

# Population genomics of post-glacial western Eurasia


<https://doi.org/10.1038/s41586-023-06865-0>

Received: 2 May 2022

Accepted: 14 November 2023

Published online: 10 January 2024

Open access

 Check for updates

Western Eurasia witnessed several large-scale human migrations during the Holocene<sup>1–5</sup>. Here, to investigate the cross-continental effects of these migrations, we shotgun-sequenced 317 genomes—mainly from the Mesolithic and Neolithic periods—from across northern and western Eurasia. These were imputed alongside published data to obtain diploid genotypes from more than 1,600 ancient humans. Our analyses revealed a ‘great divide’ genomic boundary extending from the Black Sea to the Baltic. Mesolithic hunter-gatherers were highly genetically differentiated east and west of this zone, and the effect of the neolithization was equally disparate. Large-scale ancestry shifts occurred in the west as farming was introduced, including near-total replacement of hunter-gatherers in many areas, whereas no substantial ancestry shifts happened east of the zone during the same period. Similarly, relatedness decreased in the west from the Neolithic transition onwards, whereas, east of the Urals, relatedness remained high until around 4,000 BP, consistent with the persistence of localized groups of hunter-gatherers. The boundary dissolved when Yamnaya-related ancestry spread across western Eurasia around 5,000 BP, resulting in a second major turnover that reached most parts of Europe within a 1,000-year span. The genetic origin and fate of the Yamnaya have remained elusive, but we show that hunter-gatherers from the Middle Don region contributed ancestry to them. Yamnaya groups later admixed with individuals associated with the Globular Amphora culture before expanding into Europe. Similar turnovers occurred in western Siberia, where we report new genomic data from a ‘Neolithic steppe’ cline spanning the Siberian forest steppe to Lake Baikal. These prehistoric migrations had profound and lasting effects on the genetic diversity of Eurasian populations.

Genetic diversity in west Eurasian human populations was largely shaped by three major prehistoric migrations: anatomically modern hunter-gatherers (HGs) occupying the area from around 45,000 BP (refs. 4,6); Neolithic farmers expanding from the Middle East from around 11,000 BP (ref. 4); and steppe pastoralists coming out of the Pontic Steppe around 5,000 BP (refs. 1,2). Palaeogenomic analyses have uncovered the early post-glacial colonization routes<sup>7</sup> that led to a basal ancestral dichotomy between HGs in central and western Europe and HG groups represented further east<sup>8</sup>. Western HG (WHG) ancestry appears to be derived directly from ancestry sources related to Epigravettian, Azilian and Epipalaeolithic cultures (the Villabruna cluster)<sup>9</sup>, whereas eastern HG (EHG) ancestry shows further admixture with an Upper Palaeolithic Siberian source (Ancient North Eurasian; ANE)<sup>10</sup>. The WHG ancestry composition was regionally variable in the Mesolithic populations. There is evidence for continuous local admixture in Iberian HGs<sup>11</sup>, which contrasts with the more homogeneous WHG ancestry profile in Britain and northwestern continental Europe, suggesting ancestry formation before expansion<sup>12</sup>. The timing of the ancestry admixture that formed EHG has been estimated at 13,000–15,000 BP, and the composition seems to follow a cline that is broadly correlated with geography, with Baltic and Ukrainian HGs showing more affinity to the Villabruna Upper Palaeolithic cluster ancestry, as compared with HGs in Russia, who exhibited more ANE

ancestry<sup>5,7,13,14</sup>. Genomic analyses of Mesolithic skeletal material from the Scandinavian Peninsula has revealed varied mixes of WHG and EHG ancestry among the later Mesolithic populations<sup>3,15,16</sup>.

Beyond these broad-scale characterizations, our knowledge about Mesolithic population structure and demographic admixture processes is limited, and has substantial chronological and geographical information gaps. This is partly owing to a relative paucity of well-preserved Mesolithic human skeletons older than 8,000 years, and partly because most ancient DNA studies on the Mesolithic and Neolithic periods have been restricted to individuals from Europe. The archaeological record indicates a boundary from the eastern Baltic to the Black Sea, east of which HG societies persisted for much longer than in western Europe, despite the similar distance to the distribution centre for early agriculture in the Middle East<sup>17</sup>. Components of eastern and western HG ancestry appear highly variable in this boundary region<sup>5,18,19</sup> but the wider spatiotemporal genetic implications of the east–west division are unclear. The spatiotemporal mapping of population dynamics east of Europe, including northern and central Asia during the same time period, is limited. In these regions, the term ‘Neolithic’ is characterized by cultural and economic changes including societal-network differences, changes in lithic technology and use of pottery. For instance, the Neolithic cultures of the central Asian steppe and the Russian taiga belt possessed pottery, but retained a HG economy alongside stone-blade

technology, similar to the preceding Mesolithic cultures<sup>20</sup>. A fundamental lack of data from some key regions and periods has made it difficult to gain a deeper understanding of how the neolithization differed in its timing, mechanisms and effects across northern and western Eurasia.

The transition from hunting and gathering to farming was based on domesticated plants and animals of Middle Eastern origin, and represents one of the most fundamental shifts in demography, health, lifestyle and culture in human prehistory. The neolithization process in large parts of Europe was accompanied by the arrival of immigrants of Anatolian descent<sup>21</sup>. For example, in Iberia, the Neolithic began with the abrupt spread of immigrant farmers of Anatolian–Aegean ancestry along the Mediterranean and Atlantic coasts, after which admixture with local HGs gradually took place<sup>11</sup>. Similarly, in southeastern and central Europe, farming rapidly spread with Anatolian Neolithic farmers, who were to some extent subsequently admixed with local HGs<sup>22–27</sup>. Conversely, in Britain, data suggest that there was a complete replacement of the HG population when agriculture was introduced by incoming continental farmers, without a subsequent resurgence of local HG ancestry<sup>12,28</sup>. In the east Baltic region, a markedly different neolithization trajectory occurred, with the introduction of domesticates only at the emergence of the Corded Ware complex (CWC) around 4,800 calibrated years before present (cal. BP) (refs. 18,19). Similarly, in eastern Ukraine, HGs of Mesolithic ancestry co-existed for millennia with farming groups further west<sup>5,29</sup>. These studies have all provided important regional contributions to the understanding of west Eurasian population history, but from a broader cross-continental perspective, our knowledge is still patchy.

From approximately 5,000 BP, an ancestry component related to Early Bronze Age steppe pastoralists such as the Yamnaya culture rapidly spread across Europe through the expansion of the CWC and related cultures<sup>1,2</sup>. Although previous studies have identified these large-scale migrations into Europe and central Asia, central aspects concerning the demographic processes are not resolved. Yamnaya ancestry (that is, ‘steppe’ ancestry) has been characterized broadly as a mix between EHG ancestry and Caucasus hunter-gatherer (CHG), formed in a hypothetical admixture between a ‘northern’ steppe source and a ‘southern’ Caucasus source<sup>30</sup>. However, the exact origins of these ancestry sources have not been identified. Furthermore, with a few exceptions<sup>31–33</sup>, published Yamnaya Y-chromosomal haplogroups do not match those found in Europeans after 5,000 BP, and the origin of this patrilineal lineage is also unresolved. Finally, in Europe, ‘steppe’ ancestry has hitherto been identified only in admixed form, but the origin of this admixture event and the mechanism by which the ancestry subsequently spread with the CWC have remained elusive.

To investigate these formative processes at a cross-continental scale, we sequenced the genomes of 317 radiocarbon-dated (by accelerator mass spectrometry) individuals of mainly Mesolithic and Neolithic origin, covering major parts of Eurasia. We combined these with published shotgun-sequenced data to impute a dataset of more than 1,600 diploid ancient genomes. Of the 317 sampled ancient skeletons (Fig. 1, Extended Data Fig. 1 and Supplementary Data 1), 272 were radiocarbon-dated within the project, 30 dates were derived from published literature and 15 examples were dated by archaeological context. Dates were corrected for marine and freshwater reservoir effects (Supplementary Note 4) and ranged from the Upper Palaeolithic around 25,700 cal. BP to the mediaeval period (around 1,200 cal. BP). However, 97% of the individuals ( $n = 309$ ) date to between 11,000 and 3,000 cal. BP, with a heavy focus on individuals associated with various Mesolithic and Neolithic cultures. Geographically, the 317 sampled skeletons cover a vast territory across Eurasia, from Lake Baikal to the Atlantic coast and from Scandinavia to the Middle East, deriving from contexts that include burial mounds, caves, bogs and the sea floor (Supplementary Notes 6 and 7). Broadly, we can divide our research area into three large regions: (1) central, western and northern Europe; (2) eastern Europe, including western Russia, Belarus and Ukraine; and (3) the Urals and

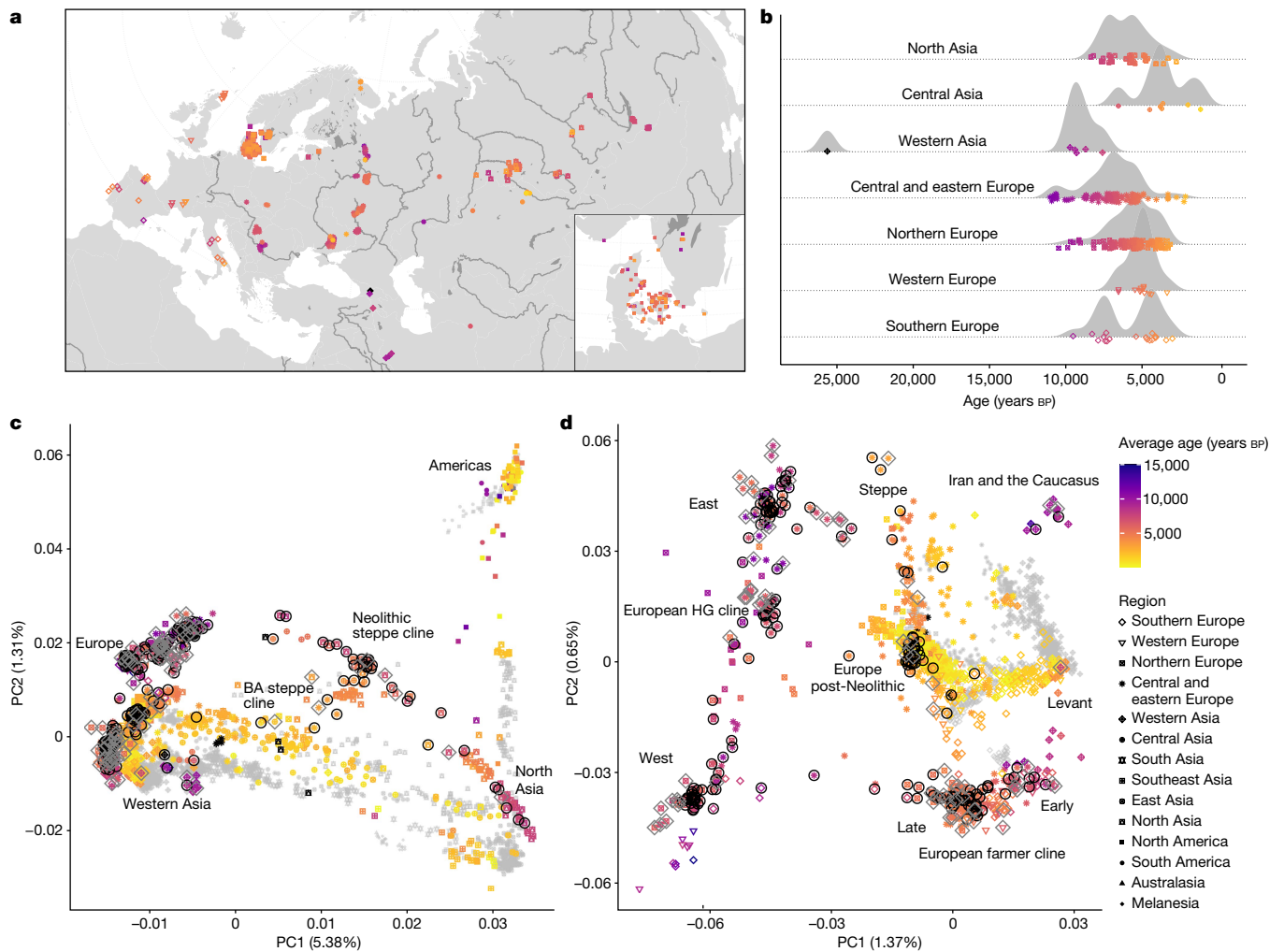
western Siberia (Supplementary Notes 6 and 7). Samples cover many of the key Mesolithic and Neolithic cultures in western Eurasia, such as the Maglemose, Ertebølle, Funnel Beaker (TRB) and Corded Ware/Single Grave cultures in Scandinavia; the Cardial in the Mediterranean; the Körös and Linear Pottery (LBK) in southeastern and central Europe; and many archaeological cultures in Ukraine, western Russia and the trans-Ural region (for example, Veretye, Lyalovo, Volosovo and Kitoi). Our sampling was particularly dense in Denmark, from where an accompanying paper presents a detailed and continuous sequence of 100 genomes spanning the Early Mesolithic to the Bronze Age<sup>34</sup>. Dense sampling was also obtained from Ukraine, western Russia and the trans-Ural region, spanning the Early Mesolithic through to the Neolithic, up to around 5,000 BP.

## Broad-scale genetic structure

Ancient DNA was extracted from either dental cementum or petrous bones, and the 317 genomes were shotgun-sequenced to a depth of coverage ranging between 0.01 $\times$  and 7.1 $\times$  (mean, 0.75 $\times$ , median, 0.26 $\times$ ), with more than 1 $\times$  coverage for 81 genomes (Supplementary Note 1). We used a computational method optimized for low-coverage data<sup>35</sup> to impute genotypes using the 1000 Genomes phased data<sup>36</sup> as a reference panel. This method was jointly applied to more than 1,300 previously published shotgun-sequenced genomes (Supplementary Data 7), resulting in a dataset of 8.5 million common single-nucleotide polymorphisms (SNPs) (with a minor allele frequency (MAF) greater than 1% and an imputation INFO score greater than 0.5) for 1,664 imputed diploid ancient genomes (Extended Data Fig. 2). For most downstream analyses,  $n = 71$  individuals were excluded because they were found to be close relatives or because the estimated contamination was greater than 5%. This resulted in 1,593 genomes, of which 1,492 were analysed as imputed (213 sequenced in this study) and 101 were analysed as pseudo-haploids owing to low coverage (less than 0.1 $\times$ ) and/or low imputation quality (average genotype probability lower than 0.98).

We conducted a broad-scale characterization of this dataset using principal component analysis (PCA) and model-based clustering (ADMIXTURE), which recapitulated previously described ancestry clines in ancient Eurasian populations at increased resolution (Fig. 1, Extended Data Fig. 1 and Supplementary Note 3d). Our imputed whole genomes allowed us to perform PCA using ancient genomes as input, instead of projecting onto a space defined by modern variation. Notably, this resulted in much higher differentiation among the ancient individuals than observed previously (Extended Data Fig. 1). This is particularly notable in a PCA of west Eurasian individuals, in which the variance explained by the first two PCs increases more than 1.5-fold, and present-day populations are confined within a small central area of the PCA space (Fig. 1d and Extended Data Fig. 1c,d). These results are consistent with the genetic differentiation between ancient Europeans being higher than is observed in present-day populations, reflecting more genetic isolation and lower effective population sizes among ancient groups.

To obtain a finer-scale characterization of genetic ancestries across space and time, we used an approach similar to the widely used ChromoPainter–FineSTRUCTURE workflow<sup>37–39</sup>. We first performed community detection on a network constructed from pairwise identity-by-descent (IBD)-sharing similarities between ancient individuals to group them into hierarchically related clusters of similar genetic ancestry (Extended Data Fig. 3 and Supplementary Note 3c). At higher levels of the hierarchy, the resulting clusters represented previously described ancestry groups reflecting broad genetic structure, such as EHGs and WHGs (‘HG\_EuropeE’ and ‘HG\_EuropeW’; Extended Data Fig. 3). Clusters at the lowest level resolved fine-scale genetic structure, grouping individuals within restricted spatiotemporal ranges and/or archaeological contexts but also revealing previously unknown connections across broader geographical areas (Extended Data Fig. 3 and



**Fig. 1 | Sample overview and broad-scale genetic structure.** **a, b**, Geographical (**a**) and temporal (**b**) distribution of the 317 ancient genomes sequenced and reported in this study. Insert shows dense sampling in Denmark<sup>34</sup>. The age and the geographical region of ancient individuals are indicated by the colour and the shape of the symbols, respectively. Colour scale for age is capped at 15,000 years; older individuals are indicated with black. Random jitter was added to geographical coordinates to avoid overplotting. **c, d**, PCA of 3,316 modern and ancient individuals from Eurasia, Oceania and the Americas (**c**), and restricted

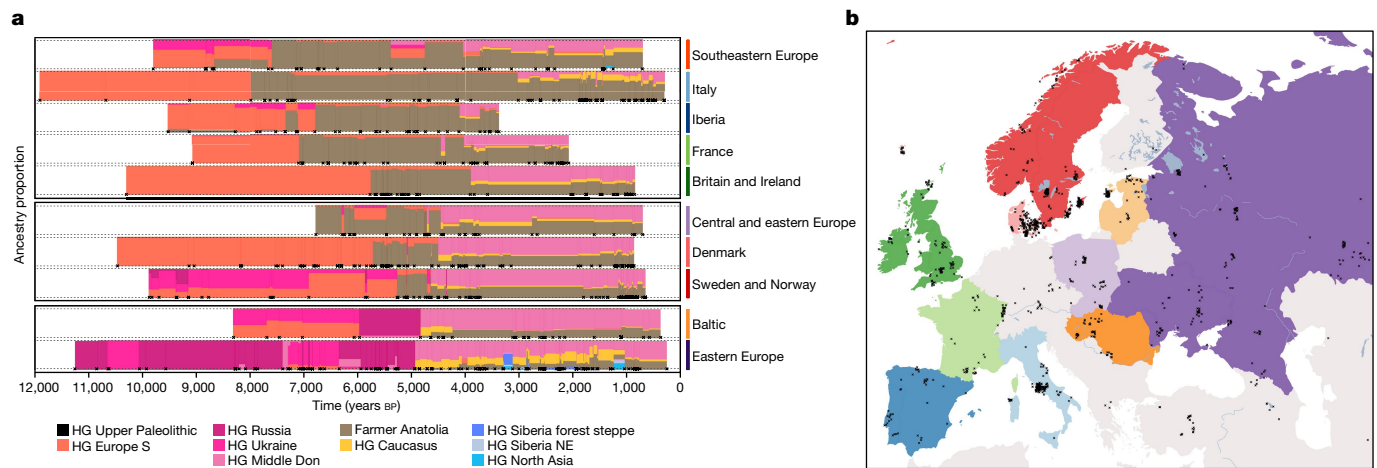
to 2,126 individuals from western Eurasia (west of the Urals) (**d**). Principal components were defined using both modern and imputed ancient ( $n = 1,492$ ) genomes passing all filters, with the remaining low-coverage ancient genomes projected. Ancient genomes sequenced in this study are indicated with black circles (imputed genomes passing all filters,  $n = 213$ ) or grey diamonds (pseudo-haploid projected genomes;  $n = 104$ ). Genomes of modern individuals are shown in grey, with population labels corresponding to their median coordinates. BA, Bronze Age.

Supplementary Note 3f). These resulting clusters were subsequently used in supervised ancestry modelling, in which sets of ‘target’ individuals were modelled as mixtures of ‘source’ groups (Methods).

### Population structure of HGs after the LGM

Our study comprises 113 shotgun-sequenced and imputed HG genomes, of which 79 were sequenced in this study. Among them, we report a 0.83× (0.83-fold coverage) genome of an Upper Palaeolithic skeleton from Kotias Klde Cave in Georgia, Caucasus (NEO283), directly dated to 26,052–25,323 cal. BP (95% confidence interval). In the PCA of all non-African individuals, this individual occupied a position distinct from those of other previously sequenced Upper Palaeolithic individuals—shifted towards west Eurasians along PC1 (Supplementary Note 3d). Using admixture graph modelling, we find that a well-fitting graph for this Caucasus Upper Palaeolithic lineage derives it as a mixture of predominantly west Eurasian Upper Palaeolithic HG ancestry (76%), with a contribution of about 24% from a ‘basal Eurasian’ ghost population, first observed in west Asian Neolithic individuals<sup>4</sup> (Supplementary

Note 3d and Supplementary Fig. 3d.16). To further explore the fine-scale structure of later European HGs, we then performed supervised ancestry modelling using sets of increasingly proximate source clusters (Extended Data Fig. 4). We replicate previous results of broad-scale genetic differentiation between HGs in eastern and western Europe after the Last Glacial Maximum (LGM)<sup>5,7</sup>. We show that the deep ancestry divisions in the Eurasian human gene pool that were established during early post-LGM dispersals<sup>7</sup> persisted throughout the Mesolithic (Extended Data Fig. 4). Using distal sets of pre-LGM HGs as sources, we modelled western HGs as predominantly derived from a source related to the herein-reported Caucasus Upper Palaeolithic individual from Kotias Klde cave (Caucasus\_25000BP), whereas eastern HGs showed varying amounts of ancestry related to a Siberian HG from Mal’ta (Malta\_24000BP; Extended Data Fig. 4a and Supplementary Data 12). Using post-LGM sources, this divide is best represented by ancestry related to southern European (Italy\_15000BP\_9000BP) and Russian (RussiaNW\_11000BP\_8000BP) HGs, respectively, corresponding to the ‘WHG’ and ‘EHG’ labels commonly used in previous studies.



**Fig. 2 | Genetic ancestry transects of western Eurasia. a**, Regional timelines of genetic ancestry compositions within the past 12,000 years in western Eurasia. Ancestry proportions in 1,012 imputed ancient genomes (representing populations west of the Urals) inferred using supervised ancestry modelling with the ‘deep’ HG ancestry source groups. Coloured bars within the timelines represent ancestry proportions for temporally consecutive individuals, with the width corresponding to their age difference. Individuals with identical age

were offset along the time axis by adding random jitter. **b**, Map highlighting geographical areas (coloured areas) for samples included in the individual regional timelines, and excavation locations (black crosses). Only shotgun-sequenced genomes were used in our study, so the exact timing of ancestry shifts might differ slightly from previous studies if they are based on different types of data from different individuals.

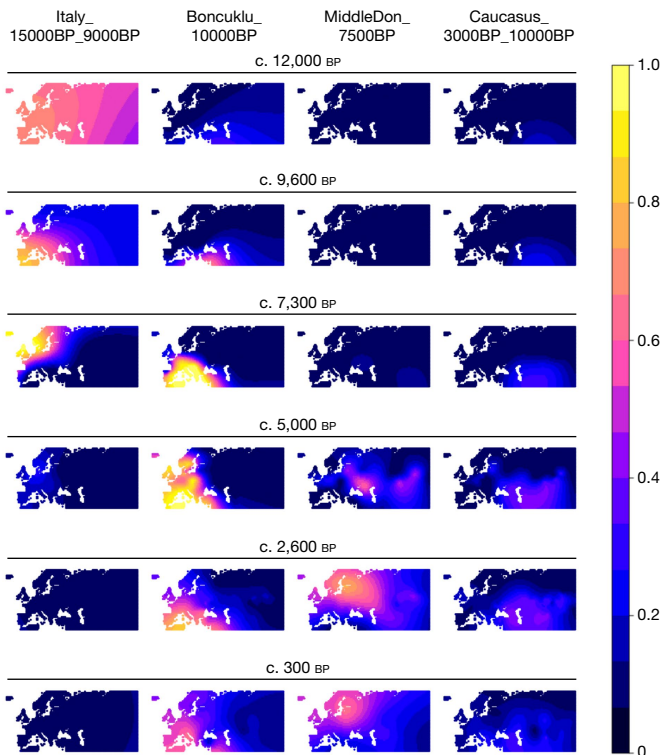
Adding extra proximate sources allowed us to further refine the ancestry composition of northern European HGs. In Denmark, our 28 sequenced and imputed HG genomes derived almost exclusively from a southern European source (Italy\_15000BP\_9000), with notable homogeneity across a 5,000-year transect<sup>34</sup> (Extended Data Fig. 4a and Supplementary Data 12). By contrast, we observed marked geographical variation in the ancestry composition of HGs from other parts of Scandinavia. Mesolithic individuals from Scandinavia were broadly modelled as mixtures with varying proportions of eastern and western HGs using distal post-LGM sources (‘hgEur1’; Extended Data Fig. 4a), as previously reported<sup>15</sup>. In Mesolithic individuals from southern Sweden, the eastern HG ancestry component was largely replaced by a southeastern European source (Romania\_8800BP) in more proximate models, making up between 60% and 70% of the ancestry (Extended Data Fig. 4a and Supplementary Data 12). Ancestry related to Russian HGs increased in a cline towards the far north, peaking at around 75% in a late HG from Tromsø (VK531; around 4,350 BP) (Extended Data Fig. 4a,c and Supplementary Data 12); this was also reflected in the fact that those individuals shared the highest IBD with northern Russian HGs (Extended Data Fig. 4d). During the late Mesolithic, we observed higher southern European HG ancestry in coastal individuals (NEO260 from Evensås and NEO679 from Skateholm) than in earlier individuals from further inland. Adding Danish HGs as a proximate source substantially improved the fit for those two individuals (‘hgEur3’; Extended Data Fig. 4b), with an estimated 58–76% of ancestry derived from Danish HGs (‘hgEur3’; Extended Data Fig. 7a and Supplementary Data 12), suggesting a population genetic link with Denmark, where this ancestry prevailed (Extended Data Fig. 4c). These results indicate that there were at least three distinct waves of northwards HG ancestry into Scandinavia: (1) a predominantly southern European source into Denmark and coastal southwestern Sweden; (2) a source related to southeastern European HGs into the Baltic and southeastern Sweden; and (3) a northwest Russian source into the far north, which then spread south along the Atlantic coast of Norway<sup>15</sup> (Extended Data Fig. 4c). These movements are likely to represent post-glacial expansions from refugial areas shared with many plant and animal species<sup>40</sup>.

On the Iberian Peninsula, the earliest individuals, including an approximately 9,200-year-old HG (NEO694) from Santa Maira (eastern Spain), sequenced in this study, showed predominantly southern European HG ancestry, with a minor contribution from Upper Palaeolithic HG

sources (Extended Data Fig. 4a). This observed Upper Palaeolithic HG ancestry source mix is likely to reflect the pre-LGM Magdalenian-related ancestry component that has previously been reported in Iberian HGs<sup>11</sup>, for which a good source population proxy is lacking in our dataset. By contrast, later individuals from northern Iberia were more similar to HGs from southeastern Europe, deriving around 30–40% of their ancestry from a source related to HGs from the Balkans in more proximate models<sup>11,41</sup> (Extended Data Fig. 4a and Supplementary Data 12). The earliest evidence for this gene flow was observed in a Mesolithic individual from El Mazo, Spain (NEO646) who was dated, calibrated and reservoir-corrected to around 8,200 BP (8,365–8,182 cal. BP; 95%) but dated slightly earlier by context<sup>42</sup> (8,550–8,330 BP). The directly dated age coincides with some of the oldest Mesolithic geometric microliths in northern Iberia, appearing around 8,200 BP at this site<sup>42</sup>. An influx of southeastern European HG-related ancestry in Ukrainian individuals after the Mesolithic (Extended Data Fig. 4a and Supplementary Data 12) suggests a similar eastward expansion in southeastern Europe<sup>5</sup>. Of note, two newly reported approximately 7,300-year-old genomes from the Middle Don River region in the Pontic-Caspian steppe (Golubaya Krinitza, NEO113 & NEO212) were found to be predominantly derived from earlier Ukrainian HGs, but with around 18–24% of their ancestry contributed from a source related to HGs from the Caucasus (Caucasus\_13000BP\_10000BP) (Extended Data Fig. 4a and Supplementary Data 12). Further lower-coverage (non-imputed) genomes from the same site project in the same PCA space (Fig. 1d) shifted away from the European HG cline towards Iran and the Caucasus. Using the linkage-disequilibrium-based method DATES<sup>43</sup>, we dated this admixture to around 8,300 BP (Supplementary Data 14). These results document genetic contact between populations from the Caucasus and the steppe region that is much earlier than previously known, providing evidence of admixture before the advent of later nomadic steppe cultures—in contrast with recent hypotheses—and further to the west than has been previously reported<sup>5,44</sup>.

### Major genetic transitions in Europe

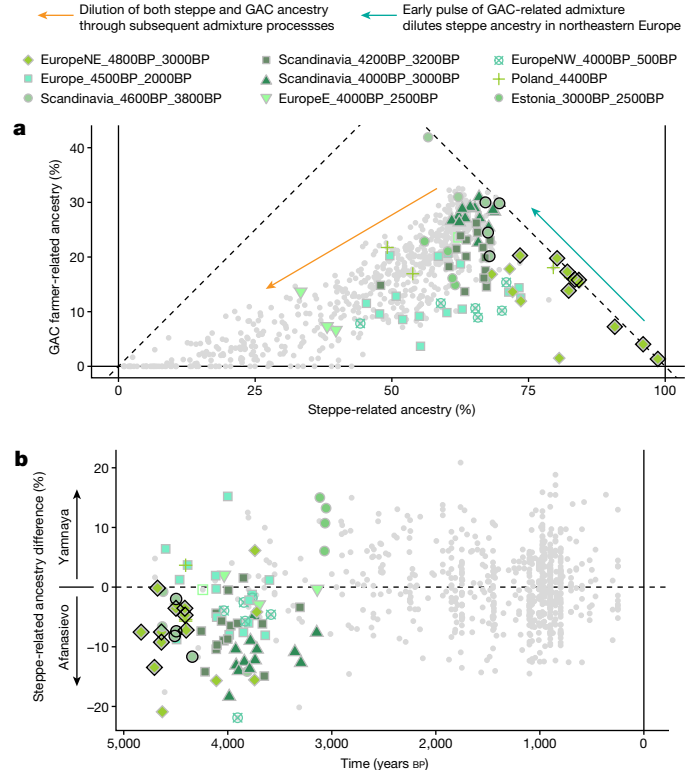
Previous ancient genomics studies have documented several episodes of large-scale population turnover in Europe within the past 10,000 years (see, for example, refs. 1,2,5,45), but the 317 genomes reported here fill important knowledge gaps. Our analyses reveal profound



**Fig. 3 | Spatiotemporal kriging analysis of major ancestries.** The temporal transects show how WHG ancestry (Italy\_15000BP\_9000BP) was replaced by Neolithic farmer ancestry (Boncuklu\_10000BP) during the Neolithic transition in Europe. Later, the steppe migrations around 5,000 cal. BP introduced both EHG (MiddleDon\_7500BP) and CHG (Caucasus\_13000BP\_10000BP) ancestry into Europe, thereby reducing Neolithic farmer ancestry.

differences in the spatiotemporal neolithization dynamics across Europe. Supervised admixture modelling (using the ‘deep’ ancestry set; Supplementary Data 11) and spatiotemporal kriging<sup>46</sup> document a broad east–west distinction along a boundary zone running from the Black Sea to the Baltic. On the western side of this ‘great divide’, the Neolithic transition is accompanied by large-scale shifts in genetic ancestry from local HGs to farmers with Anatolian-related ancestry (Boncuklu\_10000BP; Fig. 2a and Fig. 3 and Extended Data Figs. 5–7). The arrival of Anatolian-related ancestry in different regions spans an extensive time period of more than 3,000 years, from its earliest evidence in the Balkans (Lepenski Vir) at around 8,700 BP (ref. 5) to around 5,900 BP in Denmark.

Furthermore, we corroborate previous reports (for example, refs. 2,5,45,47) of widespread, low-level admixture between early European farmers and local HGs, resulting in a resurgence of HG ancestry in many regions of Europe during subsequent centuries (Extended Data Fig. 8b,c and Supplementary Data 8). The resulting estimated proportions of HG ancestry rarely exceeded 10%, with notable exceptions observed in individuals from southeastern Europe (Iron Gates) and Sweden (Pitted Ware Culture), as well as in the herein-reported Early Neolithic genomes from Portugal (western Cardial), which are estimated to contain 27%–43% Iberian HG ancestry (Iberia\_9000BP\_7000BP). The latter result, together with an estimated admixture date of just 200 years earlier (‘Iberia farmer early’ in Supplementary Data 14), suggests extensive first-contact admixture, and is in agreement with archaeological inferences derived from modelling the spread of farming across west Mediterranean Europe<sup>48</sup>. Neolithic individuals from Denmark showed some of the highest overall proportions of HG ancestry (up to around 25%), but this was mostly derived from non-local western European-related HGs (EuropeW\_13500BP\_8000BP), with only

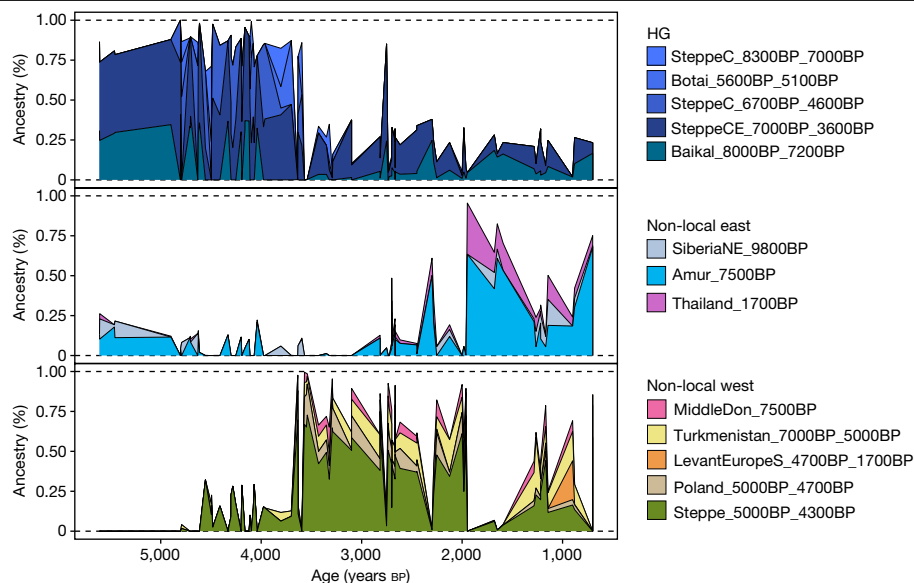


**Fig. 4 | Fine-scale structure and temporal dynamics of steppe-related ancestry during the second transition in Europe.** **a**, Correlation between the estimated proportions of steppe-related and GAC farmer-related ancestries (‘postNeol’ source set), across west Eurasian target individuals. **b**, Timeline of difference in estimated steppe-related ancestry proportions, using individuals from the genetic cluster ‘Steppe\_5000BP\_4300BP’ associated with either Yamnaya or Afanasievo cultural contexts as separate sources. Individuals from European post-Neolithic genetic clusters before 3,000 cal. BP are indicated with coloured symbols; other west Eurasian target individuals are indicated with grey symbols. Symbols with black outlines highlight early steppe-related individuals associated with either Corded Ware or related (for example, Battle Axe) cultural contexts.

a small contribution from local Danish HG groups in some individuals (Extended Data Fig. 8b and Supplementary Note 3f).

We find evidence for regional stratification in early Neolithic farmer ancestries in subsequent Neolithic groups. Specifically, southern European early farmers were found to have provided major genetic ancestry to Neolithic groups of later dates in western Europe, whereas central European early farmer ancestry was mainly observed in subsequent Neolithic groups in eastern Europe and Scandinavia (Extended Data Fig. 8e). These results are consistent with distinct migratory routes of expanding farmer populations, as previously suggested<sup>49</sup>.

On the eastern side of the great divide, no ancestry shifts can be observed during this period. In the east Baltic region<sup>50</sup>, Ukraine and western Russia, local HG ancestry prevailed until around 5,000 BP without a noticeable input of Anatolian-related farmer ancestry (Figs. 2 and 3 and Extended Data Figs. 5–7). This eastern genetic continuity is in congruence with the archaeological record, which shows the persistence of pottery-using forager groups in this wide region, and a delayed introduction of cultivation and animal husbandry by several thousand years (Supplementary Note 5). Around 5,000 BP, major demographic events unfolded on the Eurasian Steppe, resulting in steppe-related ancestry spreading rapidly both eastwards and westwards<sup>1,2</sup>, marking the end of the great population genomic divide (Figs. 3 and 6). We find that this second transition happened at a faster pace than during the neolithization, reaching most parts of Europe within an approximately



**Fig. 5 | Genetic transects east of the Urals.** Timelines of genetic ancestry compositions within the past 6,000 years east of the Urals. Shown are ancestry proportions in 148 imputed ancient genomes from this region, inferred using

supervised ancestry modelling ('postNeol' source set). Panels separate ancestry proportions from local forest steppe HGs (HG) and sources representing ancestries originating further east or west.

1,000-year time period after first appearing in the eastern Baltic region around 4,800 cal. BP (Fig. 3). In line with previous reports, we observe that by around 4,200 cal. BP, steppe-related ancestry was already dominant in individuals from Britain, France and the Iberian Peninsula<sup>12,51</sup>. Notably, because of the delayed neolithization in southern Scandinavia, these dynamics resulted in two episodes of large-scale genetic turnover in Denmark and southern Sweden within a period of roughly 1,000 years<sup>34</sup> (Fig. 3).

Although the broader effects of the steppe migrations around 5,000 cal. BP are well known, the origin of this ancestry has remained a mystery. Here we show that the steppe ancestry composition (Steppe\_5000BP\_4300BP) can be modelled as a mixture of around 65% ancestry related to herein-reported HG genomes from the Middle Don River region (MiddleDon\_7500BP) and around 35% ancestry related to HGs from Caucasus (Caucasus\_13000BP\_10000BP) (Extended Data Fig. 6 and Supplementary Data 9). Thus, Middle Don HGs, who already carried ancestry related to Caucasus HGs (Extended Data Fig. 4a), serve as a hitherto-unknown proximal source for the majority ancestry contribution into Yamnaya-related genomes. The individuals in question derive from the burial ground Golubaya Krinitsa (Supplementary Note 3). Material culture and burial practices at this site are similar to the Mariupol-type graves, which are widely found in neighbouring regions of Ukraine; for instance, along the Dnepr River. They belong to the group of complex pottery-using HGs mentioned above, but the genetic composition at Golubaya Krinitsa is different from that in the remaining Ukrainian sites (Fig. 2a and Extended Data Fig. 5). A previous study<sup>30</sup> suggested a model for the formation of Yamnaya ancestry that includes a 'northern' steppe source (EHG + CHG ancestry) and a 'southern' Caucasian Chalcolithic source (CHG ancestry), but did not identify the exact origin of these sources. The Middle Don genomes analysed here show the appropriate balance of EHG and CHG ancestry, suggesting that they are candidates for the missing northern proximate source for Yamnaya ancestry.

The dynamics of the continent-wide transition from Neolithic farmer ancestry to steppe-related ancestry also differ markedly between geographical regions. The contribution of local Neolithic ancestry to the incoming groups was high in eastern, western and southern Europe, reaching more than 50% on the Iberian Peninsula<sup>41</sup> ('postNeol' set; Extended Data Fig. 6 and Supplementary Data 10). Scandinavia,

however, shows a very different picture, with much lower contributions (less than 15%), including near-complete replacement of the local population in some regions (Extended Data Fig. 9b). Steppe-related ancestry accompanies and spreads with the formation of the CWC across Europe, and our results provide new evidence on the foundational admixture event. Individuals associated with the CWC carry a mix of steppe-related and Neolithic farmer-related ancestry; we show that the latter can be modelled as deriving exclusively from a genetic cluster associated with the Late Neolithic Globular Amphora culture (GAC) (Poland\_5000BP\_4700BP), and that this ancestry co-occurred with steppe-related ancestry across all sampled European regions (Fig. 4a and Extended Data Fig. 6). This suggests that the spread of steppe-related ancestry was predominantly mediated through groups already admixed with GAC-related farmer groups of the eastern European plains—an observation that has major implications for understanding the emergence of the CWC.

A stylistic connection between GAC and CWC ceramics has long been suggested, including the use of amphora-shaped vessels and the development of cord decoration patterns<sup>52</sup>. Moreover, shortly before the emergence of the earliest CWC groups, eastern GAC and western Yamnaya groups exchanged cultural elements in the forest-steppe transition zone northwest of the Black Sea, where GAC ceramic amphorae and flint axes were included in Yamnaya burials, and the typical Yamnaya use of ochre was included in GAC burials<sup>53</sup>, indicating close interactions between these groups. Previous ancient genomic data from a few individuals suggested that this was limited to cultural influences and not population admixture<sup>54</sup>. However, in the light of our new genetic evidence, it seems that this zone—and possibly other similar zones of contact between GAC and groups from the steppe (for example, the Yamnaya)—were key in the formation of the CWC, through which steppe-related ancestry and GAC-related ancestry co-dispersed far towards the west and the north<sup>55</sup>. This resulted in regionally diverse situations of interaction and admixture<sup>14,32</sup>, but a substantial part of the CWC dispersal happened through corridors of cultural and demic transmission that had been established by the GAC during the preceding period<sup>33,56</sup>. Differences in Y-chromosomal haplogroups between CWC and Yamnaya suggest that the currently published Yamnaya-associated genomes do not represent the most direct source for the steppe ancestry component in CWC<sup>32,33</sup>. This notion was supported by proximate

ancestry modelling using published genomes<sup>1</sup> associated with Yamnaya or Afanasievo cultural contexts as separate sources, which revealed a subtle increase in affinity for an Afanasievo-related source over a Yamnaya-related source in early individuals with European steppe ancestry before 3,000 cal. BP (Fig. 4b and Extended Data Fig. 9d). The result confirms the subtle population genetic structure in the population associated with Yamnaya or Afanasievo, showing that more dense sampling across the steppe horizon will be required to find the direct source or sources of steppe ancestry in the early CWC.

## HG resilience east of the Urals

In contrast to the considerable number of ancient HG genomes from western Eurasia that have been studied so far, genomic data from HGs east of the Urals have remained sparse. These regions are characterized by an early introduction of pottery from areas further east, and were inhabited by complex forager societies with permanent and sometimes fortified settlements<sup>20,57</sup>. Here, we substantially expand knowledge on ancient populations of this region by reporting genomic data from 38 individuals, 28 of whom date to pottery-associated HG contexts between 8,300 and 5,000 cal. BP (Supplementary Data 2). Most of these genomes form a previously only sparsely sampled<sup>13,43</sup> 'Neolithic steppe' cline that spans the Siberian forest steppe zones of the Irtysh, Ishim, Ob, and Yenisei River basins to the Lake Baikal region (Fig. 1c and Extended Data Figs. 1a and 3e). Supervised admixture modelling (using the 'deep' set of ancestry sources; Supplementary Data 9) revealed contributions from three major sources in these HGs from east of the Urals: early west Siberian HG ancestry (SteppeC\_8300BP\_7000BP) dominated in the western forest steppe; northeast Asian HG ancestry (Amur\_7500BP) was highest at Lake Baikal; and Palaeo-Siberian ancestry (SiberiaNE\_9800BP) was observed in a cline of decreasing proportions from northern Lake Baikal westwards across the forest steppe<sup>13</sup> (Extended Data Figs. 7 and 10a).

We used these Neolithic HG clusters ('postNeol' ancestry source set; Extended Data Fig. 7) as putative source groups in more proximal admixture modelling to investigate the spatiotemporal dynamics of ancestry compositions across the steppe and the Lake Baikal region after the Neolithic period. We replicate previously reported evidence for a genetic shift towards higher forest steppe HG ancestry (source SteppeCE\_7000BP\_3600BP) in Late Neolithic and Early Bronze Age (LNBA) individuals at Lake Baikal (clusters Baikal\_5600BP\_5400BP and Baikal\_4800BP\_4200BP)<sup>13,58</sup>. However, ancestry related to this cluster is also already observed at around 7,000 BP in herein-reported Neolithic HG individuals both at Lake Baikal (NEO199 and NEO200) and along the Angara river to the north (NEO843) (Extended Data Fig. 7). Both male individuals at Lake Baikal belonged to the Y-chromosome haplogroup Q1b1, characteristic of the later LNBA groups in the same region (Supplementary Note 3b and Supplementary Fig. 3b.5). Together with an early estimated admixture time (upper bound of around 7,300 cal. BP) for the LNBA groups (Supplementary Data 14), these results suggest that gene flow between HGs of Lake Baikal and those of the south Siberian forest steppe regions already occurred during the eastern Early Neolithic, consistent with archaeological interpretations of contact. In this region, bifacially flaked tools first appeared near Baikal<sup>59</sup>, from where the technique spread far to the west. We find echoes of such bifacial flaking in archaeological complexes (Shiderty 3, Borly, Sharbakty 1, Ust-Narym and so on) in northern and eastern Kazakhstan, around 6,500–6,000 cal. BP (refs. 60,61). Here, Mesolithic cultural networks with southwest Asia have also been recorded, as evidenced by pebble and flint lithics known from southwest Asia cultures<sup>62</sup>.

Genomes reported here also shed light on the genetic origins of the Early Bronze Age Okunevo Culture in the Minusinsk Basin in Southern Siberia. In contrast to previous results, we find no evidence for Lake Baikal HG-related ancestry in the Okunevo<sup>13,58</sup> when using our newly reported Siberian forest steppe HG genomes jointly with Lake

Baikal LNBA genomes as putative proximate sources. Instead, we find that they originate from the admixture of a forest steppe HG source (best modelled as a mixture of clusters Steppe\_6700BP\_4600BP and SteppeCE\_7000BP\_3600BP) and steppe-related ancestry (Steppe\_5300BP\_4000BP; Extended Data Fig. 7, set 'postBA' and Supplementary Data 11). We date the admixture with steppe-related ancestry to around 4,600 BP (Supplementary Data 14), and find it to be modelled exclusively from an Afanasievo-related source in proximate modelling separating the Yamnaya and Afanasievo steppe ancestries (Extended Data Figs. 9d and 10c,e). This is direct evidence for gene flow from peoples of the Afanasievo Culture, who were closely related to the Yamnaya and existed near Altai and Minusinsk Basin during the era of the steppe migrations<sup>1,58</sup>.

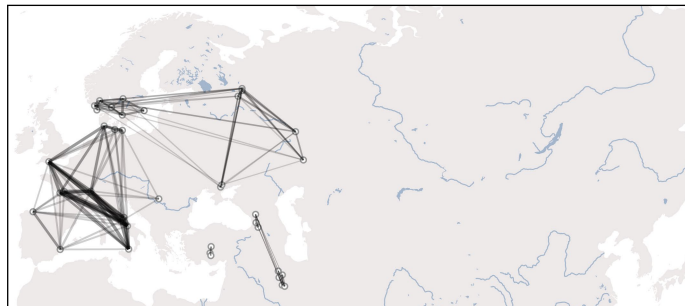
From around 3,700 cal. BP, individuals across the steppe and Lake Baikal regions show markedly different ancestry profiles (Fig. 5 and Extended Data Figs. 7 and 9b). We document a sharp increase in non-local ancestries, with only limited ancestry contributions from local HGs. The early stages of this transition are characterized by an influx of steppe-related ancestry, which decays over time from its peak of around 70% in the earliest individuals. Similar to the dynamics in western Eurasia, steppe-related ancestry is here correlated with GAC-related farmer ancestry (Poland\_5000BP\_4700BP; Fig. 5 and Extended Data Fig. 10b), recapitulating the previously documented gene flow from GAC groups into neighbouring groups of the steppe and the forest steppe, and the eastward expansion of admixed western steppe pastoralists from the Sintashta and Andronovo complexes during the Bronze Age<sup>43,63</sup>. However, GAC-related ancestry is notably absent in individuals of the Okunevo culture, and individuals with steppe ancestry after 3,700 BP show a slight excess in affinity to Yamnaya over Afanasievo in proximate modelling (Extended Data Fig. 10d), providing further support for two distinct eastward migrations of western steppe pastoralists during the early (Yamnaya-related) and later (Sintashta and Andronovo) Bronze Age. The later stages of the transition are characterized by increasing central Asian (Turkmenistan\_7000BP\_5000BP) and northeast Asian-related (Amur\_7500BP) ancestry components (Fig. 5 and Extended Data Fig. 10b). Together, these results show that deeply structured HG ancestry dominated the eastern Eurasian steppe substantially longer than in western Eurasia, before successive waves of population expansions swept across the steppe within the last 4,000 years. These included a large-scale introduction of domesticated horse lineages concomitant with new equestrian equipment and spoke-wheeled chariotry<sup>63,64</sup>, as well as the adoption of millet as a robust subsistence crop<sup>65</sup>.

## Sociocultural insights

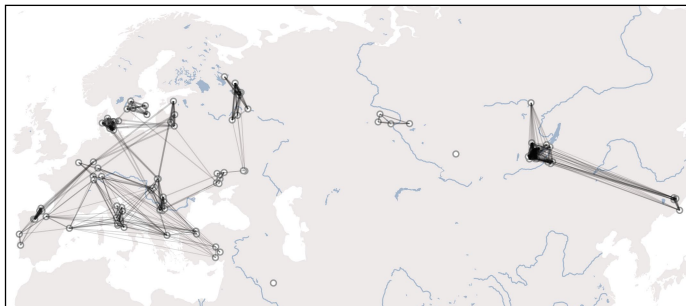
We used patterns of pairwise IBD sharing between individuals to examine our data for temporal shifts in relatedness within genetic clusters. We found clear trends of a reduction of within-cluster relatedness over time, in both western and eastern Eurasia (Extended Data Fig. 11a). This pattern is consistent with a scenario of increasing effective population sizes during this period<sup>66</sup>. Nevertheless, we observe notable differences in temporal relatedness patterns between western and eastern Eurasia, mirroring the wider difference in population dynamics discussed above. In the west, within-group relatedness changed substantially during the Neolithic transition (around 9,000–6,000 BP), in which clusters of individuals with Anatolian farmer-related ancestry show overall reduced IBD sharing compared with clusters of individuals with HG-associated ancestry (Extended Data Fig. 11a). In the east, genetic relatedness remained high until around 4,000 BP, consistent with a much longer persistence of smaller localized HG groups (Fig. 6 and Extended Data Fig. 11a).

Next, we examined the data for evidence of recent parental relatedness, by identifying individuals in which more than 50 centimorgans (cM) of their genomes was contained in long (more than 20 cM) runs

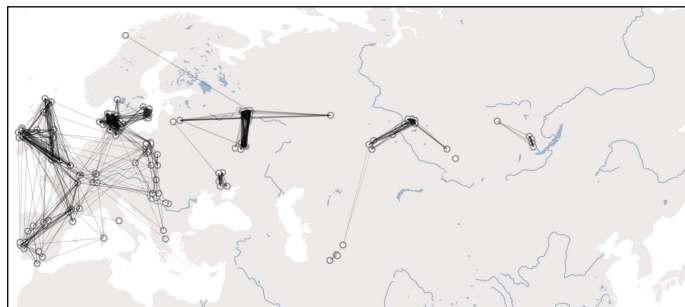
Before 9,000 cal. BP



9,000–7,000 cal. BP



7,000–5,000 cal. BP



5,000–3,000 cal. BP



**Fig. 6 | Genetic relatedness across western Eurasia.** Maps showing networks of highest IBD sharing (top 10 highest sharing per individual) during different time periods for 579 imputed genomes predating 3,000 cal. BP and located in the geographical region shown. Shading and thickness of lines are scaled to represent the amount of IBD shared between two individuals. In the earliest periods, sharing networks exhibit strong links within relatively narrow geographical regions, representing predominantly close genetic ties between small HG communities, and rarely crossing the East–West divide extending

from the Baltic to the Black Sea. From around 9,000 cal. BP onwards, a more extensive network with weaker individual ties appears in the south, linking Anatolia to the rest of Europe, as early Neolithic farmer communities spread across the continent. The period 7,000–5,000 cal. BP shows more connected subnetworks of western European and eastern/northern European Neolithic farmers, while locally connected networks of HG communities prevail on the eastern side of the divide. From c. 5,000 BP onwards the divide finally collapses, and continental-wide genetic relatedness unifies large parts of western Eurasia.

of homozygosity (ROH) segments<sup>67</sup>. We detected only 29 such individuals out of a total sample of 1,396 imputed ancient genomes from across Eurasia (Extended Data Fig. 11b). This suggests that close kin mating was not common in the regions and periods covered by our data. No obviously discernible spatiotemporal or cultural clustering were observed among the individuals with recent parental relatedness. Notably, an approximately 1,700-year-old Sarmatian individual from Temyaysovo (tem003)<sup>68</sup> was found to be homozygous for almost the entirety of chromosome 2, but without evidence of ROH elsewhere in the genome, suggesting that this is the first documented case of uniparental disomy in an ancient individual (Extended Data Fig. 11c). Among several noteworthy familial relationships (see Supplementary Fig. 3c.2), we report a Mesolithic father–son burial at Ertebølle (NEO568 and NEO569), as well as a Mesolithic mother–daughter burial at Dragsholm (NEO732 and NEO733), Denmark<sup>34</sup>.

### Formation and dissolution of the divide

We have provided evidence for the existence of a clear east–west genetic division extending from the Black Sea to the Baltic, mirroring archaeological observations, and persisting for several millennia. We show that this deep ancestry division in the Eurasian human gene pool that was established during early post-LGM dispersals<sup>7</sup> was maintained throughout the Mesolithic and Neolithic ages (Fig. 6). Accordingly, we show that the genetic effect of the Neolithic transition was highly distinct east and west of this boundary. These observations raise a series of questions related to understanding the underlying drivers.

In eastern Europe, the expansion of Neolithic farming was halted for around 3,000 years, and this delay could be linked to environmental factors, with regions east of the division having more continental climates and harsher winters, possibly less suited for Middle

Eastern agricultural practices<sup>69</sup>. Here, highly developed HG societies persisted with stable, complex and sometimes fortified settlements, long-distance exchange and large cemeteries<sup>70,71</sup>. A diet including freshwater fish is clear both from our isotopic data (Supplementary Data 2) and from analyses of lipids in pottery<sup>71</sup>. In the northern forested regions of this boundary zone, HG societies persisted until the emergence of the CWC around 5,000 cal. BP, whereas in the southern and eastern steppe regions, hunting and gathering was eventually complemented with some animal husbandry (cattle and sheep), and possibly horse herding in central Asia<sup>72</sup>. Some of these groups, such as Khvalynsk at the Volga, saw the emergence of male sodalities involved in wide-ranging trade connections of copper objects from east central Europe and the Caucasus<sup>29</sup>. Settlements were confined mainly to the flat flood plains and river valleys, whereas the steppe belt remained largely unexploited.

The eventual dissolution of this genetic, economic and social border was driven by events that unfolded in the steppe region. Here, two temporal phases of technological innovations can be observed archaeologically: the widespread dispersal of ox-drawn wheeled vehicles around 5,500 cal. BP and the later development of horse riding. Combined with possible changing environmental conditions<sup>73</sup>, this opened up the steppe as an economic zone, allowing Yamnaya groups to exploit the steppe as pastoral nomads around 5,000 cal. BP (ref. 74). Eneolithic settlements along river valleys were replaced by this new mobile economy<sup>75</sup>, which finally dissolved the great genomic boundary that had persisted in the preceding millennia (Fig. 6).

By 4,000 cal. BP, the invention of chariot warfare and the adoption of millet as a food crop allowed the final eastward expansion into central Asia and beyond by the Andronovo and related groups, with global legacies for the expansion of Indo-European languages<sup>76</sup>. Our study has provided new genetic knowledge on these steppe migrations on



two levels: we have identified a hitherto-unknown source of ancestry in HGs from the Middle Don region contributing ancestry to the steppe pastoralists, and we have documented how the later spread of steppe-related ancestry into Europe through the CWC was first mediated through peoples associated with the GAC. In a contact zone that included forested northern regions, the CWC was rapidly formed from a cultural and genetic amalgamation of steppe-groups related to the Yamnaya and the GAC groups in eastern Europe. In accordance with their mixed cultural and genetic background, the CWC practised a mixed economy, using various subsistence strategies in different environments. This flexibility would have contributed substantially to their success in settling and adapting to very different ecological and climatic settings over a very short period of time<sup>33</sup>.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-06865-0>.

- Allentoft, M. E. et al. Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
- Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
- Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
- Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–24 (2016).
- Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
- Posth, C. et al. Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a late glacial population turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).
- Posth, C. et al. Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023).
- Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
- Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
- Raghavan, M. et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
- Villalba-Mouco, V. et al. Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177 (2019).
- Brace, S. et al. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat. Ecol. Evol.* **3**, 765–771 (2019).
- de Barros Damgaard, P. et al. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).
- Saag, L. et al. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci. Adv.* **7**, eabd6535 (2021).
- Günther, T. et al. Population genomics of Mesolithic Scandinavia: investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
- Kashuba, N. et al. Ancient DNA from mastics solidifies connection between material culture and genetics of mesolithic hunter-gatherers in Scandinavia. *Commun. Biol.* **2**, 185 (2019).
- Zvelebil, M., Domanska, L. & Dennell, R. *Harvesting the Sea, Farming the Forest: The Emergence of Neolithic Societies in the Baltic Region* (Bloomsbury, 1998).
- Jones, E. R. et al. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576–582 (2017).
- Mittnik, A. et al. The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
- Kislenko, A. & Tatarintseva, N. in *Late Prehistoric Exploitation of the Eurasian Steppe* (eds Levine, M. et al.) 183–216 (McDonald Institute for Archaeological Research, 1999).
- Furholt, M. Mobility and social change: understanding the European Neolithic period after the archaeological revolution. *J. Archaeol. Res.* **29**, 481–535 (2021).
- Lipson, M. et al. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science* **361**, 92–95 (2018).
- Fernandes, D. M. et al. A genomic Neolithic time transect of hunter-farmer admixture in central Poland. *Sci. Rep.* **8**, 14879 (2018).
- Immel, A. et al. Genome-wide study of a Neolithic Wartberg grave community reveals distinct HLA variation and hunter-gatherer ancestry. *Commun. Biol.* **4**, 113 (2021).
- Jeong, C. et al. The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* **3**, 966–976 (2019).
- Nikitin, A. G. et al. Interactions between earliest Linearbandkeramik farmers and central European hunter gatherers at the dawn of European Neolithization. *Sci. Rep.* **9**, 19544 (2019).
- Gelabert, P. et al. Social and genetic diversity among the first farmers of Central Europe. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.07.07.548126> (2023).
- Cassidy, L. M. et al. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl Acad. Sci. USA* **113**, 368–373 (2016).
- Penske, S. et al. Early contact between late farming and pastoralist societies in southeastern Europe. *Nature* **620**, 358–365 (2023).
- Lazaridis, I. et al. The genetic history of the Southern Arc: a bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
- Egffjord, A. F.-H. et al. Genomic Steppe ancestry in skeletons from the Neolithic Single Grave Culture in Denmark. *PLoS One* **16**, e0244872 (2021).
- Papac, L. et al. Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci. Adv.* **7**, eabi6941 (2021).
- Heyd, V. in *Rethinking Migrations in Late Prehistoric Eurasia* (eds Fernández-Götz, M. et al.) 41–62 (Oxford Univ. Press, 2023).
- Allentoft, M. E. et al. 100 ancient genomes show repeated population turnovers in Neolithic Denmark. *Nature* <https://doi.org/10.1038/s41586-023-06862-3> (2024).
- Rubinaacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* **53**, 412 (2021).
- The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
- Leslie, S. et al. The fine-scale genetic structure of the British population. *Nature* **519**, 309–314 (2015).
- Hofmanová, Z. et al. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl Acad. Sci. USA* **113**, 6886–6891 (2016).
- Busby, G. B. et al. Admixture into and within sub-Saharan Africa. *eLife* **5**, e15266 (2016).
- Schmitt, T. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front. Zool.* **4**, 11 (2007).
- Olalde, I. et al. The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
- García-Escárzaga, A. et al. Human forager response to abrupt climate change at 8.2 ka on the Atlantic coast of Europe. *Sci. Rep.* **12**, 6481 (2022).
- Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
- Wang, C.-C. et al. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat. Commun.* **10**, 590 (2019).
- Lipson, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368 (2017).
- Racimo, F. et al. The spatiotemporal spread of human migrations during the European Holocene. *Proc. Natl Acad. Sci. USA* **117**, 9898–9000 (2020).
- Martiniano, R. et al. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet.* **13**, e1006852 (2017).
- Isern, N., Zilhão, J., Fort, J. & Ammerman, A. J. Modeling the role of voyaging in the coastal spread of the Early Neolithic in the West Mediterranean. *Proc. Natl Acad. Sci. USA* **114**, 897–902 (2017).
- Betti, L. et al. Climate shaped how Neolithic farmers and European hunter-gatherers interacted after a major slowdown from 6,100 BCE to 4,500 BCE. *Nat. Hum. Behav.* **4**, 1004–1010 (2020).
- Saag, L. et al. Extensive farming in Estonia started through a sex-biased migration from the Steppe. *Curr. Biol.* **27**, 2185–2193 (2017).
- Seguin-Orlando, A. et al. Heterogeneous hunter-gatherer and Steppe-related ancestries in Late Neolithic and Bell Beaker genomes from present-day France. *Curr. Biol.* **31**, 1072–1083 (2021).
- Furholt, M. Die Złota-Gruppe in Klempolen: Ein Beispiel für die Transformation eines Zeichensystems? *Germania* **86**, 1–28 (2008).
- Szymt, M. in *A Turning of Ages* (ed. Kadrow, S.) 443–466 (Institute of Archaeology and Ethnology, Polish Academy of Sciences, 2000).
- Tassi, F. et al. Genome diversity in the Neolithic Globular Amphorae culture and the spread of Indo-European languages. *Proc. R. Soc. B* **284**, 20171540 (2017).
- Nordqvist, K. & Heyd, V. The forgotten child of the wider Corded Ware family: Russian Fatyanovo Culture in context. *Proc. Prehist. Soc.* **86**, 65–93 (2020).
- Kristiansen, K. et al. Re-theorising mobility and the formation of culture and language among the Corded Ware Culture in Europe. *Antiquity* **91**, 334–347 (2017).
- Borzunov, V. A. The neolithic fortified settlements of the Western Siberia and Trans-Urals. *Russ. Archaeol.* **4**, 20–34 (2013).
- Yu, H. et al. Paleolithic to Bronze Age Siberians reveal connections with First Americans and across Eurasia. *Cell* **181**, 1232–1245 (2020).
- Okladnikov, A. P. *Neolit i Bronzovyy vek Pribaikaliya [Neolithic and Bronze Age of the Baikal region]* (AS USSR Publications, 1950).
- Merts, V. in *Paleodemography and Migration Processes in Western Siberia in Antiquity and the Middle Ages* (ed. Kiryushin, Y. F.) 39–42 (Altai State University, 1994).
- Merts, V. *Periodization of the Holocene Complexes of Northern and Central Kazakhstan Based on the Materials of the Multilayer Site Shiderty 3* (Thesis, Kemerovo State Univ., 2008).
- Merts, V. Neolithization processes in the Northeast Kazakhstan. *Herald Omsk Univ. Ser. Histor. Stud.* **3**, 99–109 (2018).
- de Barros Damgaard, P. et al. 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369–374 (2018).
- Librado, P. et al. The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* **598**, 634–640 (2021).
- Huang, Y. et al. The early adoption of East Asian crops in West Asia: rice and broomcorn millet in northern Iran. *Antiquity* **97**, 674–689 (2023).
- Palamara, P. F., Lencz, T., Darvasi, A. & Pe'er, I. Length distributions of identity by descent reveal fine-scale demographic history. *Am. J. Hum. Genet.* **91**, 809–822 (2012).
- Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).
- Krzywińska, M. et al. Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci. Adv.* **4**, eaat4457 (2018).

69. Matuzeviciute, G. M. The possible geographic margin effect on the delay of agriculture introduction in the East Baltic. *Eston. J. Archaeol.* **22**, 149–162 (2018).
70. Piezonka, H. *Jäger, Fischer, Topfer: Wildbeuterguppen mit Früher Keramik in Nordosteuropa im 6. und 5. Jahrtausend v. Chr.* (Habelt, 2015).
71. Oras, E. et al. The adoption of pottery by north-east European hunter-gatherers: Evidence from lipid residue analysis. *J. Archaeol. Sci.* **78**, 112–119 (2017).
72. Matuzeviciute, G. M. et al. Archaeobotanical investigations at the earliest horse herder site of Botai in Kazakhstan. *Archaeol. Anthropol. Sci.* **11**, 6243–6258 (2019).
73. Anthony, D. W. in *The Black Sea Flood Question: Changes in Coastline, Climate and Human Settlement* (eds Yanko-Hombach, V. et al.) 345–370 (Springer, 2007).
74. Trautmann, M. et al. First bioanthropological evidence for Yamnaya horsemanship. *Sci. Adv.* **9**, eade2451 (2023).
75. Anthony, D. W. et al. The Eneolithic cemetery at Khvalynsk on the Volga River. *Præhistor. Zeitschr.* **97**, 22–67 (2022).
76. Kristiansen, K., Kroonen, G. & Willerslev, E. *The Indo-European Puzzle Revisited: Integrating Archaeology, Genetics, and Linguistics* (Cambridge Univ. Press, 2023).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

**Morten E. Allentoft**<sup>1,2,113</sup>✉, **Martin Sikora**<sup>1,113</sup>✉, **Alba Refoyo-Martínez**<sup>1,113</sup>, **Evan K. Irving-Pease**<sup>1,113</sup>, **Anders Fischer**<sup>1,3,4,113</sup>, **William Barrie**<sup>5,6,113</sup>, **Andrés Ingason**<sup>1,7,113</sup>, **Jesper Stenderup**<sup>1</sup>, **Karl-Göran Sjögren**<sup>8</sup>, **Alice Pearson**<sup>6</sup>, **Bárbara Sousa da Mota**<sup>6,9</sup>, **Bettina Schulz Paulsson**<sup>3</sup>, **Alma Halgren**<sup>10</sup>, **Ruairidh Macleod**<sup>1,5,11,12</sup>, **Marie Louise Schjellerup Jørgov**<sup>13</sup>, **Fabrice Demeter**<sup>1,14</sup>, **Lasse Sørensen**<sup>15</sup>, **Poul Otto Nielsen**<sup>15</sup>, **Rasmus A. Henriksen**<sup>1</sup>, **Tharsika Vimala**<sup>1</sup>, **Hugh McColl**<sup>1</sup>, **Ashot Margaryan**<sup>16,17</sup>, **Melissa Ilardo**<sup>18</sup>, **Andrew Vaughn**<sup>19</sup>, **Morten Fischer Mortensen**<sup>19</sup>, **Anne Birgitte Nielsen**<sup>20</sup>, **Mikkel Ulfeldt Hede**<sup>21</sup>, **Niels Nørkjær Johannsen**<sup>22</sup>, **Peter Rasmussen**<sup>15</sup>, **Lasse Vinner**<sup>1</sup>, **Gabriel Renaud**<sup>23</sup>, **Aaron Stern**<sup>19</sup>, **Thies Zetner Trolle Jensen**<sup>16</sup>, **Gabriele Scorrano**<sup>1</sup>, **Hannes Schroeder**<sup>16</sup>, **Per Lysdahl**<sup>24</sup>, **Abigail Daisy Ramsøe**<sup>1</sup>, **Andrei Skorobogatov**<sup>25</sup>, **Andrew Joseph Schork**<sup>6,26</sup>, **Anders Rosengren**<sup>16</sup>, **Anthony Ruter**<sup>1</sup>, **Alan Outram**<sup>27</sup>, **Aleksey A. Timoshenko**<sup>28</sup>, **Alexandra Buzhilova**<sup>29</sup>, **Alfredo Coppa**<sup>30</sup>, **Alisa Zubova**<sup>31</sup>, **Ana Maria Silva**<sup>32,33</sup>, **Anders J. Hansen**<sup>1</sup>, **Andrey Gromov**<sup>31</sup>, **Andrey Logvin**<sup>34</sup>, **Anne Birgitte Gotfredsen**<sup>1</sup>, **Bjarne Henning Nielsen**<sup>35</sup>, **Borja González-Rabanal**<sup>36</sup>, **Carles Lalueza-Fox**<sup>37,38</sup>, **Catriona J. McKenzie**<sup>27</sup>, **Charleen Gaunitz**<sup>3</sup>, **Concepción Blasco**<sup>39</sup>, **Corina Liesau**<sup>39</sup>, **Cristina Martínez-Labarga**<sup>40</sup>, **Dmitri V. Pozdnyakov**<sup>28</sup>, **David Cuenca-Solana**<sup>41,42</sup>, **David O. Lordkipanidze**<sup>43,44</sup>, **Dmitri En'shin**<sup>45</sup>, **Domingo C. Salazar-García**<sup>46,47</sup>, **T. Douglas Price**<sup>3,48</sup>, **Dušan Boric**<sup>30,49</sup>, **Elena Kostyleva**<sup>50</sup>, **Elizaveta V. Veselovskaya**<sup>51</sup>, **Emma R. Usmanova**<sup>52,53,54,55</sup>, **Enrico Cappellini**<sup>19</sup>, **Erik Brinch Petersen**<sup>56</sup>, **Esben Kannegaard**<sup>57,14</sup>, **Francesca Radina**<sup>58</sup>, **Fulya Eylem Yediyar**<sup>1</sup>, **Henri Duday**<sup>59</sup>, **Igor Gutiérrez-Zugasti**<sup>60</sup>, **Ilya Merts**<sup>60</sup>, **Inna Potekhina**<sup>61,62</sup>, **Irina Shevnina**<sup>34</sup>, **Isin Altinkaya**<sup>1</sup>, **Jean Guislaine**<sup>63</sup>, **Jesper Hansen**<sup>64</sup>, **Joan Emili Aura Tortosa**<sup>46</sup>, **João Zilhão**<sup>33,65</sup>, **Jorge Vega**<sup>66</sup>, **Kristoffer Buck Pedersen**<sup>67</sup>, **Krzysztof Tunia**<sup>68</sup>, **Lei Zhao**<sup>1</sup>, **Liudmila N. Mylnikova**<sup>28</sup>, **Lars Larsson**<sup>69</sup>, **Laure Metz**<sup>70</sup>, **Levon Yepiskoposyan**<sup>1,72</sup>, **Lisbeth Pedersen**<sup>73</sup>, **Lucia Sarti**<sup>74</sup>, **Ludovic Orlando**<sup>75</sup>, **Ludovic Slimak**<sup>75</sup>, **Lutz Klassen**<sup>7</sup>, **Malou Blank**<sup>3</sup>, **Manuel González-Morales**<sup>41</sup>, **Mara Silvestrini**<sup>76</sup>, **Maria Vretemark**<sup>77</sup>, **Marina S. Nesterova**<sup>26</sup>, **Marina Rykun**<sup>78</sup>, **Mario Federico Rolfó**<sup>79</sup>, **Marzena Szymyt**<sup>80</sup>, **Marcin Przybyła**<sup>81</sup>, **Mauro Calatriní**<sup>74</sup>, **Mikhail Sablin**<sup>82</sup>, **Miluše Dobisiková**<sup>83</sup>, **Morten Meldgaard**<sup>84</sup>, **Morten Johansen**<sup>85</sup>, **Natalia Berezina**<sup>29</sup>, **Nick Card**<sup>86</sup>, **Nikolai A. Saveliev**<sup>87</sup>, **Olga Poshekhonova**<sup>15</sup>, **Olga Rickards**<sup>40</sup>, **Olga V. Lozovskaya**<sup>88</sup>, **Oliver Gábor**<sup>89</sup>, **Otto Christian Uldum**<sup>85,90</sup>, **Paola Aurino**<sup>91</sup>, **Pavel Kosintsev**<sup>92,93</sup>, **Patrice Courtaud**<sup>59</sup>, **Patricia Rios**<sup>39</sup>, **Peder Mortensen**<sup>94,114</sup>, **Per Lotz**<sup>95,96</sup>, **Per Persson**<sup>97</sup>, **Pernille Bangsgaard**<sup>98</sup>, **Peter de Barros Damgaard**<sup>1</sup>, **Peter Vang Petersen**<sup>15</sup>, **Pilar Prieto Martinez**<sup>99</sup>, **Piotr Włodarczak**<sup>68</sup>, **Roman V. Smolyaninov**<sup>100</sup>, **Rikke Maring**<sup>23,57</sup>, **Roberto Mendúña**<sup>66</sup>, **Ruben Badalyan**<sup>101</sup>, **Rune Iversen**<sup>86</sup>, **Ruslan Turin**<sup>25</sup>, **Sergey Vasilyev**<sup>51,102</sup>, **Sidsel Wählin**<sup>24</sup>, **Svetlana Borutskaya**<sup>29</sup>, **Svetlana Skochina**<sup>95</sup>, **Søren Anker Sørensen**<sup>95</sup>, **Søren H. Andersen**<sup>103</sup>, **Thomas Jørgensen**<sup>95</sup>, **Yuri B. Serikov**<sup>104</sup>, **Vyacheslav I. Molodin**<sup>28</sup>, **Vaclav Smrcka**<sup>105</sup>, **Victor Merts**<sup>106</sup>, **Vivek Appadurai**<sup>16</sup>, **Vyacheslav Moiseyev**<sup>31</sup>, **Yvonne Magnusson**<sup>107</sup>, **Kurt H. Kjær**<sup>1</sup>, **Niels Lynnerup**<sup>13</sup>, **Daniel J. Lawson**<sup>108</sup>, **Peter H. Sudmant**<sup>109</sup>, **Simon Rasmussen**<sup>109</sup>, **Thorfinn Sand Korneliusen**<sup>115</sup>, **Richard Durbin**<sup>61,10,115</sup>, **Rasmus Nielsen**<sup>110,115</sup>, **Olivier Delaneau**<sup>8,8,115</sup>, **Thomas Werge**<sup>16,111,115</sup>, **Fernando Racimo**<sup>115</sup>, **Kristian Kristiansen**<sup>1,3,115</sup> & **Eske Willerslev**<sup>1,5,112,115</sup>✉

<sup>1</sup>Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>2</sup>Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Perth, Western Australia, Australia. <sup>3</sup>Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden. <sup>4</sup>Sealand Archaeology, Kalundborg, Denmark. <sup>5</sup>GeoGenetics Group, Department of Zoology, University of Cambridge, Cambridge, UK. <sup>6</sup>Department of Genetics, University of Cambridge, Cambridge, UK. <sup>7</sup>Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Roskilde, Denmark. <sup>8</sup>Department of Computational Biology,

University of Lausanne, Lausanne, Switzerland. <sup>9</sup>Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland. <sup>10</sup>Department of Integrative Biology, University of California, Berkeley, CA, USA. <sup>11</sup>Research Department of Genetics, Evolution and Environment, University College London, London, UK. <sup>12</sup>Department of Archaeology, University of Cambridge, Cambridge, UK. <sup>13</sup>Laboratory of Biological Anthropology, Department of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark. <sup>14</sup>Muséum National d'Histoire Naturelle, CNRS, Université de Paris, Musée de l'Homme, Paris, France. <sup>15</sup>The National Museum of Denmark, Copenhagen, Denmark. <sup>16</sup>Section for Evolutionary Genomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>17</sup>Centre for Evolutionary Hologenomics, University of Copenhagen, Copenhagen, Denmark. <sup>18</sup>Anthropology Department, University of Utah, Salt Lake City, UT, USA. <sup>19</sup>Center for Computational Biology, University of California, Berkeley, CA, USA. <sup>20</sup>Department of Geology, Lund University, Lund, Sweden. <sup>21</sup>Tårnby Gymnasium og HF, Kastrup, Denmark. <sup>22</sup>Department of Archaeology and Heritage Studies, Aarhus University, Aarhus, Denmark. <sup>23</sup>Department of Health Technology, Section of Bioinformatics, Technical University of Denmark, Kongens Lyngby, Denmark. <sup>24</sup>Vendsyssel Historiske Museum, Hjørring, Denmark. <sup>25</sup>Terra Ltd., Voronezh, Russian Federation. <sup>26</sup>Neurogenomics Division, The Translational Genomics Research Institute (TGEN), Phoenix, AZ, USA. <sup>27</sup>Department of Archaeology, University of Exeter, Exeter, UK. <sup>28</sup>Institute of Archeology and Ethnography, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation. <sup>29</sup>Department of Anthropology, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russian Federation. <sup>30</sup>Department of Environmental Biology, Sapienza University of Rome, Rome, Italy. <sup>31</sup>Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint Petersburg, Russian Federation. <sup>32</sup>CIAS, Department of Life Sciences, University of Coimbra, Coimbra, Portugal. <sup>33</sup>UNIARQ, University of Lisbon, Lisbon, Portugal. <sup>34</sup>Kostanay Regional University A. Baitursynov, Kostanay, Kazakhstan. <sup>35</sup>Vesthimmerlands Museum, Aars, Denmark. <sup>36</sup>Grupo EvoAdapta, Departamento de Ciencias Históricas, Universidad de Cantabria, Santander, Spain. <sup>37</sup>Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain. <sup>38</sup>Natural Sciences Museum of Barcelona (MCNB), Barcelona, Spain. <sup>39</sup>Departamento de Prehistoria y Arqueología, Universidad Autónoma de Madrid, Madrid, Spain. <sup>40</sup>Department of Biology, University of Rome Tor Vergata, Rome, Italy. <sup>41</sup>Instituto Internacional de Investigaciones Prehistóricas de Cantabria, Universidad de Cantabria, Banco Santander, Gobierno de Cantabria, Santander, Spain. <sup>42</sup>Centre de Recherche en Archéologie, Archeosciences, Histoire (CRéAAH), UMR-6869 CNRS, Rennes, France. <sup>43</sup>Georgian National Museum, Tbilisi, Georgia. <sup>44</sup>Tbilisi State University, Tbilisi, Georgia. <sup>45</sup>IPND, Tyumen Scientific Centre, Siberian Branch of the Russian Academy of Sciences, Tyumen, Russian Federation. <sup>46</sup>Departament de Prehistòria, Arqueologia i Història Antiga, Universitat de València, València, Spain. <sup>47</sup>Department of Geological Sciences, University of Cape Town, Cape Town, South Africa. <sup>48</sup>Laboratory for Archaeological Chemistry, Department of Anthropology, University of Wisconsin-Madison, Madison, WI, USA. <sup>49</sup>Department of Anthropology, New York University, New York, NY, USA. <sup>50</sup>Institute of Humanities, Ivanovo State University, Ivanovo, Russian Federation. <sup>51</sup>Institute of Ethnology and Anthropology, Russian Academy of Sciences, Moscow, Russian Federation. <sup>52</sup>Saryarka Archaeological Institute, Buketov Karaganda University, Karaganda, Kazakhstan. <sup>53</sup>South Ural State University, Chelyabinsk, Russia. <sup>54</sup>A. Kh. Khalikov Institute of Archeology of the Academy of Sciences of the Republic of Tatarstan, Kazan, Russia. <sup>55</sup>Margulan Institute of Archaeology, Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Almaty, Kazakhstan. <sup>56</sup>The Saxo Institute, University of Copenhagen, Copenhagen, Denmark. <sup>57</sup>Museum Østjylland, Randers, Denmark. <sup>58</sup>Soprintendenza Archeologia Belle Arti e Paesaggio per la Città Metropolitana di Bari, Bari, Italy. <sup>59</sup>UMR 5199 PACEA, CNRS, Université de Bordeaux, Pessac, France. <sup>60</sup>A. Kh. Margulan Institute of Archaeology, Almaty, Kazakhstan. <sup>61</sup>Institute of Archaeology, National Academy of Sciences of Ukraine, Kyiv, Ukraine. <sup>62</sup>National University of Kyiv-Mohyla Academy, Kyiv, Ukraine. <sup>63</sup>Collège de France, Paris, France. <sup>64</sup>Svendborg Museum, Svendborg, Denmark. <sup>65</sup>ICREA, University of Barcelona, Barcelona, Spain. <sup>66</sup>ARGEA Consultores SL, Madrid, Spain. <sup>67</sup>Museum Sydøstdanmark, Vordingborg, Denmark. <sup>68</sup>Institute of Archaeology and Ethnology, Polish Academy of Sciences, Kraków, Poland. <sup>69</sup>Department of Archaeology and Ancient History, Lund University, Lund, Sweden. <sup>70</sup>Aix-Marseille Université, CNRS, Min. Culture, UMR 7269, LAMPEA, Maison Méditerranéenne des Sciences de l'Homme, Aix-en-Provence, France. <sup>71</sup>Institute of Molecular Biology, National Academy of Sciences, Yerevan, Armenia. <sup>72</sup>Institute of Archaeology and Ethnography, National Academy of Sciences, Yerevan, Armenia. <sup>73</sup>HistorieUdvikler, Kalundborg, Denmark. <sup>74</sup>Department of History and Cultural Heritage, University of Siena, Siena, Italy. <sup>75</sup>Centre d'Anthropologie et de Génétique de Toulouse, CNRS UMR 5500, Université Paul Sabatier, Toulouse, France. <sup>76</sup>Soprintendenza per i Beni Archeologici delle Marche, Ancona, Italy. <sup>77</sup>Västergötlands Museum, Skara, Sweden. <sup>78</sup>Cabinet of Anthropology, Tomsk State University, Tomsk, Russian Federation. <sup>79</sup>Department of History, Humanities and Society, University of Rome Tor Vergata, Rome, Italy. <sup>80</sup>Faculty of Archaeology, Adam Mickiewicz University in Poznań, Poznań, Poland. <sup>81</sup>Institute of Archaeology, Jagiellonian University, Kraków, Poland. <sup>82</sup>Zoological Institute of Russian Academy of Sciences, Saint Petersburg, Russian Federation. <sup>83</sup>Department of Anthropology, Czech National Museum, Prague, Czech Republic. <sup>84</sup>Department of Health and Nature, University of Greenland, Nuuk, Greenland. <sup>85</sup>The Viking Ship Museum, Roskilde, Denmark. <sup>86</sup>Archaeology Institute, University of Highlands and Islands, Orkney, UK. <sup>87</sup>Scientific Research Center "Baikal region", Irkutsk State University, Irkutsk, Russian Federation. <sup>88</sup>Laboratory for Experimental Traceology, Institute for the History of Material Culture of the Russian Academy of Sciences, Saint Petersburg, Russian Federation. <sup>89</sup>Janus Pannonius Museum, Pécs, Hungary. <sup>90</sup>Langelands Museum, Rudkøbing, Denmark. <sup>91</sup>Soprintendenza Archeologia, Belle Arti e Paesaggio per la provincia di Cosenza, Cosenza,

Italy. <sup>92</sup>Paleoecology Laboratory, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russian Federation. <sup>93</sup>Department of History of the Institute of Humanities, Ural Federal University, Ekaterinburg, Russian Federation. <sup>94</sup>Centre for the Study of Early Agricultural Societies, Department of Cross-Cultural and Regional Studies, University of Copenhagen, Copenhagen, Denmark. <sup>95</sup>Museum Nordsjælland, Hillerød, Denmark. <sup>96</sup>Museum Vestsjælland, Holbæk, Denmark. <sup>97</sup>Museum of Cultural History, University of Oslo, Oslo, Norway. <sup>98</sup>ArchaeoScience, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>99</sup>Department of History, University of Santiago de Compostela, Santiago de Compostela, Spain. <sup>100</sup>Lipetsk Regional Scientific Public Organisation "Archaeological Research", Lipetsk, Russian Federation. <sup>101</sup>Center for Egyptological Studies, Russian Academy of Sciences, Moscow, Russian Federation. <sup>102</sup>Russian-Armenian University, Yerevan, Armenia. <sup>103</sup>Moesgaard Museum, Højbjerg, Denmark. <sup>104</sup>Nizhny Tagil State Socio-Pedagogical Institute, Nizhny Tagil, Russia. <sup>105</sup>Institute for History

of Medicine, First Faculty of Medicine, Charles University, Prague, Czech Republic. <sup>106</sup>Centre for Archaeological Research, Toraighyrov University, Pavlodar, Kazakhstan. <sup>107</sup>Malmö Museer, Malmö, Sweden. <sup>108</sup>Institute of Statistical Sciences, School of Mathematics, University of Bristol, Bristol, UK. <sup>109</sup>Novo Nordisk Foundation Centre for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. <sup>110</sup>Wellcome Sanger Institute, Hinxton, UK. <sup>111</sup>Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. <sup>112</sup>MARUM Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Bremen, Germany. <sup>113</sup>These authors contributed equally: Morten E. Allentoft, Martin Sikora, Alba Refoyo-Martínez, Evan K. Irving-Pease, Anders Fischer, William Barrie, Andrés Ingason. <sup>114</sup>Deceased: Esben Kannegaard, Peder Mortensen. <sup>115</sup>These authors jointly supervised this work: Thorfinn Korneliussen, Richard Durbin, Rasmus Nielsen, Olivier Delaneau, Thomas Werge, Fernando Racimo, Kristian Kristiansen, Eske Willerslev. <sup>✉</sup>e-mail: morten.allentoft@curtin.edu.au; martin.sikora@sund.ku.dk; ew482@cam.ac.uk

## Methods

**Generation and authentication of ancient DNA data**

Sampling of ancient human remains was undertaken in collaboration with co-authors responsible for the curation and contextual analyses of these, and with the approval of the relevant institutions responsible for the archaeological remains (detailed in the Reporting Summary). Laboratory work was undertaken in dedicated ancient DNA clean-lab facilities (Globe Institute, University of Copenhagen) following optimized ancient DNA protocols<sup>1,77</sup> (Supplementary Note 1). Double-stranded blunt-end libraries were constructed from the extracted DNA using NEBNext DNA Prep Master Mix Set E6070 (New England Biolabs) and sequenced (80 bp and 100 bp single read) on Illumina HiSeq 2500 and 4000 platforms. Initial shallow shotgun screening identified 317 of 962 ancient samples with sufficient DNA preservation for deeper sequencing. Of these, 211 were teeth, 91 were petrous bones and 15 were sampled from long bones, ribs and cranial bones (Supplementary Data 2). Reads were mapped to the human reference genome build 37 and also to the mitochondrial genome (rCRS) alone. Mapped reads were filtered for mapping quality 30 and sorted using Picard (v.1.127) (<http://picard.sourceforge.net>) and SAMtools<sup>78</sup>. Data were merged to library level and duplicates were removed using Picard MarkDuplicates (v.1.127) and merged to sample level. Sample-level BAMs were re-aligned using GATK (v.3.3.0) and hereafter had the md-tag updated and extended BAQs calculated using samtools calmd (v.1.10)<sup>78</sup>. Read depth and coverage were determined using pysam (<https://github.com/pysam-developers/pysam>) and BEDtools (v.2.23.0)<sup>79</sup>. Post-mortem DNA damage patterns were determined using mapDamage2.0 (ref. 80). For the 317 samples we observed C-to-T deamination fractions ranging from 10.4% to 67.8%, with an average of 38.3% across all samples (Supplementary Data 1). These numbers indicate DNA-molecule degradation consistent with a millennia-scale depositional age. Three methods were used to estimate DNA contamination: two based on mitochondrial sequences<sup>81,82</sup> and one method investigating X-chromosomal data in males (ANGSD, Supplementary Note 1). All contamination estimates are reported in Supplementary Data 5 (summary values in Supplementary Data 1). On the basis of this approach, we had a total of 15 samples flagged as ‘possibly contaminated’ in our downstream analyses (Supplementary Note 1).

**Imputation of ancient genomes**

We imputed the ancient genomes in this study using the imputation and phasing tool GLIMPSE v.1.0.0 (ref. 35) and 1000 Genomes phase 3 (ref. 36) as a reference panel. We first generated genotype likelihoods at the biallelic 1000 Genomes variant sites from the bam files with bcftools v.1.10 and the command `bcftools mpileup` with parameters `-l -E -a 'FORMAT/DP' --ignore-RG`, followed by `bcftools call -Aim -C alleles`. Using GLIMPSE\_chunk, the genotype likelihood data were first split into chunks of sizes between 1 and 2 Mb with a buffer region of 200 kb at each side. We then imputed each chunk with GLIMPSE\_phase with parameters `--burn 10, --main 15 and --pbwt-depth 2`. Finally, the imputed chunks were ligated with GLIMPSE\_ligate. To validate the accuracy of the imputation, 42 high-coverage (5× to 39×) genomes, including a Neolithic trio, were downsampled for testing<sup>83</sup> (Supplementary Note 2). We evaluated imputation accuracy on the basis of depth of coverage; MAF; and ancestry and time frame of ancient genomes, using high-coverage ancient genomes<sup>83</sup>. Genomes with higher than 1× coverage provided a notably high imputation accuracy (closely matching that obtained for modern samples; Extended Data Fig. 2), except for African genomes, which had lower accuracy owing to the poor representation of this ancestry in the reference panel. Imputation accuracy was influenced by both MAF and coverage (Supplementary Fig. 2.3). We found that coverage as low as 0.1× and 0.4× was sufficient to obtain  $r^2$  imputation accuracies of 0.8 and 0.9 at common variants (MAF ≥ 10%), respectively. We conclude that ancient genomes can be imputed confidently from coverages above 0.4×, and that genome-wide aggregate

analyses relying on common SNPs (for example, PCA and admixture modelling) can be performed with a low amount of bias for genome coverage from as low as 0.1× when using specific quality control on the imputed data (although at very low coverage a bias arises towards the major allele; see Supplementary Note 2). We also tested for possible effects of bias affecting inferred ancestry components<sup>83</sup> propagating biases in individual-level pairwise analyses, using D-statistics, which indicated that imputed ancient genomes down to 0.1× coverage are not significantly affected (Supplementary Note 2).

**Demographic inference**

We determined the genetic sex of the study individuals using the ratio of reads aligning to either of the sex chromosomes ( $R_Y$  statistic)<sup>84</sup>. Y chromosomes of inferred male individuals were further analysed using phylogenetic placement<sup>85</sup>. We built a reference phylogenetic tree of 1,244 male individuals from the 1000 Genomes project with RAXML-NG (ref. 86), using the general time-reversible model including among-site rate heterogeneity and ascertainment correction (model GTR+G+ASC\_LEWIS). For each ancient sample, haploid genotypes given the positions and alleles in the reference panel were called using ‘bcftools call’ (options `-C alleles -ploidy 1 -i`). The resulting genotypes were converted to fasta format and placed onto the reference tree using EPA-ng (ref. 85). Phylogenetic placements were processed and visualized using gappa (ref. 87). To convert phylogenetic placements into haplogroup calls, we assigned each branch of the reference phylogeny to its representing haplogroup, using SNP annotations from ISOGG (v.15.73). For each ancient sample, haplogroups were then called using the most basal branch accumulating 99% of the placement weights, obtained using ‘accumulate’ in gappa. Phylogenetic analyses of reconstructed mitochondrial genomes were also undertaken using RAXML-ng (ref. 85; see Supplementary Note 3a).

To infer genetic relatedness between the study individuals, we used the allele-frequency-free inference method introduced previously<sup>88</sup>. For each pair of individuals, three relatedness estimators were calculated, R0, R1 and KING-robust (ref. 89) using the site-frequency-spectrum (SFS)-based approach. We used the realSFS method<sup>90</sup> implemented in the ANGSD package<sup>91</sup> to infer the 2D-SFS, selecting the SFS with the highest likelihood across ten replicates. We used a set of 1,191,529 autosomal transversion SNPs with MAF ≥ 0.05 from the 1000 Genomes Project<sup>36</sup> for the analysis. Previously established cut-offs<sup>89</sup> for the KING-robust estimator were applied to assign individual pairs to first-, second- or third-degree relationships. Parent-offspring relationships were distinguished from sibling relationships using R0 and R1 ratios, by requiring that  $R0 \leq 0.02$  and  $0.4 \leq R1 \leq 0.6$  to infer a parent-offspring relative pair. Individual pairs with fewer than 20,000 sites contributing to the estimators were excluded.

We generated a dataset for population genetic analysis by combining the 317 newly sequenced individuals with 1,347 previously published ancient genomes with genomic coverage higher than 0.1× generated using shotgun sequencing (Supplementary Data 7). Imputed genotype data (Supplementary Note 2) for this set of 1,664 ancient genomes were merged with genotypes of 2,504 modern individuals from the 1,000 Genomes project<sup>36</sup> used as a reference panel in the imputation. We retained only SNPs that passed the 1000 Genomes strict mask, resulting in a final dataset of 4,168 individuals genotyped at 7,321,965 autosomal SNPs (‘1000G’ dataset). As well as imputed genotypes, we also generated pseudo-haploid genotypes for each ancient individual by randomly sampling an allele from sequencing reads covering those SNPs. For population structure analyses in the context of global genetic diversity, we generated a second dataset by intersecting the ancient genotype data with SNP array data of 2,180 modern individuals from 213 worldwide populations<sup>3,4,92,93</sup> (‘HO’ dataset).

To facilitate filtering for downstream analyses, we flagged individuals to potentially exclude according to the following criteria: (i) contamination estimate greater than 5% (‘contMT5pct’, ‘contNuc5pct’;

Supplementary Note 1); (ii) autosomal coverage less than 0.1× ('lowcov'); (iii) genome-wide average imputation genotype probability less than 0.98 ('lowGpAvg'); (iv) individual is the lower-quality sample in a close relative pair ('1d\_rel', '2d\_rel'; Supplementary Note 3c). A total of 1,492 individuals (213 newly reported) passed all filters, which were used in most of the downstream analyses unless otherwise noted.

We investigated overall population structure among the dataset individuals using PCA and model-based clustering (ADMIXTURE<sup>94</sup>). We performed PCA using different subsets of individuals in the 'HO' dataset. For the PCA including only imputed diploid samples, we used GCTA (ref. 95), excluding SNPs with MAF < 0.05 in the respective panel. For PCA projecting low coverage or flagged individuals, we used smartpca (refs. 96,97) with options 'lsqproject: YES' and 'autoshrink: YES' on a fixed set of 400,186 SNPs with MAF ≥ 0.05 in non-African individuals passing all filters. We ran ADMIXTURE on a set of 1,593 ancient individuals from the '1000G' dataset, excluding individuals flagged as close relatives or with a contamination estimate greater than 5%. For the 1,492 individuals passing all filters we used imputed genotypes; the remaining 101 lower-coverage samples were represented by pseudo-haploid genotypes. We restricted the analysis to transversion SNPs with imputation INFO score ≥ 0.8 and MAF ≥ 0.05. We further performed linkage-disequilibrium pruning and filtering for missingness using plink<sup>98</sup> (options --indep-pairwise 500 50 0.4 -geno 0.8), for a final analysis set of 142,550 SNPs.

We performed admixture graph fitting (qpGraph) to investigate deep Eurasian population structure using ADMIXTOOLS2 (ref. 99). For these analyses, pairwise  $f_2$ -statistics were pre-computed from pseudo-haploid genotypes in the '1000G' dataset using the 'extract\_f2' function with 'afProd=TRUE'. We grouped individuals into populations using their membership in the genetic clusters inferred from IBD sharing (Supplementary Note 3f), with the exception of the Upper Palaeolithic European individual Kostenki 14, who was treated as a separate population (new cluster label 'Europe\_37000BP\_33000BP\_Kostenki'). We carried out admixture graph fitting using a semi-automatic iterative approach (Supplementary Note 3d).

We used IBDseq<sup>100</sup> to detect genomic segments shared IBD between all individuals in the '1000G' dataset, restricting to transversion SNPs with imputation INFO score ≥ 0.8 and MAF ≥ 0.01. We filtered the resulting IBD segments for LOD score ≥ 3 and a minimum length of 2 centimorgans (cM), and further removed regions of excess long IBD as described previously<sup>101</sup>. First, we used the GenomicRanges<sup>102</sup> package in R to calculate the total number of long IBD segments (greater than 10 cM) overlapping each position along the genome, and calculated their 3% trimmed mean and s.d. We then called regions of excess IBD if they were more than 10 trimmed s.d. from the trimmed mean, and removed any segments overlapping the excess IBD regions. For analyses of ROH we used a shorter length cut-off of 1 cM.

We performed genetic clustering of the ancient individuals using hierarchical community detection on a network of pairwise IBD-sharing similarities<sup>103</sup>. To facilitate the detection of clusters at a finer scale, we ran IBDseq (v.r1206) on a dataset restricting to ancient samples only, and applied more lenient filters of imputation INFO score > 0.5, and minimum IBD segment length of 1 cM. We constructed a weighted network of the individuals using the igraph<sup>104</sup> package in R, with the fraction of the genome shared IBD between pairs of individuals as weights. We then performed iterative community detection on this network using the Leiden algorithm<sup>105</sup> implemented in the leidenAlgo R package (v1.01; <https://github.com/kharchenkolab/leidenAlgo>). We used a resolution parameter of  $r = 0.5$  as the starting value for each level of community detection. If more than one community was detected, we split the network into the respective communities, and repeated the community detection step. If no communities were detected, we incremented the resolution parameter in steps of 0.5 until a maximum value of  $r = 3$ . The initial clustering was completed when no more communities were detected at the highest resolution parameter, across

all subcommunities. To convert the resulting hierarchy into a final clustering, we simplified the initial clustering by collapsing nodes into single clusters on the basis of observed spatiotemporal annotations of the samples. We note that the obtained clusters should not be interpreted as 'populations' in the sense of a local community of individuals, but rather as sets of individuals with shared ancestry. Although this approach is an oversimplification of the complex spatiotemporally structured populations investigated here, the obtained clusters nevertheless captured real effects, grouping individuals within restricted spatiotemporal ranges and/or archaeological contexts and recapitulating known relationships between clusters.

To circumvent some of the pitfalls of grouping individuals into discrete clusters, we used supervised ancestry modelling in which sets of 'target' individuals were modelled as mixtures of 'source' groups, selected to represent particular ancestry components. As an illustrative case, an individual of European HG ancestry with a minor contribution of Neolithic farmer admixture might be inferred to be a member of a HG genetic cluster, but will be modelled as a mixture of a HG and Neolithic farmer sources in the ancestry modelling. To estimate ancestry proportions from patterns of pairwise IBD sharing, we applied an approach akin to 'chromosome painting'<sup>106</sup>. We first inferred an IBD-based 'painting profile' for each target individual, by summing up the total amount of IBD shared with each 'donor' group (using population labels for modern donors or IBD-based genetic clusters for ancient donors), and normalizing them to the interval [0,1]. We used a leave-one-out approach<sup>38</sup> to account for the fact that recipient individuals cannot be included as donors from their own group. We then used these painting profiles in supervised modelling of target individuals as mixtures from different sets of putative source groups<sup>38,107</sup>, using non-negative least squares implemented in the R package limSolve<sup>108</sup>. We estimated standard errors of ancestry proportions using a weighted block jackknife, leaving out each chromosome in turns. A comparison of results obtained using this approach to other commonly used methods (supervised ADMIXTURE, qpAdm) is shown in Supplementary Note 3f). We focused our analyses on three panels of putative source clusters reflecting different temporal depths: 'deep', using a set of deep ancestry source groups reflecting major ancestry poles; 'postNeol', using diverse Neolithic and earlier source groups; and 'postBA', using Late Neolithic and Bronze Age source groups (Extended Data Figs. 5–7). We also used additional source sets in follow-up analyses of more restricted spatiotemporal contexts (Supplementary Data 7–13).

Finally, we aimed to infer the geographical and temporal spread of major ancestries (Supplementary Note 3e). We used a method<sup>46</sup> applying spatiotemporal ordinary kriging on latent ancestry proportion estimates from ancient and present-day genomes. This way, we obtained spatiotemporal maps reflecting the dynamics of the spread of ancestry during the transition from the Mesolithic to the Neolithic, Bronze Age, Iron Age and more recent periods. We obtained ancestry proportions estimated using ADMIXTURE<sup>109</sup> with  $K = 9$  latent ancestry clusters (Supplementary Note 3d) on a sequence dataset including both whole-genome shotgun-sequenced genomes and genomic sequences obtained through SNP capture (Supplementary Note 2, intersection with 'HO' dataset). We performed spatiotemporal kriging<sup>110</sup> of these proportions over the last 12,900 years, in intervals of 300 years, with a 5,000-point spatial grid spanning western and central Eurasia. We used the R package gstat to fit a spatiotemporal variogram via a metric covariance model, and perform ordinary kriging<sup>111</sup>. We focused on the ancestry clusters for which we could fit variogram models that were not static over time.

#### <sup>14</sup>C chronology and reservoir effects

Of the 317 individuals sequenced in this study, 272 were <sup>14</sup>C-dated in the project, 30 <sup>14</sup>C-dates were obtained from literature and 15 were dated by archaeological context (Supplementary Note 4 and Supplementary Data 2). Some individuals were dated twice. Most of the dates ( $n = 242$ )

# Article

were performed at the <sup>14</sup>CHRONO Centre laboratory at Queen's University, Belfast, following published sample pretreatment and laboratory protocols<sup>112</sup>. Additional samples were analysed by the Oxford Radiocarbon Accelerator Unit (ORAU) laboratory ( $n = 24$ ) and by the Keck-CCAMS Group ( $n = 6$ ) (see previous reports<sup>113,114</sup> for laboratory procedures). Only datings with a C/N ratio of 2.9–3.6 were accepted; both  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  collagen measurements were also performed, and were used in estimates of marine and freshwater reservoir effects (MRE and FRE, respectively) (see Supplementary Note 4 and Supplementary Data 4). Published values of MRE and FRE were used where available, but for some regions, such as sites in western Russia, a standard FRE value of 500 years was applied. A diet-weighted reservoir offset was then applied to the <sup>14</sup>C central value before calibration. Calibrations were made in Oxcal 4.4 using the Intcal20 calibration curve<sup>115</sup>. For display and calculation purposes a midpoint of the reservoir-corrected and calibrated 95% interval was calculated. Full details of the reservoir correction and calibration procedure are given in Supplementary Note 4 and the calculations are in Supplementary Table 4.1.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All adapter-trimmed sequence data (fastq) for the samples sequenced in this study are publicly available on the European Nucleotide Archive under accession PRJEB64656, together with sequence alignment map files, aligned using human build GRCh37. The full analysis dataset including both imputed and pseudo-haploid genotypes for all ancient individuals used in this study is available at <https://doi.org/10.17894/ucph.d71a6a5a-8107-4fd9-9440-bdafdf81455>. Aggregated IBD-sharing data as well as high-resolution versions of supplementary figures are available at Zenodo (<https://doi.org/10.5281/zenodo.8196989>). Previously published ancient genomic data used in this study are detailed in Supplementary Data 7, and are all already publicly available. Bioarchaeological data (including accelerator mass spectrometry results) are included in the online supplementary materials of this submission. Map figures were created using Natural Earth Data (in Figs. 1–3 and 6 and Extended Data Figs. 1, 3, 4 and 8–11.).

## Code availability

All analyses relied on available software, which has been fully referenced in the manuscript and is detailed in the relevant supplementary notes. A collection of R functions for IBD-based mixture model inference is available at [https://github.com/martinsikora/mixmodel\\_ibd](https://github.com/martinsikora/mixmodel_ibd).

- Damgaard, P. B. et al. Improving access to endogenous DNA in ancient bones and teeth. *Sci. Rep.* **5**, 11184 (2015).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
- Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).
- Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
- Sousa da Mota, B. et al. Imputation of ancient human genomes. *Nat. Commun.* **14**, 3660 (2023).
- Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
- Barbera, P. et al. EPA-ng: massively parallel evolutionary placement of genetic sequences. *Syst. Biol.* **68**, 365–369 (2019).

- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
- Czech, L., Barbera, P. & Stamatakis, A. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* **36**, 3263–3265 (2020).
- Waples, R. K., Albrechtsen, A. & Moltke, I. Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. *Mol. Ecol.* **28**, 35–48 (2019).
- Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
- Nielsen, R., Korneliusson, T., Albrechtsen, A., Li, Y. & Wang, J. SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. *PLoS One* **7**, e37558 (2012).
- Korneliusson, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).
- Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
- Pickrell, J. K. et al. The genetic prehistory of southern Africa. *Nat. Commun.* **3**, 1143 (2012).
- Shringarpure, S. S., Bustamante, C. D., Lange, K. & Alexander, D. H. Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinformatics* **17**, 218 (2016).
- Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
- Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
- Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
- Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
- Maier, R. et al. On the limits of fitting complex models of population history to f-statistics. *eLife* **12**, e85492 (2023).
- Browning, B. L. & Browning, S. R. Detecting identity by descent and estimating genotype error rates in sequence data. *Am. J. Hum. Genet.* **93**, 840–851 (2013).
- Browning, S. R. & Browning, B. L. Accurate non-parametric estimation of recent effective population size from segments of identity by descent. *Am. J. Hum. Genet.* **97**, 404–418 (2015).
- Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
- Greenbaum, G., Rubin, A., Templeton, A. R. & Rosenberg, N. A. Network-based hierarchical population structure analysis for large genomic data sets. *Genome Res.* **29**, 2020–2033 (2019).
- Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal* **1695**, 1–9 (2006).
- Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
- Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
- Hellenthal, G. et al. A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).
- Soetaert, K., Van den Meersche, K. & van Oevelen, D. limSolve: solving linear inverse models. R version 1.5.7 <https://cran.r-project.org/web/packages/limSolve/limSolve> (2009).
- Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
- Cressie, N. & Wikle, C. K. *Statistics for Spatio-Temporal Data* (John Wiley & Sons, 2015).
- Gräler, B., Pebesma, E. & Heuvelink, G. Spatio-Temporal Interpolation using gstat. *R. J.* **8**, 204 (2016).
- Reimer, P. et al. *Laboratory Protocols Used for AMS Radiocarbon Dating at the 14CHRONO Centre*. Report No. 5-2015 (English Heritage, 2015).
- Brock, F., Higham, T., Ditchfield, P. & Ramsey, C. B. Current pretreatment methods for AMS radiocarbon dating at the Oxford Radiocarbon Accelerator Unit (ORAU). *Radiocarbon* **52**, 103–112 (2010).
- Beaumont, W., Beverly, R., Southon, J. & Taylor, R. E. Bone preparation at the KCCAMS laboratory. *Nucl. Instrum. Methods Phys. Res. B* **268**, 906–909 (2010).
- Reimer, P. J. et al. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

**Acknowledgements** We acknowledge P. Bennike, who was involved in initiating this project, for her substantial contributions to its conception and to prehistoric research more broadly; she passed away in 2017. We thank L. Olsen and P. Selmer Olsen for administrative and technical assistance, respectively; the UK Biobank for access to the UK Biobank genomic resource; Illumina for collaboration; and S. Ellingvåg for assistance with sample access. E.W. thanks St John's College, Cambridge, for providing a stimulating environment of discussion and learning. The Lundbeck Foundation GeoGenetics Centre is supported by grants from the Lundbeck Foundation (R302-2018-2155 and R155-2013-16338), the Novo Nordisk Foundation (NNF18SA0035006), the Wellcome Trust (214300), the Carlsberg Foundation (CF18-0024), the Danish National Research Foundation (DNRF94, DNRF174), the University of Copenhagen (KU2016 programme) and Ferring Pharmaceuticals A/S to E.W. This research has been conducted using the UK Biobank Resource and the iPSYCH Initiative, funded by the Lundbeck Foundation (R102-A9118 and R155-2014-1724). This work was further supported by the Swedish Foundation for Humanities and Social Sciences grant (Riksbankens Jubileumsfond M16-0455:1) to K.K. M.E.A. was supported by Marie Skłodowska-Curie Actions of the EU (grant no. 300554), The Villum Foundation (grant no. 10120) and Independent Research Fund Denmark (grant no. 7027-00147B). W.B. is supported by the Hanne and Torkel Weis-Fogh Fund (Department of Zoology, University of Cambridge). A.P. is funded by the Wellcome grant WT214300; B.S.d.M and O.D. by the Swiss National Science Foundation (SFNS PPOOP3\_176977) and the European Research Council (ERC 679330); R. Macleod by an SSHRC doctoral studentship grant (G101449: 'Individual Life Histories in Long-Term Cultural Change'); G.R. by a Novo Nordisk Foundation Fellowship (gNNF20OCO062491); N.N.J. by Aarhus University Research Foundation; B.S.P. by an ERC-Starter Grant 'NEOSEA' (grant no. 949424); H.S. by a Carlsberg Foundation

Fellowship (CF19-0601); G.S. by Marie Skłodowska-Curie Individual Fellowship 'PALAEO-ENEO' (grant agreement number 751349); A. J. Schork by a Lundbeckfonden Fellowship (R335-2019-2318) and the National Institute on Aging (NIH award numbers U19AG023122, U24AG051129 and UH2AG064706); A.L. and I.S. by the Science Committee, Ministry of Education and Science of the Republic of Kazakhstan (AP08856317); B.G.-R. and M.G.-M. by the Spanish Ministry of Science and Innovation (project HAR2016-75605-R); C.M.-L. and O.R. by the Italian Ministry for the Universities (grants 2010-11 prot.2010EL8TXP\_001, 'Biological and cultural heritage of the central-southern Italian population through 30 thousand years' and 2008 prot. 2008B4J2HS\_001, 'Origin and diffusion of farming in central-southern Italy: a molecular approach'); and D.C.-S. and I.G.-Z. by the Spanish Ministry of Science and Innovation (project HAR2017-86262-P). D.C.S.-G. acknowledges funding from the Generalitat Valenciana (CIDEGENT/2019/O61) and the Spanish Government (EUR2020-112213). D.B. was supported by the NOMIS Foundation and Marie Skłodowska-Curie Global Fellowship 'CUSP' (grant no. 846856); E.R.U. by the Science Committee, Ministry of Education and Science of the Republic of Kazakhstan (AP09261083: 'Transcultural Communications in the Late Bronze Age (Western Siberia–Kazakhstan–Central Asia)'); E.C. by Villum Fonden (17649); J.E.A.T. by the Spanish Ministry of Economy and Competitiveness (HAR2013-46861-R) and Generalitat Valenciana (Aico/ 2018/125 and Aico 2020/97); and P.K. by the Russian Ministry of Science and Higher Education (Ural Federal University Program of Development within the Priority-2030 Program). P.K. also acknowledges the Museum of the Institute of Plant and Animal Ecology (UB RAS, Ekaterinburg). L.Y. acknowledges funding by the Science Committee of the Armenian Ministry of Education and Science (project 21AG-1FO25); L.O. by the ERC Consolidator Grant 'PEGASUS' (agreement no. 681605); M. Sablin by the Russian Ministry of Science and Higher Education (075-15-2021-1069); N.C. by Historic Environment Scotland; S.V. and E.V.V. by the Russian Ministry of Science and Higher Education (075-15-2022-328); and V.M. by the Science Committee, Ministry of Education and Science of the Republic of Kazakhstan (AR08856925). V.A. is supported by a Lundbeckfonden Fellowship (R335-2019-2318); P.H.S. by the National Institute of General Medical Sciences (R35GM142916); S.R. by the Novo Nordisk Foundation (NNF14CC0001); T.S.K. is funded by Carlsberg grant CF19-0712; R.D. by the Wellcome Trust (WT214300); R.N. by the National Institute of General Medical Sciences (NIH grant R01GM138634); and F. Racimo by a Villum Fonden Young Investigator Grant (no. 00025300); by a Novo Nordisk Fonden Data Science Ascending Investigator Award (NNF22OC0076816) and by the European Research Council (ERC) under the European Union's Horizon Europe programme (grant agreement No. 101077592). T.W. and V.A. are supported by the Lundbeck Foundation iPSYCH initiative (R248-2017-2003).

**Author contributions** E.W. initiated the study. M.E.A., M.S., T.S.K., R.D., R.N., O.D., T.W., F. Racimo, K.K. and E.W. led the study. M.E.A., M.S., A.F., M.M., C.L.-F., R.N., E.C., T.W., K.K. and E.W. conceptualized the study. M.E.A., M.S., H.S., L.O., T.S.K., R.D., R.N., O.D., T.W., F. Racimo, K.K. and E.W. supervised the research. M.E.A., L.O., R.D., R.N., T.W., K.H.K., K.K. and E.W. acquired

funding for research. M.E.A., A.F., J.S., K.-G.S., M.L.S.J., M.U.H., A.A.T., A.C., A.Z., A.M.S., A.J.H., A.G., A.L., B.H.N., B.G.-R., C.B., C.L., C.M.-L., D.V.P., D.C.-S., D.O.L., D.E., D.C.S.-G., D.B., E.B.P., E.K., E.V.V., E.R.U., E. Kannegaard, F. Radina, H.D., I.G.-Z., I.P., I.S., J.G., J.H., J.E.A.T., J.Z., J.V., K.B.P., K.T., L.N., L.L., L.N.M., L.Y., L.P., L. Sarti, L. Slimak, L.K., M.G.-M., M. Silvestrini, M.D., M.V., M.S.N., M.R., M.S., M.P., M.C., M. Sablin, N.C., N.S., O.P., O.R., O.V.L., P.A., P.K., P.C., P. Rios, P. Lotz, P. Lysdahl, P.M., P.P., P.B., P.d.B.D., P.V.P., P.P.M., P.W., R.V.S., R. Maring, R. Menduina, R.B., R.T., S.V., S.W., S.B., S.S., S.A.S., S.H.A., T.D.P., T.Z.T.J., Y.B.S., V.I.M., V.S., V.M., Y.M., I.M., O.G. and N.L. were involved in sample collection. M.E.A., M.S., A.R.-M., E.K.I.-P., W.B., A.I., J.S., A.P., B.S.d.M., M.I., M.M., L.V., A. Stern, C.G., F.E.Y. D.J.L., T.S.K., R.D., R.N., O.D., F. Racimo, K.H.K. and E.W. were involved in developing and applying methodology. M.E.A., J.S., C.G. and L.V. led the DNA laboratory work research component. K.-G.S., A.F. and M.E.A. led bioarchaeological data curation. M.E.A., M.S., A.R.-M., E.K.I.-P., W.B., A.I., A.P., B.S.d.M., B.S.P., A.H., R.A.H., T.V., H.M., A.M., A.V., A.B.N., P. Rasmussen, G.R., A. D. Ramsøe, A. Skorobogatov, A. J. Schork, A. Rosengren, C.J.M., I.A., L.Z., R. Maring, V.S., V.A., P.H.S., S.R., T.S.K., O.D. and F. Racimo undertook formal analyses of data. M.E.A., M.S., A.R.-M., E.K.I.-P., A.F., W.B., A.I., K.-G.S., R. Macleod, D.J.L., P.H.S., T.S.K., F. Racimo and E.W. drafted the main text (M.E.A. and M.S. led this). M.E.A., M.S., A.R.-M., E.K.I.-P., A.F., W.B., A.I., K.-G.S., A.P., B.S.d.M., B.S.P., A.H., R. Macleod, R.A.H., T.V., M.F.M., A.B.N., M.U.H., P. Rasmussen, A. J. Stern, N.N.J., H.S., G.S., A. Ramsøe, A. Skorobogatov, A. Rosengren, A.O., A.B., A.C., A.G., A.L., A.B.G., C.J.M., D.C.S.-G., E.B.P., E. Kostyleva, E.R.U., E. Kannegaard, I.G.-Z., I.P., I.S., J.G., J.H., J.E.A.T., K.H.K., L.Z., L.Y., L.P., L.K., M.B., M.G.-M., M.V., M.R., M.J., N.B., O.V.L., O.C.U., P.K., P. Lysdahl, P.B., P.W., R.V.S., R. Maring, R.B., R.I., S.V., S.W., S.B., S.H.A., T.Z.T.J., V.S., D.J.L., P.H.S., S.R., T.S.K., O.D. and F. Racimo drafted supplementary notes and materials. M.E.A., M.S., A.R.-M., E.K.I.-P., A.F., W.B., A.I., G.S., A.H., M.L.S.J., F.D., R. Macleod, L. Sørensen, P.O.N., R.A.H., T.V., H.M., A.M., N.N.J., H.S., A. Ramsøe, A. Skorobogatov, A. J. Schork, A. Ruter, A.O., B.H.N., B.G.-R., D.C.-S., D.C.S.-G., I.G.-Z., I.P., J.G., J.E.A.T., L.Z., L.O., L.K., M.G.-M., P.d.B.D., R.I., S.A.S., D.J.L., I.M., O.G., P.H.S., T.S.K., R.D., R.N., O.D., T.W., F. Racimo, K.K. and E.W. were involved in reviewing drafts and editing (M.E.A., M.S., A.F., K.-G.S., R. Macleod and E.W. led this, and subsequent finalization of the study). All co-authors read, commented on and agreed on the submitted manuscript.

**Competing interests** The authors declare no competing interests.

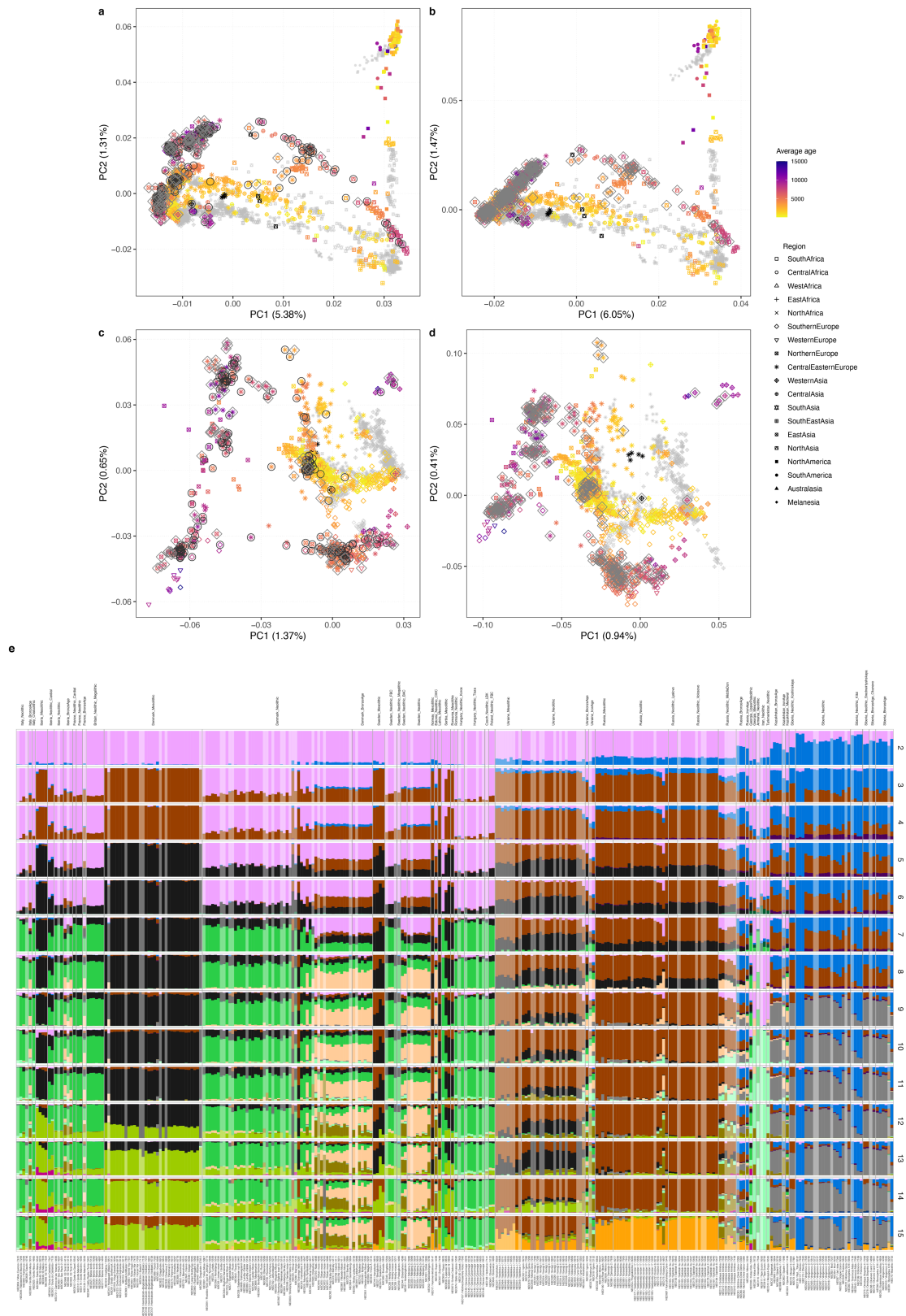
#### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-06865-0>.

**Correspondence and requests for materials** should be addressed to Morten E. Allentoft, Martin Sikora or Eske Willerslev.

**Peer review information** *Nature* thanks Benjamin Peter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

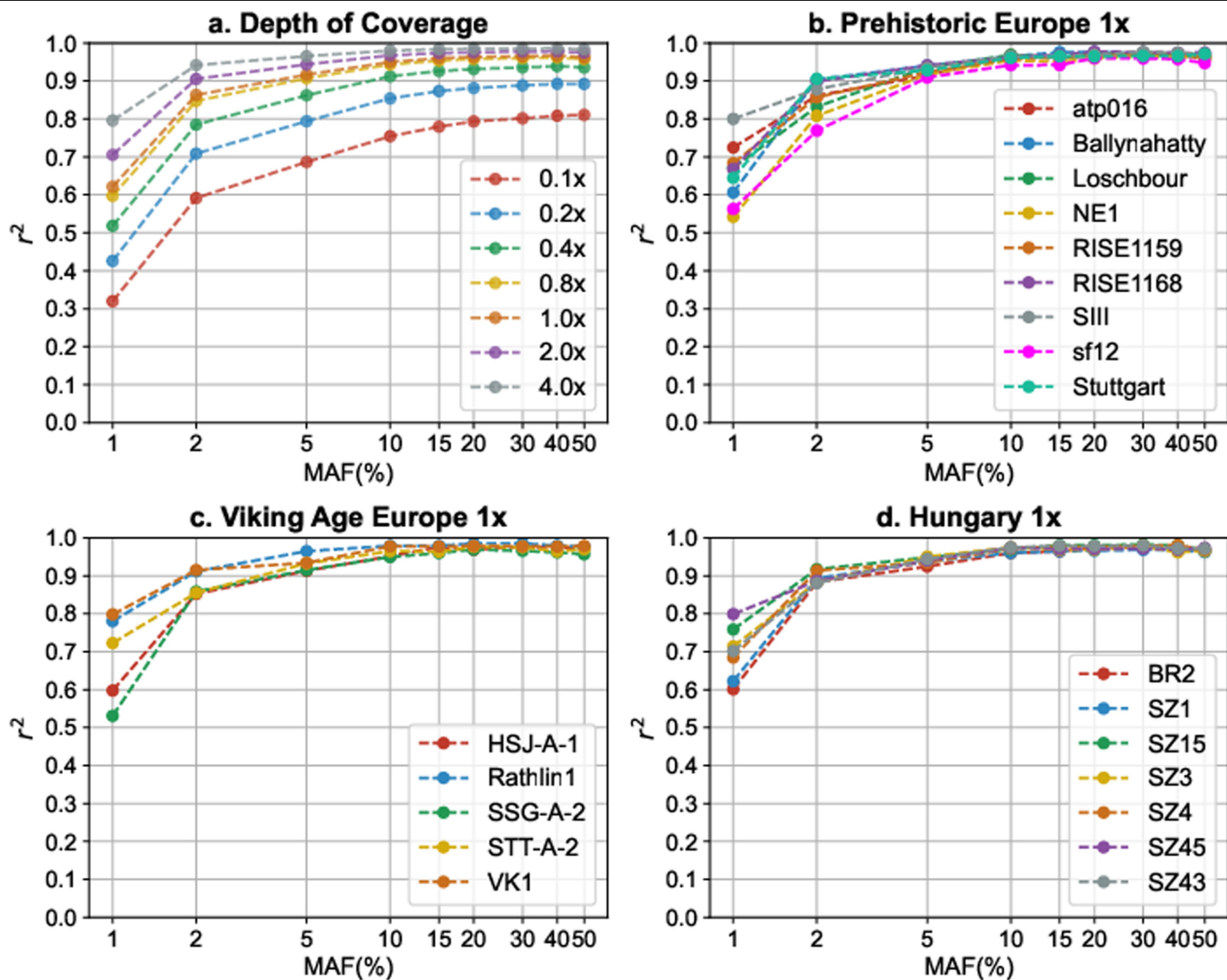
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | Genetic structure of the 317 herein-reported ancient genomes.** **a–d**, PCA of 3,316 modern and ancient individuals from Eurasia, Oceania and the Americas (**a,b**), as well as restricted to 2,126 individuals from western Eurasia (west of the Urals) (**c,d**). Shown are analyses with principal components inferred either using both modern and imputed ancient genomes passing all filters, and projecting low coverage ancient genomes (**a,c**); or only modern genomes and projecting all ancient genomes (**b,d**). Ancient genomes sequenced in this study are indicated either with black circles

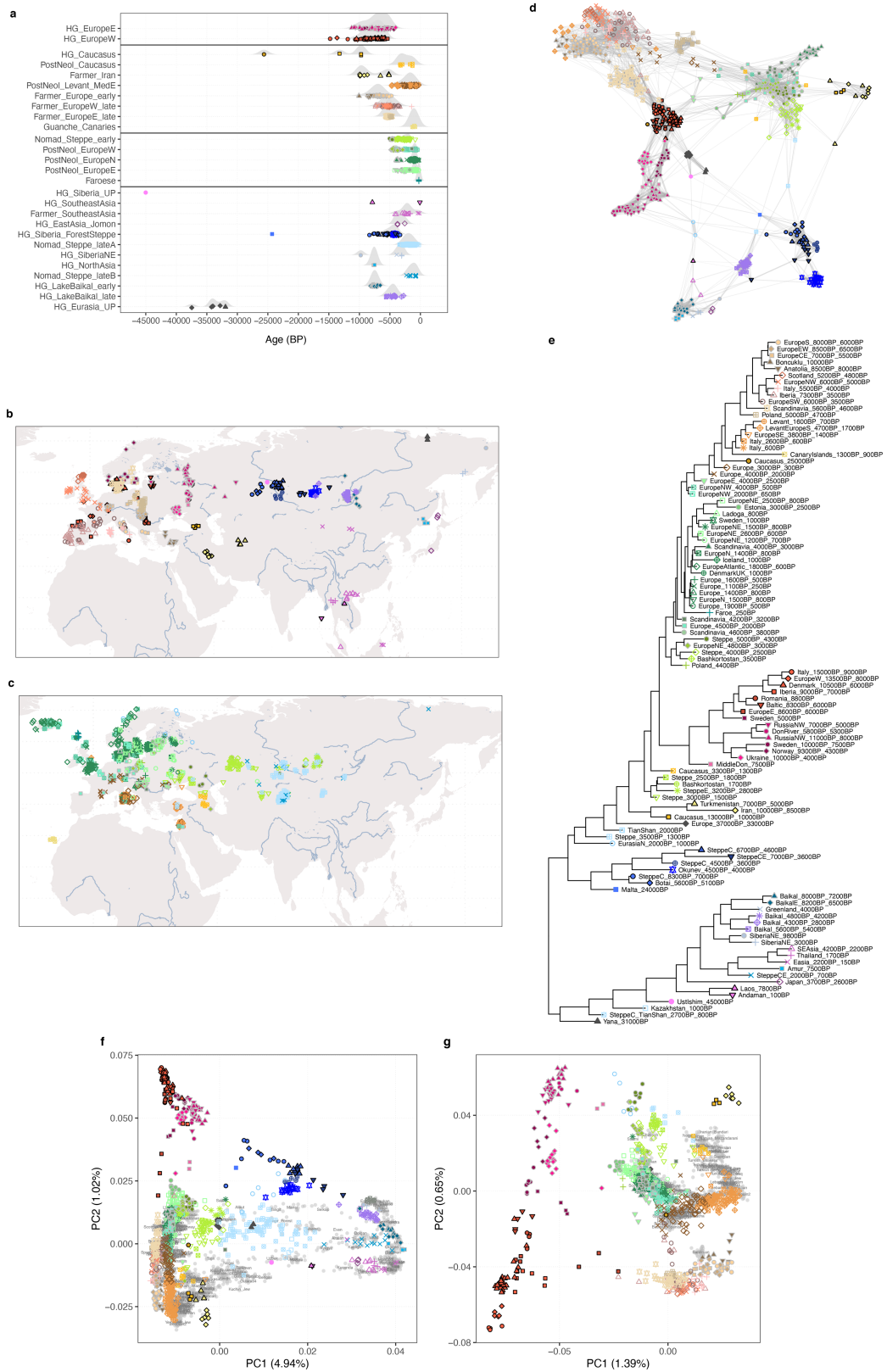
(imputed genomes) or grey diamonds (projected genomes). **e**, Model-based clustering results using ADMIXTURE for 284 newly reported genomes (excluding close relatives and individuals flagged for possible contamination). Results shown are based on ADMIXTURE runs with  $K = 2$  to  $K = 15$  on 1,593 ancient individuals, corresponding to the full set of 1,492 imputed genomes passing filters as well as 101 low coverage genomes represented by pseudo-haploid genotypes (flags “lowcov” or “lowGpAvg”, Supplementary Data 7; indicated with alpha transparency in plot).





**Extended Data Fig. 2 | Imputation accuracy of ancient DNA.** **a**, Imputation accuracy across 42 high-coverage ancient genomes when downsampled to lower depth of coverage values (see Supplementary Note 2 and Supplementary Table 2.1). **b**, Imputation accuracy for 1x depth of coverage across 9 prehistoric

European genomes; **c**, across 5 Viking age genomes; and **d**, across 7 ancient genomes from Early Medieval Hungary. In all panels, imputation accuracy is shown as the squared Pearson correlation between imputed and true genotype dosages as a function of MAF of the target variant sites.

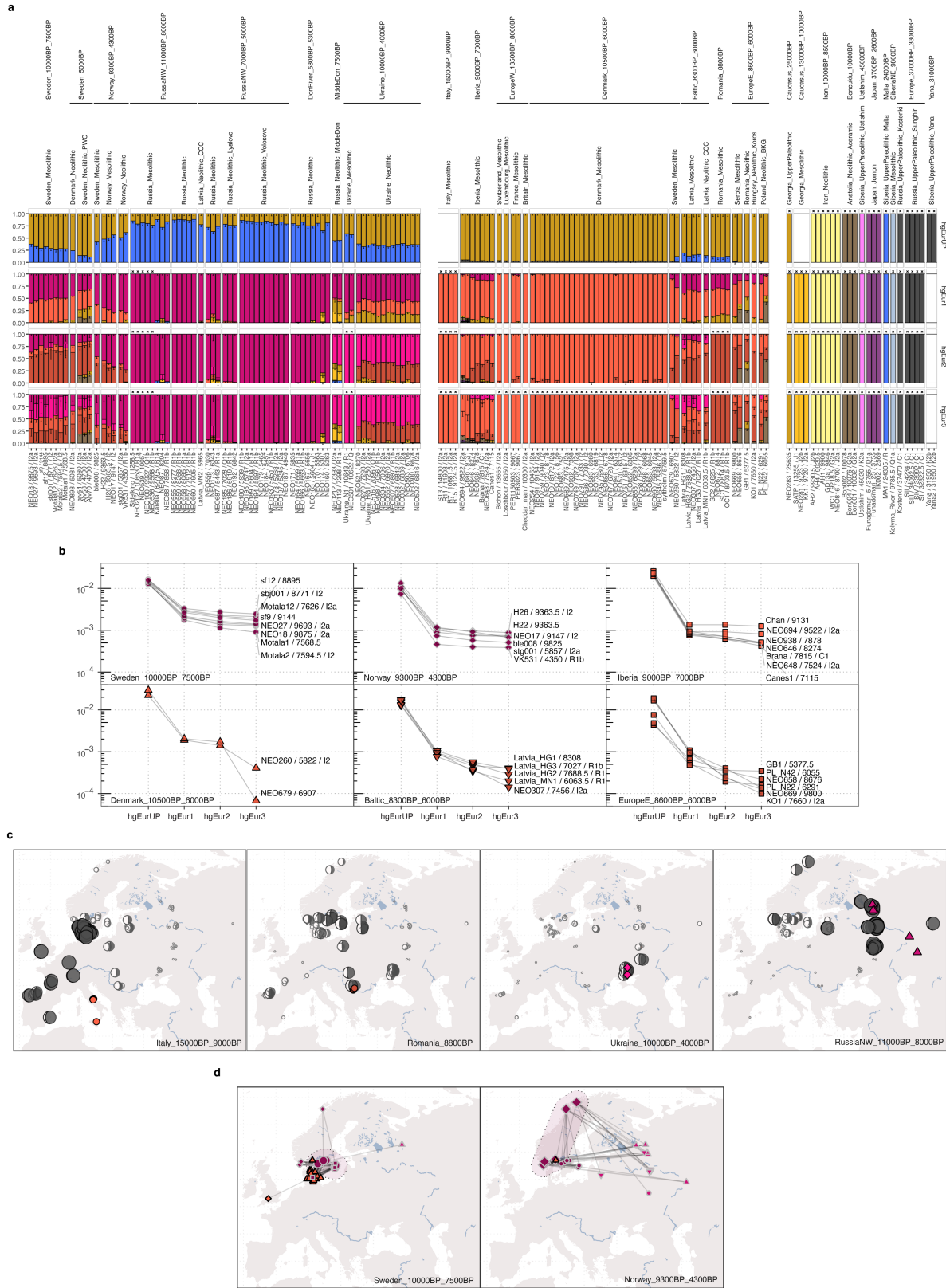


Extended Data Fig. 3 | See next page for caption.

**Extended Data Fig. 3 | Genetic clustering of ancient individuals.**

Characterization of genetic clusters for 1,401 imputed ancient individuals from Eurasia (that is, excluding 91 individuals from Africa and Americas), inferred from pairwise IBD sharing (indicated using coloured symbols throughout), **a**, Temporal distribution of clustered individuals, grouped by broad ancestry cluster. **b,c**, Geographical distribution of clustered individuals, shown for individuals predating 3,000 BP (**b**) and after 3,000 BP (**c**). **d**, Network graph of pairwise IBD sharing between 596 ancient Eurasians predating 3,000 BP, highlighting within- and between-cluster relationships. Each node represents

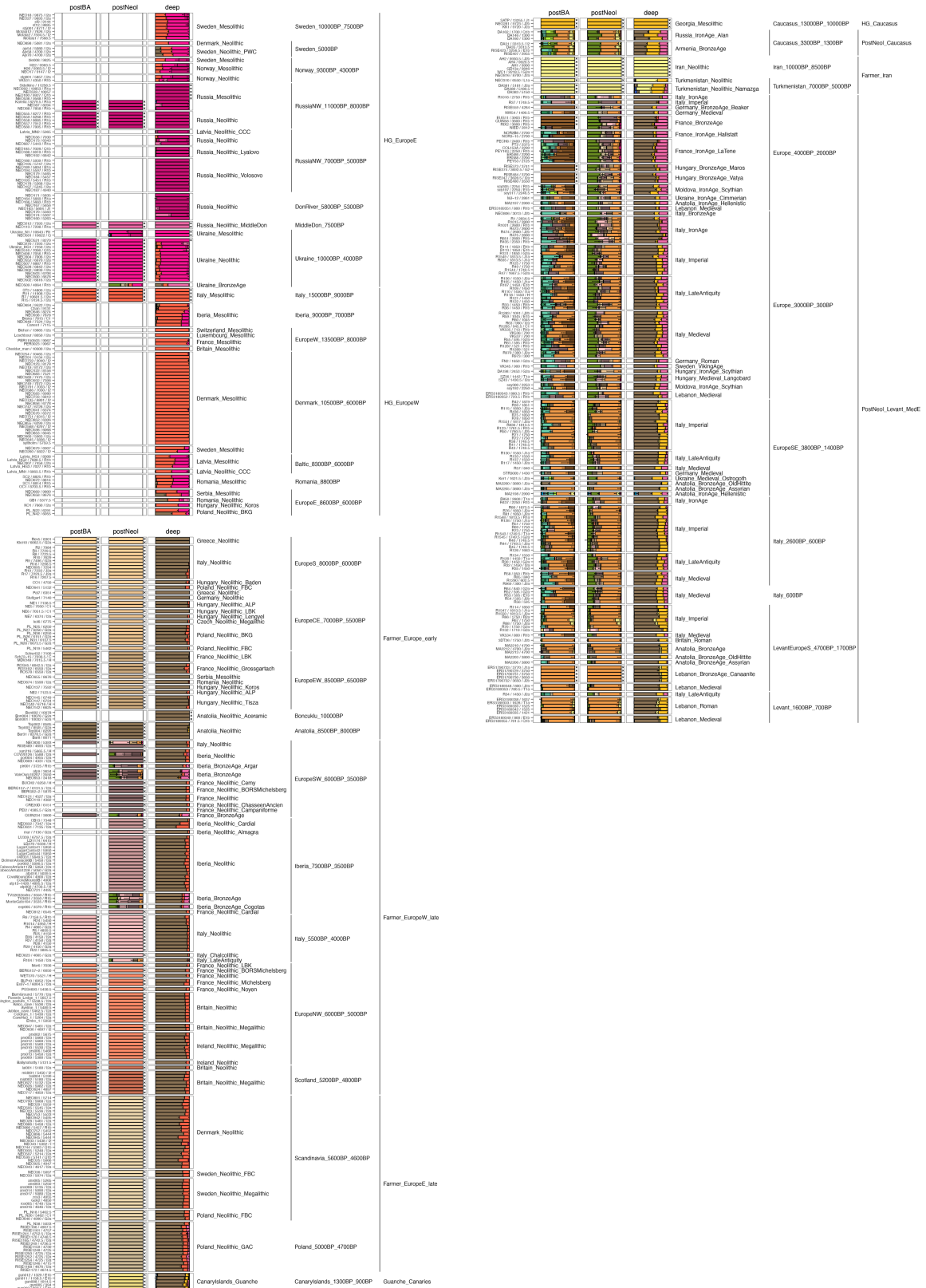
an individual, and the width of edges connecting nodes indicates the fraction of the genome shared IBD between the respective pair of individuals. Network edges were restricted to the 10 highest sharing connections for each individual, and the layout was computed using the force-directed Fruchterman-Reingold algorithm. **e**, Neighbour-joining tree showing relationships between genetic clusters, inferred using total variation distance (TVD) of IBD painting palettes. **f,g**, PCA of 3,119 Eurasian (**f**) or 2,126 west Eurasian (**g**) ancient and modern individuals ("HO" dataset).



Extended Data Fig. 4 | See next page for caption.

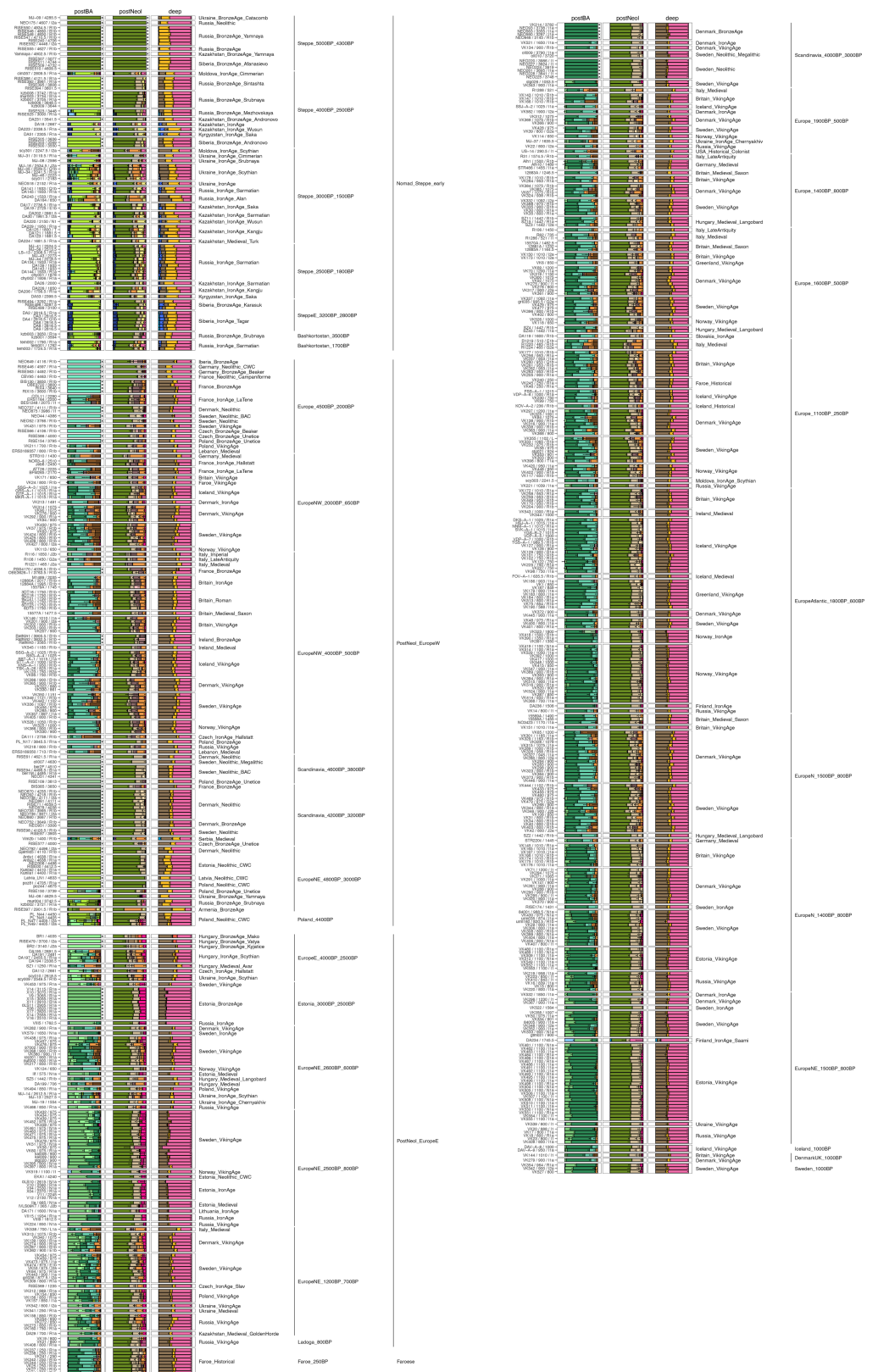
**Extended Data Fig. 4 | Genetic structure of European HGs after the LGM.**  
**a**, Supervised ancestry modelling using non-negative least squares on IBD sharing profiles. Panels show estimated ancestry proportions for target individuals from genetic clusters representing European HGs, using different sets of increasingly proximal source groups. Individuals used as sources in a particular set are indicated with black crosses and coloured bars with 100% ancestry proportion. Black lines indicate 1 standard error for the respective ancestry component. **b**, Residuals for model fit of target individuals from selected genetic clusters across different source sets. **c**, Moon charts showing

spatial distribution of ancestry proportions in European HGs deriving from four European source groups (set "hgEur2"; source origins shown with coloured symbol). Estimated ancestry proportions are indicated by both size and amount of fill of moon symbols. Note that 'Italy\_15000BP\_9000BP' and 'RussiaNW\_11000BP\_8000BP' correspond to 'WHG' and 'EHG' labels used in previous studies. **d**, Maps showing networks of highest between-cluster IBD sharing (top 10 highest sharing per individual) for individuals from two genetic clusters representing Scandinavian HGs. See Supplementary Data 1 and 7 for details of individual sample IDs presented here.



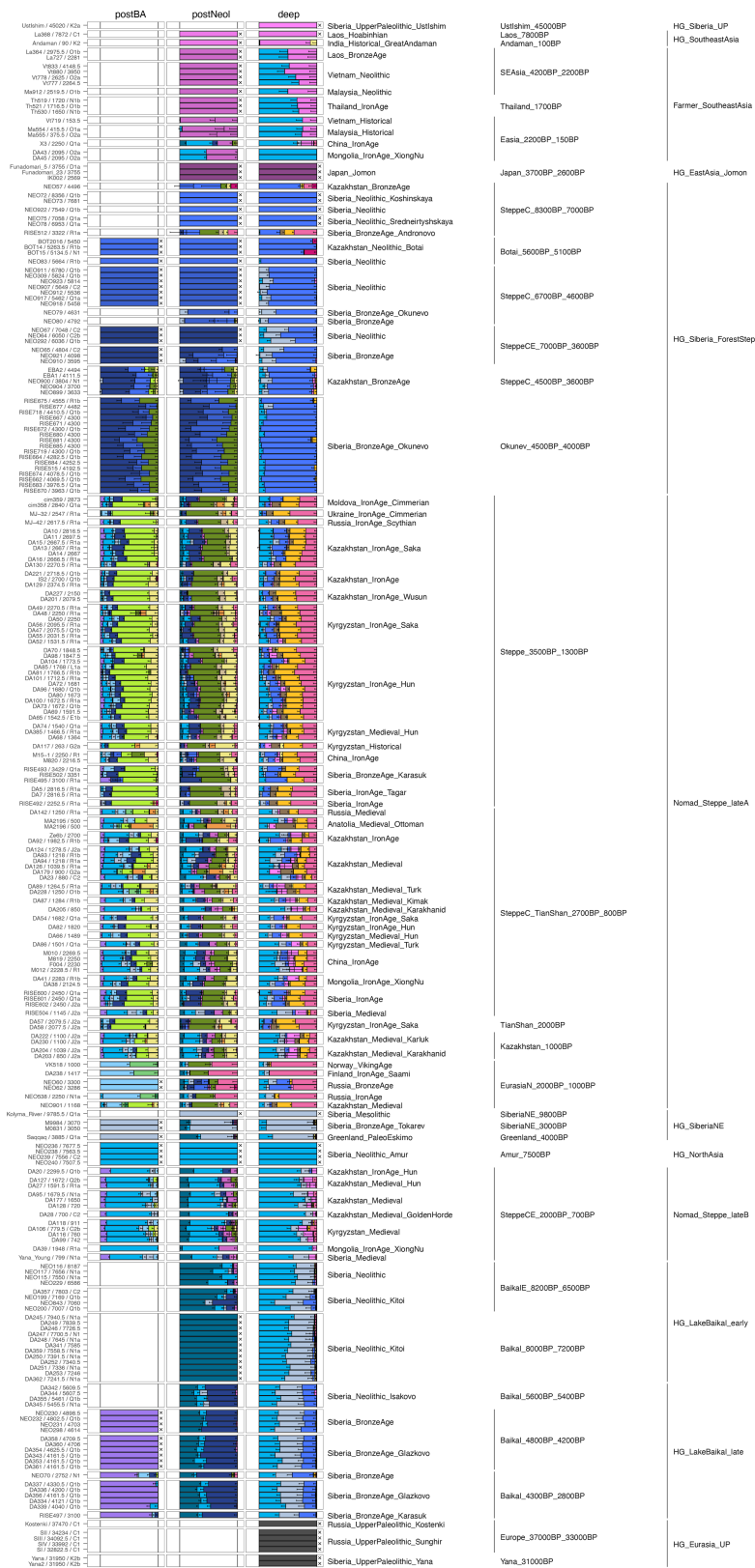
**Extended Data Fig. 5 | Ancestry modelling for HG and Neolithic farmer-associated genetic clusters.** Supervised ancestry modelling using non-negative least squares on IBD sharing profiles. Panels show estimated ancestry proportions of two global Eurasian clusters, corresponding to European HGs before 4,000 BP and individuals from Europe and western Asia from around 10,000 BP until historical times, including Anatolian-associated (Neolithic) farmers, Caucasus HGs and recent individuals with genetic affinity to the Levant. Columns show results of modelling target individuals using three

panels of increasingly distal source groups: “postBA”: Bronze Age and Neolithic source groups; “postNeol”, Bronze Age and later targets using Late Neolithic/early Bronze Age and earlier source groups; “deep”, Mesolithic and later targets using deep ancestry source groups. Individuals used as sources in a particular set are indicated with black crosses and coloured bars with 100% ancestry proportion. Black lines indicate 1 standard error for the respective ancestry component.



**Extended Data Fig. 6 | Ancestry modelling for post-Neolithic genetic clusters.** Supervised ancestry modelling using non-negative least squares on IBD sharing profiles. Panels show estimated ancestry proportions of a global Eurasian cluster corresponding to European individuals after 5,000 BP, as well as pastoralist groups from the Eurasian steppe. Columns show results of modelling target individuals using three panels of increasingly distal source

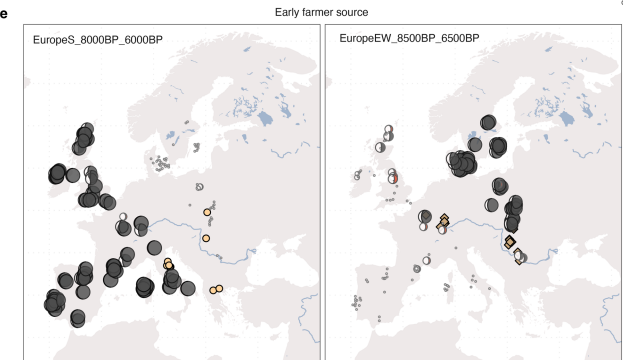
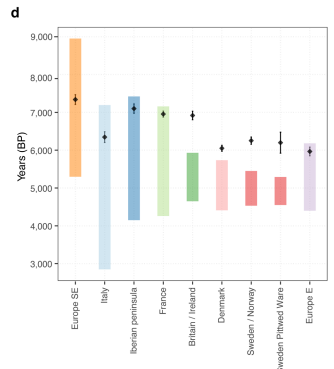
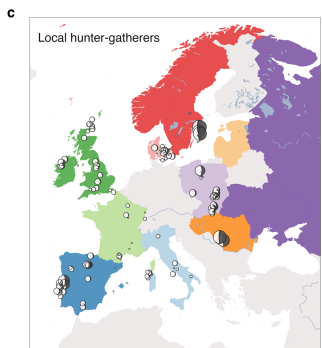
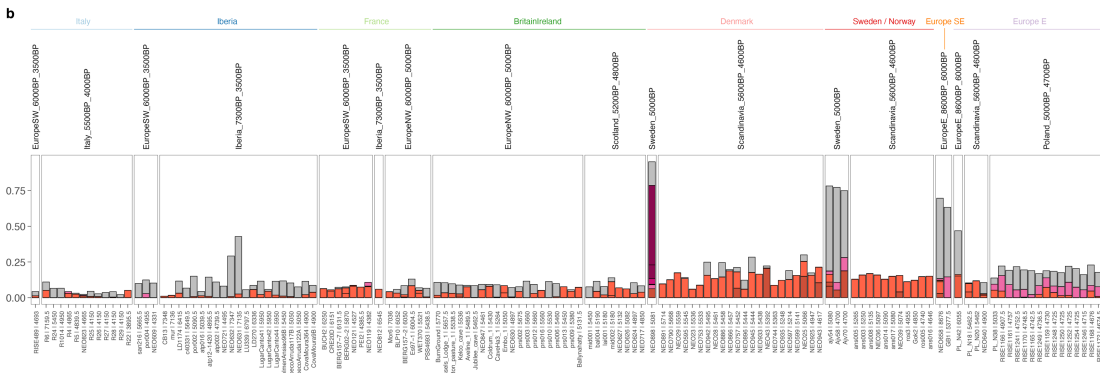
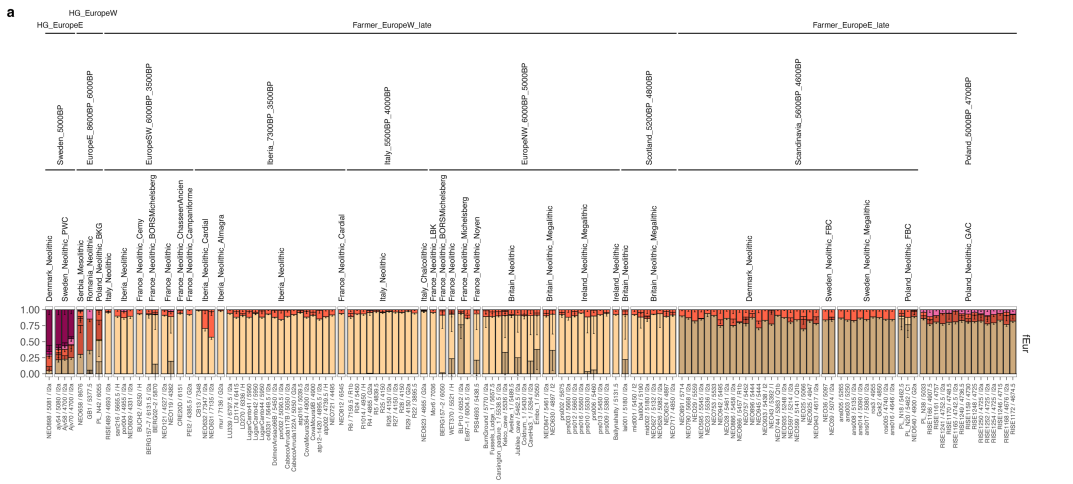
groups: “postBA”: Bronze Age and Neolithic source groups; “postNeol”, Bronze Age and later targets using Late Neolithic/early Bronze Age and earlier source groups; “deep”, Mesolithic and later targets using deep ancestry source groups. Individuals used as sources in a particular set are indicated with black crosses and coloured bars with 100% ancestry proportion. Black lines indicate 1 standard error for the respective ancestry component.



**Extended Data Fig. 7 | Ancestry modelling for genetic clusters east of the Urals.** Supervised ancestry modelling using non-negative least squares on IBDaring profiles. Panels show estimated ancestry proportions of a global Eurasian cluster corresponding to central, east and north Asian individuals with east Eurasian genetic affinities. Columns show results of modelling target individuals using three panels of increasingly distal source groups: “postBA”:

Bronze Age and Neolithic source groups; “postNeol”, Bronze Age and later targets using Late Neolithic/early Bronze Age and earlier source groups; “deep”, Mesolithic and later targets using deep ancestry source groups. Individuals used as sources in a particular set are indicated with black crosses and coloured bars with 100% ancestry proportion. Black lines indicate 1 standard error for the respective ancestry component.



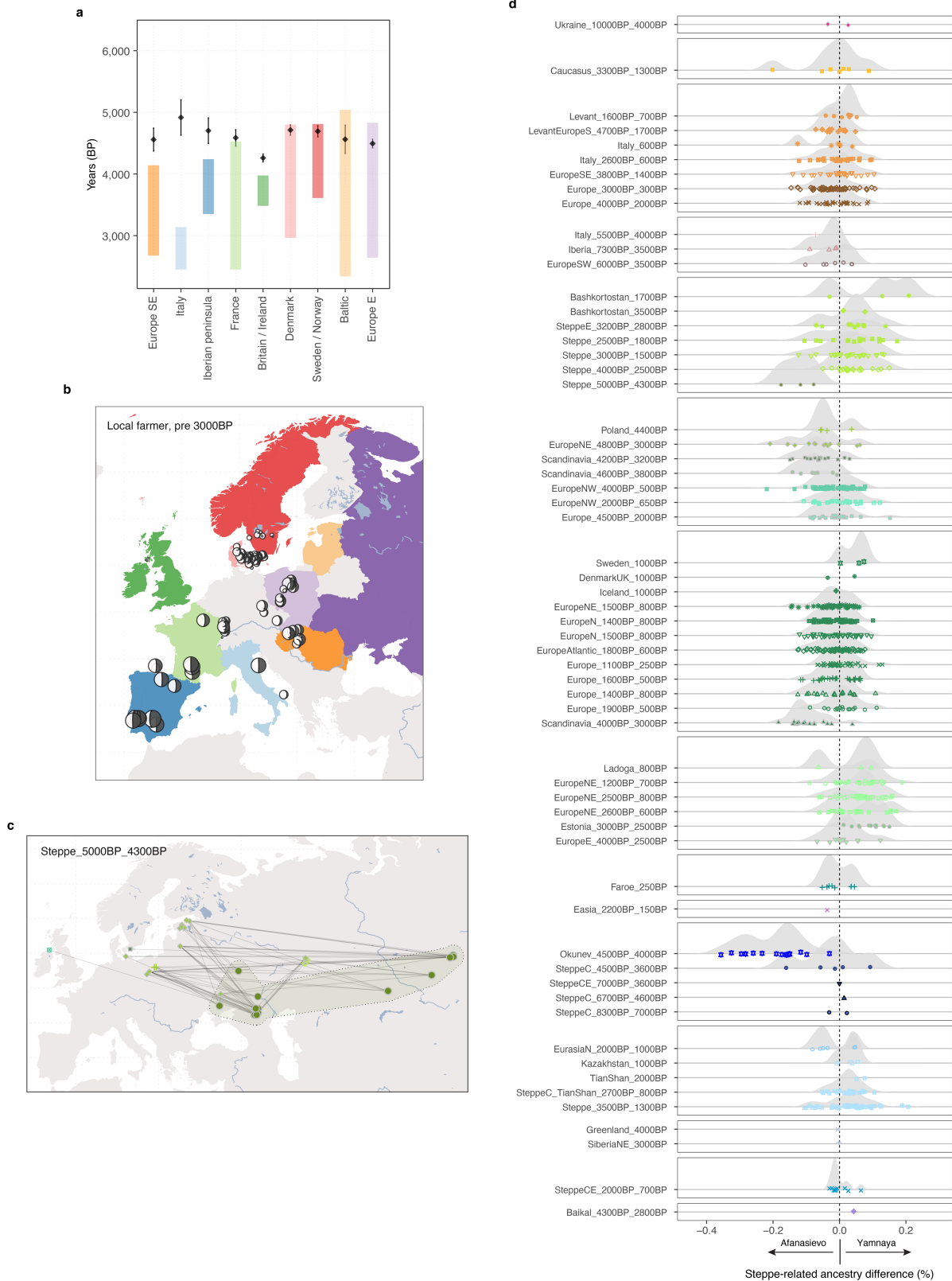


Extended Data Fig. 8 | See next page for caption.

## Article

**Extended Data Fig. 8 | Dynamics of the Neolithic transition in Europe.**  
**a,** Supervised ancestry modelling using non-negative least squares on IBD sharing profiles. Panels show estimated ancestry proportions for target individuals from genetic clusters representing European Neolithic farmer individuals (“fEur” source set). Individuals used as sources in a particular set are indicated with black crosses and coloured bars with 100% ancestry proportion. Black lines indicate 1 standard error for the respective ancestry component. **b,** Composition of HG ancestry proportions from different source groups in individuals with Neolithic farmer ancestry, shown as bar plots. Grey bars represent contributions from a source with ancestry related to local HGs. **c,** Moon charts showing spatial distribution of estimated ancestry proportions related to local HGs across Europe. Estimated ancestry proportions are

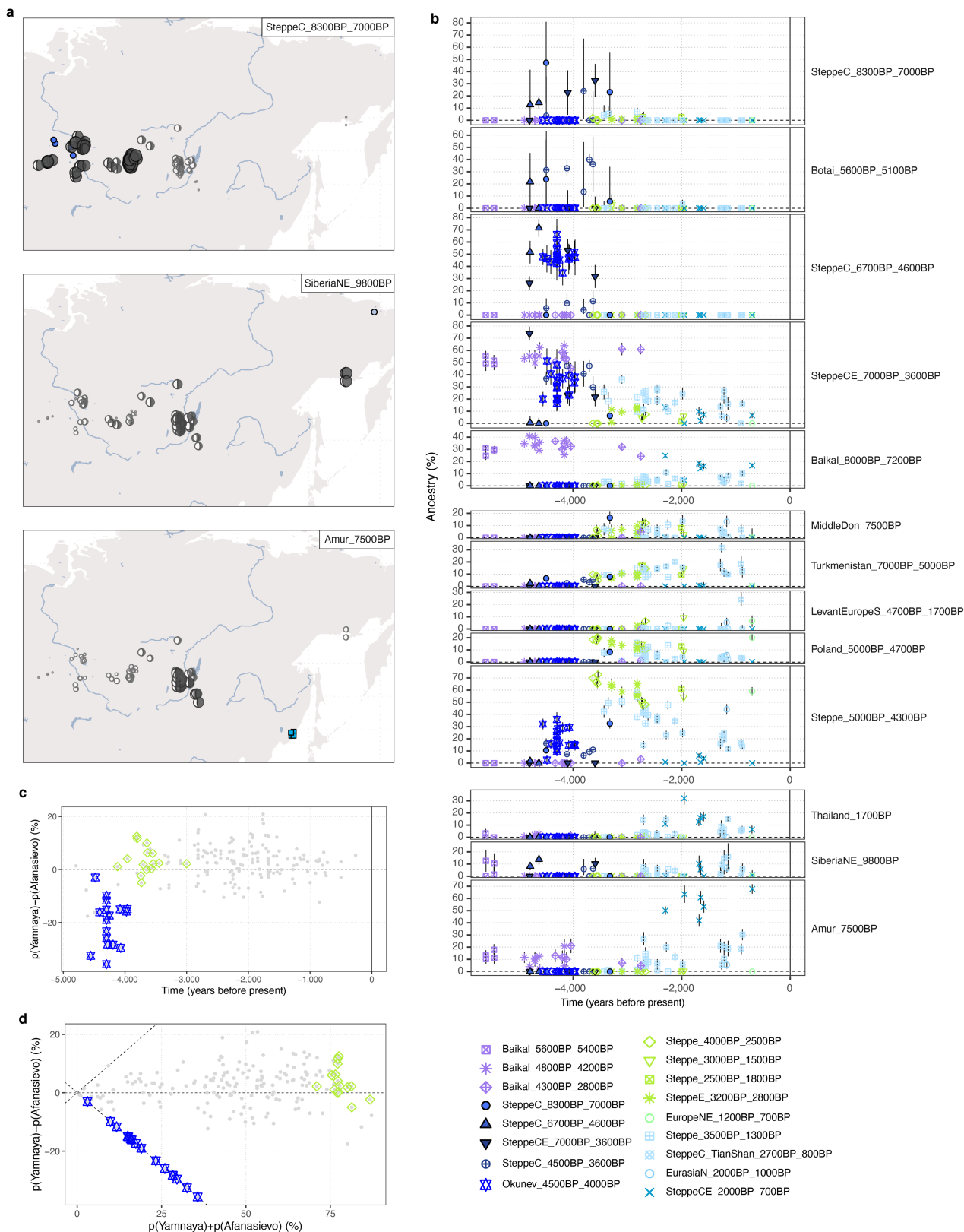
indicated by size and amount of fill of moon symbols. Coloured areas indicate the geographical extent of individuals included as local sources in the respective regions. **d,** Estimated time of admixture between local HG groups and Neolithic farmers. Black diamonds and error bars represent point estimate and standard errors of admixture time, coloured bars show temporal range of included target individuals. The time to admixture was adjusted backwards by the average age of individuals for each region. **e,** Moon charts showing spatial distribution of estimated ancestry proportions derived from genetic clusters of early Neolithic European farmers (locations indicated with coloured symbols). Estimated ancestry proportions are indicated by size and amount of fill of moon symbols. Red symbols indicate individuals where standard errors exceed the point estimates for the respective ancestry source.



**Extended Data Fig. 9 | Dynamics of the steppe transition in Europe.**

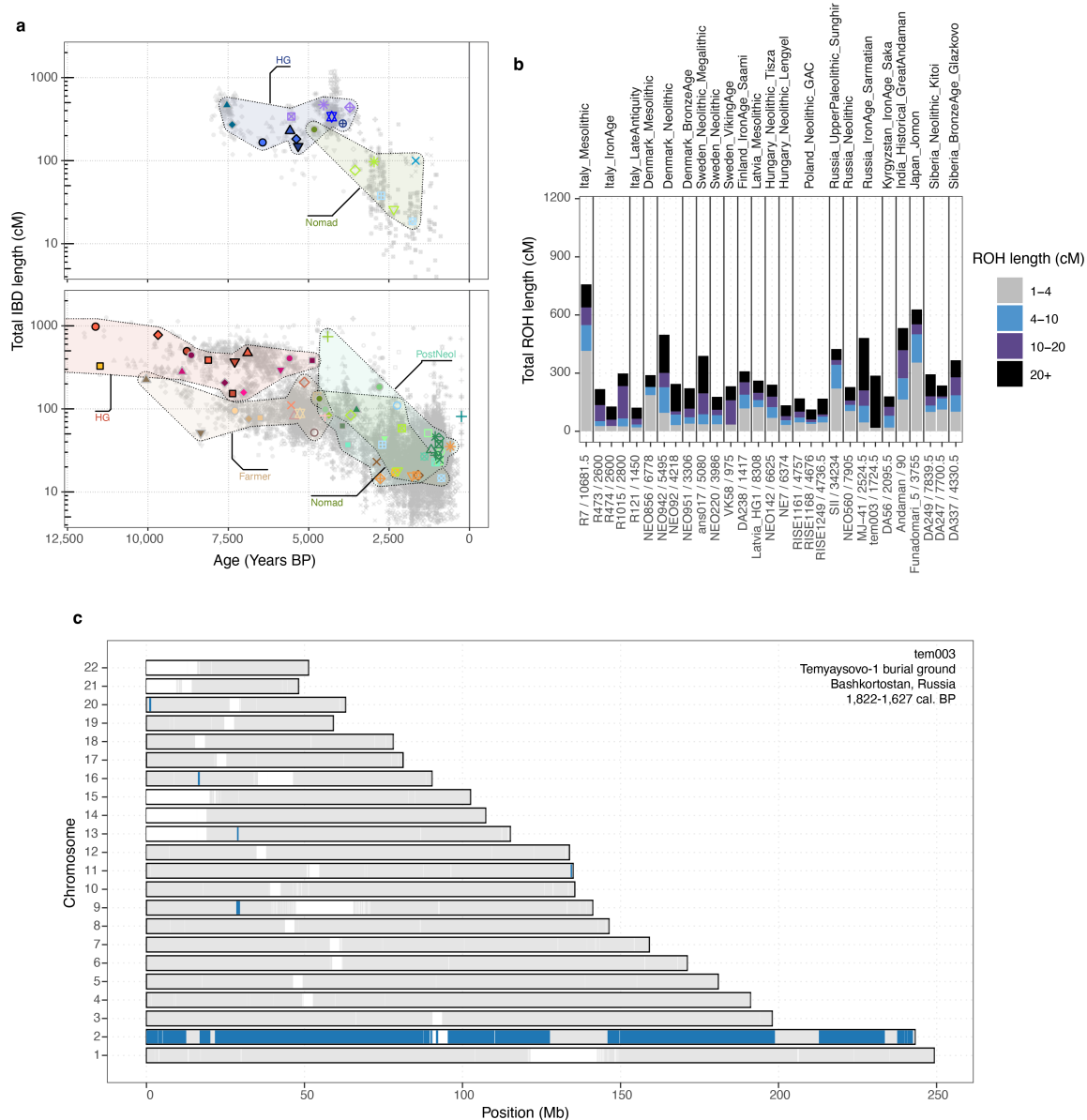
**a**, Estimated time of admixture between local HG groups and Neolithic farmers. Black diamonds and error bars represent point estimate and standard errors of admixture time, coloured bars show temporal range of included target individuals. The time to admixture was adjusted backwards by the average age of individuals for each region. **b**, Moon charts showing spatial distribution of estimated ancestry proportions related to local Neolithic farmers across Europe. Estimated ancestry proportions are indicated by size and amount of fill

of moon symbols. Coloured areas indicate the geographical extent of individuals included as local sources in the respective regions. **c**, Map showing networks of highest between-cluster IBD sharing (top 10 highest sharing per individual) for individuals from genetic cluster “Steppe\_5000BP\_4300BP” representing the major steppe ancestry source for Europeans. **d**, Distributions of difference in estimated steppe-related ancestry proportions, using individuals from the genetic cluster “Steppe\_5000BP\_4300BP”, associated with either Yamnaya or Afanasievo cultural contexts as separate sources.



**Extended Data Fig. 10 | Genetic transformations across the Eurasian steppe.** **a**, Moon charts showing spatial distribution of estimated ancestry proportions of Siberian HGs from the “deep” Siberian ancestry sources (names and locations indicated with coloured symbols). Estimated ancestry proportions are indicated by size and amount of fill of moon symbols. **b**, Timelines of ancestry proportions from “postNeol” sources in central and north Asian ancient individuals after 5,000 BP. Symbol shape and colour indicate the genetic cluster

of each individual. Black lines indicate 1 standard error. **c,d**, Difference in estimated steppe-related ancestry proportions, using individuals from genetic cluster “Steppe\_5000BP\_4300BP” associated with either Yamnaya or Afanasievo cultural contexts as separate sources, as a function of time (**c**) or total estimated steppe-ancestry proportion (**d**). Individuals from genetic clusters of individuals associated with Okunevo (blue stars) or Sintashta/Andronovo (green diamonds) contexts are indicated with coloured symbols.



**Extended Data Fig. 11 | Patterns of co-ancestry.** **a**, Panels show within-cluster genetic relatedness over time, measured as the total length of genomic segments shared IBD between individuals. Results for both measures are shown separately for individuals from western versus eastern Eurasia. Small grey dots indicate estimates for individual pairs, with larger coloured symbols indicating median values within genetic clusters. Ranges of median values for

major ancestry groups are indicated with labelled convex hulls. **b**, Distribution of ROH lengths for 29 individuals with evidence for recent parental relatedness (>50 cM total in ROHs > 20 cM). **c**, Karyogram showing genomic distribution of ROH in individual tem003, an ancient case of uniparental disomy for chromosome 2. Regions within ROH are indicated with blue colour.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Sequencing data and metadata pertaining to the sequencing of ancient genomes is managed on secure servers at the Globe Institute, University of Copenhagen, and via the Illumina Inc. BaseSpace platform. Accelerator Mass Spectrometry data and associated metadata is managed at the Department of Historical Studies, University of Gothenberg.

Data analysis

Custom scripts used to apply chromopainter from large-scale phased data are available at [https://github.com/will-camb/Nero/tree/master/scripts/cp\\_panel\\_scripts](https://github.com/will-camb/Nero/tree/master/scripts/cp_panel_scripts). All other analyses relied upon available software which has been fully referenced in the manuscript and detailed in the relevant supplementary notes. These comprise:

CASAVA (v.1.8.2)  
 AdapterRemoval (v.2.1.3)  
 Picard (v.1.127)  
 GATK (v.3.3.0)  
 Samtools (1.9)  
 Samtools calmd (v.1.10)  
 pysam (<https://github.com/pysam-developers/pysam>)  
 BEDtools (v.2.23.0)  
 mapDamage2.0 (v2.2.1)  
 BWA (0.7-17)  
 Schmutzi (VERSION?)  
 ContamMix (VERSION?)  
 ANGSD (0.938)  
 GLIMPSE (v1.0.1)

Beagle (v4.1)  
 BCFtools 1.10  
 MAFFT (v7.490)  
 RAxML-ng (v1.1.0)  
 ry\_compute (v0.4)  
 EPA-ng (0.3.8)  
 NgsRelate (v2)  
 ADMIXTURE (1.3.0)  
 smartpca  
 GCTA  
 R version 4.0  
 ADMIXTOOLS2  
 gstat (2.0-9)  
 IBDseq (version r1206)  
 GenomicRanges (3.15)  
 leidenAlg (v1.01, <https://github.com/kharchenkolab/leidenAlg>)  
 igraph (0.9.9)  
 limSolve (1.5.6)  
 Scikit-learn (0.21.2)  
 Chromopainter (0.0.4)  
 fineSTRUCTURE (0.0.5)  
 QCTOOL v2 ([https://www.well.ox.ac.uk/~gav/qctool\\_v2/](https://www.well.ox.ac.uk/~gav/qctool_v2/))  
 ArcGIS Online ([www.arcgis.com](http://www.arcgis.com))  
 OxCal v4.4  
 LRA v0.1.0 (<https://github.com/petrkunes/LRA>)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All collapsed and paired-end sequence data for novel samples sequenced in this study will be made publicly available on the European Nucleotide Archive, together with trimmed sequence alignment map files, aligned using human build GRCh37. Previously published ancient genomic data used in this study is detailed in Supplementary Table VII, and are all already publicly available. Bioarchaeological data (including stable isotopes and radiocarbon dating results) are included in the online supplementary materials of this submission.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

No living or recently deceased human research participants were affected by this study

Population characteristics

*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

Recruitment

*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

Ethics oversight

*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We sequenced 317 novel ancient human genomes from primarily Mesolithic and Neolithic individuals across Eurasia, combined with new radiocarbon dates, stable isotope data. We additionally test and apply GLIMPSE to impute 1492 previously published ancient genomes, and incorporate this data in our analyses. We undertake a suite of analyses of this data to generate new insights into population structure through time, natural selection and health-related adaptation.
Research sample	Ancient human remains (bone and dental cementum) were collected, sampled, extracted and screened by shallow shotgun sequencing (see Supplementary Note 1). This resulted in 317 samples found to be suitable for deep sequencing (of 962 initial specimens). These represent a considerable geographic range across Eurasia (from Lake Baikal to the Atlantic coast). A high sample density for Denmark (100) was undertaken to provide high resolution insights at this location. A particular focus upon individuals related to Mesolithic or Neolithic culture was made, with 300 of the individuals from the time span 11,000 cal. BP to 3,000 cal. BP.
Sampling strategy	Sampling was dependent upon the availability of ancient human remains, though a considerable transect of the Mesolithic and Neolithic periods in Eurasia was represented, together with a detailed continuous sequence of human occupation of Denmark specifically. Our assemblage is well-represented with individuals of such key archaeological complexes as the Maglemose, Ertebølle and Funnel Beaker cultures in Scandinavia, the Cardial in the Mediterranean, the Körös and Linear Pottery complexes in SE and Central Europe, and many archaeological cultures in Ukraine, western Russia, and the trans-Ural (e.g. Veretye, Lyalovo, Volosovo, Kitoi).
Data collection	Sampled remains consisted of teeth (sampled for dental cementum, N=211), petrous bones (N=91), and other bone types (N=15; long bones, ribs and cranial bones). Sampling and ancient DNA laboratory work was undertaken in dedicated ancient DNA clean lab facilities at the GeoGenetics Centre. Additionally, 272 novel accelerator mass spectrometry results were generated at the 14CHRONO laboratory, Queen's University Belfast (242 samples), at the Oxford Radiocarbon Accelerator Unit (ORAU) laboratory (24 samples) and at the Keck-CCAMS Group, Irvine, California, USA (6 samples).
Timing and spatial scale	Sample chronology was generated by radiocarbon dating, with dates corrected for marine and freshwater reservoir effects. These showed samples ranging from the Upper Palaeolithic (c. 25,700 cal. BP) to the mediaeval period (c. 1200 cal. BP). Most individuals (97%, N=309) span 11,000 cal. BP to 3,000 cal. BP, the period broadly associated with the Mesolithic and Neolithic in Eurasia. Our research area can broadly be divided into three large regions: 1) central, western and northern Europe, 2) eastern Europe including western Russia and Ukraine, and 3) the Urals and western Siberia.
Data exclusions	Samples that were screened by shallow sequencing and found to have endogenous DNA of >1% were not sequenced further and not included in analyses. This amounted to 645 specimens.
Reproducibility	Data quality and uncertainty (e.g. contamination) was accounted for in computational analyses to assess robustness of inferences, and all methods and data are made available for future replication.
Randomization	Sample groups were defined to reflect archaeological populations, based on phylogenetic inferences, temporal and geographic provenance, and cultural interpretations evidenced by archaeological contexts.
Blinding	Blinding was not applicable to this study.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input type="checkbox"/>	<input checked="" type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging



Specimen provenance

Samples analysed in this study represent a considerable transect of Eurasian Mesolithic and Neolithic archaeological sites. Institution: 176 in total. Given the highly collaborative, international scale of the sampling effort, key contact persons for each site helped facilitate the access to material and coordinate legal authorisation for accessing sample material. Samples were collected with permission from organisations holding specimens and documented via permissions or agreements letters held by the Lundbeck Foundation GeoGenetics Centre. Full details and documentation for these is maintained in a dedicated database by F.B., K.G.S and Pernille Olsen of the GeoGenetics Centre. The following list provides the site name, country, sample ID, contact and institution. Further details of each site and excavation history are provided in Supplementary Notes 5-6.

Afontova Gora, Russia. Samples: NEO102. Key contact: Mikhail Sablin. Institution: Zoological Institute of the Russian Academy of Sciences.

Aknashen, Armenia. Samples: NEO110. Key contact: Levon Yeppiskoposyan. Institution: National Academy of Sciences of Armenia, Laboratory of Ethnogenomics, Yerevan, Armenia.

Ångdala, Sweden. Samples: NEO046. Key contact: Catharina Ödman / Yvonne Magnusson. Institution: Malmö museum.

Avelebjerg, Denmark. Samples: NEO961. Key contact: Anders Fischer. Institution: Sealand Archaeology/University of Copenhagen.

Baile herculane, Romania. Samples: NEO674. Key contact: Dušan Borić. Institution: Columbia university.

Banks Tomb, UK. Samples: NEO624, NEO625, NEO626, NEO627, NEO630, NEO717. Key contact: Nick Card. Institution: Orkney Research Centre for Archaeology. Institution: ORCA.

Bazaiha, Russia. Samples: NEO070. Key contact: Mikhail Sablin. Institution: Zoological Institute of the Russian Academy of Sciences.

Biestamak, Kazakhstan. Samples: NEO057. Key contact: Emma Usmanova, Irina Shevnina, Andrey Logvin. Institution: Karaganda State University.

Bjørnsholm, Denmark. Samples: NEO751. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Bodal K, Denmark. Samples: NEO814. Key contact: Lasse Sørensen, Peter Vang, Anders Fischer. Institution: National Museum of Denmark

Bol'shoy Oleni Ostrov, Russia. Samples: NEO060, NEO061, NEO062. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Borovjanka XVII, Russia. Samples: NEO080, NEO081, NEO083. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and Ethnography (Kunstkamera).

Borreby, Denmark. Samples: NEO735, NEO737. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Bredgården, Sweden. Samples: NEO027. Key contact: Maria Vretemark. Institution: Västergötlands museum.

Bybjerg, Denmark. Samples: NEO563. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Byholm, Denmark. Samples: NEO564. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Camino de las Yeseras, Spain. Samples: NEO721. Key contact: Corina Liesau, Pilar Prieto. Institution: Universidad Autónoma de Madrid.

Dalmosegaard, Denmark. Samples: NEO886. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Dezsk, Hungary. Samples: NEO130. Key contact: Vaclav Smrcka. Institution: Institute for History of Medicine and Foreign Languages of the First Faculty of Medicine, Charles University.

Døjringe, Denmark. Samples: NEO566. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Dolgoe Ozero, Russia. Samples: NEO292. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Dösemarken, Sweden. Samples: NEO044. Key contact: Catharina Ödman / Yvonne Magnusson. Institution: Malmö museum.

Dragsholm, Denmark. Samples: NEO822. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Danish National Museum.

Dragsholm, Denmark. Samples: NEO732, NEO733. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

El Mazo, Spain. Samples: NEO646, NEO938. Key contact: Igor Gutiérrez Zugasti. Institution: Universidad de Cantabria.

Elkenøre, Denmark. Samples: NEO888. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Ertebølle, Denmark. Samples: NEO568, NEO569. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Evensås, Sweden. Samples: NEO260. Key contact: Karl-Göran Sjögren. Institution: Gothenburg university.

Falköping 5, Sweden. Samples: NEO220, NEO221, NEO223, NEO224, NEO225, NEO226, NEO227, NEO228. Key contact: Malou Blank. Institution: Gothenburg university.

Fannerup D, Denmark. Samples: NEO855. Key contact: Lutz Klassen. Institution: Museum Østjylland.

Fannerup E, Denmark. Samples: NEO570. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Fannerup F, Denmark. Samples: NEO930. Key contact: Lutz Klassen / Rikke Maring. Institution: Museum Østjylland.

Femhøve Vig, Denmark. Samples: NEO744. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Fofonovo, Russia. Samples: NEO199, NEO200, NEO201. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Fontenoce, Italy. Samples: NEO830. Key contact: Gabriele Scorrano. Institution: University of Copenhagen.

Frälsegården, Sweden. Samples: NEO259. Key contact: Karl-Göran Sjögren. Institution: Gothenburg university.

Gaudio, Italy. Samples: NEO828. Key contact: Gabriele Scorrano. Institution: University of Copenhagen.

Gazel Cave, France. Samples: NEO812, NEO813. Key contact: Patrice Cortaud. Institution: Université de Bordeaux.

Golubaya Krinitsa, Russia. Samples: NEO204, NEO207, NEO209, NEO210, NEO212. Key contact: Ruslan Turin. Institution: Karaganda

State University.  
Golubaya Krinitza, Russia. Samples: NEO113. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.  
Gorzsa Cukormajor, Hungary. Samples: NEO140, NEO142, NEO143, NEO145, NEO147, NEO148, NEO149. Key contact: Vaclav Smrcka. Institution: Institute for History of Medicine and Foreign Languages of the First Faculty of Medicine, Charles University.  
Gregorievka 1, Kazakhstan. Samples: NEO899. Key contact: Victor Merc/Emma Usmanova. Institution: Karaganda State University.  
Grøfte, Denmark. Samples: NEO571. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Grotta Delle Mura, Italy. Samples: NEO806. Key contact: Gabriele Scorrano. Institution: University of Copenhagen.  
Grotta Nisco, Italy. Samples: NEO823. Key contact: Gabriele Scorrano. Institution: University of Copenhagen.  
Gruta do Caldeirão, Portugal. Samples: NEO631, NEO632. Key contact: Joao Zilhao. Institution: Universitat de Barcelona.  
Hanaskede, Sweden. Samples: NEO018. Key contact: Maria Vretemark. Institution: Västergötlands museum.  
Havnø, Denmark. Samples: NEO941. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/University of Copenhagen.  
Hedegaard. Institution: Bislev), Denmark. Samples: NEO013. Key contact: Bjarne Nielsen.  
Henriksholm-Bøgebakken. Institution: Vedbæk), Denmark. Samples: NEO745, NEO746, NEO747, NEO748, NEO749. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Hetty Pegler's Tump, UK. Samples: NEO847. Key contact: Alan Outram. Institution: University of Exeter.  
Hindby Mosse, Sweden. Samples: NEO036, NEO038, NEO039. Key contact: Catharina Ödman / Yvonne Magnusson. Institution: Malmö museum.  
Hödmezövásáhely Kotac, Hungary. Samples: NEO137. Key contact: Vaclav Smrcka. Institution: Institute for History of Medicine and Foreign Languages of the First Faculty of Medicine, Charles University.  
Holmegård-Djursland, Denmark. Samples: NEO001. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Hove Å, Denmark. Samples: NEO946. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/University of Copenhagen.  
Hummervik, Norway. Samples: NEO017. Key contact: Per Åke Persson. Institution: UiO: Kulturhistorisk museum.  
Igren, Ukraine. Samples: NEO509, NEO516, NEO518, NEO521. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.  
Itkul', Russia. Samples: NEO063, NEO064, NEO065. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and Ethnography (Kunstkamera).  
Jorløse Mose, Denmark. Samples: NEO023. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Jørlundegaard, Denmark. Samples: NEO702. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Kainsbakke, Denmark. Samples: NEO025. Key contact: Lutz Klassen. Institution: Museum Østjylland.  
Karavaikha, Russia. Samples: NEO555, NEO556, NEO557, NEO558, NEO559, NEO560, NEO561. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.  
Kastanjegården, Sweden. Samples: NEO051. Key contact: Catharina Ödman / Yvonne Magnusson. Institution: Malmö museum.  
Klæstrup Holm, Denmark. Samples: NEO951. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/University of Copenhagen.  
Kleshnya III 1998, Ukraine. Samples: NEO278. Key contact: Inna Potekhina. Institution: Institute of Archaeology of Ukrainian National Academy of Sciences.  
Klokkehøj, Denmark. Samples: NEO580. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Koed I, Denmark. Samples: NEO583. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Koed IV, Denmark. Samples: NEO586. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Koelbjerg, Denmark. Samples: NEO254. Key contact: Jesper Hansen. Institution: Odense Bys Museer.  
Køge Sønakke, Denmark. Samples: NEO759. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Kolind, Denmark. Samples: NEO738, NEO739. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Kongemose, Denmark. Samples: NEO587. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Korsør Glasværk, Denmark. Samples: NEO589. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Korsør Nor, Denmark. Samples: NEO791. Key contact: Lasse Sørensen, Poul Otto Nielsen. Institution: National Museum of Denmark.  
Kotias Klde Cave, Georgia. Samples: NEO281, NEO283. Key contact: David O. Lordkipanidze. Institution: Georgian National Museum.  
Ksizovo 6, Russia. Samples: NEO172, NEO173, NEO174, NEO175. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.  
Kumyshanskaya Cave, Russia. Samples: NEO687. Key contact: Andrej Evteev. Institution: Moscow State University.  
Kyndeløse, Denmark. Samples: NEO878. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Læsten Mose, Denmark. Samples: NEO945. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/University of Copenhagen.  
Langø Skaldyng, Denmark. Samples: NEO853. Key contact: Anders Fischer/ Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Lendemark, Denmark. Samples: NEO896. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Lepenski Vir, Serbia. Samples: NEO669. Key contact: Dušan Borčić. Institution: Columbia university.  
Lohals Nord, Denmark. Samples: NEO029. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.  
Lollikehuse, Denmark. Samples: NEO857. Key contact: Kristian Gregersen. Institution: Statens Naturhistoriske Museum.  
Lundby-Falster, Denmark. Samples: NEO865, NEO866. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand

Archaeology/University of Copenhagen.

Lysa Gora, Ukraine. Samples: NEO262, NEO265. Key contact: Inna Potekhina. Institution: Institute of Archaeology of Ukrainian National Academy of Sciences.

Maddalena, Italy. Samples: NEO695. Key contact: Alfredo Coppa. Institution: Università di Roma Sapienza.

Madesø, Denmark. Samples: NEO752. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Magleø, Denmark. Samples: NEO590. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Mamaj Gora, Ukraine. Samples: NEO268, NEO270. Key contact: Inna Potekhina. Institution: Institute of Archaeology of Ukrainian National Academy of Sciences.

Mandrin Cave, France. Samples: NEO119, NEO120, NEO121. Key contact: Ludovic Slimak. Institution: Université Toulouse.

Mandzuli-Depe, Turkmenistan. Samples: NEO310. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Mergen' 6, Russia. Samples: NEO072, NEO073. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Minino, Russia. Samples: NEO536, NEO537, NEO538, NEO539. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.

Mora Cavorso, Italy. Samples: NEO834. Key contact: Gabriele Scorrano. Institution: University of Copenhagen.

Mosedede Mose. Institution: Karlsruhle Mose), Denmark. Samples: NEO860. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Mosedede Mose III, Denmark. Samples: NEO861. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Myrebjerg Mose, Denmark. Samples: NEO925. Key contact: Otto Christian Uldum. Institution: Langelands Museum.

Næs, Denmark. Samples: NEO792. Key contact: Lasse Sørensen, Poul Otto Nielsen. Institution: National Museum of Denmark.

Nederst, Denmark. Samples: NEO856. Key contact: Lutz Klassen. Institution: Museum Østjylland.

Neverkær Mose I, Denmark. Samples: NEO594. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Norsminde, Denmark. Samples: NEO852. Key contact: Anders Fischer/ Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Okunevo 5, Russia. Samples: NEO068, NEO077. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Okunevo 7, Russia. Samples: NEO079. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Omskaya Stoyanka II, Russia. Samples: NEO075, NEO078. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Orehoved Sejlrende, Denmark. Samples: NEO122, NEO123. Key contact: Anders Fischer, Pia Bennike, Morten Johansen. Institution: Sealand Archaeology/University of Copenhagen.

Ostrov 2, Russia. Samples: NEO076. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Pad Tokui, Russia. Samples: NEO116. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Pandebjerg, Denmark. Samples: NEO595. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Peschanitsa, Russia. Samples: NEO202. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Pogostishche, Russia. Samples: NEO554. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.

Porsmose, Denmark. Samples: NEO795. Key contact: Lasse Sørensen, Poul Otto Nielsen. Institution: National Museum of Denmark.

Protoka, Russia. Samples: NEO309. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Ravnsbjerggård, Denmark. Samples: NEO960. Key contact: Anders Fischer. Institution: Sealand Archaeology/University of Copenhagen.

Rødhals (Sejerø), Denmark. Samples: NEO645. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Røntesten, Denmark. Samples: NEO019. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Roskilde Fjord, Denmark. Samples: NEO891. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Rude, Denmark. Samples: NEO041, NEO043. Key contact: Niels Lynnerup. Institution: Department of Forensic Medicine, Copenhagen university.

Sakhtish 8, Russia. Samples: NEO184. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Sakhtish II, Russia. Samples: NEO158, NEO178, NEO179, NEO192. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Sakhtish Ila, Russia. Samples: NEO180, NEO181, NEO182, NEO183, NEO185, NEO186, NEO187, NEO188, NEO189, NEO193, NEO194, NEO195, NEO197. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Salpetermosen, Denmark. Samples: NEO028. Key contact: Niels Lynnerup/Thomas Jørgensen.

Santa Maira, Spain. Samples: NEO694. Key contact: Carles Lalueza-Fox. Institution: Pompeu Fabra.

Sao Paulo II, Portugal. Samples: NEO603, NEO609. Key contact: Ana Maria Silva. Institution: University of Coimbra.

Schela Cladovei, Romania. Samples: NEO671, NEO672. Key contact: Dušan Borić. Institution: Columbia university.

Sejerby (Sejerø), Denmark. Samples: NEO757. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Sigersdal, Denmark. Samples: NEO007. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Sigersdal Mose, Denmark. Samples: NEO753. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Sillvik, Sweden. Samples: NEO261. Key contact: Karl-Göran Sjögren. Institution: Gothenburg university.

Sjauke, Kazakhstan. Samples: NEO900. Key contact: Victor Merc/Emma Usmanova. Institution: Karaganda State University.

Sjauke 1, Kazakhstan. Samples: NEO902. Key contact: Victor Merc/Emma Usmanova. Institution: Karaganda State University.

Sjauke settlement, Kazakhstan. Samples: NEO901. Key contact: Victor Merc/Emma Usmanova. Institution: Karaganda State University.

Sjiderti 10, Kazakhstan. Samples: NEO904. Key contact: Victor Merc/Emma Usmanova. Institution: Karaganda State University.

Skateholm I, Sweden. Samples: NEO679. Key contact: Douglas Price / Lars Larsson,. Institution: Lund university.

Slonowice, Poland. Samples: NEO640, NEO641. Key contact: Piotr Włodarczyk. Institution: Polish Academy of Sciences. Institution: Instytut Archeologii i Etnologii.

Sludegård Sømose, Denmark. Samples: NEO933. Key contact: Anders Fischer / Jesper Hansen. Institution: Odense Bys museer.

Sølager, Denmark. Samples: NEO598. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Sope-2, Estonia. Samples: NEO306. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Sosnovyy Myis, Russia. Samples: NEO841, NEO843. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Stenderup Hage, Denmark. Samples: NEO943. Key contact: Anders Fischer/Niels Lynnerup.

Storelyng. Institution: Øgård boat III), Denmark. Samples: NEO597. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Storelyng (Østrup Homo II), Denmark. Samples: NEO602. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Strøby Egede, Denmark. Samples: NEO092. Key contact: Niels Lynnerup. Institution: Department of Forensic Medicine, Copenhagen university.

Strøby Grøftemark, Denmark. Samples: NEO091. Key contact: Niels Lynnerup, Anders Fischer. Institution: Department of Forensic Medicine, Copenhagen university.

Strøby Ladeplads, Denmark. Samples: NEO093. Key contact: Niels Lynnerup. Institution: Department of Forensic Medicine, Copenhagen university.

Svinninge Vejle, Denmark. Samples: NEO898. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Tepe Guran, Iran. Samples: NEO816, NEO817, NEO819. Key contact: Peder Mortensen, Pernille Bangsgaard. Institution: Statens Naturhistoriske Museum.

Tingbjerggård Vest, Denmark. Samples: NEO957. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/ University of Copenhagen.

Tissøe, Denmark. Samples: NEO942. Key contact: Anders Fischer/Niels Lynnerup. Institution: Sealand Archaeology/University of Copenhagen.

Toftum, Skanderborg, Denmark. Samples: NEO870, NEO872, NEO875, NEO876. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Toral III, Spain. Samples: NEO648, NEO649, NEO650, NEO653. Key contact: Igor Gutiérrez Zugasti. Institution: Universidad de Cantabria.

Troldebjerg, Denmark. Samples: NEO934. Key contact: Anders Fischer/Otto Uldum. Institution: Langelands Museum, Rudkøbing.

Tudse Hage, Denmark. Samples: NEO932. Key contact: Anders Fischer. Institution: Museum Vestsjælland, Museum Vestsjælland.

Tybrind Vig, Denmark. Samples: NEO683. Key contact: Anders Fischer. Institution: Sealand Archaeology/University of Copenhagen.

Tysmosen II, Denmark. Samples: NEO790. Key contact: Lasse Sørensen, Anders Fischer, Poul Otto Nielsen. Institution: National Museum of Denmark.

Ural River Beach, Russia. Samples: NEO100. Key contact: Mikhail Sablin. Institution: Zoological Institute of the Russian Academy of Sciences.

Ust'-Isha, Russia. Samples: NEO067. Key contact: Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Vængesø II, Denmark. Samples: NEO003. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Vanløse Mose II, Denmark. Samples: NEO599. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Vasøgård, Denmark. Samples: NEO815. Key contact: Lasse Sørensen. Institution: National Museum of Denmark.

Vasilevka-I, Ukraine. Samples: NEO492, NEO494, NEO496, NEO497, NEO501, NEO545, NEO549, NEO550. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.

Vasilevsky, Ukraine. Samples: NEO305. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Vasilyevskiy kordon 17, Russia. Samples: NEO160, NEO162, NEO163, NEO164, NEO166, NEO167, NEO168, NEO170, NEO171. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Vattenledningen, Sweden. Samples: NEO052. Key contact: Catharina Ödman / Yvonne Magnusson. Institution: Malmö museum.

Vedbæk Boldbaner, Denmark. Samples: NEO600. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Vedrovice, Czech Republic. Samples: NEO128. Key contact: Vaclav Smrcka. Institution: Institute for History of Medicine and Foreign Languages of the First Faculty of Medicine, Charles University.

Vengerovo-2, Russia. Samples: NEO907, NEO910, NEO911, NEO912, NEO915, NEO916, NEO917, NEO918, NEO921, NEO922, NEO923. Key contact: V. I. Molodin. Institution: Institute of Archaeology and Ethnography, Russian Academy of Science, Siberian Branch.

Vibygårds Mose, Denmark. Samples: NEO935. Key contact: Anders Fischer / Per Lotz. Institution: Værløse Museum.

Viksø Mose, Denmark. Samples: NEO601. Key contact: Niels Lynnerup, Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/ University of Copenhagen.

Vittrup Mose, Denmark. Samples: NEO033. Key contact: Anders Fischer, Pia Bennike. Institution: Sealand Archaeology/University of Copenhagen.

Vlasac, Serbia. Samples: NEO655, NEO657, NEO658, NEO677. Key contact: Dušan Borić. Institution: Columbia university.

Volnensky, Ukraine. Samples: NEO300, NEO302, NEO304. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Voloshskoe, Ukraine. Samples: NEO522, NEO527. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.

Vovnigi, Ukraine. Samples: NEO498, NEO502, NEO503, NEO507, NEO508, NEO514, NEO524, NEO528, NEO529, NEO551, NEO552, NEO553. Key contact: Alexandra Buzhilova. Institution: Lomonosov Moscow State University.

Zamostje 2, Russia. Samples: NEO087, NEO088. Key contact: Alisa Zubova/Vyacheslav Moiseyev. Institution: Peter the Great Museum of Anthropology and ethnography (Kunstkamera).

Zhindo, Russia. Samples: NEO115, NEO117. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Zvejnieki, Latvia. Samples: NEO307. Key contact: Sergey Vasilyev. Institution: Inst. of Ethnography and Anthropology, Moscow.

Specimen deposition

All specimens studied are available upon direct contact/request to the archaeologists, curators or officials responsible for their curation at the organisation where they are held.

Dating methods

272 novel radiocarbon dates were generated at the 14CHRONO laboratory, Queen's University Belfast (242 samples), at the Oxford Radiocarbon Accelerator Unit (ORAU) laboratory (24 samples) and at the Keck-CCAMS Group, Irvine, California, USA (6 samples).

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

No ethical oversight was required - all specimens were sampled following protocols designed to minimise invasive impact of sampling, with the full agreement of curators or archaeologists responsible for collections where specimens are held. Our use of shotgun whole genome data also reduces the need to resequence samples in the future (in contrast to targeting specific loci).

Note that full information on the approval of the study protocol must also be provided in the manuscript.