Genomic Structure in Europeans dating back at least 36,200 years

Andaine Seguin-Orlando[1]*, Thorfinn S. Korneliussen[1]*, Martin Sikora[1]*, Anna-Sapfo Malaspinas[1], Andrea Manica[2], Ida Moltke[3,4], Anders Albrechtsen[4], Amy Ko[5], Ashot Margaryan[1], Vyacheslav Moiseyev[6], Ted Goebel[7], Michael Westaway[8], David Lambert[8], Valeri Khartanovich[6], Jeffrey D. Wall[9], Philip R. Nigst[10,11], Robert A. Foley[1,12], Marta Mirazon Lahr[1,12]†, Rasmus Nielsen[5]†, Ludovic Orlando[1], Eske Willerslev[1]†

[1]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen, Denmark

[2]Department of Zoology, University of Cambridge, Downing St, Cambridge, CB2 3EJ, UK

[3] Department of Human Genetics, University of Chicago, 920 E. 58th Street, CLSC, Chicago, IL 60637

[4]The Bioinformatics Center, University of Copenhagen, Ole Maaløes Vej 5, 2200 København N, Denmark

[5]Department of Integrative Biology, University of California, Berkeley, California 94720, USA

[6]Department of Physical Anthropology, Kunstkamera, Peter the Great Museum of Anthropology and Etnography, Russian Academy of Sciences, 24 Srednii prospect, Vassilievskii Island, St. Peterburg, Russia.

[7]Center for the Study of the First Americans and Department of Anthropology, Texas A&M University, TAMU-4352, College Station, Texas 77845-4352, USA

[8]Griffith University, Nathan Campus, 170 Kessels Road, Nathan, Brisbane, Queensland 4111, Australia

[9]Department of Epidemiology and Biostatistics, University of California San Francisco, 185 Berry Street, Lobby 5, Suite 5700, San Francisco, CA 94107, USA

[10]Division of Archaeology, University of Cambridge, Cambridge, Downing Street, CB2 3DZ, UK

[11]Department of Human Evolution, Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Deutscher Platz 6, D-04103, Germany

[12] Leverhulme Centre for Human Evolutionary Studies, Department of Archaeology & Anthropology, University of Cambridge, Cambridge, Fitzwilliam Street, CB2 1QH, UK

*These authors contributed equally to the work and should be regarded as joint first authors.

[†] Corresponding authors: ewillerslev@snm.ku.dk, rasmus_nielsen@berkeley.edu, mbml1@cam.ac.uk

Abstract

The origin of contemporary Europeans remains contentious. We obtain a genome sequence from Kostenki 14 in European Russia dating to 38,700-36,200 years ago, one of the oldest fossils of Anatomically Modern Humans from Europe. We find that K14 shares a close ancestry with the 24,000 year old Mal'ta boy from central Siberia, European Mesolithic hunter-gatherers, some contemporary western Siberians and many Europeans, but not eastern Asians. Additionally, the Kostenki 14 genome shows evidence of shared ancestry with a population basal to all Eurasians that also relates to later European Neolithic

farmers. We find that Kostenki 14 contains more Neandertal DNA that is contained in longer tracts than present Europeans. Our findings reveal the timing of divergence of western Eurasians and East Asians to be >36,200 years ago, and that European genomic structure today dates back to the Upper Paleolithic and derives from a meta-population that at times stretched from Europe to central Asia.

One Sentence Summary:

The genome of Kostenki 14, a ~37 ka-old modern human from European Russia, reveals that European genomic structure dates back to the Upper Paleolithic.

Main Text:

The ancestors of contemporary Eurasians are believed to have left Africa some 60,000-50,000 years ago (60-50 ka) (*1*, *2*), possibly 30-40 ka later than Australo-Melanesian ancestors (*3*). Despite controversies about routes out of Africa, the first Upper Paleolithic (UP) industries of Eurasia are found in the Levant from *c.* 48 ka (*4*, *5*). Expansion into Europe took place through multiple events that by *c.* 40 ka had generated a spatially and culturally structured Anatomically Modern Human (AMH) population – from Russia (*6*), to Georgia (*7*), Bulgaria (*8*), southern Europe (*9*, *10*) and the UK (*11*). The few AMH fossils associated with these initial UP industries are morphologically variable (*9*, *12–17*). In western Eurasia, the distinctive Aurignacian toolkit, first observed at Willendorf (Austria) by 43.5 ka (*18*), becomes predominant across the earlier range by 39 ka. Although analyses of ancient human genomes have advanced our understanding of the European past, revealing contributions from Paleolithic Siberians, European Mesolithic and Near Eastern Neolithic groups to the European gene pool (*19–23*), the possible contribution of the earliest Eurasians to these later cultures and to contemporary human populations remains unknown. To investigate this, we sequenced the genome of Kostenki 14 (K14, Markina Gora, Figure 1A).

The locality of Kostenki-Borshchevo on the Middle Don River, Russia, has one of the most extensive Paleolithic records in eastern Europe. The K14 human skeleton was excavated in 1954 (*24*) and recently dated to 33,250 ± 500 radiocarbon years BP (*25*), 38.7-36.2 thousand calendar years BP (ka cal BP), in agreement with the stratigraphic position of the burial that cuts into the Campanian Ignimbrite ash layer dated to *c.* 39.3 ka cal BP (*26*). Below the skeleton there is a

distinctive early UP industry, with end scrapers, burins, prismatic cores and bone artifacts (Layer IV); the cultural layer above (Layer III) has a regionally local character (*27*, *28*) (SOM S1-S2).

We performed 13 DNA extractions from a total of 1.285 grams of the left tibia (dorsal side of the shaft), using two extraction methods based on silica purification (*29*, *30*). We first constructed 7 Illumina libraries and validated the presence of typical signatures of post-mortem DNA damage, using a fraction of DNA extracts (SOM S3). The remaining extracts were built into 63 libraries following enzymatic USER treatment to limit the impact of nucleotide mis-incorporations in downstream analyses (*31*) (SOM Table S2). Additionally, a limited fraction of two DNA extracts was purified for methylated DNA fragments using Methyl Binding Domain (MBD)-enrichment (*32*) before USER treatment and library building, for a total of 8 DNA libraries. Following stringent quality criteria for read alignment, we identified a total number of 175.2 million unique reads aligning against the human reference genome hg19, representing an average depth-of-coverage of 2.84X (SOM S4). The eight USER treated DNA libraries that exhibited limited error rate and contamination levels were selected for further analyses. This restricted the dataset to 148.9 million unique reads, representing a final depth-of-coverage of 2.42X. We exploited the fact that K14 was a male and used the heterozygosity levels present in the X chromosome to estimate overall levels of contamination around 2.0% (SOM S5-S6;Table S5). Note that the population genetics analyses results are robust to contamination of that level. In particular we replicated the main analyses with selected libraries with varying contamination levels and observed no qualitative effect on the results (see SOM S9 for details).

Mitochondrial analyses confirmed the sequence previously reported for K14 (haplogroup U2, (*33*)), which supports data authenticity. The Y chromosome belongs to haplogroup C M130, the same as in La Braña – a late Mesolithic hunter-gatherer (MHG) from northern Spain (*22*) (SOM S7).

To identify patterns of shared ancestry and admixture among K14, other ancient genomes and contemporary Eurasians (based on a single-nucleotide polymorphism (SNP) array panel of 2091 individuals from 167 populations), we carried out a series of analyses – model-based clustering and principal component analysis (PCA) - to show the contribution of diverse genetic components within K14; D-statistics to explore the affinity of K14 to pairs of populations (using Mbuti Pygmy as an outgroup); f_4 statistics to test whether a given modern population is equidistant to an ancient individual and a particular recent group (here Sardinians), given an outgroup (here Papuans); and f_3 statistics to explore both patterns of admixture ("admixture" f_3) and shared ancestry ("outgroup" f_3). Key results were also replicated using two whole-genome sequencing datasets of modern individuals from worldwide populations (*23, 34*).

Model-based clustering analyses (*35*) show that K14 has different genetic components of substantial size (Fig. 1B, SOM S10), suggesting the sharing of sets of alleles with different Eurasian groups. The largest fraction of K14's ancestry derives from a component that is maximized in European MHGs, and also predominant in contemporary northern and eastern Europeans. The genetic affinity of K14 to contemporary Europeans is also observed using "outgroup" f_3-statistics (*36*). Using Mbuti Pygmy as outgroup, we find that among a panel of 167 contemporary populations, Europeans have the greatest affinity (i.e. the largest f_3) to K14 (Figure

1C). This conclusion is also formally supported by comparing pairs of populations to K14 using the D statistics of the form D(Mbuti Pygmy, K14; Population 1, Population 2). This statistic is expected to be equal to zero if K14 is symmetrically related to Population 1 and Population 2, whereas its expectation is negative (positive) if K14 is more closely related to Population 1 (Population 2). For pairs of populations involving East Asians (Population 1) and Europeans (Population 2), K14 is always significantly closer related to Europeans (e.g. Z = 12.1, (Han, Lithuanians)), in all datasets analyzed (SOM S9;Table S7). We also confirm that these results are robust to possible contamination from a modern DNA source, by filtering for reads with a high likelihood of ancient DNA using a model-based approach (37) as well as calculating contamination-corrected D-statistics (23)(SOM S9;Figure S18).

Within Europe, northern Europeans show the closest affinity to K14, based both on the f_3 (Figure 1D) and D-statistics (e.g., Z = 6.7, for (Sardinians, Lithuanians);Table S7;Figure S16). This pattern closely resembles that of European MHGs (La Braña, Ajv58, Loschbour, Motala) and Mal'ta (MA1) (Figure S14-S15), with the exception of the latter's strong genetic affinity with Native Americans, which is unique to that individual. Furthermore, a direct comparison to ancient genomes in the "outgroup" f_3 statistics shows K14 has a higher affinity with MHGs (Loschbour, La Braña) than any other ancient individual or contemporary population (Figure S14). Together with the rare Y chromosome lineage shared with La Braña, these results provide strong evidence of shared ancestry and extensive gene flow between UP West Eurasian people related to K14, and European MHGs and their contemporary European descendants.

An interpretation of the above results would be that K14 is an early member of a lineage leading to western Eurasian MHGs, after their split from the proposed ancestral northern Eurasian lineage including MA1. However, D-statistics of the form D(Mbuti Pygmy, Modern; Ancient, K14), which test whether K14 and an ancient individual form a clade with respect to a modern population, reject this simple tree-like relationship. We find that all contemporary non-Africans, except Australo-Melanesians, are closer to either Mal'ta (MA1) or MHGs than to K14 (e.g., $Z =$ -5.3, for D(Mbuti, Han; Loschbour, K14); SOM S9;Table S10;Figure S19). This would suggest a basal position of K14 with respect to MHGs and ancient north Eurasians, which is also shown in admixture graphs using TreeMix (SOM S12;Figure S24-S25). In addition, a sizeable component of K14's ancestry observed in the model-based clustering analyses is predominant in contemporary Middle Eastern/Caucasus (ME/C) populations and Neolithic ancient genomes (NEOL) (Gok2, Iceman, Stuttgart), but absent in MA1 or MHGs (Figure 1B;Figure S20). This component has been associated with a suggested "basal Eurasian" lineage contributing to NEOL, to explain an observed increase in allele sharing between MHGs / MA1 and East Asians compared to NEOL (*21*). Since K14 shows the same pattern as NEOL, a parsimonious explanation would be that K14 also derives some ancestry from a related "basal Eurasian" lineage. Consistent with this hypothesis, we find that East Asians are equally distant to NEOL and K14 using D-statistics as described above (e.g., $Z = 0.0$, for D(Mbuti, Han; Stuttgart, K14); Table S10-S11). This suggests that the main ancestral components proposed for contemporary Europeans, including the Middle-Eastern component commonly attributed to the expansion of early farmers within Europe, were likely already genetically differentiated and related through complex gene flow by the time of K14, at least 36.2 ka ago (Figure 2).

We further investigated the relationship of K14 and the other ancient genomes to East Asian and Siberian populations using f_4 statistics f_4(Sardinian, Ancient; Modern, Papuan), which measure whether a modern population shares more alleles with contemporary Europeans or an ancient genome. We find that all Siberian and East Asians are equally distant from western MHGs (all $|Z| < 1.9$; Figure 3D; Table S12), supporting the postulated early split between East Asians and western Eurasians. In contrast to MHGs and MA1 all Siberian populations are genetically closer to contemporary Europeans (Sardinians) than to K14 ($3.1 < |Z| < 9.9$; Table S12), particularly those from the Yenisei and Ob' basins (e.g. Shors, $Z = 8.0$) (Figure 3A). Furthermore, these populations derive parts of their ancestry from a European "hunter-gatherer" (HG) component inferred in the ADMIXTURE analysis (Figure 1D; Figure S20), with populations showing higher "HG" ancestry proportion also being closer to contemporary Europeans using the f_4 statistic (Spearman $\rho = 0.96$, $p = 3.0 \times 10^{-18}$, Figure 3D; Table S13). Notably, the opposite pattern is observed with Scandinavian MHGs (Ajv58, Motala), where the same populations tend to share more alleles with MHGs than contemporary Europeans and the "HG" component is negatively correlated with f_4 (e.g. Motala $\rho = -0.85$, $p = 6.2 \times 10^{-10}$; Figure 3C, 3D). Calculating "admixture" f_3 statistics, we find significant evidence for admixture in those populations, with a variety of Siberian and European source populations. The best pair of source populations (i.e., the most negative f_3 statistic) involves Swedish MHGs (Motala) and Evens (a northeast Siberian population) (e.g. f_3(Shors; Evens, Motala) = -0.012, $Z = -9.1$)(Table S14). Altogether, these results suggest that contemporary Siberian populations from the Yenisei basin derive part of their gene pool from a Eurasian HG population that shares ancestry with K14, but is more closely related to Scandinavian MHGs than to either MA1 or western European MHGs, indicating gene

flow between their ancestors and Scandinavian Europe after K14 but prior to the Mesolithic

(36.3 > x > 7 ka BP).

Finally, we estimated levels of Neandertal ancestry in K14 using f_4-ratio statistics (*38*). Our

estimates are consistent with previous analyses (*34*) showing a Neandertal contribution lower

than 2 % for most individuals (Figure 4A). However, both La Braña and K14 show slightly

elevated levels, with an estimated 2.4 ± 0.4% in K14 (Table S15-S16). Restricting this analysis

to genomic regions without evidence for Neandertal introgressed haplotypes in contemporary

humans (*38*, *39*) results in 0% estimated ancestry for most individuals except K14, where 0.9 ±

0.4% Neandertal ancestry is still detected (Table S17-S18). The difference between K14 and

modern genomes could be caused by several factors including sampling effects and genetic drift,

natural selection as argued in (*38*, *39*), or by the effects of additional Neandertal admixture not

represented in the modern gene-pool. We next compared the size distribution of genomic tracts

of archaic hominin origin in K14 and other ancient individuals (Figure 4B), by identifying

genomic regions with high frequencies of archaic alleles at sites where all modern Africans carry

the ancestral allele. The length of Neandertal tracts was higher in K14 than in other ancient

individuals, with the longest tract totaling ~3Mb on chromosome 6 (Figure 4C). This is

consistent with K14 being closer to the time of the admixture event with Neandertals, and

carrying longer archaic tracts that have been affected by less recombination, than in the other

~11-30,000 year old younger ancient genomes. We then used the length distribution of shared

ancestry to estimate the admixture time of Neandertals and humans based on the K14 sample,

and obtained an estimate of approximately 54K years (S15). We note that genomic data from a

45,000-year-old modern human from Siberia, which was published during the review process of

this study, also shows longer segments of Neanderthal ancestry, further supporting our conclusions (*40*). Because of the divergent position of the K14 sample, we also examined if it contained any fragments of introgressed DNA from other previously un-sampled hominins. However, the distribution of tracts of divergent DNA provides no evidence for additional divergent introgressed DNA (S14).

Several studies have reported on the basal genetic distinctiveness between western Eurasian and eastern Asian populations, as well as between all Eurasians and Australo-Melanesians (*41–43*). Our results show no close genetic relationship between K14 and Australo-Melanesians, and support earlier studies that suggest Australo-Melanesians derive part of their ancestry from an early population divergence that pre-dates the separation of Europeans and East Asians (*3*). The K14 genome shows that this early UP individual was clearly part of a western Eurasian lineage that had already diverged from eastern Asians, thus establishing a minimum date for that separation at least 36.2 ka. The fact that the limited genomic information on the *c*. 40 ka Tianyuan modern human from China clusters with contemporary East Asian populations (*44*) suggests an even earlier date.

Our results further suggest that the early stages of the western Eurasian lineage were already complex (see also Figure 2). Besides its core affinities with subsequent European groups, K14 also shares alleles with European Neolithic farmers and contemporary people from the Middle East/Caucasus, which are not found in MA1 and western European MHGs, indicating genetic exchange between K14 and a Basal Eurasian Lineage (which eventually contributed to Neolithic groups) after the ancestors of MA1 and subsequent European MHGs had diverged. This implies

that early AMH populations became structured early in their history, but already in the UP contained the major genetic components found in Europeans today. As such our findings show the existence of a meta-population structure in Europe from the Upper Paleolithic onwards, remnants of which are still found today, despite migrations to and from Europe since the UP. The early UP contribution is greater among northern than southern Europeans, in agreement with the southeast to west and north gene flow cline resulting from the expansion of Neolithic famers 9-6 ka cal BP (*20*, *45*). However, descendants of the early UP population represented by K14 likely also contributed genes to western Siberian groups living around the mouth of the Yenisei River. Therefore, our findings support the view that these Uralic-speaking populations represent an ancient admixture between European and East Asian lineages. The recently proposed Holocene gene flow from East Asians into northern Europeans (*21*) can, in our view, be equally well explained by population structure of the hunter-gatherer meta-population within Europe. As such our results paint an increasingly complex picture of colonization history of Europe from the UP to today. Instead of inferring a few discrete migration events from Asia into Europe, we now see evidence that humans in Western Eurasia formed a large meta-population with gene flow in multiple directions occurring repeatedly and perhaps continuously.

References and Notes:

1. R. G. Klein, Archeology and the evolution of human behavior. *Evol. Anthropol.* **9**, 17–36 (2000).

2. P. Mellars, A new radiocarbon revolution and the dispersal of modern humans in Eurasia. *Nature*. **439**, 931–935 (2006).

3. M. Rasmussen *et al.*, An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia. *Science*. **334**, 94–98 (2011).

4. A. E. Marks, *Prehistory and Paleoenvironments in the Central Negev, Israel.* (Southern Methodist University Press, Dallas, 1983).

5. N. R. Rebollo *et al.*, New radiocarbon dating of the transition from the Middle to the Upper Paleolithic in Kebara Cave, Israel. *J. Archaeol. Sci.* **38**, 2424–2433 (2011).

6. A. Gibbons, Oldest Homo sapiens genome pinpoints neandertal input. *Science*. **343**, 1417 (2014).

7. O. Bar-Yosef, A. Belfer-Cohen, D. S. Adler, The implications of the Middle-Upper Paleolithic chronological boundary in the Caucasus to Eurasian prehistory. *Anthropologie*. **XLIV**, 49–60 (2006).

8. N. Sirakov *et al.*, Un nouveau faciès lamellaire du début du Paléolithique supérieur dans les Balkans. *Paléorient*, 131–144 (2007).

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

10. S. L. Kuhn *et al.*, The early Upper Paleolithic occupations at Uçağizli Cave (Hatay, Turkey). *J. Hum. Evol.* **56**, 87–113 (2009).

11. T. Higham *et al.*, The earliest evidence for anatomically modern humans in northwestern Europe. *Nature*. **479**, 521–524 (2011).

12. S. E. Bailey, T. D. Weaver, J.-J. Hublin, Who made the Aurignacian and other early Upper Paleolithic industries? *J. Hum. Evol.* **57**, 11–26 (2009).

13. C. A. Bergman, C. B. Stringer, Fifty years after: Egbert, an early Upper Palaeolithic juvenile from Ksar Akil, Lebanon. *Paléorient*. **15**, 99–111 (1989).

14. V. Formicola, Early Aurignatian Deciduous Incisor from Riparo Bombrini at Balzi Rossi (Grimaldi, Italy). *Riv. Antropol.* **67**, 287–292 (1989).

15. D. Henry-Gambier, C. Normand, J.-M. Pétillon, Datation radiocarbone directe et attribution culturelle des vestiges humains paléolithiques de la grotte d'Isturitz (Pyrénées-Atlantiques). *Bull. Société Préhistorique Fr.* **110**, 645–656 (2013).

16. J.-J. Hublin, The earliest modern human colonization of Europe. *Proc. Natl. Acad. Sci. USA*. **109**, 13471–13472 (2012).

17. C. Yazbeck, Le Paleolithique du Liban: Bilan critique. *Paléorient*. **30**, 111–126 (2004).

18. P. R. Nigst *et al.*, Early modern human settlement of Europe north of the Alps occurred 43,500 years ago in a cold steppe-type environment. *Proc. Natl. Acad. Sci.* **111**, 14394–14399 (2014).

19. P. Skoglund *et al.*, Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. *Science*. **336**, 466–469 (2012).

20. P. Skoglund *et al.*, Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*. **344**, 747–750 (2014).

21. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. **513**, 409–413 (2014).

22. I. Olalde *et al.*, Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature*. **507**, 225–228 (2014).

23. M. Raghavan *et al.*, Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*. **505**, 87–91 (2014).

24. A. N. Rogachev, *Aleksandrovskoe poselenie drevnekamennogo veka u sela Kostenki na Donu* (Moscow-Leningrad, 1955).

25. A. Marom, J. S. O. McCullagh, T. F. G. Higham, A. A. Sinitsyn, R. E. M. Hedges, Single amino acid radiocarbon dating of Upper Paleolithic modern humans. *Proc. Natl. Acad. Sci. USA*. **109**, 6878–6881 (2012).

26. D. M. Pyle *et al.*, Wide dispersal and deposition of distal tephra during the Pleistocene `Campanian Ignimbrite/Y5' eruption, Italy. *Quat. Sci. Rev.* **25**, 2713–2728 (2006).

27. A. A. Sinitsyn, J. F. Hoffecker, Radiocarbon dating and chronology of the Early Upper Paleolithic at Kostenki. *Quat. Int.* **152-153**, 164–174 (2006).

28. Vishnyatsky, L.B., Nehorosheve, P.E., in *The beginning of the Upper Paleolithic on the Russian Plain.* (University of California Press, Berkeley, CA, Brantingham PJ, Kuhn SL, Kerry KW., 2004), pp. 80–96.

29. N. Rohland, M. Hofreiter, Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2**, 1756–1762 (2007).

30. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci.* **110**, 15758–15763 (2013).

31. A. W. Briggs *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci.* **104**, 14616–14621 (2007).

32. G. R. Feehery *et al.*, A Method for Selectively Enriching Microbial DNA from Contaminating Vertebrate Host DNA. *PLoS ONE*. **8**, e76096 (2013).

33. J. Krause *et al.*, The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*. **464**, 894–897 (2010).

34. K. Prufer *et al.*, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. **505**, 43–49 (2014).

35. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

36. N. J. Patterson *et al.*, Ancient Admixture in Human History. *Genetics*, genetics.112.145037 (2012).

37. P. Skoglund *et al.*, Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl. Acad. Sci.* **111**, 2229–2234 (2014).

38. S. Sankararaman *et al.*, The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. **507**, 354–357 (2014).

39. B. Vernot, J. M. Akey, Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. *Science*. **343**, 1017–1021 (2014).

40. Q. Fu *et al.*, Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*. **514**, 445–449 (2014).

41. W. W. Howells, *Skull Shapes and the Map: Craniometric Analyses in the Dispersion of Modern Homo.* (Harvard University Press, Cambridge, 1989), *Peabody Museum of Archaeology and Ethnology*.

42. R. A. Foley, M. M. Lahr, Mode 3 technologies and the evolution of modern humans. *Camb. Archaeol. J.* **7**, 3–36 (1997).

43. N. A. Rosenberg *et al.*, Genetic structure of human populations. *Science*. **298**, 2381–2385 (2002).

44. Q. Fu *et al.*, DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci.* **110**, 2223–2227 (2013).

45. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes:* (Princeton University Press, Princeton, N.J., Abridged edition., 1996).
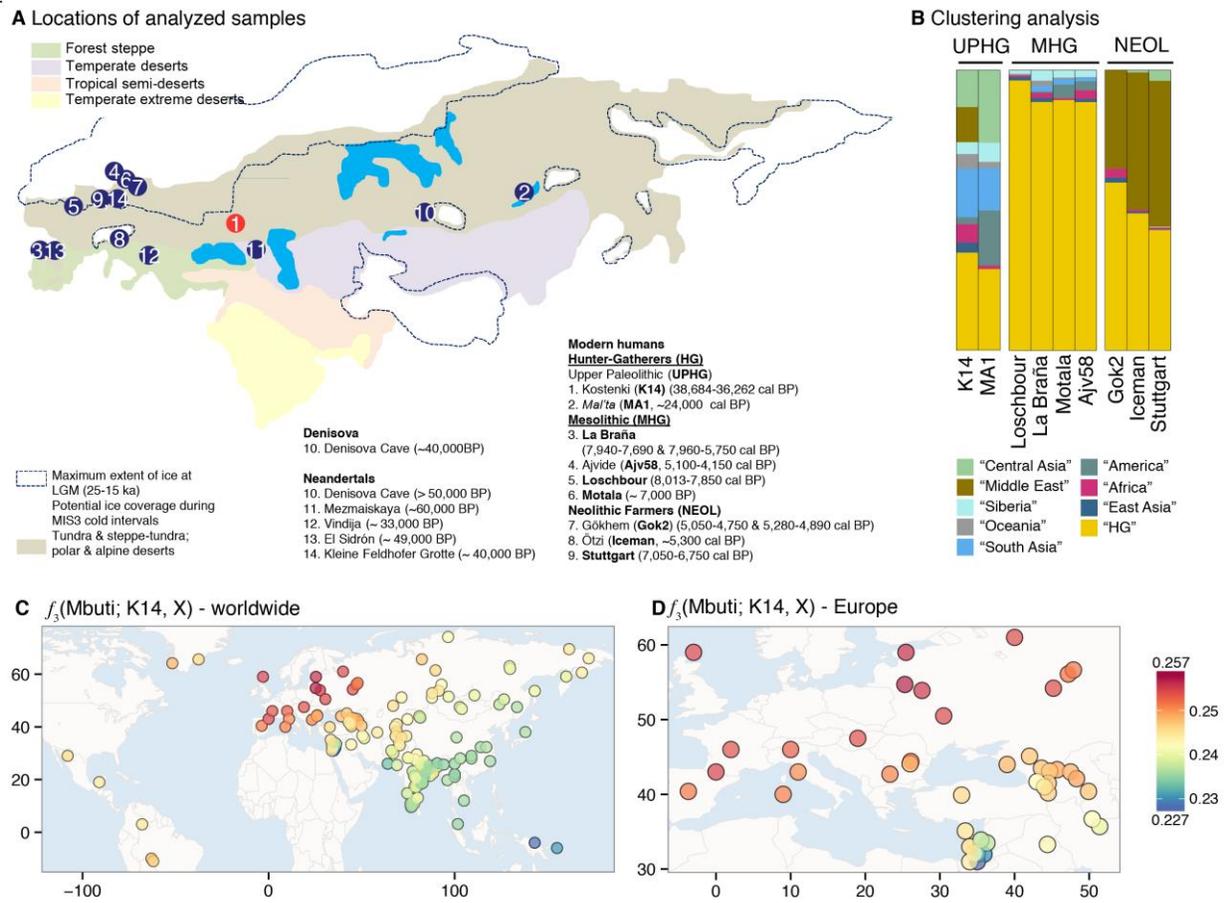
Figure 1. (A) Location of Kostenki and the samples analyzed in this study. Kostenki (K14) is shown in red while comparative ancient samples are shown in blue. **(B)** Admixture proportions for the ancient genomes assuming nine ancestral components for a clustering analysis in a set of modern worldwide populations. We labeled the components according to the modern populations in which they are maximized for all but one case: the yellow component that we label HG is maximized in eastern Europeans. UP: Upper Paleolithic, M: Mesolithic, HG: Hunter-Gatherers, NEOL: Neolithic Farmers. **(C)** Shared drift between K14 and a set of worldwide populations. For every modern population X on the map, we compute f_3(Mbuti Pygmy; K14, X). The warmer colors indicate increased shared ancestry. **(D)** Shared drift between K14 and a set of European populations. This figure is a zoom of panel (C).
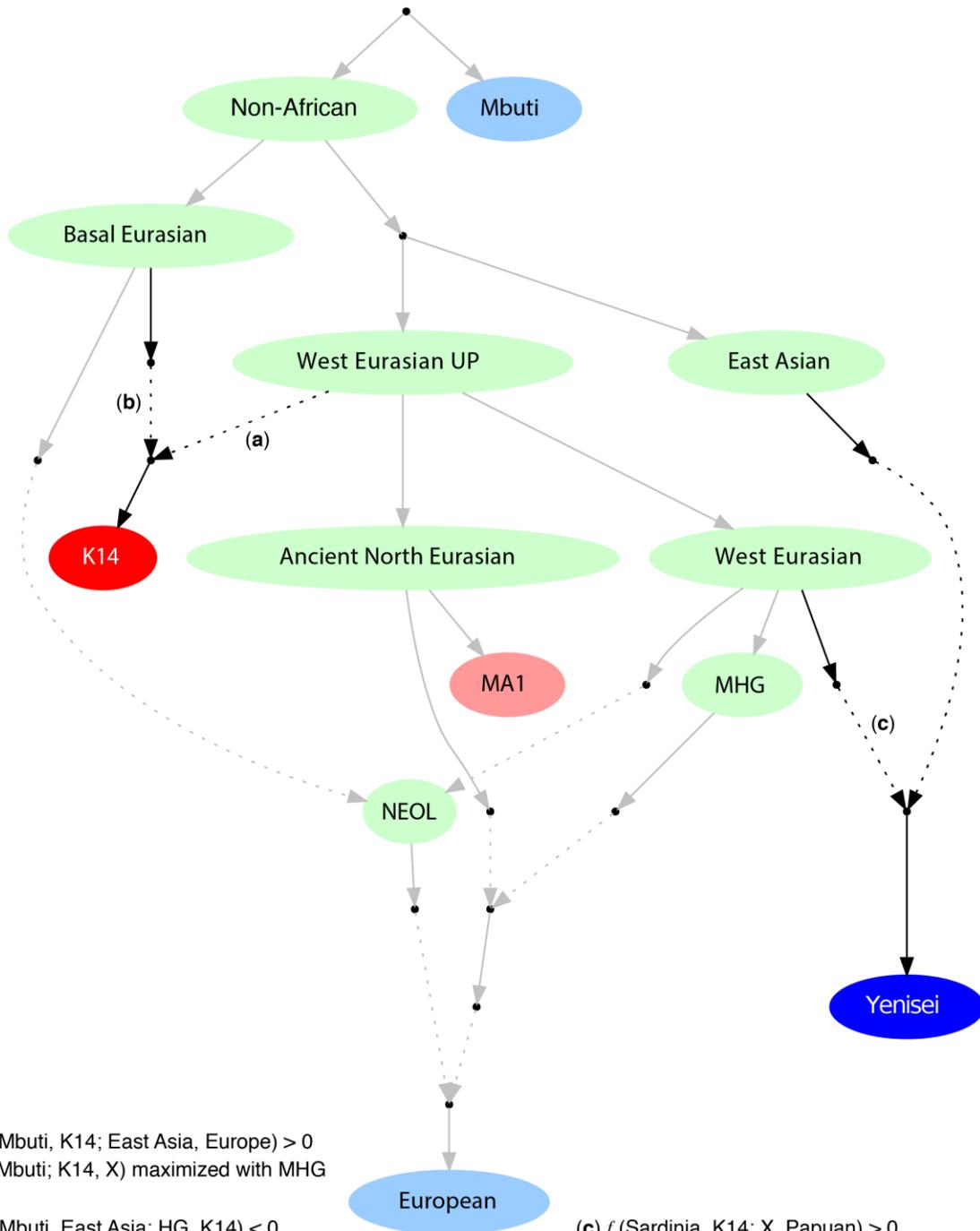
(a) D(Mbuti, K14; East Asia, Europe) > 0
    f_3(Mbuti; K14, X) maximized with MHG

(b) D(Mbuti, East Asia; HG, K14) < 0
    D(Mbuti, East Asia; NEOL, K14) = 0
    NEOL component for K14 in ADMIXTURE

(c) f_4(Sardinia, K14; X, Papuan) > 0
    f_4(Sardinia, Scandinavian MHG; X, Papuan) < 0
    f_3(X; Motala, Evens) < 0 and lowest f_3

Figure 2. Relationships of the K14 sample and MA1, MHG, NF, modern Europeans and the modern populations in the Yenisei region. This representation is a possible topology consistent with the results presented in this study in the context of the relationships described by Lazaridis et al. (*21*) for the modern European populations and Raghavan et al. (*23*) for MA1. Present day populations are colored in blue, ancient in red and ancestral populations in green. Solid lines represent descent without admixture events, and dashed lines, admixture events. Arrows do not depart from ancient samples (K14 and MA1) as they represent relationships of population ancestry. We only show the topology of the potential population tree: there is no notion of time in this representation. We also note that the tree is not the result of a model-fitting procedure, but rather a possible topology consistent with the key results of this study (indicated with lower case letters in the figure). UP: Upper Paleolithic, MHG: Mesolithic Hunter-Gatherers, NF: Neolithic Farmers.
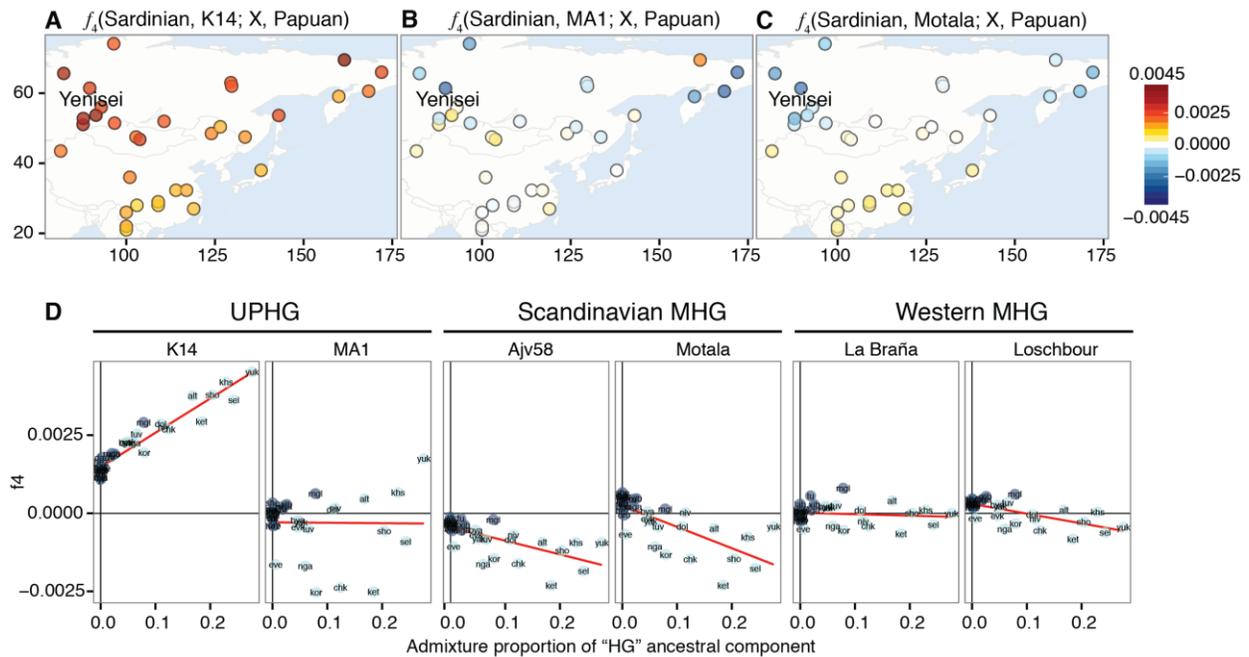
Figure 3. (A) Values of the f_4 statistic for a set of Siberian and East Asian populations and K14. We compute the f_4 statistic for a topology (Sardinian, K14; X, Papuan). Warmer values indicate departure from the topology (Sardinian, K14; X, Papuan) with increased ancestry between the modern population X and the Sardinian. The Yenisei region includes the Selkup, Shor, and Ket populations. **(B)** Values of the f_4 statistic for a set of Siberian and East Asian populations and MA1. We compute the f_4 statistic for a topology (Sardinian, MA1; X, Papuan). **(C)** Values of the f_4 statistic for a set of Siberian and East Asian populations and Scandinavian hunter gatherers (Motala). **(D)** Relationship between the "HG" admixture proportion and the f_4(Sardinian, K14; X, Papuan) shown in (A). The red lines are linear regressions for each case. MHG: Mesolithic Hunter-Gatherers.
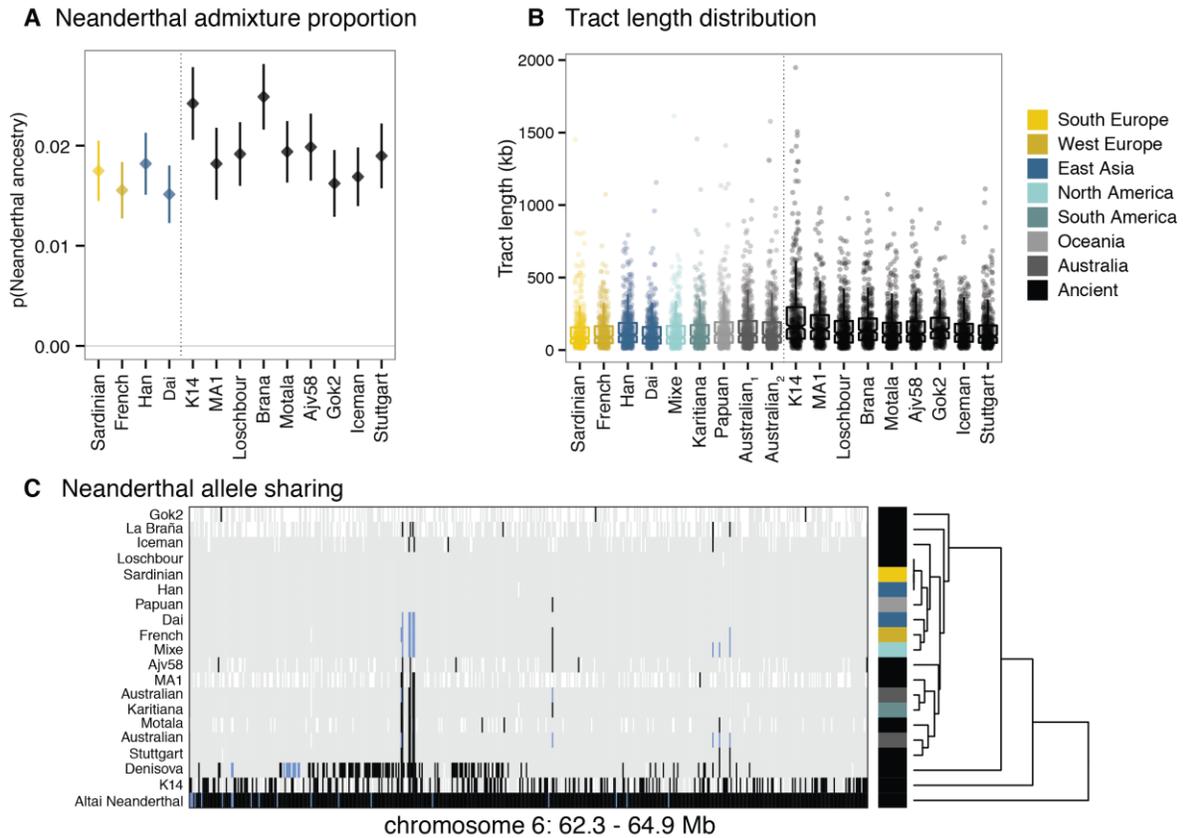
Figure 4. (A) Neandertal admixture proportions for the modern and ancient individuals from Eurasia. **(B)** Ancestry tract length distribution for tracts identified as Neandertal through a sliding window approach. The sites are ascertained to be ancestral in the African populations. For each non-African, the tracts are identified as the regions where sites are derived in Neandertal and the individual shown in X. **(C)** The longest "Neandertal haplotype" identified in K14 through a sliding window approach. Individuals were clustered using hierarchical clustering on the genotype matrix for the region. Missing data is shown in white, grey indicates homozygous ancestral, blue heterozygote and black homozygous derived.