5.1) as well as statistics like *D(Villabruna_Cluster$_1$, Villabruna_Cluster$_2$; Villabruna_Cluster$_3$, Mbuti)* (Table S5.7). We caution, however, that the libraries from these two samples are both non-UDG-treated (the majority of samples analysed in this study are represented by UDG-treated libraries), and this could artifactually be contributing to an inference of them being outgroups. *Falkenstein* is difficult to place, as it is genetically closer to *Chaudardes1* than it is to *Villabruna* (Table S5.7). We do not attempt to more finely place *Continenza* and *Bockstein* because of their low coverage. There are complexities in the population structure within the Villabruna Cluster, which we explore in more detail in Supplementary Information section 13.

Figure S5.1 Heat matrix of f_3*(X, Y; Mbuti)* for selected Villabruna Cluster samples. We restrict the analysis to samples with at least 30,000 SNPs covered at least once.

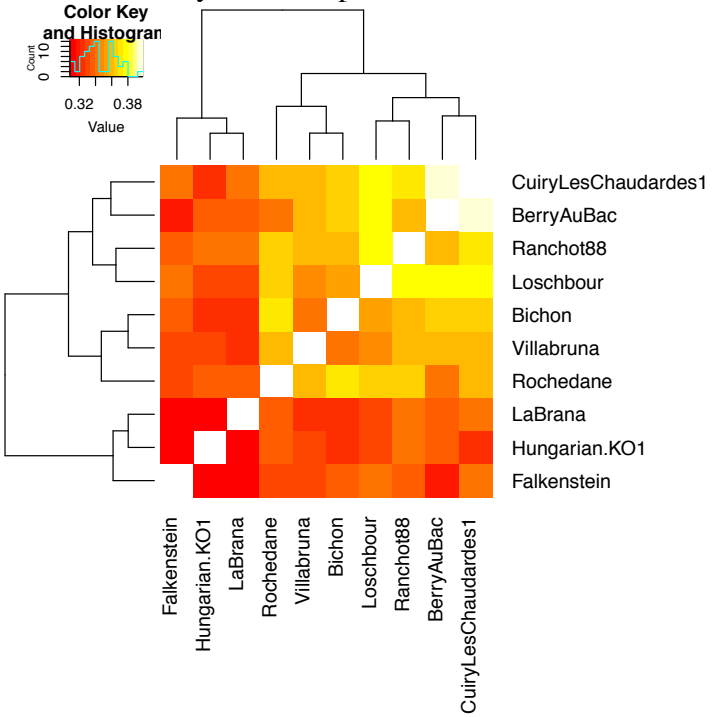


Table S5.6 Z-score of *D(X, Y; Villabruna_Cluster, Mbuti)*

D(X, Y; Villabruna, Mbuti) Villabruna: 1,215,433 SNPs								
X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-1.8	-5.3	-14.4	-29.5	-21.4	-24.2	-16.8
GoyetQ116-1	1.8	NA	-3.4	-12.4	-29.4	-19.4	-22.4	-15.2
Vestonice16	5.3	3.4	NA	-9.2	-25.1	-16.7	-18.5	-12.1
ElMiron	14.4	12.4	9.2	NA	-15.1	-6.6	-9.3	-1.8
Loschbour	29.5	29.4	25.1	15.1	NA	9.1	5.9	13.9
LaBrana1	21.4	19.4	16.7	6.6	-9.1	NA	-3	5.1
Hungarian.KO1	24.2	22.4	18.5	9.3	-5.9	3	NA	8
Motala12	16.8	15.2	12.1	1.8	-13.9	-5.1	-8	NA
Malta1	-1.8	-2.2	-6.6	-15	-31.1	-21.6	-23.8	-19.4
D(X, Y; Bichon, Mbuti) Bichon: 2,116,782 SNPs								
X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-2.4	-4.4	-16	-30.2	-24.7	-22.2	-17.1
GoyetQ116-1	2.4	NA	-2	-13.5	-28.8	-21.2	-19.7	-13.9

Vestonice16	4.4	2	NA	-12	-27.7	-19.8	-18.2	-12.7
ElMiron	16	13.5	12	NA	-15.7	-7.5	-6.8	0
Loschbour	30.2	28.8	27.7	15.7	NA	9.2	9	17
LaBrana1	24.7	21.2	19.8	7.5	-9.2	NA	0	8
Hungarian.KO1	22.2	19.7	18.2	6.8	-9	0	NA	8.2
Motala12	17.1	13.9	12.7	0	-17	-8	-8.2	NA
Malta1	-0.2	-2.5	-4.3	-15.6	-31.1	-24.5	-23	-17.1

D(X, Y; Rochedane, Mbuti) Rochedane: 237,390 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-2.1	-3.5	-14	-26.3	-17.9	-16.1	-12.6
GoyetQ116-1	2.1	NA	-1.5	-12.2	-23.1	-15.3	-13.5	-9.8
Vestonice16	3.5	1.5	NA	-10.7	-23.3	-14.5	-13.5	-10.2
ElMiron	14	12.2	10.7	NA	-11.2	-3.2	-2.6	1.6
Loschbour	26.3	23.1	23.3	11.2	NA	9.2	8.5	14.1
LaBrana1	17.9	15.3	14.5	3.2	-9.2	NA	0.7	5.1
Hungarian.KO1	16.1	13.5	13.5	2.6	-8.5	-0.7	NA	4.2
Motala12	12.6	9.8	10.2	-1.6	-14.1	-5.1	-4.2	NA
Malta1	-0.7	-3.3	-4.3	-13.5	-25.2	-17.9	-15.5	-13.6

D(X, Y; Ranchot88, Mbuti) Ranchot88: 414,863 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-3.7	-5.2	-16.8	-34.4	-23	-20.4	-14.4
GoyetQ116-1	3.7	NA	-0.9	-13.2	-29	-19	-17	-10.5
Vestonice16	5.2	0.9	NA	-12.2	-29.2	-17.8	-17.1	-10.1
ElMiron	16.8	13.2	12.2	NA	-16.8	-5.8	-4.7	2.6
Loschbour	34.4	29	29.2	16.8	NA	12.2	12.2	20.4
LaBrana1	23	19	17.8	5.8	-12.2	NA	0.9	8.2
Hungarian.KO1	20.4	17	17.1	4.7	-12.2	-0.9	NA	7.1
Motala12	14.4	10.5	10.1	-2.6	-20.4	-8.2	-7.1	NA
Malta1	0	-3.9	-5.2	-16.1	-31.5	-21.8	-20.7	-15

D(X, Y; Falkenstein, Mbuti) Falkenstein: 64,428 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-0.8	-2.6	-9.3	-20.4	-14.1	-11.2	-10.1
GoyetQ116-1	0.8	NA	-1.4	-8	-18.4	-11.7	-9.5	-8.1
Vestonice16	2.6	1.4	NA	-6.9	-17.1	-11.5	-9.2	-6.9
ElMiron	9.3	8	6.9	NA	-10.6	-5.1	-3.1	-0.2
Loschbour	20.4	18.4	17.1	10.6	NA	5.7	6.5	9.8
LaBrana1	14.1	11.7	11.5	5.1	-5.7	NA	1.4	4.3
Hungarian.KO1	11.2	9.5	9.2	3.1	-6.5	-1.4	NA	3.1
Motala12	10.1	8.1	6.9	0.2	-9.8	-4.3	-3.1	NA
Malta1	-0.8	-1.6	-3	-8.4	-19.6	-14.2	-10.7	-9.4

D(X, Y; Chaudardes1, Mbuti) Chaudardes1: 92,657 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-1.8	-4	-12	-26	-17.4	-12.4	-12.4
GoyetQ116-1	1.8	NA	-1.9	-9.7	-23.1	-14.2	-11.3	-9.9
Vestonice16	4	1.9	NA	-7.9	-20.6	-12.5	-8.7	-7.1
ElMiron	12	9.7	7.9	NA	-11.8	-3.5	-1.5	0.8
Loschbour	26	23.1	20.6	11.8	NA	8.8	10.7	13.6
LaBrana1	17.4	14.2	12.5	3.5	-8.8	NA	3.4	4.6
Hungarian.KO1	12.4	11.3	8.7	1.5	-10.7	-3.4	NA	1.6
Motala12	12.4	9.9	7.1	-0.8	-13.6	-4.6	-1.6	NA
Malta1	1.6	0.7	-1.7	-8.7	-21.7	-13.7	-9.9	-10

D(X, Y; BerryAuBac, Mbuti) BerryAuBac: 54,690 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-3.1	-2.6	-8.5	-19.9	-12.8	-12.4	-9.2
GoyetQ116-1	3.1	NA	0.1	-7	-16.9	-10.6	-9.8	-6.5
Vestonice16	2.6	-0.1	NA	-6	-17.6	-10.9	-9.6	-6.8
ElMiron	8.5	7	6	NA	-9.2	-3.9	-3.1	0.1
Loschbour	19.9	16.9	17.6	9.2	NA	6.8	6.6	10.6
LaBrana1	12.8	10.6	10.9	3.9	-6.8	NA	0.7	3.3
Hungarian.KO1	12.4	9.8	9.6	3.1	-6.6	-0.7	NA	2.5
Motala12	9.2	6.5	6.8	-0.1	-10.6	-3.3	-2.5	NA
Malta1	0.9	-1.7	-1.2	-7.5	-17.3	-12.5	-10.5	-8.1

D(X, Y; Continenza, Mbuti) Continenza: 11,717 SNPs

X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-0.4	-0.8	-4.4	-10.2	-6.1	-6	-3.1
GoyetQ116-1	0.4	NA	0.1	-4	-9.7	-4.8	-5.7	-2.8
Vestonice16	0.8	-0.1	NA	-4.4	-9.5	-5.5	-6.2	-2.9
ElMiron	4.4	4	4.4	NA	-4.7	-1.2	-1.8	1.7
Loschbour	10.2	9.7	9.5	4.7	NA	4.3	2.2	6.6
LaBrana1	6.1	4.8	5.5	1.2	-4.3	NA	-0.9	2.7
Hungarian.KO1	6	5.7	6.2	1.8	-2.2	0.9	NA	3.5
Motala12	3.1	2.8	2.9	-1.7	-6.6	-2.7	-3.5	NA

Malta1	-1.3	-1.2	-1.8	-5.1	-10.2	-7	-7	-4.9
D(X, Y; Bockstein, Mbuti) Bockstein: 21,977 SNPs								
X/Y	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Loschbour	LaBrana1	Hungarian.KO1	Motala12
Kostenki14	NA	-1.7	-0.8	-4.6	-11	-6.5	-6.6	-5.4
GoyetQ116-1	1.7	NA	0.5	-3.1	-9.6	-4.6	-6.2	-3.5
Vestonice16	0.8	-0.5	NA	-4.2	-10.2	-6	-6.9	-4.9
ElMiron	4.6	3.1	4.2	NA	-6.1	-1.7	-2.5	-0.5
Loschbour	11	9.6	10.2	6.1	NA	3.8	2.3	5.4
LaBrana1	6.5	4.6	6	1.7	-3.8	NA	-1.1	1
Hungarian.KO1	6.6	6.2	6.9	2.5	-2.3	1.1	NA	2.4
Motala12	5.4	3.5	4.9	0.5	-5.4	-1	-2.4	NA
Malta1	-0.2	-1.5	-0.6	-5.1	-10.6	-6	-7.4	-5.6

Table S5.7 Z of *D(Villabruna_Cluster$_1$, Villabruna_Cluster$_2$; Villabruna_Cluster$_3$, Mbuti)*

D(X, Y; Villabruna, Mbuti) Villabruna: 1,215,433 SNPs									
X/Y	Bichon	Rochedane	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Bichon	NA	-0.7	-0.4	-1.3	0.9	0.3	-1.2	7.5	3.9
Rochedane	0.7	NA	0.4	-0.9	0.8	0.9	-0.1	7.8	4.5
Ranchot88	0.4	-0.4	NA	-0.2	0.2	0.7	-1	8.2	3.9
Falkenstein	1.3	0.9	0.2	NA	-1	-0.1	0.2	6.8	3.7
Chaudardes1	-0.9	-0.8	-0.2	1	NA	0.5	-2.1	4.7	2.1
BerryAuBac	-0.3	-0.9	-0.7	0.1	-0.5	NA	-1	4.2	1.7
Loschbour	1.2	0.1	1	-0.2	2.1	1	NA	9.1	5.9
LaBrana1	-7.5	-7.8	-8.2	-6.8	-4.7	-4.2	-9.1	NA	-3
Hungarian.KO1	-3.9	-4.5	-3.9	-3.7	-2.1	-1.7	-5.9	3	NA
D(X, Y; Bichon, Mbuti) Bichon: 2,116,782 SNPs									
X/Y	Villabruna	Rochedane	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-3.4	-1.5	-2.8	-0.1	-1	-4.2	4.4	4.7
Rochedane	3.4	NA	2.3	0.2	2.4	0.7	0.1	8.1	7.4
Ranchot88	1.5	-2.3	NA	-1.2	0.7	-0.3	-2.6	6.6	6.2
Falkenstein	2.8	-0.2	1.2	NA	-0.3	-0.2	0.6	6.4	5.4
Chaudardes1	0.1	-2.4	-0.7	0.3	NA	0.1	-2.6	4.1	3.3
BerryAuBac	1	-0.7	0.3	0.2	-0.1	NA	-0.5	4.2	4.3
Loschbour	4.2	-0.1	2.6	-0.6	2.6	0.5	NA	9.2	9
LaBrana1	-4.4	-8.1	-6.6	-6.4	-4.1	-4.2	-9.2	NA	0
Hungarian.KO1	-4.7	-7.4	-6.2	-5.4	-3.3	-4.3	-9	0	NA
D(X, Y; Rochedane, Mbuti) Rochedane: 237,390 SNPs									
X/Y	Villabruna	Bichon	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-2.7	-0.8	-1.3	0.6	2.9	-0.6	7.9	7.8
Bichon	2.7	NA	1.6	0.8	2.4	3	2.7	10.1	9.2
Ranchot88	0.8	-1.6	NA	-0.9	0.1	2.2	-0.1	7.7	5.9
Falkenstein	1.3	-0.8	0.9	NA	0.4	-0.6	2	3.6	4.1
Chaudardes1	-0.6	-2.4	-0.1	-0.4	NA	0.6	-2.4	3.2	1.5
BerryAuBac	-2.9	-3	-2.2	0.6	-0.6	NA	-2.6	1.4	1.1
Loschbour	0.6	-2.7	0.1	-2	2.4	2.6	NA	9.2	8.5
LaBrana1	-7.9	-10.1	-7.7	-3.6	-3.2	-1.4	-9.2	NA	0.7
Hungarian.KO1	-7.8	-9.2	-5.9	-4.1	-1.5	-1.1	-8.5	-0.7	NA
D(X, Y; Ranchot88, Mbuti) Ranchot88: 700833 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.1	-1.1	-2.2	-2.8	-0.3	-6.1	5	4.5
Bichon	1.1	NA	-0.8	-2	-1.9	-0.3	-5.2	5.9	6.2
Rochedane	1.1	0.8	NA	-2	-0.8	0	-3.9	6	4.8
Falkenstein	2.2	2	2	NA	0.4	-1	-0.5	4.4	4.5
Chaudardes1	2.8	1.9	0.8	-0.4	NA	0.5	-1.5	5.2	4
BerryAuBac	0.3	0.3	0	1	-0.5	NA	-2.1	3	3.3
Loschbour	6.1	5.2	3.9	0.5	1.5	2.1	NA	12.2	12.2
LaBrana1	-5	-5.9	-6	-4.4	-5.2	-3	-12.2	NA	0.9
Hungarian.KO1	-4.5	-6.2	-4.8	-4.5	-4	-3.3	-12.2	-0.9	NA
D(X, Y; Falkenstein, Mbuti) Falkenstein: 64,428 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.5	-0.4	-2.1	-3.4	-0.3	-2	3.4	3.3
Bichon	1.5	NA	0.6	-0.9	-0.3	0.7	-0.9	5	4.4
Rochedane	0.4	-0.6	NA	-1.2	-0.1	-1	-1.2	1.5	2.2
Ranchot88	2.1	0.9	1.2	NA	0.4	0.8	0.5	4.7	4.4
Chaudardes1	3.4	0.3	0.1	-0.4	NA	1.1	0.1	1.8	2.2
BerryAuBac	0.3	-0.7	1	-0.8	-1.1	NA	0.2	1	1.5
Loschbour	2	0.9	1.2	-0.5	-0.1	-0.2	NA	5.7	6.5
LaBrana1	-3.4	-5	-1.5	-4.7	-1.8	-1	-5.7	NA	1.4
Hungarian.KO1	-3.3	-4.4	-2.2	-4.4	-2.2	-1.5	-6.5	-1.4	NA
D(X, Y; Chaudardes1, Mbuti) Chaudardes1: 92,657 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1	-0.2	-3	-2.3	-1	-4.7	3.4	5.4
Bichon	1	NA	0.2	-2.7	0	-0.6	-4.1	4.8	6.2
Rochedane	0.2	-0.2	NA	-0.9	-0.5	-0.8	-2.5	2.9	2.9
Ranchot88	3	2.7	0.9	NA	0.1	-1.3	-0.8	5.3	6.4
Falkenstein	2.3	0	0.5	-0.1	NA	1	0.3	2.9	3.9
BerryAuBac	1	0.6	0.8	1.3	-1	NA	0.1	2.5	2.3
Loschbour	4.7	4.1	2.5	0.8	-0.3	-0.1	NA	8.8	10.7

LaBrana1	-3.4	-4.8	-2.9	-5.3	-2.9	-2.5	-8.8	NA	3.4
Hungarian.KO1	-5.4	-6.2	-2.9	-6.4	-3.9	-2.3	-10.7	-3.4	NA
D(X, Y; BerryAuBac, Mbuti) BerryAuBac: 54,690									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.2	2.2	-1.1	-0.1	-1.5	-3.4	3.2	3.4
Bichon	1.2	NA	2.5	0	1	-0.7	-2	4.4	4.2
Rochedane	-2.2	-2.5	NA	-2.4	-0.4	-1.3	-3.7	0.8	0.7
Ranchot88	1.1	0	2.4	NA	1.9	-1.8	-1.1	3.6	3
Falkenstein	0.1	-1	0.4	-1.9	NA	-0.2	-0.5	1.5	1.2
Chaudardes1	1.5	0.7	1.3	1.8	0.2	NA	1.3	3.7	2.6
Loschbour	3.4	2	3.7	1.1	0.5	-1.3	NA	6.8	6.6
LaBrana1	-3.2	-4.4	-0.8	-3.6	-1.5	-3.7	-6.8	NA	0.7
Hungarian.KO1	-3.4	-4.2	-0.7	-3	-1.2	-2.6	-6.6	-0.7	NA
D(X, Y; Loschbour, Mbuti) Loschbour: 2,091,584 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	LaBrana1	KO1
Villabruna	NA	-3	-0.6	-5.2	-2.2	-2.6	-2.4	4.2	4.6
Bichon	3	NA	2.5	-2.8	-1.5	-1.4	-1.5	7.2	7
Rochedane	0.6	-2.5	NA	-3.7	-3.2	-0.3	-1.1	5.2	5.2
Ranchot88	5.2	2.8	3.7	NA	1.1	0.7	1	10	9.8
Falkenstein	2.2	1.5	3.2	-1.1	NA	0.2	-0.6	5.7	5
Chaudardes1	2.6	1.4	0.3	-0.7	-0.2	NA	1.1	6.9	6.8
BerryAuBac	2.4	1.5	1.1	-1	0.6	-1.1	NA	6	4.8
LaBrana1	-4.2	-7.2	-5.2	-10	-5.7	-6.9	-6	NA	0.1
Hungarian.KO1	-4.6	-7	-5.2	-9.8	-5	-6.8	-4.8	-0.1	NA
D(X, Y; LaBrana1, Mbuti) LaBrana1: 1,884,745 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	Loschbour	KO1
Villabruna	NA	-3.2	0.1	-3.6	-3.4	-1.4	-0.9	-5.7	3.4
Bichon	3.2	NA	2	-0.6	-1.6	0.7	0.4	-1.8	5.7
Rochedane	-0.1	-2	NA	-1.6	-2	-0.2	-0.6	-4.2	3.4
Ranchot88	3.6	0.6	1.6	NA	0.3	0.5	0.7	-2.1	6.6
Falkenstein	3.4	1.6	2	-0.3	NA	1.1	0.5	-0.4	5
Chaudardes1	1.4	-0.7	0.2	-0.5	-1.1	NA	1	-2.2	5.2
BerryAuBac	0.9	-0.4	0.6	-0.7	-0.5	-1	NA	-1.2	2.7
Loschbour	5.7	1.8	4.2	2.1	0.4	2.2	1.2	NA	7.9
Hungarian.KO1	-3.4	-5.7	-3.4	-6.6	-5	-5.2	-2.7	-7.9	NA
D(X, Y; Hungarian.KO1, Mbuti) Hungarian.KO1: 1,410,303 SNPs									
X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	Loschbour	LaBrana1
Villabruna	NA	0.7	2.9	0.4	-0.2	3.5	1.6	-1.4	6.4
Bichon	-0.7	NA	2.1	0.1	-0.6	2.9	0	-1.8	5.5
Rochedane	-2.9	-2.1	NA	-1.1	-1.7	1.4	-0.4	-3.1	2.9
Ranchot88	-0.4	-0.1	1.1	NA	-0.1	2.5	-0.2	-1.7	5.4
Falkenstein	0.2	0.6	1.7	0.1	NA	1.6	-0.3	-1.9	3.4
Chaudardes1	-3.5	-2.9	-1.4	-2.5	-1.6	NA	0.3	-4	1.8
BerryAuBac	-1.6	0	0.4	0.2	0.3	-0.3	NA	-2.1	2
Loschbour	1.4	1.8	3.1	1.7	1.9	4	2.1	NA	7.4
LaBrana1	-6.4	-5.5	-2.9	-5.4	-3.4	-1.8	-2	-7.4	NA
D(X, Y; Villabruna, Mbuti) Villabruna: 1,215,433 SNPs									
X/Y	Bichon	Rochedane	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Bichon	NA	-0.7	-0.4	-1.3	0.9	0.3	-1.2	7.5	3.9
Rochedane	0.7	NA	0.4	-0.9	0.8	0.9	-0.1	7.8	4.5
Ranchot88	0.4	-0.4	NA	-0.2	0.2	0.7	-1	8.2	3.9
Falkenstein	1.3	0.9	0.2	NA	-1	-0.1	0.2	6.8	3.7
Chaudardes1	-0.9	-0.8	-0.2	1	NA	0.5	-2.1	4.7	2.1
BerryAuBac	-0.3	-0.9	-0.7	0.1	-0.5	NA	-1	4.2	1.7
Loschbour	1.2	0.1	1	-0.2	2.1	1	NA	9.1	5.9
LaBrana1	-7.5	-7.8	-8.2	-6.8	-4.7	-4.2	-9.1	NA	-3
Hungarian.KO1	-3.9	-4.5	-3.9	-3.7	-2.1	-1.7	-5.9	3	NA
D(X, Y; Bichon, Mbuti) Bichon: 2,116,782 SNPs									
X/Y	Villabruna	Rochedane	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-3.4	-1.5	-2.8	-0.1	-1	-4.2	4.4	4.7
Rochedane	3.4	NA	2.3	0.2	2.4	0.7	0.1	8.1	7.4
Ranchot88	1.5	-2.3	NA	-1.2	0.7	-0.3	-2.6	6.6	6.2
Falkenstein	2.8	-0.2	1.2	NA	-0.3	-0.2	0.6	6.4	5.4
Chaudardes1	0.1	-2.4	-0.7	0.3	NA	0.1	-2.6	4.1	3.3
BerryAuBac	1	-0.7	0.3	0.2	-0.1	NA	-0.5	4.2	4.3
Loschbour	4.2	-0.1	2.6	-0.6	2.6	0.5	NA	9.2	9
LaBrana1	-4.4	-8.1	-6.6	-6.4	-4.1	-4.2	-9.2	NA	0
Hungarian.KO1	-4.7	-7.4	-6.2	-5.4	-3.3	-4.3	-9	0	NA
D(X, Y; Rochedane, Mbuti) Rochedane: 237,390 SNPs									
X/Y	Villabruna	Bichon	Ranchot88	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-2.7	-0.8	-1.3	0.6	2.9	-0.6	7.9	7.8
Bichon	2.7	NA	1.6	0.8	2.4	3	2.7	10.1	9.2
Ranchot88	0.8	-1.6	NA	-0.9	0.1	2.2	-0.1	7.7	5.9
Falkenstein	1.3	-0.8	0.9	NA	0.4	-0.6	2	3.6	4.1
Chaudardes1	-0.6	-2.4	-0.1	-0.4	NA	0.6	-2.4	3.2	1.5
BerryAuBac	-2.9	-3	-2.2	0.6	-0.6	NA	-2.6	1.4	1.1
Loschbour	0.6	-2.7	0.1	-2	2.4	2.6	NA	9.2	8.5
LaBrana1	-7.9	-10.1	-7.7	-3.6	-3.2	-1.4	-9.2	NA	0.7
Hungarian.KO1	-7.8	-9.2	-5.9	-4.1	-1.5	-1.1	-8.5	-0.7	NA
D(X, Y; Ranchot88, Mbuti) Ranchot88: 700,833 SNPs									

X/Y	Villabruna	Bichon	Rochedane	Falkenstein	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.1	-1.1	-2.2	-2.8	-0.3	-6.1	5	4.5
Bichon	1.1	NA	-0.8	-2	-1.9	-0.3	-5.2	5.9	6.2
Rochedane	1.1	0.8	NA	-2	-0.8	0	-3.9	6	4.8
Falkenstein	2.2	2	2	NA	0.4	-1	-0.5	4.4	4.5
Chaudardes1	2.8	1.9	0.8	-0.4	NA	0.5	-1.5	5.2	4
BerryAuBac	0.3	0.3	0	1	-0.5	NA	-2.1	3	3.3
Loschbour	6.1	5.2	3.9	0.5	1.5	2.1	NA	12.2	12.2
LaBrana1	-5	-5.9	-6	-4.4	-5.2	-3	-12.2	NA	0.9
Hungarian.KO1	-4.5	-6.2	-4.8	-4.5	-4	-3.3	-12.2	-0.9	NA

D(X, Y; Falkenstein, Mbuti) Falkenstein: 64,428 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Chaudardes	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.5	-0.4	-2.1	-3.4	-0.3	-2	3.4	3.3
Bichon	1.5	NA	0.6	-0.9	-0.3	0.7	-0.9	5	4.4
Rochedane	0.4	-0.6	NA	-1.2	-0.1	-1	-1.2	1.5	2.2
Ranchot88	2.1	0.9	1.2	NA	0.4	0.8	0.5	4.7	4.4
Chaudardes1	3.4	0.3	0.1	-0.4	NA	1.1	0.1	1.8	2.2
BerryAuBac	0.3	-0.7	1	-0.8	-1.1	NA	0.2	1	1.5
Loschbour	2	0.9	1.2	-0.5	-0.1	-0.2	NA	5.7	6.5
LaBrana1	-3.4	-5	-1.5	-4.7	-1.8	-1	-5.7	NA	1.4
Hungarian.KO1	-3.3	-4.4	-2.2	-4.4	-2.2	-1.5	-6.5	-1.4	NA

D(X, Y; Chaudardes1, Mbuti) Chaudardes1: 92,657 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Berry	Loschbour	LaBrana1	KO1
Villabruna	NA	-1	-0.2	-3	-2.3	-1	-4.7	3.4	5.4
Bichon	1	NA	0.2	-2.7	0	-0.6	-4.1	4.8	6.2
Rochedane	0.2	-0.2	NA	-0.9	-0.5	-0.8	-2.5	2.9	2.9
Ranchot88	3	2.7	0.9	NA	0.1	-1.3	-0.8	5.3	6.4
Falkenstein	2.3	0	0.5	-0.1	NA	1	0.3	2.9	3.9
BerryAuBac	1	0.6	0.8	1.3	-1	NA	0.1	2.5	2.3
Loschbour	4.7	4.1	2.5	0.8	-0.3	-0.1	NA	8.8	10.7
LaBrana1	-3.4	-4.8	-2.9	-5.3	-2.9	-2.5	-8.8	NA	3.4
Hungarian.KO1	-5.4	-6.2	-2.9	-6.4	-3.9	-2.3	-10.7	-3.4	NA

D(X, Y; BerryAuBac, Mbuti) BerryAuBac: 54,690 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Loschbour	LaBrana1	KO1
Villabruna	NA	-1.2	2.2	-1.1	-0.1	-1.5	-3.4	3.2	3.4
Bichon	1.2	NA	2.5	0	1	-0.7	-2	4.4	4.2
Rochedane	-2.2	-2.5	NA	-2.4	-0.4	-1.3	-3.7	0.8	0.7
Ranchot88	1.1	0	2.4	NA	1.9	-1.8	-1.1	3.6	3
Falkenstein	0.1	-1	0.4	-1.9	NA	-0.2	-0.5	1.5	1.2
Chaudardes1	1.5	0.7	1.3	1.8	0.2	NA	1.3	3.7	2.6
Loschbour	3.4	2	3.7	1.1	0.5	-1.3	NA	6.8	6.6
LaBrana1	-3.2	-4.4	-0.8	-3.6	-1.5	-3.7	-6.8	NA	0.7
Hungarian.KO1	-3.4	-4.2	-0.7	-3	-1.2	-2.6	-6.6	-0.7	NA

D(X, Y; Loschbour, Mbuti) Loschbour: 2,091,584 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	LaBrana1	KO1
Villabruna	NA	-3	-0.6	-5.2	-2.2	-2.6	-2.4	4.2	4.6
Bichon	3	NA	2.5	-2.8	-1.5	-1.4	-1.5	7.2	7
Rochedane	0.6	-2.5	NA	-3.7	-3.2	-0.3	-1.1	5.2	5.2
Ranchot88	5.2	2.8	3.7	NA	1.1	0.7	1	10	9.8
Falkenstein	2.2	1.5	3.2	-1.1	NA	0.2	-0.6	5.7	5
Chaudardes1	2.6	1.4	0.3	-0.7	-0.2	NA	1.1	6.9	6.8
BerryAuBac	2.4	1.5	1.1	-1	0.6	-1.1	NA	6	4.8
LaBrana1	-4.2	-7.2	-5.2	-10	-5.7	-6.9	-6	NA	0.1
Hungarian.KO1	-4.6	-7	-5.2	-9.8	-5	-6.8	-4.8	-0.1	NA

D(X, Y; LaBrana1, Mbuti) LaBrana1: 1,884,745 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	Loschbour	KO1
Villabruna	NA	-3.2	0.1	-3.6	-3.4	-1.4	-0.9	-5.7	3.4
Bichon	3.2	NA	2	-0.6	-1.6	0.7	0.4	-1.8	5.7
Rochedane	-0.1	-2	NA	-1.6	-2	-0.2	-0.6	-4.2	3.4
Ranchot88	3.6	0.6	1.6	NA	0.3	0.5	0.7	-2.1	6.6
Falkenstein	3.4	1.6	2	-0.3	NA	1.1	0.5	-0.4	5
Chaudardes1	1.4	-0.7	0.2	-0.5	-1.1	NA	1	-2.2	5.2
BerryAuBac	0.9	-0.4	0.6	-0.7	-0.5	-1	NA	-1.2	2.7
Loschbour	5.7	1.8	4.2	2.1	0.4	2.2	1.2	NA	7.9
Hungarian.KO1	-3.4	-5.7	-3.4	-6.6	-5	-5.2	-2.7	-7.9	NA

D(X, Y; Hungarian.KO1, Mbuti) Hungarian.KO1: 1,410,303 SNPs

X/Y	Villabruna	Bichon	Rochedane	Ranchot88	Falkenstein	Chau	Berry	Loschbour	LaBrana1
Villabruna	NA	0.7	2.9	0.4	-0.2	3.5	1.6	-1.4	6.4
Bichon	-0.7	NA	2.1	0.1	-0.6	2.9	0	-1.8	5.5
Rochedane	-2.9	-2.1	NA	-1.1	-1.7	1.4	-0.4	-3.1	2.9
Ranchot88	-0.4	-0.1	1.1	NA	-0.1	2.5	-0.2	-1.7	5.4
Falkenstein	0.2	0.6	1.7	0.1	NA	1.6	-0.3	-1.9	3.4
Chaudardes1	-3.5	-2.9	-1.4	-2.5	-1.6	NA	0.3	-4	1.8
BerryAuBac	-1.6	0	0.4	0.2	0.3	-0.3	NA	-2.1	2
Loschbour	1.4	1.8	3.1	1.7	1.9	4	2.1	NA	7.4
LaBrana1	-6.4	-5.5	-2.9	-5.4	-3.4	-1.8	-2	-7.4	NA

"Mal'ta cluster" (these individuals all lived 24,000-17,000 BP in Siberia)

Samples from the sites of Mal'ta (*Malta1*) and Afontova Gora (*AfontovaGora2* and *AfontovaGora3*)—both east of the Ural Mountains—descend from a common ancestral population relative to the other samples, but also show important sub-structure (Table S5.8):

- *Malta1* shares more drift with *AfontovaGora3* than with pre-Neolithic Europeans.
- *AfontovaGora3* appears to derive from a lineage of the Mal'ta Cluster that contributed more to some later human populations than did the lineage leading to *Malta1* itself, as: (i) *Karelia* shares more alleles with *AfontovaGora3* than with *Malta1*, and (ii) Native Americans share more alleles with *AfontovaGora3* than with *Malta1*.
- *AfontovaGora2* is not genetically closer to *AfontovaGora3* than it is to *Malta1*. Thus, there is no evidence of an "Afontova Gora Cluster".

Table S5.8 *D(X, Y; Z, Mbuti)* with *Malta1* and *AfontovaGora3* always in the statistic.

X	Y	Z	Mbuti	D value	Z score	Sites used
Malta1	AfontovaGora3	UstIshim	Mbuti	0.0008	1.3	509641
Malta1	AfontovaGora3	Oase1	Mbuti	0.0003	0.4	116250
Malta1	AfontovaGora3	Kostenki14	Mbuti	0.0011	1.7	486631
Malta1	AfontovaGora3	GoyetQ116-1	Mbuti	0.0009	1.4	401152
Malta1	AfontovaGora3	Vestonice16	Mbuti	0.0010	1.5	398604
Malta1	AfontovaGora3	Ostuni1	Mbuti	0.0006	0.7	189230
Malta1	AfontovaGora3	KremsWA3	Mbuti	0.0004	0.5	120508
Malta1	AfontovaGora3	ElMiron	Mbuti	0.0001	0.1	354861
Malta1	AfontovaGora3	Villabruna	Mbuti	-0.0004	-0.6	459597
Malta1	AfontovaGora3	Ranchot88	Mbuti	-0.0001	-0.2	223016
Malta1	AfontovaGora3	Loschbour	Mbuti	0.0004	0.7	505247
Malta1	AfontovaGora3	LaBrana1	Mbuti	-0.0008	-1.3	489381
Malta1	AfontovaGora3	Hungarian.KO1	Mbuti	-0.0005	-0.8	354548
Malta1	AfontovaGora3	Motala12	Mbuti	-0.0005	-0.4	55428
Malta1	AfontovaGora3	Karelia	Mbuti	-0.0055	-7.5	444390
Malta1	AfontovaGora3	AfontovaGora2	Mbuti	0.0016	1.2	45578
Malta1	AfontovaGora3	Stuttgart	Mbuti	0.0000	-0.1	502197
Malta1	AfontovaGora3	French	Mbuti	-0.0006	-1.4	510782
Malta1	AfontovaGora3	Sardinian	Mbuti	0.0001	0.2	510785
Malta1	AfontovaGora3	Han	Mbuti	0.0002	0.4	510784
Malta1	AfontovaGora3	Dai	Mbuti	0.0003	0.6	510787
Malta1	AfontovaGora3	Karitiana	Mbuti	-0.0030	-5.3	510778
Malta1	UstIshim	AfontovaGora3	Mbuti	0.0247	28.6	509641
Malta1	Oase1	AfontovaGora3	Mbuti	0.0266	22.9	116250
Malta1	Kostenki14	AfontovaGora3	Mbuti	0.0192	23.8	486631
Malta1	GoyetQ116-1	AfontovaGora3	Mbuti	0.0181	20.9	401152
Malta1	Vestonice16	AfontovaGora3	Mbuti	0.0180	21.0	398604
Malta1	Ostuni1	AfontovaGora3	Mbuti	0.0179	18.2	189230
Malta1	KremsWA3	AfontovaGora3	Mbuti	0.0180	16.0	120508
Malta1	ElMiron	AfontovaGora3	Mbuti	0.0179	20.8	354861
Malta1	Villabruna	AfontovaGora3	Mbuti	0.0165	20.0	459597
Malta1	Ranchot88	AfontovaGora3	Mbuti	0.0174	18.7	223016
Malta1	Loschbour	AfontovaGora3	Mbuti	0.0167	21.1	505247
Malta1	LaBrana1	AfontovaGora3	Mbuti	0.0162	20.0	489381
Malta1	Hungarian.KO1	AfontovaGora3	Mbuti	0.0161	19.4	354548
Malta1	Motala12	AfontovaGora3	Mbuti	0.0116	8.9	55428
Malta1	Karelia	AfontovaGora3	Mbuti	0.0037	4.2	444390
Malta1	AfontovaGora2	AfontovaGora3	Mbuti	0.0011	0.8	45578
Malta1	Stuttgart	AfontovaGora3	Mbuti	0.0212	27.7	502197
Malta1	French	AfontovaGora3	Mbuti	0.0173	25.3	510782
Malta1	Sardinian	AfontovaGora3	Mbuti	0.0202	29.1	510785
Malta1	Han	AfontovaGora3	Mbuti	0.0222	31.0	510784
Malta1	Dai	AfontovaGora3	Mbuti	0.0220	31.2	510787
Malta1	Karitiana	AfontovaGora3	Mbuti	0.0112	14.9	510778
UstIshim	AfontovaGora3	Malta1	Mbuti	-0.0239	-26.8	509641
Oase1	AfontovaGora3	Malta1	Mbuti	-0.0263	-23.1	116250
Kostenki14	AfontovaGora3	Malta1	Mbuti	-0.0180	-20.8	486631
GoyetQ116-1	AfontovaGora3	Malta1	Mbuti	-0.0172	-18.6	401152
Vestonice16	AfontovaGora3	Malta1	Mbuti	-0.0170	-19.2	398604
Ostuni1	AfontovaGora3	Malta1	Mbuti	-0.0173	-17.5	189230
KremsWA3	AfontovaGora3	Malta1	Mbuti	-0.0176	-16.4	120508

ElMiron	AfontovaGora3	Malta1	Mbuti	-0.0178	-20.1	354861
Villabruna	AfontovaGora3	Malta1	Mbuti	-0.0169	-20.0	459597
Ranchot88	AfontovaGora3	Malta1	Mbuti	-0.0175	-18.5	223016
Loschbour	AfontovaGora3	Malta1	Mbuti	-0.0163	-20.4	505247
LaBrana1	AfontovaGora3	Malta1	Mbuti	-0.0171	-21.0	489381
Hungarian.KO1	AfontovaGora3	Malta1	Mbuti	-0.0166	-19.9	354548
Motala12	AfontovaGora3	Malta1	Mbuti	-0.0121	-9.1	55428
Karelia	AfontovaGora3	Malta1	Mbuti	-0.0091	-10.9	444390
AfontovaGora2	AfontovaGora3	Malta1	Mbuti	0.0005	0.4	45578
Stuttgart	AfontovaGora3	Malta1	Mbuti	-0.0212	-26.7	502197
French	AfontovaGora3	Malta1	Mbuti	-0.0180	-25.9	510782
Sardinian	AfontovaGora3	Malta1	Mbuti	-0.0202	-28.0	510785
Han	AfontovaGora3	Malta1	Mbuti	-0.0220	-28.6	510784
Dai	AfontovaGora3	Malta1	Mbuti	-0.0218	-28.8	510787
Karitiana	AfontovaGora3	Malta1	Mbuti	-0.0142	-19.4	510778

"Satsurblia Cluster" (these individuals all lived 13,000-10,000 BP in the Caucasus)

The *Satsurblia* and *Kotias*[8] samples have several features that are distinctive from those of the other samples in this study (Table S5.9):

- *Satsurblia* shares more drift with *Kotias* than with pre-Neolithic Europeans.
- *ElMiron* shares more drift with *Kotias* (9,700 BP) than with *Satsurblia* (13,300 BP).

The relationships of these samples to the ancient Europeans who are the focus of this study is investigated in more detail in Supplementary Information section 12.

Table S5.9 *D(X, Y; Z, Mbuti)* where *Satsurblia* and *Kotias* are always in the statistic.

				All sites			Transversions only		
X	**Y**	**Z**	**Mbuti**	**D value**	**Z score**	**Sites used**	**D value**	**Z score**	**Sites used**
Satsurblia	Kotias	UstIshim	Mbuti	0.0002	0.4	1415190	0.0001	0.2	766712
Satsurblia	Kotias	Oase1	Mbuti	-0.0007	-1.1	189526	-0.0009	-1.0	82749
Satsurblia	Kotias	Kostenki14	Mbuti	-0.0007	-1.2	1152869	-0.0004	-0.8	553455
Satsurblia	Kotias	GoyetQ116-1	Mbuti	-0.0012	-1.9	558196	-0.0014	-1.6	130649
Satsurblia	Kotias	Vestonice16	Mbuti	-0.0005	-0.8	625570	-0.0002	-0.3	217427
Satsurblia	Kotias	Ostuni1	Mbuti	-0.0003	-0.4	246277	-0.0006	-0.5	71079
Satsurblia	Kotias	KremsWA3	Mbuti	-0.0010	-1.2	157117	-0.0014	-1.0	35085
Satsurblia	Kotias	ElMiron	Mbuti	-0.0019	-3.4	529583	-0.0028	-4.1	181778
Satsurblia	Kotias	HohleFels49	Mbuti	-0.0003	-0.3	43680	-0.0028	-1.1	9354
Satsurblia	Kotias	Villabruna	Mbuti	-0.0011	-2.0	801136	-0.0013	-2.2	307477
Satsurblia	Kotias	Bichon	Mbuti	-0.0002	-0.5	1407253	0.0000	0.0	759382
Satsurblia	Kotias	Ranchot88	Mbuti	-0.0017	-2.4	272886	-0.0016	-1.4	59040
Satsurblia	Kotias	Loschbour	Mbuti	-0.0007	-1.3	1407104	-0.0007	-1.4	763252
Satsurblia	Kotias	LaBrana1	Mbuti	-0.0014	-2.7	1257041	-0.0013	-2.3	659801
Satsurblia	Kotias	Hungarian.KO1	Mbuti	-0.0007	-1.2	957314	-0.0012	-2.0	511190
Satsurblia	Kotias	Motala12	Mbuti	-0.0016	-3.0	1260562	-0.0014	-2.6	679159
Satsurblia	Kotias	Karelia	Mbuti	-0.0006	-1.1	1186256	-0.0004	-0.8	627984
Satsurblia	Kotias	Malta1	Mbuti	-0.0002	-0.4	970714	-0.0002	-0.3	514145
Satsurblia	Kotias	AfontovaGora3	Mbuti	0.0003	0.3	190536	-0.0002	-0.1	42150
Satsurblia	Kotias	Stuttgart	Mbuti	-0.0005	-1.0	1397826	-0.0004	-0.8	759544
Satsurblia	Kotias	French	Mbuti	-0.0011	-2.8	1417761	-0.0011	-2.6	767709
Satsurblia	Kotias	Sardinian	Mbuti	-0.0010	-2.6	1417775	-0.0009	-2.3	767719
Satsurblia	Kotias	Han	Mbuti	-0.0003	-0.8	1417774	-0.0003	-0.7	767720
Satsurblia	Kotias	Dai	Mbuti	-0.0005	-1.4	1417780	-0.0006	-1.4	767723
Satsurblia	Kotias	Karitiana	Mbuti	-0.0006	-1.3	1417731	-0.0008	-1.7	767698
UstIshim	Kotias	Satsurblia	Mbuti	-0.0177	-24.0	1415190	-0.0162	-22.3	766712
Oase1	Kotias	Satsurblia	Mbuti	-0.0210	-23.0	189526	-0.0177	-17.0	82749
Kostenki14.sg	Kotias	Satsurblia	Mbuti	-0.0137	-19.9	1272240	-0.0122	-17.8	678547
Kostenki14	Kotias	Satsurblia	Mbuti	-0.0140	-20.2	1152869	-0.0125	-18.4	553455
GoyetQ116-1	Kotias	Satsurblia	Mbuti	-0.0152	-19.2	558196	-0.0152	-15.5	130649
Vestonice16	Kotias	Satsurblia	Mbuti	-0.0143	-19.6	625570	-0.0133	-16.4	217427
Ostuni1	Kotias	Satsurblia	Mbuti	-0.0140	-16.2	246277	-0.0128	-10.9	71079
KremsWA3	Kotias	Satsurblia	Mbuti	-0.0143	-15.1	157117	-0.0138	-9.0	35085
ElMiron	Kotias	Satsurblia	Mbuti	-0.0137	-18.1	529583	-0.0125	-14.8	181778
HohleFels49	Kotias	Satsurblia	Mbuti	-0.0143	-10.9	43680	-0.0115	-4.3	9354
Villabruna	Kotias	Satsurblia	Mbuti	-0.0116	-16.6	801136	-0.0107	-14.2	307477
Bichon	Kotias	Satsurblia	Mbuti	-0.0107	-15.8	1407253	-0.0094	-14.0	759382
Ranchot88	Kotias	Satsurblia	Mbuti	-0.0127	-14.9	272886	-0.0107	-8.3	59040
Loschbour	Kotias	Satsurblia	Mbuti	-0.0105	-16.1	1407104	-0.0096	-15.0	763252
LaBrana1	Kotias	Satsurblia	Mbuti	-0.0120	-17.9	1257041	-0.0107	-16.1	659801
Hungarian.KO1	Kotias	Satsurblia	Mbuti	-0.0107	-15.5	957314	-0.0100	-14.3	511190
Motala12	Kotias	Satsurblia	Mbuti	-0.0106	-16.4	1260562	-0.0096	-14.5	679159
Karelia	Kotias	Satsurblia	Mbuti	-0.0096	-14.0	1186256	-0.0086	-12.3	627984
Malta1	Kotias	Satsurblia	Mbuti	-0.0118	-16.6	970714	-0.0106	-14.6	514145

AfontovaGora3	Kotias	Satsurblia	Mbuti	-0.0103	-11.4	190536	-0.0093	-6.6	42150
Stuttgart	Kotias	Satsurblia	Mbuti	-0.0100	-14.9	1397826	-0.0091	-13.3	759544
French	Kotias	Satsurblia	Mbuti	-0.0093	-16.5	1417761	-0.0087	-15.3	767709
Sardinian	Kotias	Satsurblia	Mbuti	-0.0103	-18.3	1417775	-0.0096	-16.8	767719
Han	Kotias	Satsurblia	Mbuti	-0.0162	-26.6	1417774	-0.0147	-24.6	767720
Dai	Kotias	Satsurblia	Mbuti	-0.0164	-26.7	1417780	-0.0150	-24.8	767723
Karitiana	Kotias	Satsurblia	Mbuti	-0.0144	-22.5	1417731	-0.0134	-21.1	767698
UstIshim	Satsurblia	Kotias	Mbuti	-0.0179	-24.5	1415190	-0.0162	-22.3	766712
Oase1	Satsurblia	Kotias	Mbuti	-0.0203	-21.6	189526	-0.0168	-15.7	82749
Kostenki14.sg	Satsurblia	Kotias	Mbuti	-0.0133	-19.2	1272240	-0.0120	-17.3	678547
Kostenki14	Satsurblia	Kotias	Mbuti	-0.0134	-19.4	1152869	-0.0120	-17.4	553455
GoyetQ116-1	Satsurblia	Kotias	Mbuti	-0.0140	-17.3	558196	-0.0138	-13.7	130649
Vestonice16	Satsurblia	Kotias	Mbuti	-0.0138	-18.2	625570	-0.0131	-15.7	217427
Ostuni1	Satsurblia	Kotias	Mbuti	-0.0136	-16.1	246277	-0.0122	-10.2	71079
KremsWA3	Satsurblia	Kotias	Mbuti	-0.0133	-14.2	157117	-0.0124	-7.8	35085
ElMiron	Satsurblia	Kotias	Mbuti	-0.0118	-15.6	529583	-0.0097	-11.0	181778
HohleFels49	Satsurblia	Kotias	Mbuti	-0.0140	-10.3	43680	-0.0087	-3.2	9354
Villabruna	Satsurblia	Kotias	Mbuti	-0.0105	-14.9	801136	-0.0093	-12.4	307477
Bichon	Satsurblia	Kotias	Mbuti	-0.0104	-15.6	1407253	-0.0094	-14.4	759382
Ranchot88	Satsurblia	Kotias	Mbuti	-0.0110	-12.4	272886	-0.0090	-6.8	59040
Loschbour	Satsurblia	Kotias	Mbuti	-0.0098	-15.2	1407104	-0.0089	-14.1	763252
LaBrana1	Satsurblia	Kotias	Mbuti	-0.0106	-15.3	1257041	-0.0094	-13.6	659801
Hungarian.KO1	Satsurblia	Kotias	Mbuti	-0.0100	-14.6	957314	-0.0088	-12.4	511190
Motala12	Satsurblia	Kotias	Mbuti	-0.0090	-13.1	1260562	-0.0082	-12.1	679159
Karelia	Satsurblia	Kotias	Mbuti	-0.0090	-13.6	1186256	-0.0082	-12.0	627984
Malta1	Satsurblia	Kotias	Mbuti	-0.0115	-15.8	970714	-0.0104	-14.3	514145
AfontovaGora3	Satsurblia	Kotias	Mbuti	-0.0106	-11.5	190536	-0.0091	-6.1	42150
Stuttgart	Satsurblia	Kotias	Mbuti	-0.0095	-14.2	1397826	-0.0087	-13.0	759544
French	Satsurblia	Kotias	Mbuti	-0.0082	-14.1	1417761	-0.0077	-13.2	767709
Sardinian	Satsurblia	Kotias	Mbuti	-0.0093	-15.7	1417775	-0.0087	-14.6	767719
Han	Satsurblia	Kotias	Mbuti	-0.0159	-26.4	1417774	-0.0144	-24.0	767720
Dai	Satsurblia	Kotias	Mbuti	-0.0159	-26.5	1417780	-0.0144	-24.0	767723
Karitiana	Satsurblia	Kotias	Mbuti	-0.0139	-21.7	1417731	-0.0127	-19.9	767698

Population affinities of other ancient samples

We briefly discuss the population affiliations of the remaining samples that had at least 10,000 SNPs covered. When relevant, we use *D*-statistics of the form *D(X, Y; Test, Mbuti)* to explore how each *Test* sample in turn relates to other samples (*X*, *Y*).

- *Oase1:* This has been the subject of a published paper[3].

- *Kostenki14:* This sample has been the subject of a published paper[9]. We present a separate note analyzing this sample's affinities in Supplementary Information section 9.

- *GoyetQ116-1:* We present a separate note analyzing this sample's affinities in Supplementary Information section 10.

- *Cioclovina1*, *Kostenki12 and Muierii2:* As shown in Table S5.10, these samples all have evidence of shared genetic drift with present-day West Eurasians (unlike *Oase1*[3]), as documented by the statistic *D(UstIshim/Han, other pre-Neolithic European; Cioclovina1/Kostenki12/Muierii2, Mbuti)*, which give Z<-3 scores (Table S5.10). *Muierii2* and *Kostenki12* are possibly closer to *Kostenki14*, *Ostuni1* and *Vestonice16* than to other samples (Table S5.10). However, there is too little data to infer more refined relationships.

Table S5.10 Z-score of *D(X, Y; Cioclovina/Kostenki12/Muierii2 , Mbuti)*

D(X, Y; Muierii2, Mbuti) Muierii2: 98,618 SNPs												
X/Y	Han	Ust	Oase1	Kost14	Q116-1	Vestonice16	Ostuni1	Miron	Villabruna	Losch	Brana	Malta1
Han	NA	-0.5	0.1	-12.8	-9.3	-11.3	-8.2	-9.4	-11.1	-12.3	-10	-6.6
UstIshim	0.5	NA	0.8	-10.8	-7.9	-9.3	-8	-7.8	-8.7	-9.6	-8.3	-4.8
Oase1	-0.1	-0.8	NA	-6.3	-4.6	-6.4	-4.5	-4.6	-5.6	-5.1	-4.6	-3.6
Kostenki14	12.8	10.8	6.3	NA	2.3	0.5	0.8	2.5	2.1	2	3.3	5.5
GoyetQ116-1	9.3	7.9	4.6	-2.3	NA	-0.9	-0.9	0.6	-0.5	-0.6	0.4	2.5
Vestonice16	11.3	9.3	6.4	-0.5	0.9	NA	-0.1	2.4	1.1	1.5	2.5	3.9
Ostuni1	8.2	8	4.5	-0.8	0.9	0.1	NA	2.6	0.6	1.5	1.5	2.5
ElMiron	9.4	7.8	4.6	-2.5	-0.6	-2.4	-2.6	NA	-0.7	-0.5	0.7	1.6

Villabruna	11.1	8.7	5.6	-2.1	0.5	-1.1	-0.6	0.7	NA	0	1.7	3.3
Loschbour	12.3	9.6	5.1	-2	0.6	-1.5	-1.5	0.5	0	NA	1.6	3.4
LaBrana1	10	8.3	4.6	-3.3	-0.4	-2.5	-1.5	-0.7	-1.7	-1.6	NA	2
Hungarian.KO1	7.9	6	3.7	-4.4	-1.7	-3.6	-3.9	-2	-2.1	-2.2	-1.2	-0.4
Malta1	6.6	4.8	3.6	-5.5	-2.5	-3.9	-2.5	-1.6	-3.3	-3.4	-2	NA
D(X, Y; Cioclovina1, Mbuti) Cioclovina1: 12,784 SNPs												
X/Y	Han	Ust	Oase1	Kost14	Q116-1	Vestonice16	Ostuni1	Miron	Villabruna	Losch	Brana	Malta1
Han	NA	-0.3	0	-4	-4.1	-6.5	-5.2	-5.5	-6.1	-5.7	-4.6	-4
UstIshim	0.3	NA	-0.2	-3.6	-3.8	-5.6	-3.5	-4.4	-5.2	-4.7	-3.9	-3.3
Oase1	0	0.2	NA	-2.6	-1.8	-2.4	-1.8	-2.4	-1.5	-2.2	-1.7	-1.3
Kostenki14	4	3.6	2.6	NA	-0.5	-2	-0.5	-1.3	-1.4	-0.3	-0.5	-0.5
GoyetQ116-1	4.1	3.8	1.8	0.5	NA	-1.9	-1.7	-0.9	-0.9	-0.4	0.3	-0.4
Vestonice16	6.5	5.6	2.4	2	1.9	NA	0.4	0.8	0.9	2	1.8	0.9
Ostuni1	5.2	3.5	1.8	0.5	1.7	-0.4	NA	-0.1	0.4	1.6	1.9	0.9
ElMiron	5.5	4.4	2.4	1.3	0.9	-0.8	0.1	NA	0.9	1.6	0.9	1.6
Villabruna	6.1	5.2	1.5	1.4	0.9	-0.9	-0.4	-0.9	NA	1.6	1.5	0.8
Loschbour	5.7	4.7	2.2	0.3	0.4	-2	-1.6	-1.6	-1.6	NA	0.2	0
LaBrana1	4.6	3.9	1.7	0.5	-0.3	-1.8	-1.9	-0.9	-1.5	-0.2	NA	0
Hungarian.KO1	2.8	2.3	2.3	-0.6	-1.2	-3	-1.1	-1.6	-1.3	-0.7	-1	-1.4
Malta1	4	3.3	1.3	0.5	0.4	-0.9	-0.9	-1.6	-0.8	0	0	NA
D(X, Y; Kostenki12, Mbuti) Kostenki12: 61,228 SNPs												
X/Y	Han	Ust	Oase1	Kost14	Q116-1	Vestonice16	Ostuni1	Miron	Villabruna	Losch	Brana	Malta1
Han	NA	0	1.1	-11.1	-6.8	-10.3	-8.4	-8.2	-8	-9	-7.2	-4.6
UstIshim	0	NA	1.8	-10.2	-6.3	-9.7	-7.6	-7.2	-7.1	-7.8	-6.8	-4
Oase1	-1.1	-1.8	NA	-6.7	-3.1	-5.9	-4.9	-4.8	-4.8	-4.2	-3.3	-3.2
Kostenki14	11.1	10.2	6.7	NA	3.1	0.1	-0.1	1.9	2.3	3.3	3.5	4.6
GoyetQ116-1	6.8	6.3	3.1	-3.1	NA	-3.1	-2.9	-1.3	-0.8	-0.5	-0.3	1.5
Vestonice16	10.3	9.7	5.9	-0.1	3.1	NA	-0.4	2.2	2.5	2.9	3.3	4.7
Ostuni1	8.4	7.6	4.9	0.1	2.9	0.4	NA	1.8	1.8	2.2	3.3	3.3
ElMiron	8.2	7.2	4.8	-1.9	1.3	-2.2	-1.8	NA	0.8	1.2	1.4	2.1
Villabruna	8	7.1	4.8	-2.3	0.8	-2.5	-1.8	-0.8	NA	0.5	0.6	1.7
Loschbour	9	7.8	4.2	-3.3	0.5	-2.9	-2.2	-1.2	-0.5	NA	0.5	1.6
LaBrana1	7.2	6.8	3.3	-3.5	0.3	-3.3	-3.3	-1.4	-0.6	-0.5	NA	1.1
Hungarian.KO1	6.7	5.4	3.7	-3	0.7	-2.2	-2.1	-1	-0.3	-0.2	0	1.2
Malta1	4.6	4	3.2	-4.6	-1.5	-4.7	-3.3	-2.1	-1.7	-1.6	-1.1	NA

References

1 Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87-91, doi:10.1038/nature12736 (2014).

2 Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449, doi:10.1038/nature13810 (2014).

3 Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, doi:10.1038/nature14558 (2015).

4 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).

5 Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225-228, doi:10.1038/nature12960 (2014).

6 Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nature communications* **5**, 5257, doi:10.1038/ncomms6257 (2014).

7 Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, doi:10.1038/nature14317 (2015).

8 Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature communications* **6**, 8912, doi:10.1038/ncomms9912 (2015).

9 Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).

Section 6

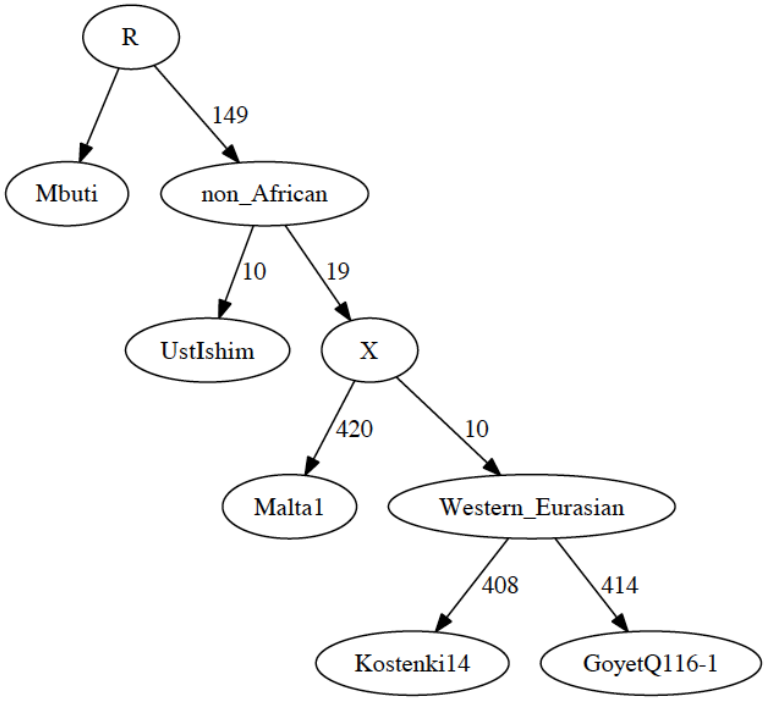
Admixture Graph Modeling of high coverage ancient genomes

Strategy

We began with the Admixture Graph (Figure S6.1) in Supplementary Information section 8, which relates the high coverage *Mbuti* (African outgroup), *UstIshim*, *Malta1*, *Kostenki14*, and *GoyetQ116-1*. This tree fits the allele frequency correlation patterns in the data to within the limits of our resolution, in the sense that the f_2-, f_3-, and f_4-statistics among all possible pairs, triples, and quadruples of populations match the observed values within three standard errors (using an empirical standard error computed with a Block Jackknife).

We proceeded by attempting to fit representative high coverage (>1×) samples—*Vestonice16*, *ElMiron*, *Villabruna*, *Loschbour*—into the Admixture Graph in turn, starting with the oldest and moving forward in time. We evaluated whether each tested model was a fit to the data, again by testing whether the predicted values of all the f_2-, f_3-, and f_4-statistics among all possible pairs, triples, and quadruples of populations matched the observed values, and assessing the significance of the difference using a Block Jackknife. In each of the Admixture Graphs shown in this note, the labels on the solid edges give the estimated genetic drift in f_2-units of squared frequency difference (parts per thousand). The labels on the dotted edges give mixture proportions.

Figure S6.1: Base model. This uses 324,336 SNPs covered in all populations.

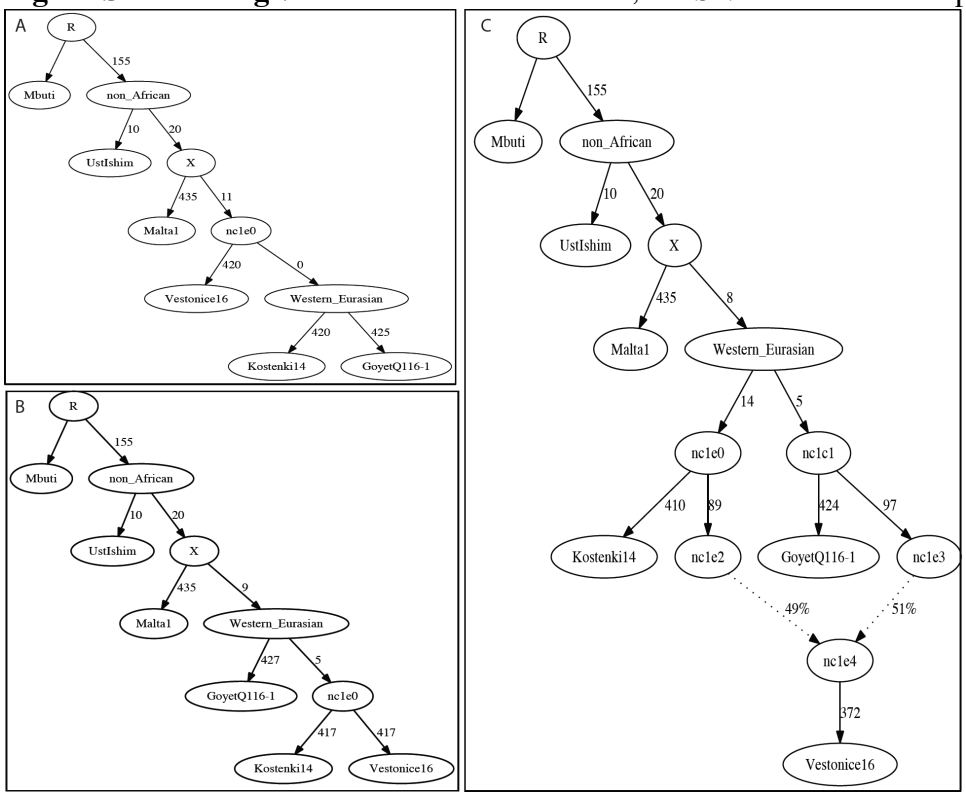


Vestonice16 (~31,000 BP)

We added *Vestonice16* to all possible nodes of Figure S6.1 either as a simple branch without mixture, or as a mixture between two branches. Altogether, we identified 3 models that fit the data (maximum |Z|<3). We show these in Figure S6.2:

- Figure S6.2A shows a fit in which *Vestonice16*, *Kostenki14* and *GoyetQ116-1* have an unresolved splitting order: a trifurcation.
- Figure S6.2B shows *Vestonice16* as a clade with *Kostenki14*, with *GoyetQ116-1* as an outgroup to both of them.
- Figure S6.2C shows *Vestonice16* mixed of lineages related to *Kostenki14* and *GoyetQ116-1*.

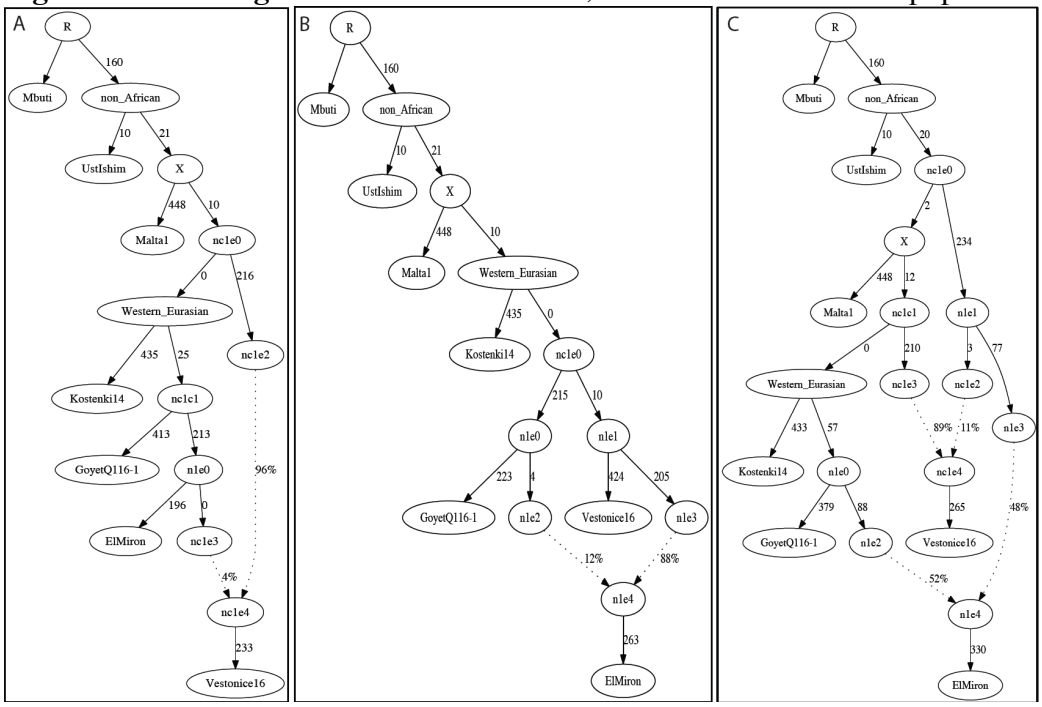
Figure S6.2 Adding *Vestonice16*. This uses 247,637 SNPs covered in all populations.



***ElMiron* (~19,000 BP)**

We added *ElMiron* to all possible nodes of the 3 models that fit the data for *Vestonice16* (Figure S6.2). We added it either as a simple branch without mixture, or as a mixture of two branches. Figure S6.3 shows the three models that fit. Figure S6.3A shows *ElMiron* as a clade with *GoyetQ116-1* and *Vestonice16* as mixed. Figure S6.3B shows *Vestonice16* as not mixed and *ElMiron* is mixed. Figure S6.3C shows both as mixed.

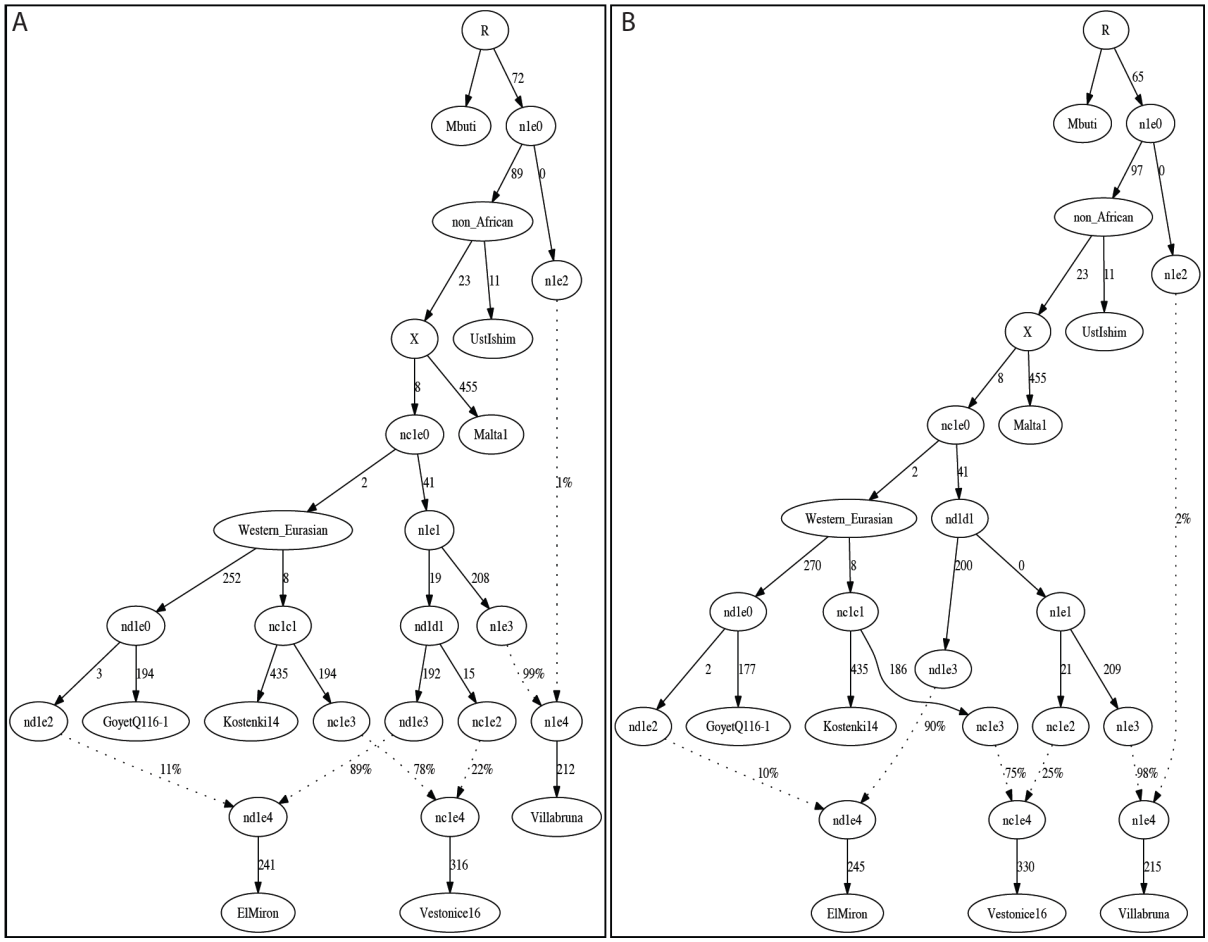
Figure S6.3 Adding *ElMiron*. This uses 186,469 SNPs covered in all populations.



***Villabruna* (~14,000 BP)**

We added *Villabruna* to all possible branches of the 3 models that fit the data for *Vestonice16* and *ElMiron* (Figure S6.3), again either as a simple branch or as a mixture between two branches. A total of nine models fit the data with at least three admixture events. We highlight two in Figure S6.4 (the other 7 models are qualitatively similar, with slight differences in the insertion points for the admixing lineages).

Figure S6.4 Adding *Villabruna*. This uses 181,563 SNPs covered in all populations.



The models that fit the data for *Villabruna* have the shared features that:

- *Vestonice16* is a mixture of lineages related to *Kostenki14* and *Villabruna*.
- *ElMiron* is a mixture of lineages related to *GoyetQ116-1* and a lineage that contributed most of the ancestry of *Villabruna*.

In both models of Figure S6.4, *Villabruna*, too, is inferred to be admixed, with 1-2% of its ancestry deriving from a deep Eurasian branch that split before the separation of *UstIshim* from all other Eurasians (drift distance ~0.09). This lineage is inferred to derive from earlier than the founder of all non-Africans, and is inferred to be more drifted than the "Basal Eurasian" lineage[1]. To investigate this unexpected signal, we removed the deep ancestry lineage, and found a single outlier: f_4*(Mbuti, UstIshim; GoyetQ116-1, Villabruna)*, which is significantly different from expectation (|Z|=3.1). When we compute the statistic *D(Mbuti, UstIshim; GoyetQ116-1, Villabruna)*, however, we obtain Z-scores of -2.2 (all SNPs) and -1.0 (transversions only). Given the number of statistics we are computing to test the fit of this model to the data, and the failure to replicate the signal for transversions only, we do not view a single outlier at |Z|>3 as a strong rejection of the model of no deep ancestry in

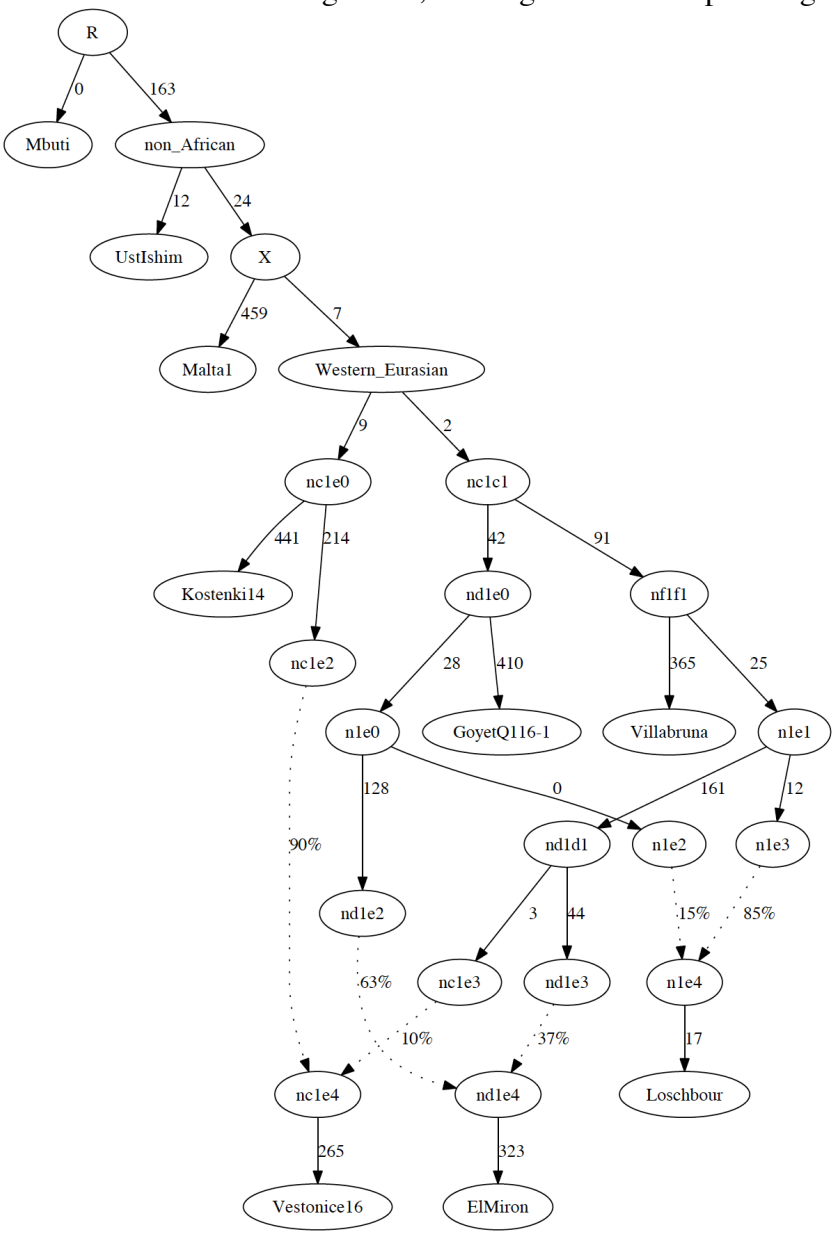
Villabruna. *Villabruna* is also artifact-prone in some ways because it is one of the few non-UDG-treated samples in the study. For subsequent analyses, we therefore do not model this deep ancestry signal in *Villabruna* (at the cost of tolerating slightly poorer fits).

***Loschbour* (~8,000 BP)**

We added *Loschbour* to all possible branches of the nine models that fit the data for *Vestonice16*, *ElMiron*, and *Villabruna*, removing in each case the deep branching lineage contributing *Villabruna*. A total of five models fit the data, after allowing for some maximum |Z|-scores slightly larger than 3 due to the deep ancestry signal in *Villabruna* (see above).

Figure S6.5 shows one passing model according to these criteria, which posits *Loschbour* as a mixture of two lineages whose closest relatives are *ElMiron* (more distantly *GoyetQ116-1*) and *Villabruna*. We use this model to explore the relationships of lower coverage samples in Supplementary Information section 7.

Figure S6.5 Adding *Loschbour*. This uses 179,232 SNPs covered in all populations. This is an alternate version of Figure 4a, which gives branch-specific genetic drift estimates.

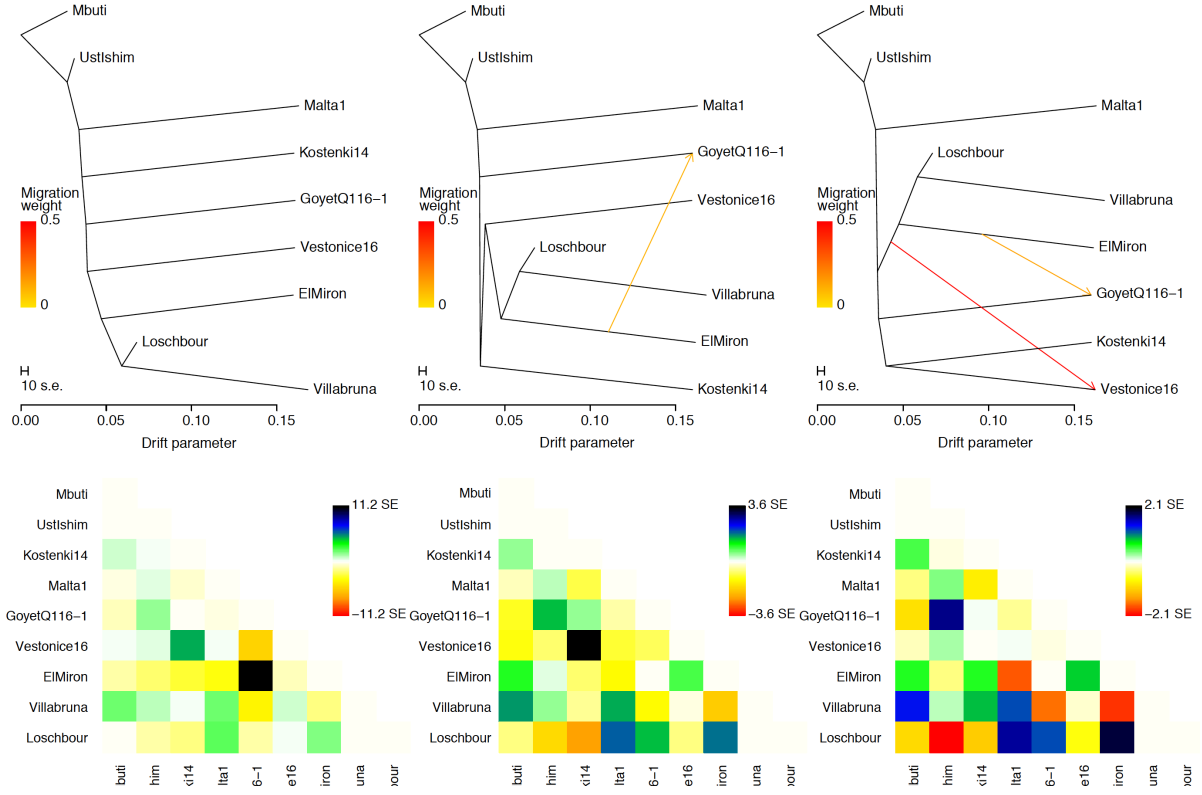


Treemix

As a complementary analysis to study the historical relationships among the samples, we used *TreeMix*[2]. Although *TreeMix* uses similar allele frequency correlation statistics as does *qpGraph*[3] to fit models to data, it takes a very different approach to exploring graph topologies. In *qpGraph*, we manually explored alternative topologies, adding in samples in a specified chronological order. In contrast, with *TreeMix*, samples are all fitted at once. We were curious to see if the two approaches gave similar solutions

We applied *TreeMix* to the same individuals shown as in Figure S6.5. The best fitting model without mixture is strongly inconsistent with the data, with some f_2-statistics among pairs of populations as high as |Z|=11.2 standard errors from expectation (Figure S6.6). When one mixture event is allowed, *Treemix* infers gene flow between the *ElMiron* and *GoyetQ116-1* lineages (Figure S6.6). With two mixture events, *TreeMix* infers gene flow between the *Vestonice16* lineage and *Kostenki14* (Figure S6.6). These two major admixture events are qualitatively similar to those that emerge from the *qpGraph* based manual analysis, although the directionalities of mixture are different.

Figure S6.6 *TreeMix* results for 0, 1 and 2 admixture events. The bottom plots shows residuals, expressed as the number of standard errors.



Summary and caveat

We have used the principle of parsimony to build these models. We make no claim that the models are exact. Indeed, we think that the models are likely to be wrong in some details. However, we hypothesize that the models may capture some important qualitative feature of the shared history of these samples:

- *Malta1* fits as an outgroup to *Vestonice16*, *Kostenki14*, *GoyetQ116-1*, and *ElMiron*, consistent with the results of Supplementary Information section 8.

- *Vestonice16* fits as a mixture of (primarily) *Kostenki14*-related ancestry, with a lesser amount of *Villabruna*-related ancestry. *GoyetQ116-1* is an outgroup.

- *ElMiron* fits as a mixture of (mostly) *GoyetQ116-1*-related ancestry, and a lesser amount of *Villabruna*-related ancestry.

- *Loschbour* fits as a mixture of (mostly) *Villabruna*-related ancestry, and a lesser amount of *GoyetQ116-1*-related ancestry.

In the rest of this study, we use the model of Figure S6.5 (which is the same as Figure 4a in the main text) as a null hypothesis that fits the allele frequency correlations. In Supplementary Information section 7, we employ this model to determine positions in the Admixture Graph for fitting lower coverage samples.

References

1 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).
2 Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS genetics* **8**, e1002967, doi:10.1371/journal.pgen.1002967 (2012).
3 Patterson, N. J. *et al.* Ancient Admixture in Human History. *Genetics*, doi:10.1534/genetics.112.145037 (2012).

Section 7

Admixture Graph based assignment of ancestry

Overview

We used the Admixture Graph model of Figure 4a in the main text—which is a good fit to the allele frequency correlation patterns of eight ancient samples and a sub-Saharan African outgroup—as a framework to test the positioning of other samples.

Confirming the position in the Admixture Graph of samples with >0.1x coverage

For a subset of samples, we have high enough sequence coverage to be able to directly test whether they fit in the Admixture Graph of Figure 4a. For the samples that fit, we can also, in some instances, use the Admixture Graph to estimate mixture proportions.

Vestonice16, Ostuni1 and KremsWA3 are interchangeable in the Admixture Graph

In the Admixture Graph of Figure 4a, *Vestonice16* is modeled as a mixture of 90% ancestry from a lineage related to *Kostenki14* and 10% ancestry related to *Villabruna*. We were interested in whether other samples from the Věstonice Cluster (Supplementary Information section 5) fit in the same position in the Admixture Graph. We therefore replaced *Vestonice16* with *Ostuni1* and *KremsWA3*—the two other high coverage samples in the Věstonice Cluster—to test whether the Admixture Graph model fits them as well.

We began by replacing *Vestonice16* in this Admixture Graph with *Ostuni1*, and found that we obtained a reasonable fit. We observed a single outlier statistic that is |Z|=3.4 standard errors from expectation, but note that this is the same statistic discussed in Supplementary Information section 6: f_4*(Mbuti, UstIshim; GoyetQ116-1, Villabruna)*. This is not too concerning given the number of hypotheses tested, and the fact that the deep ancestry signal in Villabruna—one of the few non-UDG-treated samples in our study—becomes non-significant when restricting to transversions which are not affected by characteristic ancient DNA damage (Supplementary Information section 6).

We also replaced *Vestonice16* with *KremsWA3*. This resulted in two outliers, f_4*(Mbuti, Malta1; ElMiron, Loschbour)* (|Z|=3.2) and f_4*(UstIshim, Malta1; ElMiron, Loschbour)* (|Z|=3.1). We view this as a tolerable fit given the number of hypotheses tested. Table S7.1 summarizes the quality of each of the fits.

Table S7.1. Interchangeability of samples in the Věstonice cluster. We switch the position of *Vestonice16* with other samples and show parameters of the fits.

Sample	SNPs analysed	Coverage	Kostenki14 branch	Villabruna branch	#outliers |Z|>3	Maximum |Z|-score
Vestonice16	179,232	1.31	90%	10%	1	3.1
Ostuni1	87,879	0.25	83%	17%	1	3.4
KremsWA3	55,190	0.11	84%	16%	2	3.2

El Mirón Cluster samples are interchangeable in the Admixture Graph

In the Admixture Graph of Figure 4a, *ElMiron* is modeled as a mixture of 63% ancestry related to *GoyetQ116-1*, and 37% ancestry related (deeply) to *Villabruna*.

We replaced *ElMiron* with the pool of non-Iberian El Mirón Cluster samples (ElMiron_NI_C). This resulted in a fit, with a larger proportion of ancestry from the *GoyetQ116-1* lineage (Table S7.2). This finding is consistent with Supplementary

Information section 10, which shows formally that the proportion of *GoyetQ116-1* related ancestry is significantly higher in the ElMiron_N_CI than in *ElMiron*.

Table S7.2. Interchangeability of samples in the El Mirón Cluster. We switch the position of *ElMiron* with ElMiron_NI_C and show parameters of the fits.

Sample	SNPs analysed	Coverage	GoyetQ116-1 branch	Villabruna branch	# outliers |Z|>3	Maximum |Z|-score
ElMiron	179,232	1.01	63%	37%	1	3.1
ElMiron_NI_C	54,997	0.22	80%	20%	0	<3

Loschbour, LaBrana1, Bichon and Rochedane are interchangeable in the Admixture Graph

Loschbour fits in the Admixture Graph of Figure 4a as a mixture of 16% *GoyetQ116-1* and 84% *Villabruna*-related ancestry.

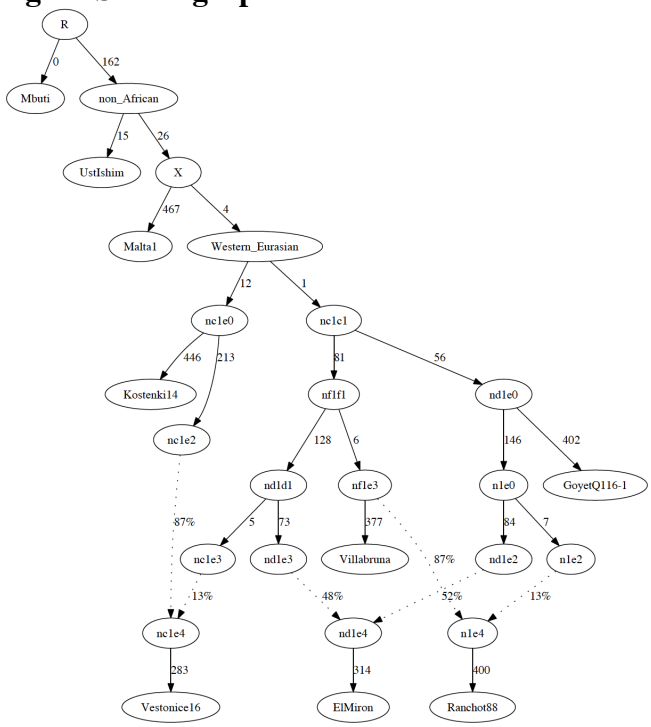
We replaced *Loschbour* with *LaBrana1*, and obtained a similar fit (Table S7.3). For *Rochedane* and *Bichon* we obtained a fit to the data without requiring any ancestry from the *GoyetQ116-1* branch, suggesting that *Rochedane* and *Bichon* may both be unadmixed descendants of the same population to which *Villabruna* belonged.

Table S7.3. Interchangeability of samples in the Villabruna Cluster. *LaBrana1, Bichon* and *Rochedane* can all be exchanged with *Loschbour* and produce a good fit.

Villabruna Cluster sample	SNPs analysed	Coverage	GoyetQ116-1 branch	Villabruna branch	# outliers |Z|>3	Maximum |Z|-score
Loschbour	179,232	20	16%	84%	1	3.1
LaBrana1	176,901	3.3	20%	80%	1	3.2
Bichon	181,406	8.1	0%	100%	2	3.1
Rochedane	50,560	0.13	0%	100%	0	<3

We finally replaced *Rochedane* with *Ranchot88*. The fit is worse in the same position in the graph (maximum |Z|-score = 4.1). However, if we change the topology so the *Villabruna*-related ancestry of *Ranchot88* is from a lineage that is a direct sister group to the *Villabruna* lineage—only a modest qualitative change in topology—we obtain a good fit (Figure S7.1).

Figure S7.1 A graph that fits *Ranchot88*. This uses 98,757 SNPs.

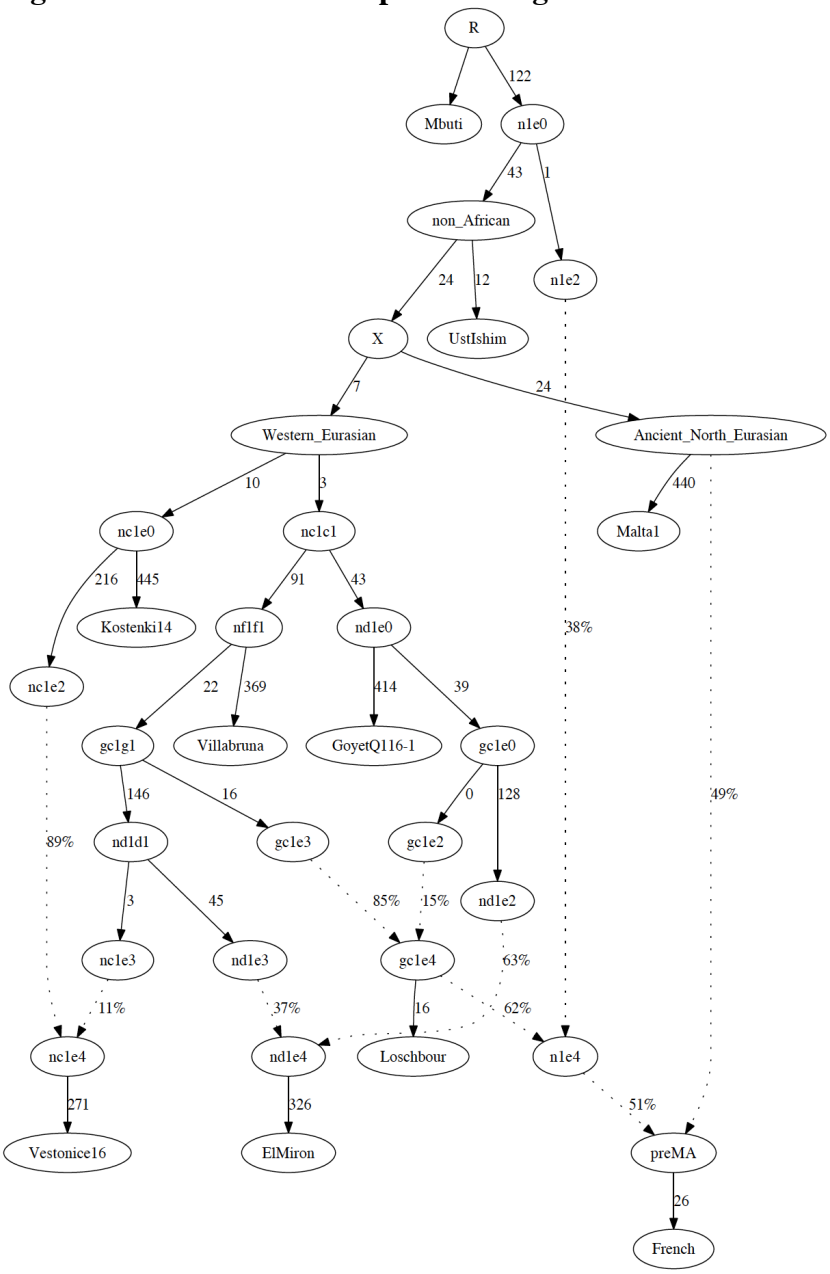


Position in the Admixture Graph of low coverage samples

For a large number of samples in this study, there are relatively few SNPs covered, reflecting the fact that we restricted to damaged sequences for our main analyses to reduce the danger of contamination. Combined with the fact that only 179,232 SNPs are covered in all nine samples included in the Admixture Graph of Figure 4a—less than 10% of the targeted SNP positions—this meant that we had very little information to use for placing low coverage samples onto the Admixture Graph.

To address the problem of limited data, we carried out Admixture Graph analysis of all samples that are not part of the skeleton admixture graph of Figure 4a, without restricting to damaged sequences. While this introduces contamination into a number of samples, we address this by assuming that the contamination is from a present-day European source, and then modeling this contamination into the Admixture Graph by using present-day *French* as a surrogate for the contamination (Figure S7.2).

Figure S7.2 Admixture Graph modeling of *French*. This uses 179,232 SNPs.



In this Admixture Graph, we model *French* as a mixture of three lineages inspired by the model of ref. [1]: (a) "Basal Eurasian" ancestry that diverged from the ancestors of *UstIshim* and *Loschbour* before they separated from each other; (b) ancestry from a sister group to *Malta1*, and (c) ancestry from a sister group to *Loschbour* (Figure S7.2). This Admixture Graph is a fit to the data in the sense that there are only two *f*-statistics for which the observed and predicted values are |Z|>3 standard errors from expectation: f_4*(Mbuti, Ust-Ishim; GoyetQ116-1, Villabruna)* (|Z|=3.1) and f_4*(UstIshim, French; GoyetQ116-1, Villabruna)* (|Z|=3.2). We make no claim that the modeling of the *French* in Figure S7.2 is correct. Indeed, it is almost certainly incorrect, as the ancestral relationships of present-day Europeans to pre-Neolithic Europeans are known to be complicated[1]. Our philosophy here is not to build an accurate model for *French*, but instead to model them into the graph as a "nuisance ancestry source" to adjust for biases from contamination. We model each ancient sample in turn as a mixture of a chosen point in the Admixture Graph and this nuisance ancestry source.

Table S7.4 gives the maximum absolute Z-score between the observed and predicted values for any f_2, f_3, or f_4-statisic relating samples in the graph, when we replace each of the non-*French* samples in the Figure S7.2, with a *Test* sample modeled as harboring contaminating ancestry from a lineage closely related to *French*. We make several observations:

- *Cioclovina1, Kostenki12* and *Muierii2* are closer to *Kostenki14* or *Vestonice16* than to other pre-Neolithic Europeans (especially the early branching *GoyetQ116-1*), consistent with the results in Supplementary Information section 5.

- *Vestonice13, Vestonice15, Vestonice14, Vestonice43, KremsWA3, Ostuni1, Paglicci133* and *Pavlov1* cluster with *Vestonice16*, consistent with the evidence from Supplementary Information section 5 that they are in the "Věstonice Cluster". They have a somewhat weaker affinity to *Kostenki14*, consistent with the Admixture Graph of Figure 4a, which has *Vestonice16* as primarily derived from a lineage related to *Kostenki14*. *Ostuni2* shows evidence of having a possibly closer affinity to *Kostenki14* than to *Vestonice16*, so we do not consider this sample to be confidently placed (Table S7.4).

- *GoyetQ376-19, GoyetQ53-1* and *GoyetQ56-16* are not possible to assign, since all three have many possible placements (Table S7.4).

- *Brillenhohle, Burkhardtshohle, GoyetQ-2, HohleFels79, HohleFels49,* and *Rigney1* are related to *ElMiron*, consistent with the finding of Supplementary Information section 5 that they are in the "El Mirón Cluster".

- *BerryAuBac, Bockstein, Continenza, Chaudardes1, Falkenstein, Hungarian.KO1, Iboussieres39, LaBrana1, LesCloseaux13, Ofnet, Ranchot88, Bichon* and *Rochedane* are closest to *Villabruna* and *Loschbour*, as expected from being in the "Villabruna Cluster". The present analysis gives us the resolution we need to conclude that *Iboussieres39*, too, is in the Villabruna Cluster.

Table S7.4. Fit of the Admixture Graph when each sample in the row is fit into the position of the sample in the column. For the samples in the rows (ordered chronologically), we use all sequences instead of restricting to damaged sequences. We model each sample as having a portion of their ancestry derived from a lineage that is a sister-group to the *French* to adjust for the contamination that we expect from undamaged sequences. Each cell reports the maximum absolute Z score for all possible *f*-statistics relating the samples in the Admixture Graph. We highlight in yellow all |Z|≤4 scores, except where there are no such |Z|-scores in which case we highlight the smallest in the row. "Min SNPs" refers to the minimum number of SNPs used in Admixture Graph fits of the sample on this row.

Sample ID	Min SNPs	UstIshim	Malta1	GoyetQ116-1	Kostenki14	Vestonice16	ElMiron	Villabruna	Loschbour
Cioclovina1	16642	5.3	3.9	5.0	3.7	4.3	5.5	5.5	5.5
Kostenki12	12996	7.2	4.8	5.2	<3	<3	7.3	7.2	7.2
Muierii2	133585	11.2	8.5	7.3	3.7	4.7	11.8	11.2	11.2
Vestonice13	150689	23.9	20.8	18.5	16.5	3.7	20.7	24.5	24.8
Vestonice15	70226	20.5	17.9	16.2	13.9	4.0	17.8	20.7	21.2
Vestonice14	31694	5.8	5.8	5.5	3.9	4.0	5.7	6.0	6.0
Vestonice43	150210	15.0	11.3	10.3	6.2	3.4	11.6	15.0	15.3
Pavlov1	105279	13.3	11.9	10.1	8.5	3.9	12.5	14.3	14.6
Paglicci133	15242	5.8	6.1	5.6	3.8	3.1	5.8	5.8	5.8
KremsWA3	49637	17.9	14.3	12.6	10.3	3.2	13.9	17.4	17.6
Ostuni2	31456	6.9	7.2	7.8	5.5	7.1	7.8	7.8	7.8
Ostuni1	169076	16.8	14.6	12.3	7.9	3.3	15.9	17.7	18.0
Paglicci108	3409	3.3	3.1	3.7	3.0	<3	3.4	3.5	3.8
GoyetQ53-1	15729	<3	<3	<3	<3	<3	<3	<3	<3
GoyetQ376-19	28263	4.7	4.8	4.0	4.4	4.1	3.9	4.8	4.8
GoyetQ56-16	14747	3.9	3.9	3.0	3.6	3.3	3.1	4.7	4.6
HohleFels79	54032	21.2	20.2	12.1	17.9	14.0	6.5	17.3	17.4
HohleFels49	69602	24.3	24.3	12.6	19.0	15.4	7.0	18.8	19.0
Rigney1	33743	19.3	19.3	10.1	15.9	13.2	3.0	14.7	14.7
GoyetQ2	74642	22.9	22.0	11.9	19.6	15.9	7.5	20.5	20.1
Brillenhohle	21200	12.4	12.4	6.9	11.0	8.2	3.5	10.9	10.9
Burkhardtshohle	36476	15.0	15.0	6.7	12.0	9.9	5.6	14.2	13.7
Bichon	179080	24.7	24.7	22.4	24.7	20.6	14.4	3.9	3.2
Rochedane	49861	20.2	20.2	16.4	20.2	15.9	11.1	3.2	<3
Continenza	26349	18.2	18.2	15.2	18.2	12.6	9.0	3.4	3.4
Iboussieres39	6487	6.9	6.9	5.6	7.1	5.4	<3	<3	<3
Ranchot88	145660	29.6	29.6	24.6	29.6	23.5	16.2	6.3	4.3
LesCloseaux13	7090	5.9	5.9	5.9	5.9	5.6	4.4	3.7	<3
Falkenstein	63695	23.9	23.9	18.5	23.9	18.5	12.0	5.4	3.6
Bockstein	43064	16.1	16.1	14.5	16.1	13.7	9.5	3.4	3.1
Ofnet	29705	8.4	8.3	8.7	8.4	8.7	5.0	3.8	4.3
CuiryLesChaudardes1	58788	21.3	21.3	20.8	21.3	17.4	12.9	4.6	3.2
LaBrana1	174613	21.3	21.3	18.0	21.3	17.6	9.8	7.4	7.0
Hungarian.KO1	124921	18.9	18.9	18.9	18.9	16.0	12.8	3.9	3.6
BerryAuBac	12882	13.2	13.2	10.5	13.2	11.4	7.5	<3	<3

References

1 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).

Section 8

No evidence for Basal Eurasian ancestry in pre-Neolithic Europeans

Overview

Seguin-Orlando et al. analysed 2.8-fold coverage shotgun sequencing data from *Kostenki14*[1] and suggested that this individual had "Basal Eurasian" ancestry from the lineage that contributed also to Early European Farmers like *Stuttgart* from the *Linearbandkeramik* culture[2]. Here, we present compelling evidence against this hypothesis.

Replication of the statistics in Seguin-Orlando et al. in our data

The evidence for Seguin-Orlando et al.'s claim is summarized in Figure 2A of their study, in which they cite a "Middle East component for *Kostenki14* in clustering analysis" and two statistics of the form *D(Kostenki14, X; Y, Mbuti)*:

$$D(\text{Kostenki14, North Eurasian Hunter Gatherer; East Asians, Mbuti}) << 0$$
$$D(\text{Kostenki14, Early European Farmers; East Asians, Mbuti}) \sim 0$$

The first of these statistics was significantly negative when X = a pre-Neolithic North Eurasian (either *Loschbour* or *Malta1*), and Y = *Han*. This is inconsistent with the hypothesis that *Han* is an outgroup to a clade including *Kostenki14* and X. We replicate this statistic when X is any sample in the "Mal'ta Cluster" (Malta_C), or any sample in the "Villabruna Cluster," (*Villabruna, Rochedane, Loschbour, LaBrana1* and *Hungarian.KO1*) whether in capture data or shotgun data, and whether restricting analysis to transversion SNPs (Table S8.1) or analyse all sites (Table S8.2). We also replicated the second statistic.

Table S8.1 Z-score of *D(Kostenki14, X; Y, Mbuti) restricting to transversions*

X/Y	Karitiana	Han	Dai	Onge	UstIshim	Oase1	GoyetQ116-1	Vestonice_CE_C	Vestonice_I_C	ElMiron	ElMiron_NI_C	Villabruna	Loschbour	LaBrana1	Hungarian.KO1	Malta_C
Capture data																
UstIshim	2.2	-1	-0.9	-1.4	NA	1.8	8.2	11.5	11.1	10.2	7.6	11.1	10.6	10.4	10.9	7.3
Oase1	3.6	2.3	2.7	1.8	3.5	NA	6.7	7.3	7.1	7	5.3	8.5	10.6	8.8	8	6.4
GoyetQ116-1	0	-2	-2.1	-1.1	-1.3	0	NA	1.3	-1.5	-8.6	-11.5	-0.9	-3.6	-4.4	-1.5	-1.2
Vestonice_CE_C	-0.1	-0.2	0.7	0.8	0.4	0.3	-0.7	NA	-6.2	-4.2	-2.3	-4.8	-5	-2.8	-3.8	0.3
Vestonice_I_C	0.7	-0.9	-0.1	0.1	1.2	-0.8	-3.2	-5.8	NA	-4.7	-2	-5	-3.4	-2.3	-3.1	0.2
ElMiron	-0.8	-1.3	-0.6	-0.7	0.8	0	-7.6	-2.1	-3.7	NA	-18.4	-11.7	-14.3	-15.5	-9.1	-0.4
ElMiron_NI_C	-0.2	-1.4	-1.3	-0.3	0	0.2	-10.5	0.2	-1.4	-19.2	NA	-5.1	-10.2	-12.2	-5.5	-1.2
Villabruna	-1.2	-1.1	-0.4	0.4	0.4	-0.4	-0.6	-3.2	-3.6	-11.9	-5.3	NA	-27	-19.4	-22.6	-1.1
Rochedane	0.1	0.3	-0.4	0.2	1.6	-0.6	-1.8	-1.2	-1.4	-6.9	-5.8	-16	-18.3	-10.4	-10.7	-0.4
Ranchot88	-1.2	-1.7	-1.2	0.1	1	0.6	-1.8	-2.6	-4	-11.9	-7	-19.5	-24.3	-16.5	-14.8	-2.5
Loschbour	-3.2	-4.2	-3	-2.5	0.1	1.8	-2.1	-1.2	-1.3	-14	-10	-25.7	NA	-24.3	-23.4	-2.7
LaBrana1	-2.9	-3.6	-2.5	-2.1	0.5	1.5	-2.8	1	-1	-14.8	-10.8	-18.8	-23.3	NA	-19	-1.8
Hungarian.KO1	-3.3	-4	-2.9	-1.3	1.8	1.2	0.3	0	-0.6	-7.4	-4.3	-20.8	-22.6	-18.7	NA	-2.6
Malta_C	-12.3	-4.8	-5.5	-3.9	-0.3	0.1	2.3	4.7	4.3	1.9	2.2	2.8	0.3	0.2	-1.3	NA
Stuttgart	2.5	0.1	0.4	2.1	3.3	3.1	5.7	6.2	6.1	3.1	3.7	0	-1.9	-0.6	-3.6	3.7
Shotgun data																
UstIshim	2	-0.8	-0.9	-1.4	NA	1.4	8.5	11.6	11	8.9	8.3	9.7	10.4	10	11.2	6.8
Oase1	4.2	2.6	3	2.3	4.1	NA	6.6	8	6.5	7.1	5.6	7.6	11.2	8.7	8.9	6.9
GoyetQ116-1	0.4	-1.7	-1.5	-0.8	-0.8	-0.7	NA	1.1	-0.8	-9.2	-10.1	-1.5	-3.5	-4.1	-1	-0.9
Vestonice_CE_C	1.2	1.1	1.8	2.1	1.4	0.4	0.2	NA	-5.6	-3.5	-2.2	-4.2	-4.2	-2	-2.8	-0.1
Vestonice_I_C	1.1	-0.7	0.4	1.1	2	-1.2	-1.6	-5.5	NA	-4.9	-1.8	-5	-3.2	-2.7	-2.3	0.2
ElMiron	0.1	-1.1	-0.3	0	0.9	-0.4	-7	-1.7	-2.8	NA	-18.5	-11.7	-14.9	-15.5	-8.7	-0.8
ElMiron_NI_C	0.4	-0.6	0	-0.2	0.7	-1.5	-9.2	-0.8	-1.4	-19.5	NA	-5.4	-9.9	-10.9	-5.1	-0.5
Villabruna	-0.5	-0.5	0.2	1	0.9	-0.7	0.3	-3	-3	-12.1	-4.8	NA	-26.6	-19.2	-22	-1.3
Rochedane	0.2	0.3	0.2	0.5	2.2	-1.4	-0.1	-1	-1	-7.2	-4.6	-15.9	-17.3	-10.4	-9.8	0.4
Ranchot88	-0.5	-1.7	-0.9	-0.3	1.2	-0.1	-1.8	-2.6	-3.8	-12	-6.2	-20.6	-24.3	-16.2	-15.2	-1.9
Loschbour	-3.2	-4.3	-3.1	-2.2	-0.1	1	-1.8	-1	-1	-15	-9	-26.5	NA	-24.2	-23.5	-2.8
LaBrana1	-2.5	-3.2	-2.2	-1.7	0.4	1.2	-2.1	1.5	-0.8	-15.6	-10.4	-19.7	-23.5	NA	-18.6	-1.7
Hungarian.KO1	-3.2	-4.3	-3.2	-1.2	1.8	0	0.4	0.2	-0.2	-8.4	-3.5	-21.4	-22.8	-19	NA	-3.2
Malta_C	-11.2	-4.6	-5.1	-3.3	-0.4	-0.8	2.8	4.3	4.6	1	2.6	1.6	0.5	0.3	-1.2	NA
Stuttgart	2.5	0.1	0.4	2.1	2.9	2.7	6.1	6.3	6.8	1.9	4.4	-1	-2	-0.7	-3.4	3.5

Table S8.2 Z-score of *D(Kostenki14, X; Y, Mbuti)* for all SNPs

X/Y	Karitiana	Han	Dai	Onge	UstIshim	Oase1	Q116-1	Vestonice_CE_C	Vestonice_I_C	ElMiron	ElMiron_NI_C	Villabruna	Loschbour	LaBrana1	Hungarian.KO1	Malta_C
Capture data																
	2.5	-1.3	-1.4	-1.7	NA	1.1	9.1	16.3	13.4	11.2	9.9	11.6	11.5	11.5	11.8	7.9
Oase1	5.6	2.9	3.2	2.9	4.1	NA	10.5	12.1	10.6	10.6	9.3	11.5	13.1	11.5	10.2	9.6
GoyetQ116-1	0.3	-2.7	-2.8	-1.9	-1.7	-1.2	NA	0.8	-0.6	-11	-15.7	-1.8	-4.2	-5.7	-1.6	-1.3
Vestonice_CE_C	-0.5	-1.5	-0.8	0.5	0.9	0.6	-3	NA	-10.1	-5.7	-3.9	-6.9	-6.5	-5.8	-5.4	-0.8
Vestonice_I_C	0.4	-1.7	-0.6	-0.3	1.5	-0.8	-1.7	-8.6	NA	-4.7	-3.5	-5.7	-4.2	-3.8	-3.2	0.4
ElMiron	-1.4	-2.2	-1.8	-1.1	0.2	-0.4	-9.8	-2.5	-3.5	NA	-25.9	-14.4	-16.9	-19.4	-10.8	-0.6
ElMiron_NI_C	-0.7	-1.9	-1.5	-0.8	-0.2	0.5	-13.6	-0.1	-2.5	-26.8	NA	-7.6	-12.9	-16.9	-8.4	-1.3
Villabruna	-1.5	-1.3	-1.1	0.6	0.6	-0.6	-0.5	-3.2	-3.9	-13.5	-6.8	NA	-29	-21.8	-26.7	-1.8
Rochedane	-0.1	-1	-1.1	0.1	1.5	-0.2	-1.4	-0.2	-1.8	-12.5	-7.7	-23.2	-24.8	-18	-17.3	-0.9
Ranchot88	-1.4	-1.6	-1.3	-0.2	2.3	-0.4	-2.4	-1.8	-3.8	-15.4	-10.1	-25.8	-33	-23.9	-22.4	-1.6
Loschbour	-3.7	-4.3	-3.2	-2.5	0.2	1.1	-3.1	-2.5	-2.3	-16.4	-13.5	-29.5	NA	-27	-26.3	-3.1
LaBrana1	-3.3	-3.8	-2.9	-2.1	0.8	0.5	-3.7	0.2	-1.8	-18.6	-15.5	-21.4	-25.8	NA	-21	-2.2
Hungarian.KO1	-3.5	-4	-3	-1.1	2.3	0.4	1	0.3	-0.4	-9.1	-6.2	-24.2	-24.2	-19.8	NA	-3
Malta_C	-14.3	-5.3	-6.1	-3.9	-0.1	-0.7	2.6	6.8	5.1	2.5	2.5	1.9	0.7	0.6	-1.4	NA
Stuttgart	2.7	0.4	0.7	2.5	3.9	2.8	6.7	9.3	7.8	3.4	5	-0.7	-1.5	-0.2	-4	4.3
Shotgun data																
UstIshim	2.1	-1.4	-1.5	-1.8	NA	0.5	8.8	15.1	13.2	10.4	9.8	10.3	11	11	11.6	7.5
Oase1	5.8	3.1	3.4	3.1	4.6	NA	10.3	12.6	10.9	11.2	10.3	10.7	13.5	11.8	11.1	9.6
GoyetQ116-1	-0.1	-2.9	-3	-2	-1.9	-2.1	NA	-0.1	-0.6	-11.8	-15.5	-2.6	-4.7	-5.8	-2	-1.6
Vestonice_CE_C	0.6	-0.8	-0.4	0.8	1.4	-0.3	-2.8	NA	-10.2	-5.6	-3.6	-7	-6.1	-5.9	-4.8	-0.8
Vestonice_I_C	0.4	-1.7	-0.5	0.2	1.7	-1.6	-1.6	-8.7	NA	-5	-3.1	-6.1	-4.4	-4.1	-3.1	0.2
ElMiron	-0.9	-1.9	-1.5	-0.8	0.4	-0.8	-9.6	-2.9	-3.2	NA	-26.4	-14.9	-17.2	-19.8	-10.8	-0.7
ElMiron_NI_C	-0.3	-1.4	-0.7	-0.3	0.2	-0.2	-12.8	-0.4	-2.4	-26.9	NA	-7.9	-12.7	-16.5	-7.5	-1.4
Villabruna	-1.3	-1.3	-1	0.8	0.5	-1.6	-0.5	-3.7	-3.8	-13.9	-6.7	NA	-29.2	-21.9	-26.6	-1.9
Rochedane	-0.7	-1.5	-1.4	-0.3	1	-0.9	-0.5	-0.6	-1.9	-13.1	-7.5	-24.2	-24.8	-18.7	-17.6	-0.8
Ranchot88	-2.2	-2.2	-1.8	-0.4	1.9	-1.3	-2.5	-2.9	-3.6	-15.6	-9.7	-26.5	-33	-23.8	-22.7	-1.7
Loschbour	-3.9	-4.6	-3.4	-2.5	-0.2	0	-3.5	-3	-2.7	-17.3	-13.6	-30.6	NA	-26.9	-26.3	-3.2
LaBrana1	-3.4	-3.9	-3	-2.1	0.4	-0.2	-3.9	-0.5	-2.2	-19.2	-15.9	-22.4	-26.1	NA	-21	-2.4
Hungarian.KO1	-3.7	-4.4	-3.3	-1.2	2	-1	0.5	-0.4	-0.5	-10	-6.1	-24.6	-24.7	-20	NA	-3.5
Malta_C	-13.5	-5.1	-5.8	-3.6	-0.2	-1.5	2.5	5.6	5.2	2.1	2.4	0.9	0.7	0.5	-1.7	NA
Stuttgart	2.5	0.2	0.4	2.3	3.4	1.9	6.3	8.2	7.7	2.5	5.1	-1.7	-1.9	-0.5	-4.2	4

Kostenki14 patterns are general in Upper Palaeolithic Europeans

The present study allows computation of these statistics in a much larger series of Upper Palaeolithic European samples. This analysis reveals that the observed signals interpreted as Basal Eurasian ancestry are not unique to *Kostenki14*, and are also seen in later samples until around 14,000 years ago. Only after around 14,000 years ago (from Villabruna onwards) do samples with genetic affinities like those in the Loschbour cluster appear.

(1) When *X* = a European from 39,000-14,000 BP, the statistic *D(Kostenki14, X; Han/Dai/Karitiana, Mbuti)* is ~ 0 (Table S8.1 and Table S8.2).

(2) When *W* = a European from 39,000-14,000 BP including *Kostenki14*, *D(W, Early European Farmers; East Asians, Mbuti)* ~ 0 (Table S8.3).

Table S8.3 *D*(X, Stuttgart; Han, Mbuti) showing that East Asians do not always share more alleles with pre-Neolithic Europeans than with Neolithic Europeans

				Transversion sites only			All sites		
X	Stuttgart	Han	Mbuti	D value	Z score	SNPs used	D value	Z score	SNPs used
UstIshim	Stuttgart	Han	Mbuti	0.0005	1	1105740	0.0008	1.5	2019942
Oase1	Stuttgart	Han	Mbuti	-0.0015	-2.2	122034	-0.002	-3.2	274308
Kostenki14	Stuttgart	Han	Mbuti	0	0.1	805172	0.0002	0.4	1651428
Kostenki14.sg	Stuttgart	Han	Mbuti	0.0001	0.1	975742	0.0001	0.2	1811868
GoyetQ116-1	Stuttgart	Han	Mbuti	0.0018	2.8	189027	0.0018	3.3	797550
Vestonice_CE_	Stuttgart	Han	Mbuti	-0.0003	-0.6	108097	0.0005	1.1	321529
Vestonice_I_C	Stuttgart	Han	Mbuti	0.0011	1.8	118730	0.0009	1.8	412907
ElMiron	Stuttgart	Han	Mbuti	0.0008	1.5	268429	0.0013	2.7	764796
ElMiron _NI_C	Stuttgart	Han	Mbuti	0.0014	2.2	108687	0.0012	2.4	450103
Villabruna	Stuttgart	Han	Mbuti	0.0006	1.4	451676	0.0008	1.8	1151986
Rochedane	Stuttgart	Han	Mbuti	0.0007	0.8	50355	0.0013	2.4	225677
Ranchot88	Stuttgart	Han	Mbuti	0.0025	3.3	86407	0.0016	3	394773
Loschbour	Stuttgart	Han	Mbuti	0.0021	4.6	1100117	0.0024	5.2	2007432
LaBrana1	Stuttgart	Han	Mbuti	0.0016	3.7	954247	0.0019	4.5	1797590
Hungarian.KO1	Stuttgart	Han	Mbuti	0.002	4.6	728993	0.0021	4.9	1353157
Malta_C	Stuttgart	Han	Mbuti	0.0024	5.1	755470	0.0027	5.9	1455444

These results show that whatever history is driving the signals identified in ref. [1] is not unique to *Kostenki14*. Two scenarios can potentially explain these observations:

(1) There is Basal Eurasian ancestry in *Kostenki14* (consistent with ref . [1]) as well as in all Europeans prior to some Villabruna Cluster samples. After that, a population replacement occurred so that some European hunter-gatherers had little or no Basal Eurasian ancestry.

(2) There is no Basal Eurasian ancestry in *Kostenki14*, and instead, gene flow occurred between the ancestors of East Asians and Europeans after 14,000 years ago. (Independently, there would need to be gene flow between the ancestors of East Asians and *Malta1* to explain its affinities.) In this scenario, the observation that *D(Kostenki14, Early European Farmers; East Asians, Mbuti)* ~ 0, highlighted in ref. [1], does not reflect Basal Eurasian ancestry in *Kostenki14*, and instead a balancing of Basal Eurasian ancestry in Early European Farmers (which biases the statistic positively), and gene flow between the ancestors of Early European Farmers and East Asians (biases the statistic negatively).

The strongest evidence of Basal Eurasian ancestry comes from *UstIshim* and *Oase1*, and does not support Basal Eurasian ancestry in *Kostenki14*

The Basal Eurasian ancestry in *Kostenki14* was initially proposed when only one reference Eurasian lineage (*Malta1*) that was not West Eurasian in affinity was available. However, data from two additional Eurasian lineages distinct from West Eurasians and East Asians have since become available: *Oase1*[3] and *UstIshim*[4] (Table S8.4). This makes it possible to distinguish scenarios (1) and (2) outlined in the previous section.

The initial statistic that motivated the Basal Eurasian hypothesis[2] was of the form: *D(North Eurasian hunter-gatherer, Early European Farmer; East Asian, Outgroup)* >> 0. A key observation of Seguin-Orlando et al.[1] was that while this statistic was greater than 0 when the North Eurasian hunter-gatherer was a Mesolithic European or *Malta1*[2], it gave a qualitatively different result (consistent with 0) when the North Eurasian hunter-gatherer was *Kostenki14* (see also Table S8.1, Table S8.2). When *Oase1* or *UstIshim* is used in the statistic in place of East Asians, however, the statistic is greater than 0 through the whole period; we no longer see a qualitative change comparing early to late pre-Neolithic Europeans (Table S8.4).

Table S8.4 *D(European hunter-gatherers, Stuttgart; X, Mbuti).*

pre-Neolithic European	X = Han				X = Oase1				X = Ust-Ishim			
	Transversions		All sites		Transversions		All sites		Transversions		All sites	
	D value	Z score	D value	Z score	D value	Z score	D value	Z score	D value	Z score	D value	Z score
Kostenki14	0	0.1	0.0002	0.4	0.0024	3.1	0.0019	2.8	0.0019	3.3	0.0025	3.9
Kostenki14.sg	0.0001	0.1	0.0001	0.2	0.002	2.7	0.0013	1.9	0.0018	2.9	0.0021	3.4
GoyetQ116-1	0.0018	2.8	0.0018	3.3	0.0052	4.1	0.0034	4.2	0.0035	4.6	0.004	5.9
Vestonice_CE_C	-0.0003	-0.6	0.0005	1.1	0.0022	1.6	0.0022	2.5	0.001	1.4	0.0017	3.1
Vestonice_I_C	0.0011	1.8	0.0009	1.8	0.0046	3.7	0.003	3.5	0.0023	3.2	0.0019	3.1
ElMiron	0.0008	1.5	0.0013	2.7	0.0035	3.7	0.0026	3.6	0.0016	2.5	0.0025	4.3
ElMiron _NI_C	0.0014	2.2	0.0012	2.4	0.0053	4	0.0019	2.4	0.0034	4.5	0.0027	4.8
Villabruna	0.0006	1.4	0.0008	1.8	0.0031	4	0.0025	3.9	0.0018	3.3	0.0022	3.9
Rochedane	0.0007	0.8	0.0013	2.4	0.0054	2.5	0.0037	3.4	0.0022	2.2	0.0023	3.5
Ranchot88	0.0025	3.3	0.0016	3	0.005	3.1	0.003	3.4	0.0025	2.8	0.0014	2.3
Loschbour	0.0021	4.6	0.0024	5.2	0.0014	2.1	0.0015	2.4	0.0018	3.3	0.0023	4.2
LaBrana1	0.0016	3.7	0.0019	4.5	0.0013	1.8	0.0016	2.4	0.0015	3	0.0019	3.5
Hungarian.KO1	0.002	4.6	0.0021	4.9	0.002	2.4	0.0017	2.4	0.0005	0.9	0.0009	1.5
Malta_C	0.0024	5.1	0.0027	5.9	0.0028	3.7	0.0025	3.9	0.0021	3.8	0.0024	4.4

A second statistic informative about Basal Eurasian ancestry—which was not possible to compute at the time when the Basal Eurasian hypothesis was initially proposed[2]—also takes advantage of *Oase1* or of *UstIshim*, and is of the form:

$D(Test_1, Test_2; Oase1/UstIshim, Mbuti) \sim 0$

As shown in Table S8.5, this statistic is consistent with 0 whenever we use a pair of pre-Neolithic samples as the *Test* samples, thus providing no evidence of Basal Eurasian ancestry in these samples. In other words, there is no evidence in pre-Neolithic samples for any ancestry from a lineage that diverged before the separation of *Ust'-Ishim*. In contrast, when *Stuttgart* is used in place of one of the pre-Neolithic samples, we observe attraction of it to *Mbuti*, confirming the distinct signal of Basal Eurasian ancestry in this sample (Table S8.5).

Table S8.5 Z-score of *D(Eurasian$_1$, Eurasian$_2$, Oase1/UstIshim/Han, Mbuti).* This table is for all sites. We add *Stuttgart* to show the Basal Eurasian signal.

X/Y	Kostenki14	GoyetQ116-1	Vestonice_CE_C	Vestonice_I_C	ElMiron	ElMiron _NI_C	Villabruna	Loschbour	LaBrana1	Hungarian.KO1	Malta_C	Stuttgart
Oase1												
Kostenki14	NA	-1.2	0.6	-0.8	-0.4	0.5	-0.6	1.1	0.5	0.4	-0.7	2.8
GoyetQ116-1	1.2	NA	0.5	0.8	0.8	1.5	1	1.7	1.4	1.4	0.7	4.2
Vestonice_CE_C	-0.6	-0.5	NA	-1	0.1	0.9	-0.8	0.9	0.4	1	0.2	2.5
Vestonice_I_C	0.8	-0.8	1	NA	1.4	1.2	0.4	1.7	0.5	1.2	-0.4	3.5
ElMiron	0.4	-0.8	-0.1	-1.4	NA	-0.3	-0.1	1.5	1.6	1.1	0.5	3.6
ElMiron _NI_C	-0.5	-1.5	-0.9	-1.2	0.3	NA	-0.9	0.4	0.3	-0.3	-0.7	2.4
Villabruna	0.6	-1	0.8	-0.4	0.1	0.9	NA	1.7	1.3	0.5	0.1	3.9
Loschbour	-1.1	-1.7	-0.9	-1.7	-1.5	-0.4	-1.7	NA	-0.2	-0.8	-2	2.4
LaBrana1	-0.5	-1.4	-0.4	-0.5	-1.6	-0.3	-1.3	0.2	NA	-0.6	-1.1	2.4
Hungarian.KO1	-0.4	-1.4	-1	-1.2	-1.1	0.3	-0.5	0.8	0.6	NA	-0.8	2.4
Malta_C	0.7	-0.7	-0.2	0.4	-0.5	0.7	-0.1	2	1.1	0.8	NA	3.9
Stuttgart	-2.8	-4.2	-2.5	-3.5	-3.6	-2.4	-3.9	-2.4	-2.4	-2.4	-3.9	NA
Ust'-Ishim												
Kostenki14	NA	-1.7	0.9	1.5	0.2	-0.2	0.6	0.2	0.8	2.3	-0.1	3.9
GoyetQ116-1	1.7	NA	3.9	3.2	2.1	2	2.2	2.3	3.1	3.8	1.8	5.9
Vestonice_CE_C	-0.9	-3.9	NA	-0.6	-2.5	-1.2	-1.2	-1.1	0.3	0.9	-1.5	3.1
Vestonice_I_C	-1.5	-3.2	0.6	NA	-1.1	-1.9	-1	-1	-0.6	0.8	-1.6	3.1
ElMiron	-0.2	-2.1	2.5	1.1	NA	0	0.7	0.4	1.3	2.2	-0.3	4.3
ElMiron _NI_C	0.2	-2	1.2	1.9	0	NA	1	0.7	1.6	2.7	0.5	4.8
Villabruna	-0.6	-2.2	1.2	1	-0.7	-1	NA	-0.5	0.6	2.2	-0.5	3.9
Loschbour	-0.2	-2.3	1.1	1	-0.4	-0.7	0.5	NA	0.9	3	-0.3	4.2
LaBrana1	-0.8	-3.1	-0.3	0.6	-1.3	-1.6	-0.6	-0.9	NA	2.1	-0.8	3.5
Hungarian.KO1	-2.3	-3.8	-0.9	-0.8	-2.2	-2.7	-2.2	-3	-2.1	NA	-3.1	1.5
Malta_C	0.1	-1.8	1.5	1.6	0.3	-0.5	0.5	0.3	0.8	3.1	NA	4.4
Stuttgart	-3.9	-5.9	-3.1	-3.1	-4.3	-4.8	-3.9	-4.2	-3.5	-1.5	-4.4	NA
Han												
Kostenki14	NA	-2.7	-1.5	-1.7	-2.2	-1.9	-1.3	-4.3	-3.8	-4	-5.3	0.4
GoyetQ116-1	2.7	NA	2	1.1	0.8	0.9	1.5	-1.1	-0.5	-1.2	-2.6	3.3
Vestonice_CE_C	1.5	-2	NA	-1.8	-2.4	-1.1	-0.4	-3.8	-3	-2.6	-4.8	1.1
Vestonice_I_C	1.7	-1.1	1.8	NA	-0.8	-0.4	0.8	-2.5	-2.5	-2.3	-4.1	1.8
ElMiron	2.2	-0.8	2.4	0.8	NA	0.3	1.5	-2	-1.3	-1.8	-3.3	2.7
ElMiron _NI_C	1.9	-0.9	1.1	0.4	-0.3	NA	1	-2.7	-1.5	-1.9	-3.1	2.4
Villabruna	1.3	-1.5	0.4	-0.8	-1.5	-1	NA	-3.9	-2.9	-3.3	-4	1.8
Loschbour	4.3	1.1	3.8	2.5	2	2.7	3.9	NA	1	0.5	-0.8	5.2
LaBrana1	3.8	0.5	3	2.5	1.3	1.5	2.9	-1	NA	-0.3	-1.8	4.5
Hungarian.KO1	4	1.2	2.6	2.3	1.8	1.9	3.3	-0.5	0.3	NA	-1.3	4.9
Malta_C	5.3	2.6	4.8	4.1	3.3	3.1	4	0.8	1.8	1.3	NA	5.9
Stuttgart	-0.4	-3.3	-1.1	-1.8	-2.7	-2.4	-1.8	-5.2	-4.5	-4.9	-5.9	NA

Gene flow among eastern non-Africans, the Mal'ta Cluster & pre-Neolithic Europeans

What is driving the statistics reported in Seguin-Orlando et al.[1]? Table S8.5 shows that the two statistics highlighted in that study—*D(Kostenki14, Loschbour/LaBrana1/Malta1; Y, Mbuti)* << 0 and *D(Kostenki14, Early European Farmer; Y, Mbuti)* ~ 0—are only observed when *Y* = Eastern non-Africans. When *Y* = *Oase1* or *UstIshim*, the signals are, as expected from Basal Eurasian ancestry, only seen in Early European Farmers.

The only way to explain these patterns is a history of gene flow between the ancestors of eastern non-Africans on the one hand, and the ancestors of three groups:

(a) A subset of Villabruna Cluster samples
(b) Early European farmers
(c) Mal'ta Cluster.

Such gene flow would induce a negative bias in two key statistics highlighted in Seguin-Orlando et al.[1] Note that for this to explain the data, three separate gene flow events are not required. Supplementary Information section 11 and Figure 4b document a link between (a) and (b), so as few as two gene flow events may be needed. Understanding the exact gene flow history responsible for these patterns is difficult with the ancient DNA sample series available here, but is an important question to address in future work.

References

1 Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).
2 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).
3 Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, doi:10.1038/nature14558 (2015).
4 Fu, Q. *et al.* Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445-449, doi:10.1038/nature13810 (2014).

Section 9

Malta1 is an outgroup to Upper Palaeolithic Europeans after 37,000 years ago

Malta1 is consistent with being an outgroup to Upper Palaeolithic Europeans

We studied the relationship of the ~24,000-year-old *Malta1* individual to pre-Neolithic mainland Europeans. Table S9.1 shows statistics of the form *D(X, Y; Malta1, Mbuti)*. For all samples younger than *Oase1*, we observe that $D(Test_1, Test_2; Malta1, Mbuti) \sim 0$. This is consistent with *Malta1* being an outgroup to pre-Neolithic mainland Europeans from the date of *Kostenki14* onward. When *X* is represented by *UstIshim, Oase1* or *Han*, we observe a significantly negative value, indicating that *Malta1* is not an outgroup to a clade including them and later Europeans (red).

Table S9.1 Z-score of *D(X, Y; Malta1, Mbuti)*

X/Y	Date	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	HohleFels49	Villabruna	Rochedane	Ranchot88	Loschbour	LaBrana1	Hungarian.KO1
Han		NA	3.1	6.3	-5.7	-7.2	-7.6	-5.9	-6.2	-2.8	-7.4	-7.2	-7.6	-9	-8.1	-9.1
UstIshim	48370-43070	-3.1	NA	3.5	-7.6	-8.7	-8.7	-6.6	-7.4	-3.6	-8.5	-8.3	-8.5	-10	-9.4	-10
Oase1	41761-37615	-6.3	-3.5	NA	-9.4	-10.2	-9.7	-6.1	-9.3	-2	-9.5	-8.5	-7.2	-11.9	-9.5	-11.3
Kostenki14	38650-37845	5.7	7.6	9.4	NA	-1.3	-1.5	-0.1	-0.1	-0.5	-1.1	-0.5	-0.9	-2.7	-1.7	-2.6
GoyetQ116-1	35160-34430	7.2	8.7	10.2	1.3	NA	-0.2	0.6	1.6	1.3	0.3	0.2	-0.4	-0.5	0.3	-0.8
Vestonice16	~31155	7.6	8.7	9.7	1.5	0.2	NA	1.2	1.4	1.5	0.7	0	-0.6	-1	0.2	-0.9
Ostuni1	27730-27530	5.9	6.6	6.1	0.1	-0.6	-1.2	NA	0.6	0.7	-0.6	-1.4	-1.1	-2.1	-1.4	-2.4
ElMiron	18735-18600	6.2	7.4	9.3	0.1	-1.6	-1.4	-0.6	NA	1.4	-0.4	-1.4	-0.7	-2.5	-1.3	-2.5
HohleFels49	16250-15568	2.8	3.6	2	0.5	-1.3	-1.5	-0.7	-1.4	NA	0	0.1	-0.1	-2	-0.7	-1.7
Villabruna	14075-13905	7.4	8.5	9.5	1.1	-0.3	-0.7	0.6	0.4	0	NA	-0.6	-0.4	-2	-0.4	-2.2
Rochedane	13090-12830	7.2	8.3	8.5	0.5	-0.2	0	1.4	1.4	-0.1	0.6	NA	0.3	-0.7	-0.2	-0.8
Ranchot88	10235-9933	7.6	8.5	7.2	0.9	0.4	0.6	1.1	0.7	0.1	0.4	-0.3	NA	-1.1	0.5	-1.6
Loschbour	8170-7940	9	10	11.9	2.7	0.5	1	2.1	2.5	2	2	0.7	1.1	NA	1.4	0
LaBrana1	7940-7690	8.1	9.4	9.5	1.7	-0.3	-0.2	1.4	1.3	0.7	0.4	0.2	-0.5	-1.4	NA	-0.9
Hungarian.KO1	7731-7596	9.1	10	11.3	2.6	0.8	0.9	2.4	2.5	1.7	2.2	0.8	1.6	0	0.9	NA

Contradiction of the *Malta1* ancestry model proposed in Seguin-Orlando et al.[1]

Statistics of the form *D(Malta1, X; Y, Mbuti)* are negative (usually highly significantly) if *(X, Y)* is any pair of pre-Neolithic mainland Europeans from *Kostenki14* until *Loschbour* and *LaBrana1* (Table S9.2). This implies that all pre-Neolithic Europeans all the way back to the time of *Kostenki14* share alleles with each other that are not shared with *Malta1*, as predicted by the hypothesis above that *Malta1* is an outgroup to Upper Palaeolithic Europeans.

These results directly contradict the findings of Seguin-Orlando et al.[1] who suggested a model in which *Kostenki14* is an outgroup to a clade that includes (*Malta1*, *Loschbour* and *LaBrana1*) (Figure 2 of that study). If that is true, then statistics of the form *D(Malta1, Loschbour/LaBrana1; Kostenki14, Mbuti)* are predicted to be consistent with 0. However, these statistics are significantly negative (Z=-3.1 for *Loschbour*, and Z=-2.2 for *LaBrana1*). In other words, there are more shared alleles between *Kostenki14* and Mesolithic Europeans, than between *Kostenki14* and *Malta1,* contradicting the model of Seguin-Orlando et al.[1]

No ancestral connection between *Malta1* and samples from the Věstonice Cluster

We were struck by the fact that the most significantly negative *D(Malta1, X; Kostenki14, Mbuti)* values were observed when *X* is any sample from the "Vestonice Cluster": *Vestonice16* gives Z=-6.1 and *Ostuni1* gives Z=-3.9 (Table S9.2). This is opposite to the positive bias in Z-scores that would be expected from shared ancestry between *Malta1* and

samples in the Věstonice Cluster. Shared ancestry between Malta1 and the Věstonice Cluster would be predicted if the assignment of both to the Gravettian culture was due to movements of people (the Mal'ta site yielded Venus figurines stylistically similar to Gravettian sites such as Dolní Věstonice in Europe, thousands of kilometers to the west). Thus, if the cultural similarity that led to the assignment of *Malta1* to the Gravettian culture is not a coincidence, it is likely to reflect communication of ideas rather than movements of people.

Table S9.2 Z-score of *D(Malta 1, X; Y, Mbuti)*

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	HohleFels49	Villabruna	Rochedane	Ranchot88	Loschbour	LaBrana1	KO1
Han	NA	0.5	0.5	10.9	9.4	11.8	9	9.1	4.4	10.8	10.5	10.4	9.9	9.4	10.6
UstIshim	3.6	NA	1.7	7.7	6.9	8.4	8	7.8	4.1	9.6	9.2	11.4	10.8	10.5	13.4
Oase1	7.1	4.9	NA	9.3	8.2	10.4	6.4	8.9	2.7	9.1	7.3	7.5	13.8	10.1	11.8
Kostenki14	5.3	0.4	0.3	NA	-1.9	-4.6	-3.9	-2.1	-1.3	-1.8	-0.7	0	-0.4	-0.4	1.6
GoyetQ116-1	2.6	-1.3	-1.6	-3.2	NA	-2.6	-3.3	-11.9	-8.1	-2.2	-3.3	-3.9	-4.9	-6.1	0
Vestonice16	5	0.2	0.5	-6.1	-3	NA	-13.8	-7.5	-1.9	-6.6	-4.3	-5.2	-5.6	-4.1	-3.3
Ostuni1	3.2	1.5	0	-3.9	-2.5	-11.9	NA	-7	-2.4	-6.8	-3	-5.5	-4.7	-4.7	-2
ElMiron	3.4	0.8	-0.8	-2.4	-10.6	-6.3	-6.6	NA	-14	-15	-13.5	-16.1	-17.1	-20.1	-8.6
HohleFels49	2	0.4	0.6	-1.7	-6.7	-0.5	-1.7	-12.5	NA	-3.7	-3.9	-4.8	-6	-6.3	-2
Villabruna	4.2	1	-0.2	-2.9	-1.9	-5.6	-7.3	-15	-3.6	NA	-23.6	-25.3	-29.4	-22.1	-23
Rochedane	3.4	0.7	-1	-1.3	-3.2	-4.2	-4.3	-15.1	-3.7	-24.6	NA	-21	-24.8	-17.8	-15.1
Ranchot88	3.5	2.8	-0.2	-1	-4.1	-5.4	-6.3	-15.7	-4.6	-25.3	-20.9	NA	-31.6	-23.2	-19.1
Loschbour	0.9	0.6	1.8	-3.1	-5.1	-6.1	-6.7	-18.5	-7.4	-31.1	-25.2	-31.5	NA	-27.2	-25.5
LaBrana1	1.9	1.1	0.8	-2.2	-5.5	-3.8	-6.3	-19.9	-6.9	-21.6	-17.9	-21.8	-27	NA	-19.4
Hungarian.KO1	1.4	3.3	0.4	-1	-0.8	-4	-4.6	-11	-3.5	-23.8	-15.5	-20.7	-25.9	-21.3	NA

New model: An Admixture Graph that jointly fits *Malta1*, *Kostenki14* and *GoyetQ116-1*

Figure S9.1 shows that a model with no admixture events is consistent with the joint data for *Mbuti, UstIshim, Malta1, Kostenki14* and *Goyet116-1*, in the sense that there are no significant deviations between observed and predicted *f*-statistics (maximum |Z|=2.5, not significant correcting for multiple hypothesis testing). No other simple model fits the data

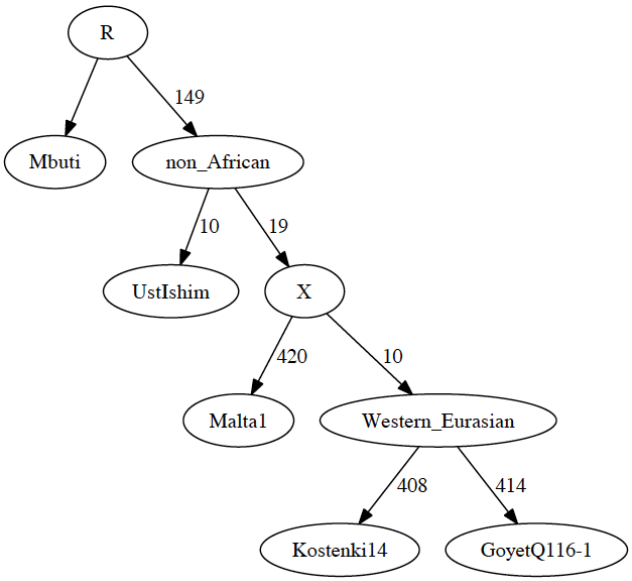


Figure S9.1 A parsimonious phylogeny with no admixture events. This is the only model that can jointly fit *Malta1, Kostenki14, UstIshim* and *GoyetQ116-1* without invoking any admixture events. (This is used as the base model for Supplementary Information section 6, and is reproduced there as Figure S6.1.)

References

1 Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).

Section 10

A genetic link between *GoyetQ116-1* and the El Mirón Cluster

Outgroup f_3-statistics show affinity of *GoyetQ116-1* to the El Mirón Cluster

In the matrix of pairwise $f_3(X, Y; Mbuti)$ (Figure 3A and Extended Data Figure 2), *GoyetQ116-1*, which dates to around 35,000 BP, shares substantial genetic drift with all the El Mirón Custer samples *ElMiron*, *HohleFels79*, *HohleFels49*, *Rigney1*, *GoyetQ-2* and *Burkhardtshohle*, which date to the period 19,000-14,000 BP. No other Upper Palaeolithic Eurasian—including *Oase1*, *Kostenki14*, *Malta1*, and multiple samples in the Věstonice Cluster—shows this affinity to the El Mirón Cluster.

D-statistics suggest a genetic link between *GoyetQ116-1* and the El Mirón Cluster

We used *D*-statistics of the form *D(X, Y; GoyetQ116-1, Mbuti)* to test formally if samples from the El Mirón Cluster are consistent with sharing more alleles with *GoyetQ116-1* as suggested by outgroup f_3 analysis (Table S10.1). We find that Z-scores of the statistic *D(El Mirón Cluster sample, other pre-Neolithic Europeans; GoyetQ116-1, Mbuti)* are all significant positive ($19.1 \geq Z \geq 3.3$), confirming the genetic affinity. The only exception to this pattern is *HohleFels79*, which is limited in its dataset size (11,211 SNPs).

Table S10.1. Z-score of *D(X, Y; Q116-1, Mbuti) for all sites*

D(X, Y; GoyetQ116-1, Mbuti): 846,983 SNPs X/Y	UstIshim	Oase1	Kostenki14	Vestonice16	Ostuni1	Villabruna	Loschbour	LaBrana	KO1	ElMiron
UstIshim	NA	3.6	-9.1	-9.7	-9.4	-9.2	-12.5	-12.9	-8.4	-18.1
Oase1	-3.6	NA	-10.5	-11.5	-9.2	-10.1	-13.2	-13.5	-9.6	-17.2
Kostenki14	9.1	10.5	NA	-1.2	-1.8	-0.5	-3.1	-3.7	1	-9.8
Vestonice13	8.7	7.6	2.5	1.6	-0.4	2.2	-0.2	-0.7	3.1	-5
Vestonice15	5.5	5.2	1.4	-0.5	-0.3	1.5	-0.6	0.5	1.5	-3.5
Vestonice43	7.9	7.2	1.7	0.9	-0.2	0.8	-1.6	-1.2	2	-6.8
Vestonice16	9.7	11.5	1.2	NA	-0.7	0.8	-2	-2.3	2.1	-8.7
Ostuni1	9.4	9.2	1.8	0.7	NA	1.4	-1.2	-1.9	1.8	-7.1
ElMiron	18.1	17.2	9.8	8.7	7.1	11.1	8.4	7.8	10.5	NA
HohleFels79	6.6	2.2	2.6	3.9	1.1	4.4	2.5	3.5	3.1	1.2
HohleFels49	12.3	9.1	6.9	7	5.3	8.9	6.7	6.6	6.8	1.7
Rigney1	10	7.3	5.9	5.6	3.4	4.5	4.2	3.3	4.7	0.6
GoyetQ-2	16.5	8.9	10.5	8.2	6.5	10.7	9.7	8.1	9.3	2.4
Burkhardtshohle	12.7	8.1	8.7	6.8	5.3	7	6.7	5.2	6.9	1.4
HohleFels79_nq	17.5	13.2	10.2	9	7.9	11.7	9.9	9.1	11.7	4.6
HohleFels49_nq	17.8	14.7	10.7	9.2	9	12.1	10.3	9.7	11.7	3.9
Rigney1_nq	15.3	11.9	9.7	8.2	5.8	8.4	7.6	6.2	8.2	1.3
GoyetQ-2_nq	21.9	18.9	14.6	13.2	11.4	14.8	13.9	12.7	15.3	4.6
Villabruna	9.2	10.1	0.5	-0.8	-1.4	NA	-3.7	-4.1	0.6	-11.1
Rochedane	9.9	8.4	1.4	0.5	-0.1	1.7	-1.6	-1.5	2.2	-8.4

To increase the power of this analysis in the face of limited data for some El Mirón Cluster individuals, we repeated the analysis using a version of the dataset that does not restrict to damaged sequences for *HohleFels49*, *HohleFels79*, *Rigney1*, or *GoyetQ-2* (suffix "_nq"). The signal becomes even stronger ($21.9 \geq Z \geq 6.2$) for each of these four individuals (Table S10.1). Contamination by present-day Europeans is expected to make this signal weaker, suggesting that the strengthening of the signal is not due to contamination.

To confirm that these patterns are not artifacts of ancient DNA damage, we restricted to transversion SNPs, while at the same time retaining as many SNPs as possible for analysis by grouping *HohleFels79*, *HohleFels49*, *Rigney1*, *GoyetQ-2*, *Burkhardtshohle* and *Brillenhohle* into an "ElMiron_NI_C" cluster (Supplementary Information section 5) (Table S10.2). Z-scores of the form *D(ElMiron_NI_C, other hunter-gatherers; GoyetQ116-1, Mbuti)* remain significantly positive ($16.5 \geq Z \geq 3.3$).

Table S10.2. Z-score of *D(X, Y; Q116-1, Mbuti) restricting to transversion SNPs*

X/Y	UstIshim	Oase1	Kostenki14	Vestonice_CE_C	Vestonice_I_C	ElMiron_NI_C	Villabruna	Loschbour	LaBrana	KO1	ElMiron
UstIshim	NA	1.2	-8.2	-6.4	-8.5	-17.6	-8.7	-10.8	-11.1	-7.6	-16.1
Oase1	-1.2	NA	-6.7	-4.9	-5.9	-11	-5.6	-8.1	-7.5	-5.4	-10.5
Kostenki14	8.2	6.7	NA	-0.7	-3.2	-10.5	-0.6	-2.1	-2.8	0.3	-7.6
Vestonice_CE_C	6.4	4.9	0.7	NA	-2	-8.2	0.1	-2.6	-1.8	0.5	-6.6
Vestonice_I_C	8.5	5.9	3.2	2	NA	-7.4	1.5	0.6	-0.1	2.3	-4.8
ElMiron	16.1	10.5	7.6	6.6	4.8	-4.8	7.7	6.9	5.6	7	NA
ElMiron_NI_C	17.6	11	10.5	8.2	7.4	NA	10.6	10.8	9.4	9.3	4.8
Villabruna	8.7	5.6	0.6	-0.1	-1.5	-10.6	NA	-2.4	-2.9	-0.3	-7.7
Rochedane	6.4	5.4	1.8	1.7	-1	-5	0.8	-0.1	-1.2	0.4	-4

A striking observation is that *GoyetQ116-1* shares more alleles with the sample pool ElMiron_NI_C than with *ElMiron*, the oldest sample in the El Mirón Cluster (last column of Table S10.1 and of Table S10.2). These results suggest that there was a gradient of *GoyetQ116-1* relatedness within the El Mirón Cluster.

Estimating mixture without a full phylogenetic model

In order to estimate the proportion of mixture from two potential reference populations in a target individual, we use the *qpAdm* method[1]. This method leverages the fact that if a *Test* population is a mixture of *N* different ancestral populations that are clades with reference populations from which we have genetic data, we can write:

$$f_4(Test, O_i; O_j, O_k) \approx \sum_{i=1}^{N} \alpha_i f_4(Ref_i, O_i; O_j, O_k)$$

Here, O_i, O_j and O_k are *Outgroup* populations. For the method to work, each of the *N* strands of ancestry in the *Test* population needs to be phylogenetically more closely related to the ancestry in one of the *Reference* populations than it is to any of the ancestry in any of the *Outgroup* populations.

We can formally test the fit of this model to the data—taking into account the covariance of these f_4-statistics—and thus produce a single P-value for fit using a Hotelling T^2 test. We can then estimate mixture proportions based on the fitted weights of the α_i coefficients.

ElMiron

We used *qpAdm*[1] to estimate the proportion of mixture in the target individual (*ElMiron*), given data from reference populations that we propose to be sister groups or descendent groups to the two mixing populations.

Briefly, we divide the populations into two sets. The *left* set *L* consists of the proposed admixed population (*ElMiron*) and *Reference* populations (*GoyetQ116-1*, *Villabruna*,

Ranchot88, LaBrana1, Loschbour). The *right* set *R* consists of 6 worldwide populations excluding Europeans (*Ulchi, Mbuti, Ami, Tubalar, Kinh, Onge*); these are *Outgroups* that provide leverage for discerning different components of ancestry among the *Left*. The *Outgroups* are chosen to represent a range of non-European ancestries, thus maximizing leverage for teasing apart lineages within Europe.

We used *qpAdm* to estimate the rank of the matrix of F_4-statistics $M(l, r) = F_4(l_x, l; r_x, r)$ where l and r are the left and right population sets. This matrix [2] has rank N-1 if there are N waves of migration. *qpAdm* also outputs a mixture proportion.

Rank 1 is not excluded (p=0.9) with the best model being a mixture of lineages related (perhaps anciently) to *GoyetQ116-1* and *LaBrana1*. *ElMiron* can be modeled as having 49±13% *GoyetQ116-1* and 51±13% *LaBrana1* ancestry. (In Supplementary Information section 6 using full Admixture Graph modeling, we obtain a point estimate of 63% *GoyetQ116-1* related ancestry, consistent with this model.) When we add a present-day *Iraqi_Jew* into *Outgroups* to represent a Near Eastern population, the standard error narrows to 10%. We note that *LaBrana1* and *Iraqi_Jew* post-date *ElMiron*, but this is not a problem for the method. What is important for *qpAdm* to be effective is that these populations have the phylogenetic positioning relative to *ElMiron* that is specified by the *qpAdm* model, which can be true even if they are relatively recent members of these populations.

ElMiron_NI_C
We tested *GoyetQ116-1, Villabruna, Loschbour, LaBrana1,* and *Ranchot88* as *Left* populations and *Ulchi, Mbuti, Ami, Tubalar, Kinh and Onge* as *Right* populations. The method infers 69±11% GoyetQ116-1 and 32±11% *LaBrana1*-related ancestry. (In Supplementary Information section 6 using Admixture Graph modeling, we obtain a consistent point estimate of 80% *GoyetQ116-1* related ancestry.) When we add *Iraqi_Jew* into the *Right* populations, the standard errors narrow to 7%.

Conclusion
Some populations that lived ~35,000 years ago in northwest Europe (i.e. *GoyetQ116-1*) have an unambiguous link to populations of the El Mirón Cluster that lived around 19,000-14,000 BP. Of course, the link might not involve direct descent from a population that lived near Goyet Cave. Instead, it may be mediated by populations related to *GoyetQ116-1* that lived elsewhere.

References
1 Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature,* doi:doi:10.1038/nature14317 (2015).
2 Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370-374, doi:10.1038/nature11258 (2012).

Section 11

Gene flow linking the Villabruna Cluster and the Near East

We investigated the relationship between pre-Neolithic Europeans and present-day as well as ancient populations using statistics of the form:

$$D(European_1, European_2; Test, Mbuti)$$

Here, the *Test* populations are Native Americans, East Asians, Oceanians and Near Eastern populations from the Simons Genome Diversity Project (SGDP) panel.

Affinities of pre-Neolithic Europeans to the Near East

When neither of the two pre-Neolithic Europeans analysed in the statistic is in the Villabruna Cluster—that is, both are older than about 14,000 BP—they tend to be symmetrically related to populations outside Europe including present-day and ancient Near Easterners. However, when one lived prior to the Villabruna Cluster (e.g. *Vestonice16, ElMiron, Kostenki14, KremsWA3*, and *GoyetQ116-1*) and the other is in the Villabruna Cluster (e.g. *BerryAuBac, Bichon, CuiryLesChaudardes1, Falkenstein, Hungarian.KO1, LaBrana1, Loschbour, Ranchot88, Rochedane* and *Villabruna*), there is a distinct attraction of the Villabruna Cluster samples to Near Eastern populations (Figure 4b; Extended Data Figure 3). Table S11.1 shows the statistics when the Near Eastern population is *Iraqi_Jew*.

There are several possible explanations for these findings. One is gene flow between relatives of Near Easterners and pre-Neolithic Europeans after ~14,000 years ago, beginning with the Villabruna Cluster. A second is population substructure in Europe. In this scenario, after post-glacial re-peopling of Europe, the balance of ancestry could have shifted toward populations that were more closely related to Near Easterners. In either case, however, major population turnovers must have occurred.

The affinity of pre-Neolithic Europeans to Near Easterners beginning around 14,000 years ago is distinct from the affinity to East Asians in Mesolithic Europeans

Seguin-Orlando et al.[1] documented that statistics of the form *D(Kostenki14; Mesolithic Europeans; East Asians, Outgroup)* are significantly less than 0. In Supplementary Information section 8, we show that this is likely due to gene flow between the ancestors of East Asians and the ancestors of Mesolithic Europeans. These patterns are evident in Figure 4b and Extended Data Figure 3, which show that a subset of Villabruna Cluster samples including Mesolithic Europeans show a significant affinity to East Asians. This pattern does not go hand-in-hand with the affinity to the Near East that is present in all Villabruna Cluster samples, and thus the two signals must therefore reflect at least two distinct historical events.

Table S11.1 ***D(European$_1$, European$_2$; Iraqi_Jew, Mbuti).*** Europeans after around 14,000 years ago in the Villabruna Cluster began to have a significant affinity to Near Easterners.

	European$_1$=Kostenki14			European$_1$=GoyetQ116-1			European$_1$=Vestonice16			European$_1$=ElMiron		
European$_2$	**D**	**Z**	**SNPs**	**D**	**Z**	**SNPs**	**D**	**Z**	**SNPs**	**D**	**Z**	**SNPs**
GoyetQ116-1	-0.0001	-0.2	778,865	n/a	n/a	n/a	0.0003	0.8	879,269	0.0014	2.5	541,039
Kostenki14	n/a	n/a	n/a	0.0001	0.2	778,865	n/a	n/a	n/a	0.0016	3	745,568
Kostenki12	0.0008	0.9	55,817	0.0013	1.4	47,155	0.0015	0.6	615,864	0.0021	2.1	44,385
Muierii2	0.0002	0.3	89,779	-0.0003	-0.4	71,687	-0.0006	1.5	49,078	0.0013	1.6	73,066
Paglicci133	0.0002	0.2	75,733	0.0008	0.8	61,511	-0.0002	-0.8	79,159	0.0015	1.7	52,021
Vestonice13	0	-0.1	128,690	-0.001	-1.3	101,433	-0.0008	-0.3	58,743	0.001	1.4	102,285
Vestonice43	0	0	147,703	-0.0002	-0.3	116,524	0.0002	-1.3	111,268	0.0016	2.4	123,622
Pavlov1	-0.0011	-1.1	50,750	-0.0007	-0.7	47,910	-0.0006	0.4	132,591	0.0018	1.8	46,263
Vestonice16	-0.0004	-0.8	879,269	-0.0003	-0.6	615,864	n/a	-0.7	49,523	0.0012	2.5	603,758
Ostuni1	-0.0009	-1.6	338,407	-0.0011	-1.7	278,753	-0.0007	-1.3	301,276	0.0007	1.3	278,203
KremsWA3	-0.0009	-1.5	219,623	-0.0003	-0.5	189,593	-0.0001	-0.2	181,849	0.0008	1.2	161,774
ElMiron	-0.0016	-3	745,568	-0.0014	-2.5	541,039	-0.0012	-2.5	603,758	n/a	n/a	n/a
HohleFels49	-0.001	-1.1	56,741	0.0002	0.3	49,569	-0.0009	-1	55,683	0.0018	2.2	52,016
GoyetQ2	-0.0025	-2.8	67,154	-0.0013	-1.5	57,936	-0.0007	-0.8	55,614	-0.0006	-0.6	49,812
Villabruna	-0.0038	-7.8	1,125,277	-0.0038	-7	718,424	-0.0033	-6.9	797,720	-0.002	-4.4	696,121
Bichon	-0.0037	-7.9	1,669,947	-0.0037	-6.9	810,452	-0.0032	-6.9	912,465	-0.0018	-4.1	776,355
Rochedane	-0.0034	-5.4	218,932	-0.0033	-5.1	189,843	-0.0031	-5	181,626	-0.0016	-2.7	161,882
Ranchot88	-0.0036	-6.3	381,831	-0.0038	-6.4	351,382	-0.0034	-6.7	337,147	-0.0022	-4.4	306,065
Falkenstein	-0.0029	-3.3	58,169	-0.0025	-2.8	49,262	-0.0023	-2.7	55,226	-0.0021	-2.7	51,576
Chaudardes1	-0.004	-4.7	86,086	-0.0043	-5	74,202	-0.0035	-4.3	71,180	-0.0027	-3.5	63,128
Loschbour	-0.0043	-9.3	1,660,854	-0.0042	-8.2	802,734	-0.004	-8.5	904,305	-0.0027	-6.4	768,832
LaBrana1	-0.0034	-7.4	1,557,333	-0.0034	-6.3	776,310	-0.0031	-6.8	875,649	-0.0018	-4.1	748,847
Hungarian.KO1	-0.0043	-9.3	1,137,698	-0.0048	-8.7	559,581	-0.004	-8.6	629,253	-0.003	-6.5	536,668
BerryAuBac	-0.0042	-4.2	50,065	-0.0036	-3.5	46,237	-0.003	-2.9	44,315	-0.0017	-1.6	40,189

References

1 Seguin-Orlando, A. *et al.* Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113-1118, doi:10.1126/science.aaa0114 (2014).

Section 12

Population affinities of the Satsurblia Cluster

Overview

Here we describe how two individuals from the Satsurblia Cluster[1]—the Upper Palaeolithic *Satsurblia* and the Mesolithic *Kotias*—relate to other samples in this study.

Satsurblia Cluster samples have Basal Eurasian ancestry

We computed statistics of the form: *D(Satsurblia/Kotias, Pre-Neolithic Europeans; UstIshim/Oase1, Mbuti)* (Table S12.1). These statistics are significantly negative, different from the pattern seen when both of the first two samples are pre-Neolithic Europeans (Supplementary Information section 8). These results suggest that Satsurblia Cluster samples, unlike pre-Neolithic people from more western parts of Europe, harbored ancestry from a "Basal Eurasian" lineage. Specifically, they harbored ancestry from a lineage that split from European hunter-gatherers, *UstIshim* and *Oase1* before those groups separated[2].

Table S12.1. Z-score of *D(Satsurblia/Kotias, X; Y, Mbuti)*

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Villabruna	Bichon	Loschbour	LaBrana1	Hungarian.KO1	Malta1	Motala12	Stuttgart
D(Satsurblia, X; Y, Mbuti) Satsurblia: 1,460,368 SNPs															
Han	NA	-4.6	-4.4	3.7	0.2	3.1	1.9	7.1	6.5	5.6	4.1	5.5	2.7	6.6	11.9
UstIshim	-1.7	NA	-2	1.8	-0.3	1.8	2.7	6.2	7.4	7.1	5.8	9	4.8	8.3	11.9
Oase1	2.7	1.4	NA	5.5	2.5	5.2	4.9	7.6	8.8	10	7.1	8.8	6.7	10.1	14.5
Kostenki14	-0.3	-3.7	-3.3	NA	-9.8	-11.5	-8.9	-5.3	-3.9	-4.4	-4.6	-2.3	-2.4	-1.3	4.8
GoyetQ116-1	-3.1	-5.1	-3.7	-8.9	NA	-10	-18	-6.5	-6.9	-8.1	-10.1	-3.5	-3.9	-4.8	3.2
Vestonice16	-1.1	-4.1	-3.5	-11.6	-10.6	NA	-13.4	-9.8	-8.2	-9.2	-8.3	-7	-3.7	-5.3	2.4
ElMiron	-2.7	-3.3	-3.6	-9.1	-18.7	-13.4	NA	-17.7	-20	-21.1	-24.5	-11.6	-3	-12.2	0.3
Villabruna	-1.8	-3.5	-3.6	-9.1	-11	-14.1	-22.2	NA	-31.6	-33.2	-26.7	-26.8	-4.2	-22.8	-3.8
Bichon	-4.2	-3.7	-3.4	-9.2	-12.1	-12.6	-24.9	-31.3	NA	-32.9	-30.2	-25.5	-5.4	-21.5	-4.1
Loschbour	-5.2	-4.1	-2.5	-9.9	-13.9	-13.9	-27.6	-34.8	-34.6	NA	-32.9	-27.7	-5.6	-24.4	-5.8
LaBrana1	-4.5	-3.1	-2.4	-8.3	-13.9	-11.3	-28.9	-26.1	-29.2	-31.1	NA	-21.3	-4	-19.4	-2.7
Hungarian.KO1	-4.8	-1.2	-3.4	-7.2	-9.3	-11.6	-17.8	-27.6	-26.2	-27	-24.2	NA	-5	-21.1	-6.1
Malta1	-5.4	-4.2	-3.2	-5.7	-7.9	-6.6	-6.2	-3.7	-3.4	-3.3	-4.2	-3.2	NA	-8.1	5
Motala12	-4.3	-3.1	-2.5	-6.5	-10.3	-9.7	-16.3	-22.1	-21	-22.8	-20.2	-19.8	-10	NA	-4.5
Stuttgart	-0.3	-0.1	-0.9	-2	-4.1	-4.5	-6.6	-6.4	-5.1	-6.3	-5.9	-6.5	1.9	-5.5	NA
Kotias	-0.8	0.4	-1.1	-1.2	-1.9	-0.8	-3.4	-2	-0.5	-1.3	-2.7	-1.2	-0.4	-3	-1
D(Kotias, X; Y, Mbuti) Kotias: 2,133,968 SNPs															
Han	NA	-4.9	-3.8	4.7	2.1	4.2	5.1	8.7	7.2	6.9	6.5	6.9	3.2	9.5	12.6
UstIshim	-1.2	NA	-1.6	2.7	0.9	2.2	4.7	8.1	7.9	8.4	7.9	10.4	5	10.6	13
Oase1	3.6	1.4	NA	5.8	4.1	5.4	7	8.1	10.2	10.8	8.8	10.3	7.2	11.1	14.4
Kostenki14	-0.1	-4.4	-3	NA	-8.6	-11.1	-6.9	-4.1	-3.8	-3.4	-3.4	-1.8	-2.2	1.1	6
GoyetQ116-1	-2.8	-6.3	-3.9	-9.2	NA	-10.3	-17.1	-5.6	-6.1	-8.2	-9.6	-3.9	-3.9	-2.7	3.9
Vestonice16	-0.6	-4.8	-3.4	-11.8	-9.9	NA	-11.7	-9.7	-7.9	-8.8	-7	-6.9	-3.7	-3.6	3.3
ElMiron	-2.4	-5	-3.5	-9.3	-18.3	-13.7	NA	-18.2	-18.7	-20.7	-23.8	-12.6	-2.8	-10.4	0.8
Villabruna	-1.4	-4.4	-4	-9.1	-9.7	-13.7	-21.4	NA	-30.5	-33	-26.4	-26.5	-3.7	-20.1	-3.2
Bichon	-3.9	-4.3	-2.8	-8.9	-10.6	-12.7	-23	-31	NA	-31.3	-29.1	-24.7	-5.2	-18.2	-3.1
Loschbour	-4.7	-4.5	-2.5	-9.3	-12.7	-14.1	-24.3	-33.6	-31.7	NA	-31.4	-27.3	-5.1	-21.3	-4.7
LaBrana1	-4	-4	-2.4	-8.6	-13.3	-11.7	-27.4	-26.3	-28.8	-31.4	NA	-22.6	-4.4	-17.4	-1.9
Hungarian.KO1	-4.3	-2	-2.7	-7.7	-8.8	-11.8	-16.5	-28.1	-25.8	-27.6	-24.4	NA	-5.3	-18.8	-5.4
Malta1	-5.5	-4.9	-3.8	-5.7	-7.2	-6.9	-4	-2.4	-3.8	-2.9	-2.9	-3	NA	-6.2	5.4
Motala12	-3.8	-3.3	-2.4	-6.1	-9.1	-9.8	-14.4	-21.4	-21.3	-22.4	-19.8	-19.8	-9.6	NA	-3.5
Stuttgart	0.4	-0.6	-0.4	-1.3	-2.6	-3.9	-4	-5.2	-4.6	-5.3	-4	-5.7	2.1	-3.1	NA
Satsurblia	0.8	-0.4	1.1	1.2	1.9	0.8	3.4	2	0.5	1.3	2.7	1.2	0.4	3	1

Satsurblia Cluster samples have West Eurasian as well as Basal Eurasian ancestry

Satsurblia Cluster samples have substantial amounts of Basal Eurasian ancestry. If they had entirely Basal Eurasian ancestry, however, a prediction would be that they would be no more closely related to pre-Neolithic Europeans than they are to *Han*, *UstIshim*, or *Oase1*. This is not the case. Table S12.2 shows that statistics of the form *D(Ust-Ishim/Han/Oase1, pre-Neolithic Europe; Satsurblia/Kotias, Mbuti)* are all significantly negative (Z<<-3 scores; first

three rows of Table S12.2). The only way to explain this is if Satsurblia Cluster samples harbor a mix of Basal Eurasian and West Eurasian ancestry.

Table S12.2 Z-score of *D(X, Y; Satsurblia/Kotias, Mbuti)*

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	ElMiron	Villabruna	Bichon	Loschbour	LaBrana1	Hungarian.KO1	Malta1	Motala12	Stuttgart
D(X, Y; Satsurblia, Mbuti) Satsurblia: 1,460,368 SNPs															
Han	NA	2.8	6.6	-4.1	-3.1	-4.2	-4.6	-9.2	-10.6	-10.8	-8.3	-10.1	-8.1	-11.2	-12.3
UstIshim	-2.8	NA	3.5	-6.1	-5.3	-6.2	-6.1	-9.9	-11.1	-11.1	-9	-10.4	-9.4	-11.7	-12
Oase1	-6.6	-3.5	NA	-8.4	-6.2	-8.6	-8.7	-11.1	-12	-11.9	-9.3	-11.3	-10.2	-12.5	-15.1
Kostenki14	4.1	6.1	8.4	NA	0.8	-0.3	-0.4	-4.2	-5.8	-6.2	-3.6	-5.1	-3.4	-5.9	-6.6
GoyetQ116-1	3.1	5.3	6.2	-0.8	NA	-0.6	-0.8	-4.3	-5.7	-6.1	-3.7	-5.7	-3.8	-5.5	-6.9
Vestonice16	4.2	6.2	8.6	0.3	0.6	NA	0.1	-3.8	-5	-5.6	-2.9	-4.4	-3	-4.8	-6.6
ElMiron	4.6	6.1	8.7	0.4	0.8	-0.1	NA	-3.9	-5.1	-5.7	-3.3	-5.5	-3.2	-4.4	-7.1
Villabruna	9.2	9.9	11.1	4.2	4.3	3.8	3.9	NA	-1.5	-2.1	0.4	-1.8	0.4	-1.2	-2.8
Bichon	10.6	11.1	12	5.8	5.7	5	5.1	1.5	NA	-0.2	2.3	0	2.1	0.1	-1.1
Loschbour	10.8	11.1	11.9	6.2	6.1	5.6	5.7	2.1	0.2	NA	2.7	0.6	2.2	0.4	-0.9
LaBrana1	8.3	9	9.3	3.6	3.7	2.9	3.3	-0.4	-2.3	-2.7	NA	-2.1	-0.3	-2	-3.2
Hungarian.KO1	10.1	10.4	11.3	5.1	5.7	4.4	5.5	1.8	0	-0.6	2.1	NA	1.7	0.1	-1.2
Malta1	8.1	9.4	10.2	3.4	3.8	3	3.2	-0.4	-2.1	-2.2	0.3	-1.7	NA	-2.3	-2.9
Motala12	11.2	11.7	12.5	5.9	5.5	4.8	4.4	1.2	-0.1	-0.4	2	-0.1	2.3	NA	-1.1
Stuttgart	12.3	12	15.1	6.6	6.9	6.6	7.1	2.8	1.1	0.9	3.2	1.2	2.9	1.1	NA
Kotias	26.6	24	23	20.2	19.2	19.6	18.1	16.6	15.8	16.1	17.9	15.5	16.6	16.4	14.9
D(X, Y; Kotias, Mbuti) Kotias: 2,133,968 SNPs															
Han	NA	3.9	7.4	-4.9	-4.7	-5	-7.5	-10.9	-11.3	-11.7	-10.7	-11.6	-8.5	-13.9	-12.9
UstIshim	-3.9	NA	3.3	-8	-7.5	-7.7	-10.1	-13.3	-12.4	-12.9	-12.2	-13.3	-9.9	-14.4	-14.5
Oase1	-7.4	-3.3	NA	-8.3	-8	-8.7	-10.9	-12.6	-13	-13.3	-11.6	-12.3	-10.7	-13.1	-15.3
Kostenki14	4.9	8	8.3	NA	-0.6	-0.2	-2.3	-5.7	-5.3	-6.6	-5	-6.5	-3.6	-7.6	-7.2
GoyetQ116-1	4.7	7.5	8	0.6	NA	0.1	-1.7	-4.6	-4.6	-5.2	-3.7	-5.2	-3	-6.5	-6.6
Vestonice16	5	7.7	8.7	0.2	-0.1	NA	-2	-4.9	-5	-6	-4.9	-5.6	-3.3	-7.1	-7.2
ElMiron	7.5	10.1	10.9	2.3	1.7	2	NA	-2.5	-3.4	-4	-3.1	-4.1	-1.1	-4.2	-5.2
Villabruna	10.9	13.3	12.6	5.7	4.6	4.9	2.5	NA	-0.2	-1.2	0.2	-1.3	1.3	-2.1	-2
Bichon	11.3	12.4	13	5.3	4.6	5	3.4	0.2	NA	-1.2	0.2	-1.1	1.6	-2.5	-1.7
Loschbour	11.7	12.9	13.3	6.6	5.2	6	4	1.2	1.2	NA	1.5	0.2	2.5	-1.3	-0.7
LaBrana1	10.7	12.2	11.6	5	3.7	4.9	3.1	-0.2	-0.2	-1.5	NA	-1.2	1.3	-2.4	-2.2
Hungarian.KO1	11.6	13.3	12.3	6.5	5.2	5.6	4.1	1.3	1.1	-0.2	1.2	NA	2.4	-1.5	-0.7
Malta1	8.5	9.9	10.7	3.6	3	3.3	1.1	-1.3	-1.6	-2.5	-1.3	-2.4	NA	-4	-3.2
Motala12	13.9	14.4	13.1	7.6	6.5	7.1	4.2	2.1	2.5	1.3	2.4	1.5	4	NA	0.5
Stuttgart	12.9	14.5	15.3	7.2	6.6	7.2	5.2	2	1.7	0.7	2.2	0.7	3.2	-0.5	NA
Satsurblia	26.4	24.5	21.6	19.4	17.3	18.2	15.6	14.9	15.6	15.2	15.3	14.6	15.8	13.1	14.2

The Satsurblia and Villabruna Clusters are not particularly closely related

What is the nature of the West Eurasian genetic affinities in the Satsurblia Cluster samples?

We observe significantly positive statistics of the form *D(Villabruna Cluster, Early Upper Palaeolithic Europeans; Satsurblia Cluster, Mbuti)*, showing that Satsurblia Cluster samples share more alleles with Villabruna Cluster samples—for example, *Villabruna*, *Bichon*, *Loschbour*, *LaBrana1*, and *Hungarian.KO1*—than with Early Upper Palaeolithic Europeans (*Kostenki14*, *GoyetQ116-1*, *Vestonice16* and *ElMiron*) (Table S12.2). This suggests the possibility that a component of the non-Basal Eurasian ancestry in the Satsurblia Cluster may be related to the ancestry that appears in our European sample series with the Villabruna Cluster. In other words, migrations involving populations related to the Satsurblia Cluster could be responsible for the genetic link between the Near East and the Villabruna Cluster (Supplementary Information section 11).

It is important to emphasize, first of all, that Satsurblia Cluster can not be the direct source of the Near Eastern affinity that appears in our European sample series from *Villabruna* onward (Figure 4b), as Satsurblia Cluster samples have substantial Basal Eurasian ancestry, whereas Villabruna Cluster samples do not (Supplementary Information section 8).

To explore the relationship between the non-Basal Eurasian ancestry in the *Satsurblia* cluster and the Near Eastern related ancestry in the Villabruna Cluster in more detail, we fit

Satsurblia into the Admixture Graph of Supplementary Information section 6 that includes *Mbuti, UstIshim, Malta1, GoyetQ116-1, Kostenki14, Vestonice16,* and *ElMiron* (Figure S6.3). In all fitting models, *Satsurblia* is inferred to harbor ~32% ancestry from a Basal Eurasian lineage that branched before *UstIshim* (Extended Data Figure 4). These results are consistent with the finding that Satsurblia Cluster samples have Basal Eurasian ancestry[1], while also rejecting the previous model in which *Satsurblia* is of entirely Basal Eurasian ancestry[1]. Our fitted model specifies that the majority of ancestry in *Satsurblia* is from a West Eurasian, lineage providing an explanation both for why *Satsurblia* has Basal Eurasian ancestry, while at the same time sharing alleles with all Europeans beginning with *Kostenki14*.

All three of the best fitting models in Extended Data Figure 4 specify that the majority ancestry component in *Satsurblia* branched very deeply in the tree of West Eurasian populations, forming a clade with *Malta1*. Further evidence for a deep connection to *Malta1* and populations with admixture of *Malta1*-related ancestry comes from the observation in Table S12.2 that *D(Motala12/Malta1, Early Upper Palaeolithic Europeans; Satsurblia Cluster, Mbuti)* is significantly positive. In a simple model of this type, a prediction is that statistics of the form *D(Villabruna Cluster, Early Upper Palaeolithic Europeans; Malta Cluster, Mbuti)*, would be significantly positive, as *Malta1* would share more alleles with Villabruna Cluster samples than with Early Upper Palaeolithic Europeans. However, we do not detect any such a signal (Supplementary Information section 9).

Regardless of whether a population closely related to *Satsurblia* is responsible for the affinity of Villabruna Cluster samples to the Near East, there is evidence that a new lineage with affinities to present-day Near Easterners spread across Europe at this time. The evidence for this spread is that the genetic affinity of pre-Neolithic Europeans to Near Easterners abruptly increases with the appearance Villabruna Cluster, with no earlier European sample showing as strong an affinity despite sharing large amounts of genetic drift with the Villabruna Cluster (Figure 4b). An important direction for future work is to analyse more ancient DNA samples from southeastern Europe and western Asia in order to understand the history of the migratory events that our data show must have occurred around this time.

References

1 Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature communications* **6**, 8912, doi:10.1038/ncomms9912 (2015).

2 Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413, doi:10.1038/nature13673 (2014).

Section 13

Population structure in the Villabruna Cluster

Villabruna Cluster samples harbor variable proportions of at least 3 ancestral lineages

We co-analysed Villabruna Cluster samples with earlier European hunter-gatherers and present-day populations. The goal of this analysis was to begin to understand the complexity of the relationship of the Villabruna Cluster samples to other samples.

We used *qpWave*[1] to compute f_4-statistic vectors of the form:

$$f_4(Left_h, Left_i; Right_j, Right_k) \quad \text{(Equation S13.1)}$$

Left (Villabruna Cluster samples with >0.1x coverage):
Villabruna, Bichon, Rochedane, Ranchot88, Loschbour, LaBrana1, Hungarian.KO1

Right (pre-Villabruna Cluster samples at >0.1x, and present-day humans):
UstIshim, GoyetQ116-1, Kostenki14, Malta1, Vestonice16, ElMiron, Stuttgart, Ami, Bougainville, Chukchi, Eskimo_Sireniki, Eskimo_Naukan, French, Han, Ju_hoan_North, Karitiana, Kharia, Mbuti, Onge, Papuan, She, Ulchi, Yoruba, Iraqi_Jew.

The papers that introduced *qpWave* showed that if the *Left* populations are mixtures – in various proportions – of *N* sources differently related to the *Right* populations, then all f_4-statistics of the form of Equation S13.1 will be consistent with being a linear combination of *N* vectors of statistics that correspond to the mixing populations, and the matrix will have rank N-1[1,2]. We can test whether this is the case using a Hotelling's T^2 test[1].

Applying this test to our data, we reject rank 0 at high statistical significance ($P=1.1\times10^{-13}$), and we also reject rank 1 (P=0.026). Rank 2 is consistent with the data to within the limits of our resolution (P=0.61). Thus, the Villabruna Cluster populations derive from a mixture of at least three ancestral populations that are differentially related to the *Right* set.

We obtained consistent results when we removed *Rochedane* with its modest coverage (its removal brought the total number of SNPs with coverage in all samples from 36,372 to 114,473). Again, we reject rank 0 at high significance ($P=2.4\times10^{-21}$), and we also reject rank 1 (P=0.020). Rank 2 is consistent with the data (P=0.41). This again supports the model of few as three ancestral populations, although the truth could of course be more.

Hints about the ancestral populations that contributed to the Villabruna Cluster

The *qpWave* analysis is able to document that at least three sources are necessary to account for the allele frequency correlation patterns in Villabruna Cluster samples. However, we did not succeed at convincingly modeling the ancient relationships among these sources. We nevertheless found some hints about the history relating these ancient samples, and we discuss several relevant observations.

The common thread that binds together Villabruna Cluster samples

Table S13.1 documents an attraction of the *Villabruna* genome to all other samples in the cluster, especially after ~14,000 years ago. This reflects the common thread that binds together the Villabruna Cluster samples, and is the reason for the designation of these samples as a cluster with at least one shared source of distinctive ancestry.

Table S13.1. Z-score of *D(X, Y; Villabruna, Mbuti)*. *Villabruna* shares more drift with Villabruna Cluster samples post-dating it in Europe than with European samples pre-dating it.

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	HohleFels49	Rochedane	Ranchot88	Loschbour	LaBrana1	Hungarian.KO1
Han	NA	0.4	3.3	-13.2	-14	-18.7	-18.8	-26.9	-11.3	-38	-39.9	-45.5	-36.6	-39.8
UstIshim	-0.4	NA	2.5	-11.6	-12	-16.1	-15.4	-23.7	-10.5	-34.1	-35.5	-39.8	-31.6	-32.5
Oase1	-3.3	-2.5	NA	-11.5	-11.1	-14.4	-12.9	-20.4	-5.5	-22	-28.3	-34.9	-27.5	-28.3
Kostenki14	13.2	11.6	11.5	NA	-1.8	-5.3	-6	-14.4	-4.4	-23.2	-25.8	-29.5	-21.4	-24.2
GoyetQ116-1	14	12	11.1	1.8	NA	-3.4	-4.3	-12.4	-3.7	-23.6	-24.8	-29.4	-19.4	-22.4
Vestonice16	18.7	16.1	14.4	5.3	3.4	NA	-1.8	-9.2	-1.8	-19.7	-22	-25.1	-16.7	-18.5
Ostuni1	18.8	15.4	12.9	6	4.3	1.8	NA	-7.6	-0.3	-16.1	-17.2	-21.4	-12.4	-14.9
ElMiron	26.9	23.7	20.4	14.4	12.4	9.2	7.6	NA	4.3	-11.4	-13	-15.1	-6.6	-9.3
HohleFels49	11.3	10.5	5.5	4.4	3.7	1.8	0.3	-4.3	NA	-5.6	-9.8	-14	-6.4	-9
Bichon	41.7	37.3	32.0	27.7	24.9	21.8	18.9	12.8	11.0	-0.7	-0.4		7.5	3.9
Rochedane	38	34.1	22	23.2	23.6	19.7	16.1	11.4	5.6	NA	0.4	-0.1	7.8	4.5
Ranchot88	39.9	35.5	28.3	25.8	24.8	22	17.2	13	9.8	-0.4	NA	-1	8.2	3.9
Loschbour	45.5	39.8	34.9	29.5	29.4	25.1	21.4	15.1	14	0.1	1	NA	9.1	5.9
LaBrana1	36.6	31.6	27.5	21.4	19.4	16.7	12.4	6.6	6.4	-7.8	-8.2	-9.1	NA	-3
Hungarian.KO1	39.8	32.5	28.3	24.2	22.4	18.5	14.9	9.3	9	-4.5	-3.9	-5.9	3	NA

Variable proportions of GoyetQ116-1-related ancestry in Villabruna Cluster samples

Despite their shared thread of ancestry, a different set of *D*-statistics show directly that Villabruna Cluster samples do not form clade with respect to all later European hunter-gatherers. In particular, statistics of the form *D(Villabruna, X; Y, Mbuti)* show that El Mirón cluster samples and *GoyetQ116-1* share more alleles with a subset of Villabruna Cluster samples (*Rochedane, Ranchot88, Loschbour and LaBrana1*) than they do with *Villabruna* itself (Table S13.2). These results document a variable fraction of ancestry from the European *GoyetQ116-1* lineage within the Villabruna Cluster, as is also shown in the fitted Admixture Graphs in Supplementary Information section 6, and the variable mixture proportion estimates in Supplementary Information section 7.

We observed that *Hungarian.KO1* does not show the same affinity to *ElMiron* and *GoyetQ116-1* as do the other Villabruna Cluster (Table S13.2). We were concerned that these results might be an artifact of the fact that the libraries that were used to build *Hungarian.KO1* were not UDG-treated. However, this does not appear to explain the observations, as the pattern replicates when restricting to transversion SNPs that are not affected by ancient DNA damage (Table S13.3). These results suggest the possibility that *Hungarian.KO1* may have less *ElMiron* and *GoyetQ116-1* affinity than does *Villabruna.*

Table S13.2. Z-score of *D(Villabruna, X; Y, Mbuti)* for all sites.

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	HohleFels49	Rochedane	Ranchot88	Loschbour	LaBrana1	Hungarian.KO1
Bichon	-0.4	1	0.1	-1.3	0.9	1.1	-2.3	-2.4	-2.7	-1.1	-3	-3.2	0.7	-3
Rochedane	-0.8	0.5	-0.7	0.6	-1.7	0.6	-0.2	-3.1	-1	NA	-1.1	-0.6	0.1	2.9
Ranchot88	-1.2	1.1	-0.5	-0.1	-3	0.6	-0.3	-3.2	-3.6	-0.8	NA	-5.2	-3.6	0.4
Loschbour	-3.9	-0.5	1.7	-0.3	-3.7	-0.6	1.7	-4.1	-3.8	-0.6	-6.1	NA	-5.7	-1.4
LaBrana1	-2.9	0.6	1.3	0.9	-4.1	2.9	2.6	-4.1	-3	7.9	5	4.2	NA	6.4
Hungarian.KO1	-3.3	2.2	0.5	1.6	0.6	2.1	3.4	5	1.1	7.8	4.5	4.6	3.4	NA

Gene flow between relatives of Han and relatives of some Villabruna Cluster samples

Table S13.4 shows that *Han* shares more alleles with a subset of Villabruna Cluster samples (*Bichon, Loschbour, La Brana1, Hungarian.KO1*) than it does with another subset of Villabruna Cluster samples (including *Villabruna* itself) as well as non-Villabruna Cluster Europeans. The variability in the Villabruna Cluster in their degree of allele sharing with East Asians documents another complexity of the history of the Villabruna Cluster (Figure 4b).

Table S13.3. *Z*-score of *D(Villabruna, X; Y, Mbuti)* for transversion sites.

X/Y	Han	UstIshim	Oase1	Kostenki14	GoyetQ116-1	Vestonice16	Ostuni1	ElMiron	HohleFels49	Rochedane	Ranchot88	Loschbour	LaBrana	Hungarian_KO1
Bichon	-2.7	-0.6	0.8	0.4	-1.2	1.5	1.7	-1.7	-1.2	-2	0.7	-2.3	-2.3	0.1
Rochedane	-0.5	0	-0.7	-0.8	-0.8	-0.8	0.4	-3.2	-1.4	NA	-0.8	-1.1	0	0.6
Ranchot88	-1.2	0.8	-0.8	-1.1	-2	-0.4	-0.6	-2.9	-1.8	-1.1	NA	-3.3	-3.1	-0.5
Loschbour	-3.9	-0.7	1.8	0.9	-2.4	0.6	2.7	-3.2	-1.7	0.8	-2.6	NA	-4.7	-0.8
LaBrana1	-2.6	0	1.7	1.2	-2.9	2.7	2.7	-3.1	-1.1	5.7	5.1	3.5	NA	5.7
Hungarian_KO1	-3.8	1.5	0.4	1.4	-0.3	1.5	2.6	3.3	0.4	3.4	3.7	3.7	2	NA

Table S13.4. *D(X, Y; Han, Mbuti)* for all sites. Han Chinese share more alleles with a subset of Villabruna Cluster samples (Bichon, Loschbour, LaBrana1, and Hungarian.K01) than they do with early European hunter-gatherers.

X/Y	Villabruna	Loschbour	LaBrana1	Hungarian.KO1	Malta1
UstIshim	0.2	-2.9	-2.2	-2.4	-3.6
Oase1	-3.7	-6.1	-5.2	-6.1	-7.1
Kostenki14	-1.3	-4.3	-3.8	-4	-5.3
Kostenki14.sg	-1.3	-4.6	-3.9	-4.4	-5.1
GoyetQ116-1	1.5	-1.1	-0.5	-1.2	-2.6
Vestonice16	-0.9	-3.9	-3.4	-3.4	-5
Ostuni1	1	-2	-1.9	-1.9	-3.2
ElMiron	1.5	-2	-1.3	-1.8	-3.4
HohleFels49	0	-2.5	-0.8	-1.8	-2
Villabruna	NA	-3.9	-2.9	-3.3	-4.2
Bichon	3	-1.1	0	-0.5	-1.9
Rochedane	0.8	-2	-0.9	-1.1	-3.4
Ranchot88	1.2	-1.9	-1.4	-2.1	-3.5
Loschbour	3.9	NA	1	0.5	-0.9
LaBrana1	2.9	-1	NA	-0.3	-1.9
Hungarian.KO1	3.3	-0.5	0.3	NA	-1.4
Malta1	4.2	0.9	1.9	1.4	NA

Summary

We have documented complexity in the ancestry of Villabruna Cluster samples, with variability in the degree of allele sharing with *ElMiron/GoyetQ116-1*, and independently with East Asians. We are confident that at least three sources of ancestry contribute to Villabruna Cluster samples, but at present we do not have a good model for the mixture.

References

1 Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488**, 370-374, doi:10.1038/nature11258 (2012).

2 Patterson, N. J. *et al.* Ancient Admixture in Human History. *Genetics*, doi:10.1534/genetics.112.145037 (2012).