POPULATION GENOMICS OF STONE AGE EURASIA

Morten E. Allentoft[1,2]*§, Martin Sikora[1]*§, Alba Refoyo-Martínez[1]§, Evan K. Irving-Pease[1]§, Anders Fischer[3,4]§, William Barrie[5]§, Andrés Ingason[6,1]§, Jesper Stenderup[1], Karl-Göran Sjögren[3], Alice Pearson[7], Barbara Mota[8,9], Bettina Schulz Paulsson[3], Alma Halgren[10], Ruairidh Macleod[1,5,11], Marie Louise Schjellerup Jørkov[12], Fabrice Demeter[1,13], Maria Novosolov[1], Lasse Sørensen[14], Poul-Otto Nielsen[14], Rasmus H.A. Henriksen[1], Tharsika Vimala[1], Hugh McColl[1], Ashot Margaryan[15,16], Melissa Ilardo[17], Andrew Vaughn[18], Morten Fischer Mortensen[14], Anne Birgitte Nielsen[19], Mikkel Ulfeldt Hede[20], Peter Rasmussen[14], Lasse Vinner[1], Gabriel Renaud[21], Aaron Stern[18], Theis Zetner Trolle Jensen[15], Niels Nørkjær Johannsen[22], Gabriele Scorrano[1], Hannes Schroeder[15], Per Lysdahl[23], Abigail Daisy Ramsøe[1], Andrei Skorobogatov[24], Andrew Joseph Schork[6,25], Anders Rosengren[6,1], Anthony Ruter[1], Alan Outram[26], Aleksey A. Timoshenko[27], Alexandra Buzhilova[28], Alfredo Coppa[29], Alisa Zubova[30], Ana Maria Silva[31,59], Anders J. Hansen[1], Andrey Gromov[30], Andrey Logvin[32], Anne Birgitte Gotfredsen[1], Bjarne Henning Nielsen[33], Borja González-Rabanal[34], Carles Lalueza-Fox[35], Catriona J. McKenzie[26], Charleen Gaunitz[1], Concepción Blasco[36], Corina Liesau[36], Cristina Martinez-Labarga[37], Dmitri V. Pozdnyakov[27], David Cuenca-Solana[38,39], David O. Lordkipanidze[40,41], Dmitri En'shin[42], Domingo C. Salazar-García[43,44], T. Douglas Price[45], Dušan Borić[29,46], Elena Kostyleva[47], Elizaveta V. Veselovskaya[48], Emma R. Usmanova[49,50], Enrico Cappellini[15], Erik Brinch Petersen[51], Esben Kannegaard[52], Francesca Radina[53], Fulya Eylem Yediay[1], Henri Duday[54], Igor Gutiérrez-Zugasti[38], Inna Potekhina[55,56], Irina Shevnina[32], Isin Altinkaya[1], Jean Guilaine[57], Jesper Hansen[58], Joan Emili Aura Tortosa[43], João Zilhão[59,60], Jorge Vega[61], Kristoffer Buck Pedersen[62], Krzysztof Tunia[63], Lei Zhao[1], Liudmila N. Mylnikova[27], Lars Larsson[64], Laure Metz[65], Levon Yeppiskoposyan[66,93], Lisbeth Pedersen[67], Lucia Sarti[68], Ludovic Orlando[69], Ludovic Slimak[65], Lutz Klassen[52], Malou Blank[3], Manuel González-Morales[38], Mara Silvestrini[70], Maria Vretemark[71], Marina S. Nesterova[27], Marina Rykun[72], Mario Federico Rolfo[73], Marzena Szmyt[74], Marcin Przybyła[75], Mauro Calattini[68], Mikhail Sablin[76], Miluše Dobisíková[77], Morten Meldgaard[78], Morten Johansen[79], Natalia Berezina[28], Nick Card[80], Nikolai A. Saveliev[81], Olga Poshekhonova[42], Olga Rickards[37], Olga V. Lozovskaya[82], Otto Christian Uldum[79], Paola Aurino[83], Pavel Kosintsev[84,85], Patrice Courtaud[54], Patricia Ríos[36], Peder Mortensen[86], Per Lotz[87,88], Per Åke Persson[89], Pernille Bangsgaard[90], Peter de Barros Damgaard[1], Peter Vang Petersen[14], Pilar Prieto Martinez[91], Piotr Włodarczak[63], Roman V. Smolyaninov[92], Rikke Maring[22,52], Roberto Menduiña[61], Ruben Badalyan[91], Rune Iversen[51], Ruslan Turin[24], Sergey Vasilyiev[27], Sidsel Wåhlin[23], Svetlana Borutskaya[28], Svetlana Skochina[42], Søren Anker Sørensen[87], Søren H. Andersen[94], Thomas Jørgensen[87], Yuri B. Serikov[95], Vyacheslav I. Molodin[27], Vaclav Smrcka[96], Victor Merz[97], Vivek Appadurai[6], Vyacheslav Moiseyev[30], Yvonne Magnusson[98], Kurt H. Kjær[1], Niels Lynnerup[12], Daniel J. Lawson[99], Peter H. Sudmant[10,18], Simon Rasmussen[100], Thorfinn Korneliussen[1]@, Richard Durbin[7,101]@, Rasmus Nielsen[10,1]@, Olivier Delaneau[8]@, Thomas Werge[1,6,102]@, Fernando Racimo[1]@, Kristian Kristiansen[1,3]@, Eske Willerslev[1,5,101,103]*@

Affiliations

[1]Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Copenhagen, Denmark. [2]Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Perth, Australia. [3]Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden. [4]Sealand Archaeology, Gl. Roesnaesvej 27, 4400 Kalundborg, Denmark. [5]GeoGenetics Group, Department of Zoology, University of Cambridge, Cambridge, UK. [6]Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Roskilde, Denmark. [7]Department of Genetics, University of Cambridge, Cambridge, UK. [8]Department of Computational Biology, University of Lausanne, Switzerland. [9]Swiss Institute of Bioinformatics, University of Lausanne, Switzerland. [10]Department of Integrative Biology, University of California, Berkeley, USA. [11]Research department of Genetics, Evolution and Environment, University College London, London, UK. [12]Laboratory of Biological Anthropology, Department of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark. [13]Muséum national d'Histoire naturelle, CNRS, Université de Paris, Musée de l'Homme, Paris, France. [14]The National Museum of Denmark, Ny Vestergade 10, Copenhagen, Denmark. [15]Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. [16]Centre for Evolutionary Hologenomics, University of Copenhagen, Copenhagen, Denmark. [17]Anthropology Department, University of Utah, USA. [18]Center for Computational Biology, University of California, Berkeley, USA. [19]Department of Geology, Lund University, Lund, Sweden. [20]Tårnby Gymnasium og HF, Kastrup, Denmark. [21]Department of Health Technology, Section of Bioinformatics, Technical University of Denmark, Kongens Lyngby, Denmark. [22]Department of Archaeology and Heritage Studies, Aarhus University, Aarhus, Denmark. [23]Vendsyssel Historiske Museum, DK-9800 Hjørring, Denmark. [24]Terra Ltd., Letchik Zlobin St. 20, Voronezh, 394055, Russian Federation. [25]Neurogenomics Division, The Translational Genomics Research Institute (TGEN), Phoenix, AZ, USA. [26]Department of Archaeology, University of Exeter, Exeter, UK. [27]Institute of Archaeology and Ethnography, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation. [28]Research Institute and Museum of Anthropology, Lomonosov Moscow State University, Mokhovaya str. 11, Moscow, Russian Federation. [29]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy. [30]Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint Petersburg, Russian Federation. [31]CIAS, Department of Life Science, University of

Coimbra, Coimbra, Portugal. [32]Kostanay Regional University A. Baitursynov, Kostanay, Kazakhstan. [33]Vesthimmerlands Museum, Søndergade 44, Aars, Denmark. [34]Grupo EvoAdapta, Departamento de Ciencias Históricas, Universidad de Cantabria, Santander, Spain. [35]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain. [36]Departamento de Prehistoria y Arqueología Department, Universidad Autónoma de Madrid, Madrid, Spain. [37]Department of Biology, University of Rome "Tor Vergata", Rome, Italy. [38]Instituto Internacional de Investigaciones Prehistóricas de Cantabria, Universidad de Cantabria, Santander, Spain. [39]Centre de Recherche en Archéologie, Archeosciences, Histoire (CReAAH), UMR-6566 CNRS, Rennes, France. [40]Georgian National Museum, Tbilisi, Georgia. [41]Tbilisi State University, Tbilisi, Georgia. [42]IPND, Tyumen Scientific Centre, Siberian Branch of the Russian Academy of Sciences, Tyumen, Russian Federation. [43]Departament de Prehistòria, Arqueologia i Història Antiga, Universitat de València, València, Spain. [44]Department of Geological Sciences, University of Cape Town, Cape Town, South Africa. [45]Laboratory for Archaeological Chemistry, Department of Anthropology, University of Wisconsin-Madison, Madison, USA. [46]Department of Anthropology, New York University, New York, USA. [47]Institute of Humanities, Ivanovo State University, Ivanovo, Russian Federation. [48]Institute of Ethnology and Anthropology, Russian Academy of Sciences, Moscow, Russian Federation. [49]Saryarka Archaeological Institute, Buketov Karaganda University, Karaganda, Kazakhstan. [50]South Ural State University, Chelyabinsk, Russia. [51]The Saxo Institute, University of Copenhagen, Copenhagen, Denmark. [52]Museum Østjylland, Stemannsgade 2, Randers, Denmark. [53]Soprintendenza Archeologia Belle Arti e Paesaggio per la Città Metropolitana di Bari, Via Pier l'Eremita, 25, 70122, Bari, Italy. [54]UMR 5199 PACEA, CNRS, Université de Bordeaux, 33615 Pessac, France. [55]Institute of Archaeology, National Academy of Sciences of Ukraine, Kyiv, Ukraine. [56]National University of Kyiv-Mohyla Academy, Kyiv, Ukraine. [57]Collège de France, 75231 Paris cedex 05, France. [58]Odense City Museums, Overgade 48, Odense, Denmark. [59]UNIARQ, University of Lisbon, Lisbon, Portugal. [60]ICREA, University of Barcelona, Barcelona, Spain. [61]ARGEA Consultores SL, C. de San Crispín, Madrid, Spain. [62]Museum Sydøstdanmark, Algade 97, 4760 Vordingborg, Denmark. [63]Institute of Archaeology and Ethnology, Polish Academy of Sciences, Kraków, Poland. [64]Department of Archaeology and Ancient History, Lund University, Lund, Sweden. [65]CNRS UMR 5608, Toulouse Jean Jaurès University, Maison de la Recherche, 5 Allées Antonio Machado, 31058 Toulouse, Cedex 9, France. [66]Institute of Molecular Biology, National Academy of Sciences, Yerevan, Armenia. [67]HistorieUdvikler, Gl. Roesnaesvej 27, DK-4400 Kalundborg, Denmark. [68]Department of history and cultural heritage, University of Siena, Siena, Italy. [69]Centre d'Anthropobiologie et de Génomique de Toulouse, CNRS UMR 5288, Université Paul Sabatier, Toulouse, France. [70]Soprintendenza per i Beni Archeologici delle Marche, Via Birarelli 18, 60100, Ancona, Italy. [71]Västergötlands museum, Stadsträdgården, Skara, Sweden. [72]Cabinet of Anthropology, Tomsk State University, Tomsk, Russian Federation. [73]Department of History, Humanities and Society, University of Rome "Tor Vergata", Rome, Italy. [74]Institute for Eastern Research, Adam Mickiewicz University in Poznań, Poznań, Poland. [75]Institute of Archaeology, Jagiellonian University, Ul. Gołębia 11, 31-007, Kraków, Poland. [76]Zoological Institute of Russian Academy of Sciences, Universitetskaya nab. 1, 199034, St. Petersburg, Russian Federation. [77]Department of Anthropology, Czech National Museum, Prague, Czech Republic. [78]Department of Health and Nature, University of Greenland, Greenland. [79]The Viking Ship Museum, Vindeboder 12, Roskilde, Denmark. [80]Archaeology Institute, University of Highlands and Islands, Scotland, UK. [81]Scientific Research Center "Baikal region", Irkutsk State University; 1, K. Marx st., Irkutsk, 664003, Russian Federation. [82]Laboratory for Experimental Traceology, Institute for the History of Material Culture of the Russian Academy of Sciences, Dvortsovaya nab., 18, 191186, St. Petersburg, Russian Federation. [83]Soprintendenza Archeologia, Belle Arti e Paesaggio per la provincia di Cosenza, Cosenza, Italy. [84]Paleoecology Laboratory, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russian Federation. [85]Department of History of the Institute of Humanities, Ural Federal University, Ekaterinburg, Russian Federation. [86]Centre for the Study of Early Agricultural Societies, Department of Cross-Cultural and Regional Studies, University of Copenhagen, 2300 Copenhagen, Denmark. [87]Museum Nordsjælland, Frederiksgade 9, 3400 Hillerød. [88]Museum Vestsjælland, Klosterstræde 18, 4300 Holbæk, Denmark. [89]Museum of Cultural History, University of Oslo, P.O. Box 6762. St. Olavs Plass NO-0130 Oslo, Norway. [90]ArchaeoScience, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. [91]Department of History, University of Santiago de Compostela, Spain. [92]Lipetsk Regional Scientific Public Organisation "Archaeological Research", Lipetsk, Russian Federation. [92]Institute of Archaeology and Ethnography, National Academy of Sciences, Yerevan, Armenia. [93]Russian-Armenian University, Yerevan, Armenia. [94]Moesgaard Museum, Moesgård Allé 15, Højbjerg, Denmark. [95]Nizhny Tagil State Socio-Pedagogical Institute, Nizhny Tagil, Russia. [96]Institute for History of Medicine, First Faculty of Medicine, Charles University, Prague, Czech Republic. [97]Centre for Archaeological Research Toraighyrov University, Pavlodar, Kazakhstan. [98]Malmö Museer, Malmöhusvägen 6, Malmö, Sweden. [99]Institute of Statistical Sciences, School of Mathematics, University of Bristol, Bristol, UK. [100]Novo Nordisk Foundation Centre for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen N, Denmark. [101]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK. [102]Department of Clinical Medicine, University of Copenhagen, 2200 Copenhagen N, Denmark. [103]MARUM, University of Bremen, Bremen, Germany.

* Corresponding authors; email: morten.allentoft@curtin.edu.au, martin.sikora@sund.ku.dk, ew482@cam.ac.uk
§ Joint first authors

@ Joint last authors

Summary

The transitions from foraging to farming and later to pastoralism in Stone Age Eurasia (c. 11-3 thousand years before present, BP) represent some of the most dramatic lifestyle changes in human evolution. We sequenced 317 genomes of primarily Mesolithic and Neolithic individuals from across Eurasia combined with radiocarbon dates, stable isotope data, and pollen records. Genome imputation and co-analysis with previously published shotgun sequencing data resulted in >1600 complete ancient genome sequences offering fine-grained resolution into the Stone Age populations. We observe that: 1) Hunter-gatherer groups were more genetically diverse than previously known, and deeply divergent between western and eastern Eurasia. 2) We identify hitherto genetically undescribed hunter-gatherers from the Middle Don region that contributed ancestry to the later Yamnaya steppe pastoralists; 3) The genetic impact of the Neolithic transition was highly distinct, east and west of a boundary zone extending from the Black Sea to the Baltic. Large-scale shifts in genetic ancestry occurred to the west of this "Great Divide", including an almost complete replacement of hunter-gatherers in Denmark, while no substantial ancestry shifts took place during the same period to the east. This difference is also reflected in genetic relatedness within the populations, decreasing substantially in the west but not in the east where it remained high until c. 4,000 BP; 4) The second major genetic transformation around 5,000 BP happened at a much faster pace with Steppe-related ancestry reaching most parts of Europe within 1,000-years. Local Neolithic farmers admixed with incoming pastoralists in eastern, western, and southern Europe whereas Scandinavia experienced another near-complete population replacement. Similar dramatic turnover-patterns are evident in western Siberia; 5) Extensive regional differences in the ancestry components involved in these early events remain visible to this day, even within countries. Neolithic farmer ancestry is highest in southern and eastern England while Steppe-related ancestry is highest in the Celtic populations of Scotland, Wales, and Cornwall (this research has been conducted using the UK Biobank resource); 6) Shifts in diet, lifestyle and environment introduced new selection pressures involving at least 21 genomic regions. Most such variants were not universally selected across populations but were only advantageous in particular ancestral backgrounds. Contrary to previous claims, we find that selection on the FADS regions, associated with fatty acid metabolism, began before the Neolithisation of Europe. Similarly, the lactase persistence allele started increasing in frequency before the expansion of Steppe-related groups into Europe and has continued to increase up to the present. Along the genetic cline separating Mesolithic hunter-gatherers from Neolithic farmers, we find significant correlations with trait associations related to skin disorders, diet and lifestyle and mental health status, suggesting marked phenotypic differences between these groups with very different lifestyles. This work provides new insights into major transformations in recent human evolution, elucidating the complex interplay between selection and admixture that shaped patterns of genetic variation in modern populations.

Introduction

The transition from hunting and gathering to farming represents one of the most dramatic shifts in lifestyle and diet in human evolution with lasting effects on the modern world. For millions of years our ancestors relied on hunting and foraging for survival but c.12,000 years ago in the Fertile

Crescent of the Near East, plant cultivation and animal husbandry were developed[1–3]. This ultimately resulted in a more sedentary lifestyle accompanied by increasing population sizes and higher social complexity. Expanding populations and the adoption of herding, carried farming practices into Europe and parts of SW Asia in the following millennia, and farming was also developed independently in other parts of the World. Today, 50% of the Earth's habitable land is used for agriculture and very few hunter-gatherers remain[4,5]. Understanding the changes to the human gene pool during this shift from hunter-gathering to farming between the Mesolithic and Neolithic periods is central to understanding ourselves and the events that led to a major transformation of our planet.

While the Neolithisation process has been studied extensively with ancient DNA (aDNA) technology, several key questions remain unaddressed. Population movements during the Neolithic can be traced in the gene pools across the European continent as farming was introduced from the Near East. Several regional studies have testified to varying degrees of reproductive interaction with local Mesolithic groups, ranging from genetic continuity[6] to gradual population admixture[7–10] to almost complete replacement[11]. However, our knowledge of the population structure in the Mesolithic period and how it was formed is limited, partly because of a paucity of data from skeletons older than 8,000 years, compromising resolution into subsequent demographic transitions. Moreover, the spatiotemporal mapping of population dynamics east of Europe, including Siberia, Central- and North Asia during the same time period remains patchy. In these regions the 'Neolithic' typically refers to new forms of lithic material culture, and/or the presence of ceramics[12]. For instance, the Neolithic cultures of the Central Asian Steppe possessed pottery, but retained a hunter-gatherer economy alongside stone blade technology similar to the preceding Mesolithic cultures[13]. The archaeological record testifies to a boundary, ranging from the eastern Baltic to the Black Sea, east of which hunter-gatherer societies persist for much longer than in western Europe[14]. The population genomic implications of this "Great Divide" is, however, largely unknown. Southern Scandinavia represents another enigma in the Neolithisation debate[15]. The introduction of farming reached a 1,000-year standstill at the doorstep to Southern Scandinavia before finally progressing into Denmark around 6,000 BP. It is not known what caused this delay and whether the transition to farming in Denmark, was facilitated by the migration of people (demic diffusion), similar to the rest of Europe[11,16,17] or mostly involved cultural diffusion[18,19]. Starting at around 5,000 BP, a new ancestry component emerged on the eastern European plains associated with Yamnaya Steppe pastoralists culture and swept across Europe mediated through expansion of the Corded Ware complex (CWC) and related cultures[20,21]. The genetic origin of the Yamnaya and the fine-scale dynamics of the formation and expansion of the CWC are largely unresolved questions of central importance to clarify the formation of the present day European gene pool.

Rapid dietary changes and expansion into new climate zones represent shifts in environmental exposure, impacting the evolutionary forces acting on the gene pool. The Neolithisation can therefore be considered as a series of large-scale selection pressures imposed on humans from around 12,000 years ago. Moreover, close contact with livestock and higher population densities have likely enhanced exposure and transmission of infectious diseases, introducing new challenges to our survival[22,23]. While signatures of selection can be identified from patterns of genetic diversity in extant populations[24,25], this can be challenging in species such as humans, which show very wide geographic distributions and have thus been exposed to highly diverse and changing local environments through space and time. In the complex mosaic of ancestries that constitute a modern human genome any putative signatures of selection may therefore misrepresent the timing and magnitude of the actual event unless we can use ancient DNA to chart the individual ancestry components back into the evolutionary past.

To investigate these formative processes in Eurasian prehistory, we conducted the largest ancient DNA study to date on human Stone Age skeletal material. We sequenced low-coverage genomes of 317 radiocarbon-dated (AMS) primarily Mesolithic and Neolithic individuals, covering major parts of Eurasia. We combined these with published shotgun-sequenced data to impute a dataset of >1600 diploid ancient genomes. Genomic data from 100 AMS-dated individuals from Denmark supported detailed analyses of the Stone Age population dynamics in Southern Scandinavia. When combined with genetically-predicted phenotypes, proxies for diet ($\delta^{13}C/\delta^{15}N$), mobility ($^{87}Sr/^{86}Sr$) and vegetation cover (pollen) we could connect this with parallel shifts in phenotype, subsidence and landscape. To test for traces of divergent selection in health and lifestyle-related genetic variants, we used the imputed ancient genomes to reconstruct polygenic risk scores for hundreds of complex traits in the ancient Eurasian populations. Additionally, we used a novel chromosome painting technique based on tree sequences, in order to model ancestry-specific allele frequency trajectories through time. This allowed us to identify many new phenotype-associated genetic variants with hitherto unknown evidence for positive selection in Eurasia throughout the Holocene.

Results/Discussion

Samples and data

In this study we present genomic data from 317 ancient individuals (Fig 1, Extended data fig. 2, Supplement Table I). A total of 272 were radiocarbon dated within the project, while 39 dates were derived from literature and 15 were dated by archaeological context. Dates were corrected for marine and freshwater reservoir effects (Supplementary Note 8) and ranged from the Upper Palaeolithic (UP) c. 25,700 calibrated years before present (cal. BP) to the mediaeval period (c. 1200 cal. BP). However, 97% of the individuals (N=309) span 11,000 cal. BP to 3,000 cal. BP, with a heavy focus on individuals associated with various Mesolithic and Neolithic cultures.

Geographically, the sampled skeletons cover a vast territory across Eurasia, from Lake Baikal to the Atlantic coast, from Scandinavia to the Middle East, and they derive from a variety of contexts, including burial mounds, caves, bogs and the seafloor (Supplementary Notes 6-7). Broadly, we can divide our research area 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. Our samples cover many of the key Mesolithic and Neolithic cultures in Western Eurasia, such as the Maglemose and Ertebølle cultures in Scandinavia, the Cardial in the Mediterranean, the Körös and Linear Pottery (LBK) in SE and Central Europe, and many archaeological cultures in Ukraine, western Russia, and the trans-Ural (e.g. Veretye, Lyalovo, Volosovo, Kitoi). Our sampling was particularly dense in Denmark from where we present a detailed and continuous sequence of 100 genomes spanning from the early Mesolithic to the Bronze Age. Dense sample sequences were also obtained from Ukraine, Western Russia, and the trans-Ural, spanning from the Early Mesolithic through the Neolithic, up to c. 5,000 BP.

We extracted ancient DNA from tooth cementum or petrous bone and shotgun sequenced the 317 genomes to a depth of genomic coverage ranging from 0.01X to 7.1X (mean = 0.75X, median = 0.26X), with 81 individuals having >1X coverage. Using a new imputation method designed for low-coverage sequencing data[26], we performed genotype imputation based on the 1,000 Genomes phased data as a reference panel. We also imputed >1,300 previously published shotgun-sequenced ancient genomes. This resulted in a "raw" dataset containing 8.5 million common Single Nucleotide Polymorphisms (SNPs) (>1% MAF and imputation info score > 0.5) from 1,664 imputed diploid ancient genomes. This number includes 42 high-coverage ancient genomes (Table S2.1, Supplementary Note 2) that were down-sampled to values between 0.1X and 4X for validation.

This demonstrated that 1-fold genome coverage provides remarkably high imputation accuracy (r^2>0.95 at common variants with MAF above 5%) and closely matches what is obtained for modern samples (Extended Fig. 1A-D). African genomes, however, exhibit lower imputation accuracy as a result of the poor representation of this ancestry in the reference panel. For European genomes, this translates into genotyping error rates usually below 5% for the most challenging genotypes to impute (heterozygous genotypes or with two copies of the non-reference allele; Supplementary Fig. S2.1-S2.2). Imputation accuracy also depends on minor allele frequency and genomic coverage (Supplementary Fig. S2.3). We find that coverage values as low as 0.1x and 0.4X are sufficient to obtain r^2 imputation accuracy of 0.8 and 0.9 at common variants (MAF>=10%), respectively. As further validation, we increased genomic coverage to 27.5X, 18.9X and 5.4X on a previously published trio (mother, father, son) from the Late Neolithic mass burial at Koszyce in Poland [27]. This allowed for a validation of imputed genotypes and haplotypes using Mendel's rules of inheritance. We obtained Mendelian error rates from 0.1% at 4X to 0.55% at 0.1X (Extended Fig. 1E). Similarly, we obtained switch error rates between 2% and 6%. Altogether, our validation analysis showed that ancient European genomes can be imputed confidently from coverages above 0.4X and highly valuable data can still be obtained with coverages as low as 0.1X when using specific QC on the imputed data, although at very low coverage a bias arise towards the major allele (see Supplementary Note 2). We filtered out samples with poor coverage or variant sites with low MAF in downstream analyses depending on the specific data quality requirements. For most analyses we use a subset of 1,492 imputed ancient genomes (213 sequenced in this study) after filtering individuals with very low coverages (<0.1X) and/or low imputation quality (average genotype probability < 0.8) and close relatives. This dataset allows us to characterise the ancient cross-continental gene pools and the demographic transitions with unprecedented resolution.

We performed broad-scale characterization of this dataset using principal component analysis (PCA) and model-based clustering (ADMIXTURE), recapitulating and providing increased resolution into previously described ancestry clines in ancient Eurasian populations (Fig. 1; Extended Data Fig. 2; Supplementary Note 3d). Strikingly, inclusion of the imputed ancient genomes in the inference of the principal components reveals much higher variance among the ancient groups than previously anticipated using projection onto a PC-space inferred from modern individuals alone (Extended Data Fig. 2). This is particularly notable in a PCA of West Eurasian individuals, where genetic variation among all present-day populations is confined within a small central area of the PCA (Extended Data Fig. 2C, D). These results are consistent with much higher genetic differentiation between ancient Europeans than present-day populations reflecting lower effective population sizes and genetic isolation among ancient groups.

To obtain a finer-scale characterization of genetic ancestries across space and time, we assigned imputed ancient individuals to genetic clusters by applying hierarchical community detection on a network of pairwise identity-by-descent (IBD)-sharing similarities[28] (Extended Data Fig. 3; Supplementary Note 3c). The obtained clusters capture fine-scale genetic structure corresponding to shared ancestry within particular spatiotemporal ranges and/or archaeological contexts, and were used as sources and/or targets in supervised ancestry modelling (Extended Data Fig. 4; Supplementary Note 3i). We focus our subsequent 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 Fig. 4).

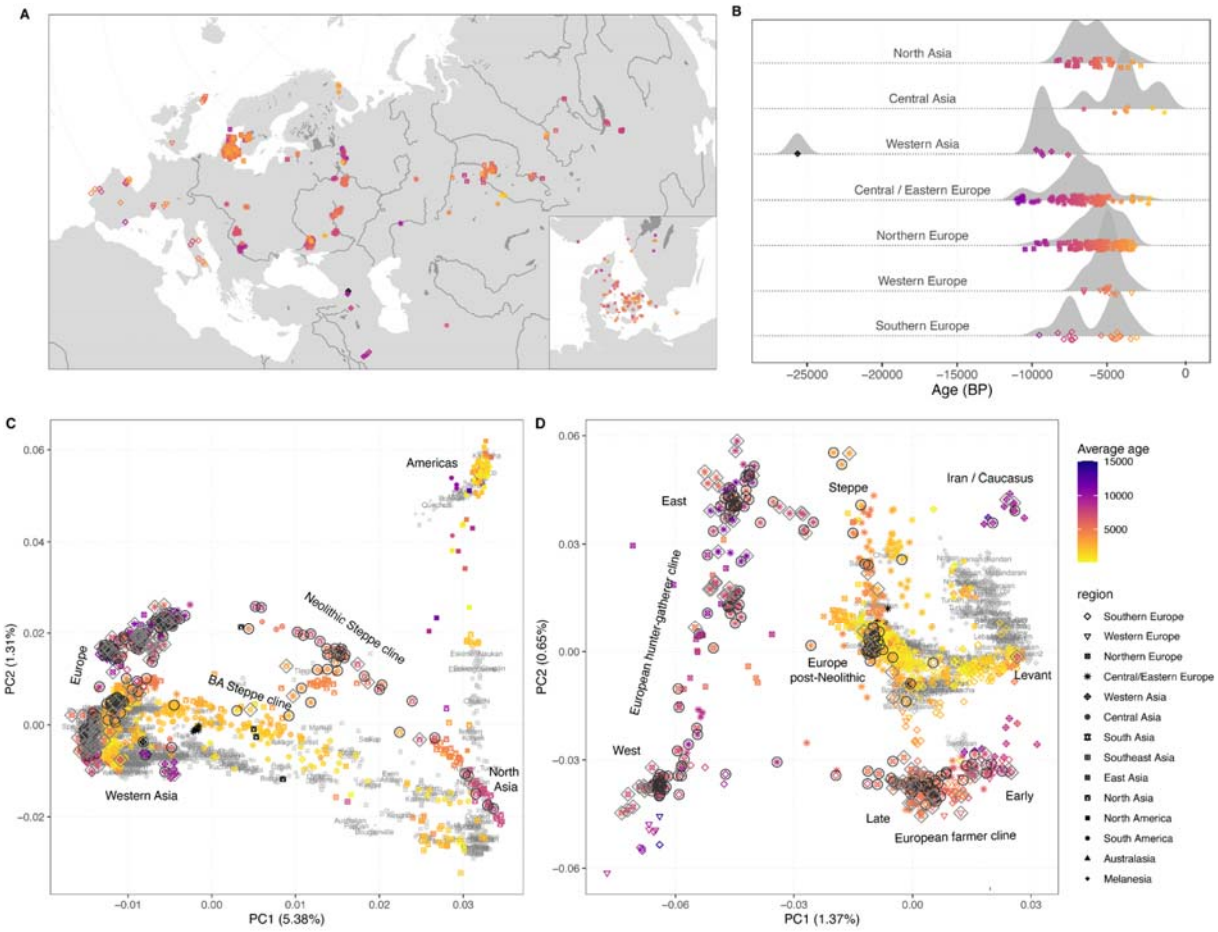


Fig 1. Sample overview and broad scale genetic structure. (A), (B) Geographic and temporal distribution of the 317 ancient genomes reported here. Age and geographic region of ancient individuals are indicated by plot symbol colour and shape, respectively. Random jitter was added to geographic coordinates to avoid overplotting. (C), (D) Principal component analysis of 3,316 modern and ancient individuals from Eurasia, Oceania, and the Americas (C), as well as restricted to 2,126 individuals from western Eurasia (west of Urals) (D). Principal components were defined using both modern and imputed ancient 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.

Deep population structure of western Eurasians

Our study comprises the largest genomic dataset on European hunter-gatherers to date, including 113 imputed hunter-gatherer genomes of which 79 were sequenced in this study. Among them, we report a 0.83X genome of an Upper Palaeolithic (UP) skeleton from Kotias Klde Cave in Georgia, Caucasus (NEO283), directly dated to 26,052 - 25,323 cal BP (95%). In the PCA of all non-African individuals, it occupies a position distinct from other previously sequenced UP individuals, shifted towards west Eurasians along PC1 (Supplementary Note 3d). Using admixture graph modelling, we find that this Caucasus UP lineage derives from a mixture of predominantly West Eurasian UP hunter-gatherer ancestry (76%) with ~24% contribution from a "basal Eurasian" ghost population, first observed in West Asian Neolithic individuals[29] (Extended Data Fig. 5A). Models attempting to reconstruct major post-LGM clusters such as European hunter-gatherers and Anatolian farmers without contributions from this Caucasus UP lineage provided poor admixture graph fits or were rejected in *qpAdm* analyses (Extended Data Fig. 5B,C). These results thus suggest a central role of

the descendants related to this Caucasus UP lineage in the formation of later West Eurasian populations, consistent with recent genetic data from the nearby Dzudzuana Cave, also in Georgia[30].

We performed supervised admixture modelling using a set of twelve possible source clusters representing Mesolithic hunter-gatherers from the extremes of the HG cline, as well as temporal or geographical outgroups of deep Eurasian lineages (Fig 2A). We replicate previous results of broad-scale genetic structure correlated to geography in European hunter-gatherers after the LGM[17], while also revealing novel insights into their fine-scale structure. Ancestry related to southern European hunter-gatherers (source: Italy_15000BP_9000 BP) predominates in western Europe. This includes Denmark, where our 28 sequenced and imputed hunter-gatherer genomes derive almost exclusively from this cluster, with remarkable homogeneity across a 5,000 year transect (Fig. 3A). In contrast, hunter-gatherer individuals from the eastern and far northern reaches of Europe show the highest proportions of Russian hunter-gatherer ancestry (source: RussiaNW_11000BP_8000BP; Fig. 2B, D), with genetic continuity until ~5,000 BP in Russia. Ancestry related to Mesolithic hunter-gatherer populations from Ukraine (source: Ukraine_10000BP_4000BP) is carried in highest proportions in hunter-gatherers from a geographic corridor extending from south-eastern Europe towards the Baltic and southern Scandinavia. Swedish Mesolithic individuals derive up to 60% of their ancestry from that source (Fig. 2C). Our results thus indicate northwards migrations of at least three distinct waves of hunter-gatherer ancestry into Scandinavia: a predominantly southern European source into Denmark; a source related to Ukrainian and south-eastern European hunter-gatherers into the Baltic and southern Sweden; and a northwest Russian source into the far north, before venturing south along the Atlantic coast of Norway[31] (Fig. 2). These movements are likely to represent post glacial expansions from refugia areas shared with many plant and animal species[32,33].

Despite the major role of geography in shaping European hunter-gatherer structure, we also document more complex local dynamics. On the Iberian Peninsula, the earliest individuals, including a ~9,200-year-old hunter-gatherer (NEO694) from Santa Maira (eastern Spain), sequenced in this study, show predominantly southern European hunter-gatherer ancestry with a minor contribution from UP hunter-gatherer sources (Fig. 3). In contrast, later individuals from Northern Iberia are more similar to hunter-gatherers from eastern Europe, deriving ~30-40% of their ancestry from a source related to Ukrainian hunter-gatherers[34,35]. The earliest evidence for this gene flow is observed in a Mesolithic individual from El Mazo, Spain (NEO646) that was dated, calibrated and reservoir-corrected to c. 8,200 BP (8365-8182 cal BP, 95%) but context-dated to slightly older (8550-8330 BP, see[36]). The younger date coincides with some of the oldest Mesolithic geometric microliths in northern Iberia, appearing around 8,200 BP at this site[36]. In southern Sweden, we find higher amounts of southern European hunter-gatherer ancestry in late Mesolithic coastal individuals (NEO260 from Evensås; NEO679 from Skateholm) than in the earlier Mesolithic individuals from further inland, suggesting either geographic genetic structure in the Swedish Mesolithic population or a possible eastward expansion of hunter-gatherers from Denmark, where this ancestry prevailed (Fig. 3). An influx of southern European hunter-gatherer-related ancestry in Ukrainian individuals after the Mesolithic (Fig. 3) suggests a similar eastwards expansion in south-eastern Europe[17]. Interestingly, two herein reported ~7,300-year-old imputed genomes from the Middle Don River region in the Pontic-Caspian steppe (Golubaya Krinitsa, NEO113 & NEO212) derive ~20-30% of their ancestry from a source cluster of hunter-gatherers from the Caucasus (Caucasus_13000BP_10000BP) (Fig. 3). Additional lower coverage (non-imputed) genomes from the same site project in the same PCA space (Fig. 1D), shifted away from the European hunter-gatherer cline towards Iran and the Caucasus. Our results thus document genetic contact between populations from the Caucasus and the Steppe region as early as 7,300 years ago, providing documentation of continuous admixture prior to the advent of later nomadic

Steppe cultures, in contrast to recent hypotheses, and also further to the west than previously reported[17,37].

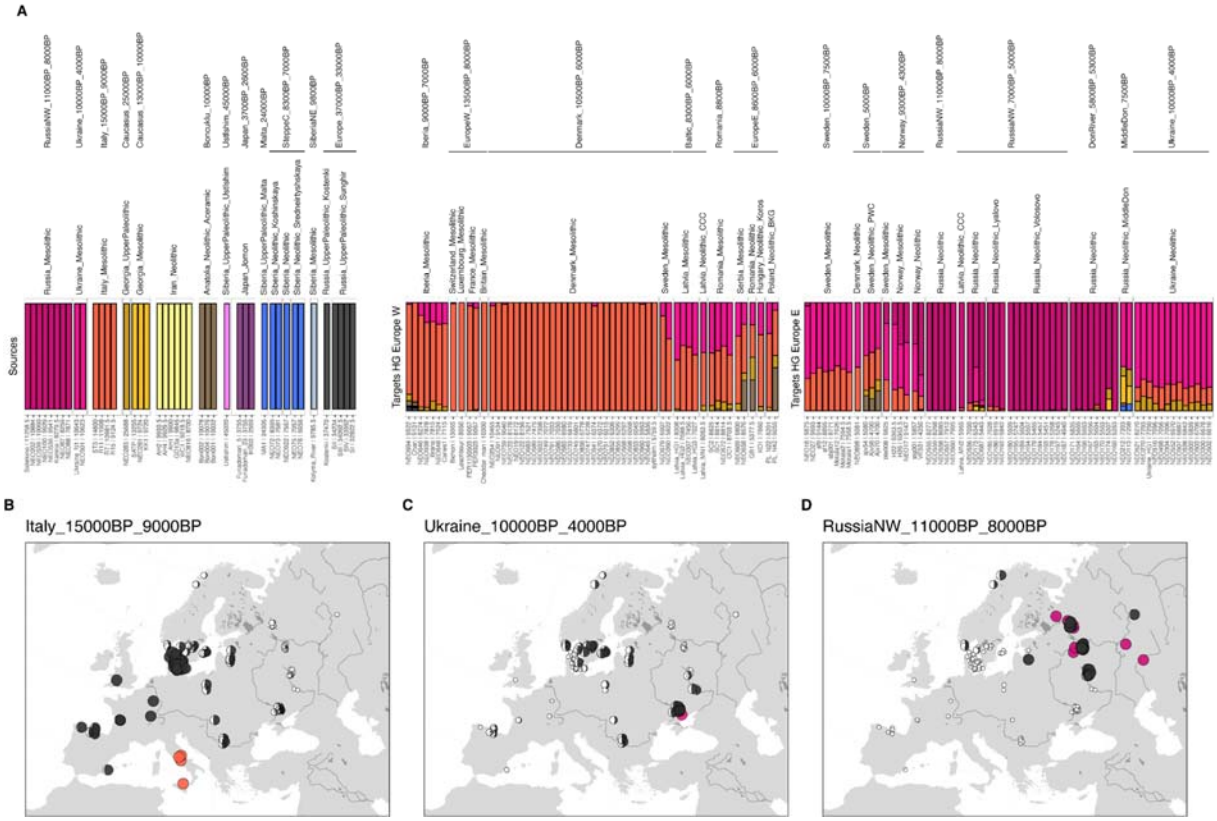


Fig 2. Genetic structure of European hunter-gatherers (A) Ancestry proportions in 113 imputed ancient genomes representing European hunter-gatherer contexts (right) estimated from supervised non-negative least squares analysis using deep Eurasian source groups (left). Individuals from target groups are grouped by genetic clusters. (B)-(D) Moon charts showing spatial distribution of ancestry proportions in European hunter-gatherers deriving from three deep Eurasian source groups; Italy_15000BP_9000BP; Ukraine_10000BP_4000BP; RussiaNW_11000BP_8000BP (source origins shown with coloured symbol). Estimated ancestry proportions are indicated by both size and amount of fill of moon symbols.

Major genetic transitions in Europe

Previous ancient genomics studies have documented multiple episodes of large-scale population turnover in Europe within the last 10,000 years[6,11,14,16,17,20,21,34,38–41]. The 317 genomes reported here fill important knowledge gaps, particularly in northern and eastern Europe, allowing us to track the dynamics of these events at both continental and regional scales.

Our analyses reveal profound differences in the spatiotemporal Neolithisation dynamics across Europe. Supervised admixture modelling ("deep" set) and spatiotemporal kriging[42] 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 hunter-gatherers to Neolithic farmers with Anatolian-related ancestry (Boncuklu_10000BP; Fig. 3; Extended Data Fig. 4, 6). The arrival of Anatolian-related ancestry in different regions spans an extensive time period of over 3,000 years, from its earliest evidence in the Balkans (Lepenski Vir) at ~8,700 BP[17] to c. 5,900 BP in Denmark. On the eastern side of this divide, no ancestry shifts can be observed during this period. In the East Baltic region (see also[43]),

Ukraine and Western Russia local hunter-gatherer ancestry prevails until ~5,000 BP without noticeable input of Neolithic Anatolian-related farmer ancestry (Fig. 3; Extended Data Fig. 4, 6). This Eastern genetic continuity is in remarkable congruence with the archaeological record showing persistence of pottery-using hunter-gatherer-fisher groups in this wide region, and delayed introduction of cultivation and husbandry by several thousand years (Supplementary Note 5).

From approximately 5,000 BP, an ancestry component appears on the eastern European plains in Early Bronze Age Steppe pastoralists associated with the Yamnaya culture and it rapidly spreads across Europe through the expansion of the Corded Ware complex (CWC) and related cultures[20,21]. We demonstrate that this "steppe" ancestry (Steppe_5000BP_4300BP) can be modelled as a mixture of ~65% ancestry related to herein reported hunter-gatherer genomes from the Middle Don River region (MiddleDon_7500BP) and ~35% ancestry related to hunter-gatherers from Caucasus (Caucasus_13000BP_10000BP) (Extended Data Fig. 4). Thus, Middle Don hunter-gatherers, who already carry ancestry related to Caucasus hunter-gatherers (Fig. 2), serve as a hitherto unknown proximal source for the majority ancestry contribution into Yamnaya 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 hunter-gatherers mentioned above, but the genetic composition at Golubaya Krinitsa is different from the remaining Ukrainian sites (Fig 2A, Extended Data Fig. 4). We find that the subsequent transition of the Late Neolithic and Early Bronze Age European gene pool happened at a faster pace than during the Neolithisation, reaching most parts of Europe within a ~1,000-year time period after first appearing in eastern Baltic region ~4,800 BP (Fig. 3). In line with previous reports we observe that beginning c. 4,200 BP, steppe-related ancestry was already dominant in samples from France and the Iberian peninsula, while it reached Britain only 400 years later[11,38,44]. Strikingly, because of the delayed Neolithisation in Southern Scandinavia these dynamics resulted in two episodes of large-scale genetic turnover in Denmark and southern Sweden within a 1,000-year period (Fig. 3).

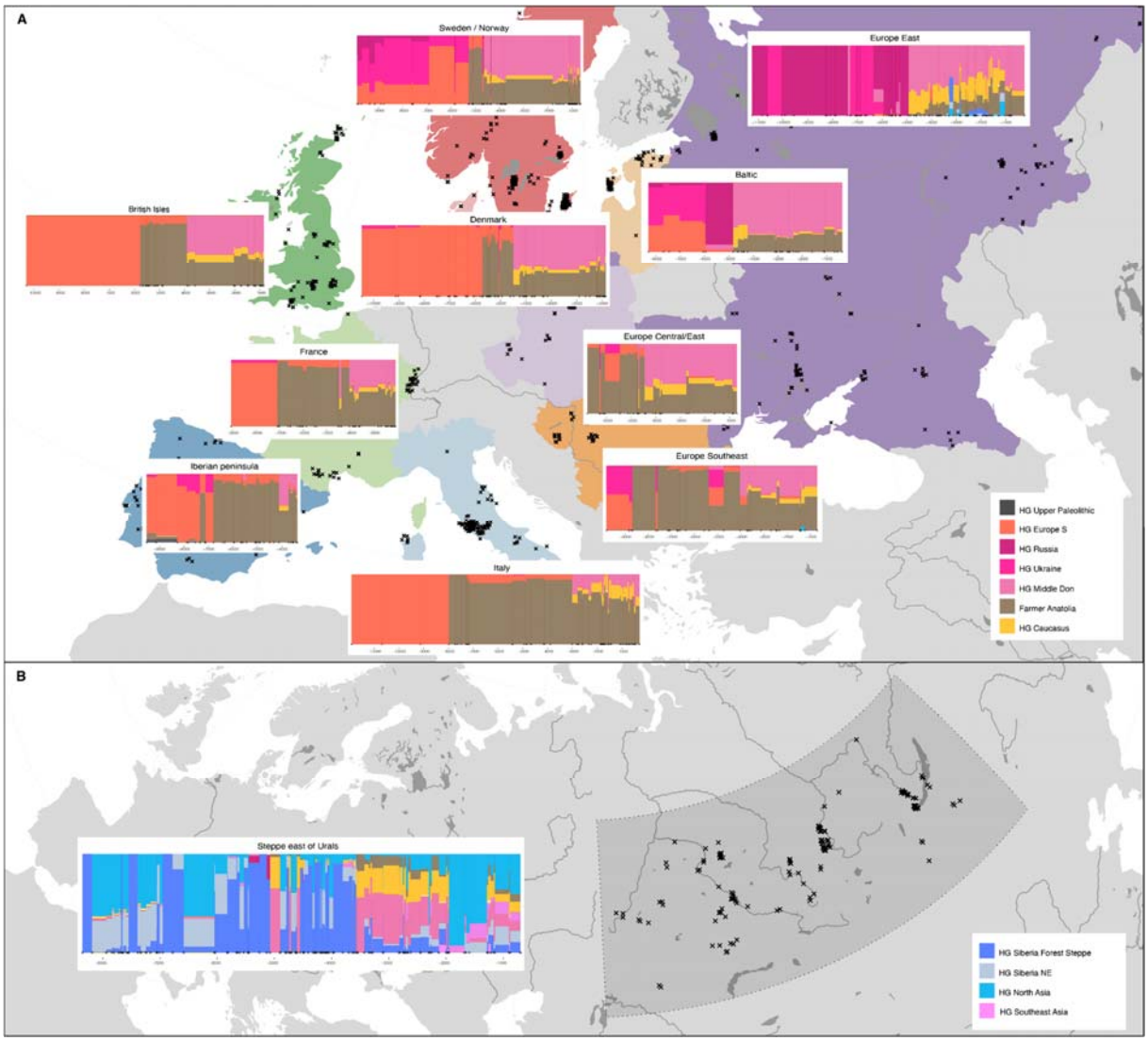


Fig. 3. Genetic transects of Eurasia. Regional timelines of genetic ancestry compositions within the past 15,000 years in western Eurasia (top) and the Eurasian Steppe belt east of the Urals (bottom). Ancestry proportions in 972 imputed ancient genomes from these regions (covering c. 12,000 BP to 500 BP), inferred using supervised admixture modelling with "deep" hunter-gatherer ancestry source groups. Geographic areas included in timelines are indicated with fill colour (west Eurasia) and grey shading (eastern Steppe region). Excavation locations of the ancient skeletons are indicated with black crosses. 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, ages. We note that the inclusion of only shotgun-sequenced samples may affect the exact timing of events in some regions from where such data are sparse.

We next investigated fine-grained ancestry dynamics underlying these transitions. We replicate previous reports[11,16,17,21,41,45,46] of widespread, but low-level admixture between Neolithic farmers and local hunter-gatherers resulting in a resurgence of HG ancestry in many regions of Europe during the middle and late Neolithic (Extended Data Fig. 7). Estimated hunter-gatherer ancestry proportions among early Neolithic people rarely exceed 10%, with notable exceptions observed in individuals from south-eastern Europe (Iron Gates), Sweden (Pitted Ware Culture) as well as herein reported early Neolithic genomes from Portugal (western Cardial), estimated to harbour 27% – 43% Iberian hunter-gatherer ancestry (Iberia_9000BP_7000BP). The latter result, suggesting extensive first-contact admixture, is in agreement with archaeological inferences derived from modelling the

spread of farming along west Mediterranean Europe[47]. Individuals associated with Neolithic farming cultures from Denmark show some of the highest overall hunter-gatherer ancestry proportions (up to ~25%), mostly derived from Western European-related hunter-gatherers (EuropeW_13500BP_8000BP) supplemented with marginal contribution from local Danish groups in some individuals (Extended Data Fig. 7D; Supplementary Note 3f). We estimated the timing of the admixture using the linkage-disequilibrium-based method DATES[48] at ~6,000 BP. Both lines of evidence thus suggest that a significant part of the hunter-gatherer admixture observed in Danish Neolithic individuals occurred already before the arrival of the incoming Neolithic people in the region (Extended Data Fig. 7), and further imply Central Europe as a key region in the resurgence of HG ancestry. Interestingly, the genomes of two ~5,000-year-old Danish male individuals (NEO33, NEO898) were entirely composed of Swedish hunter-gatherer ancestry, and formed a cluster with Pitted Ware Culture (PWC) individuals from Ajvide on the Baltic island of Gotland (Sweden)[49–51]. Of the two individuals, NEO033 also displays an outlier Sr-signature (Fig. 4), potentially suggesting a non-local origin matching his unusual ancestry. Overall, our results demonstrate direct contact across the Kattegat and Öresund during Neolithic times (Extended Data Fig. 3, 4), in line with archaeological finds from Zealand (east Denmark) showing cultural affinities to PWC on the Swedish west coast[52–55].

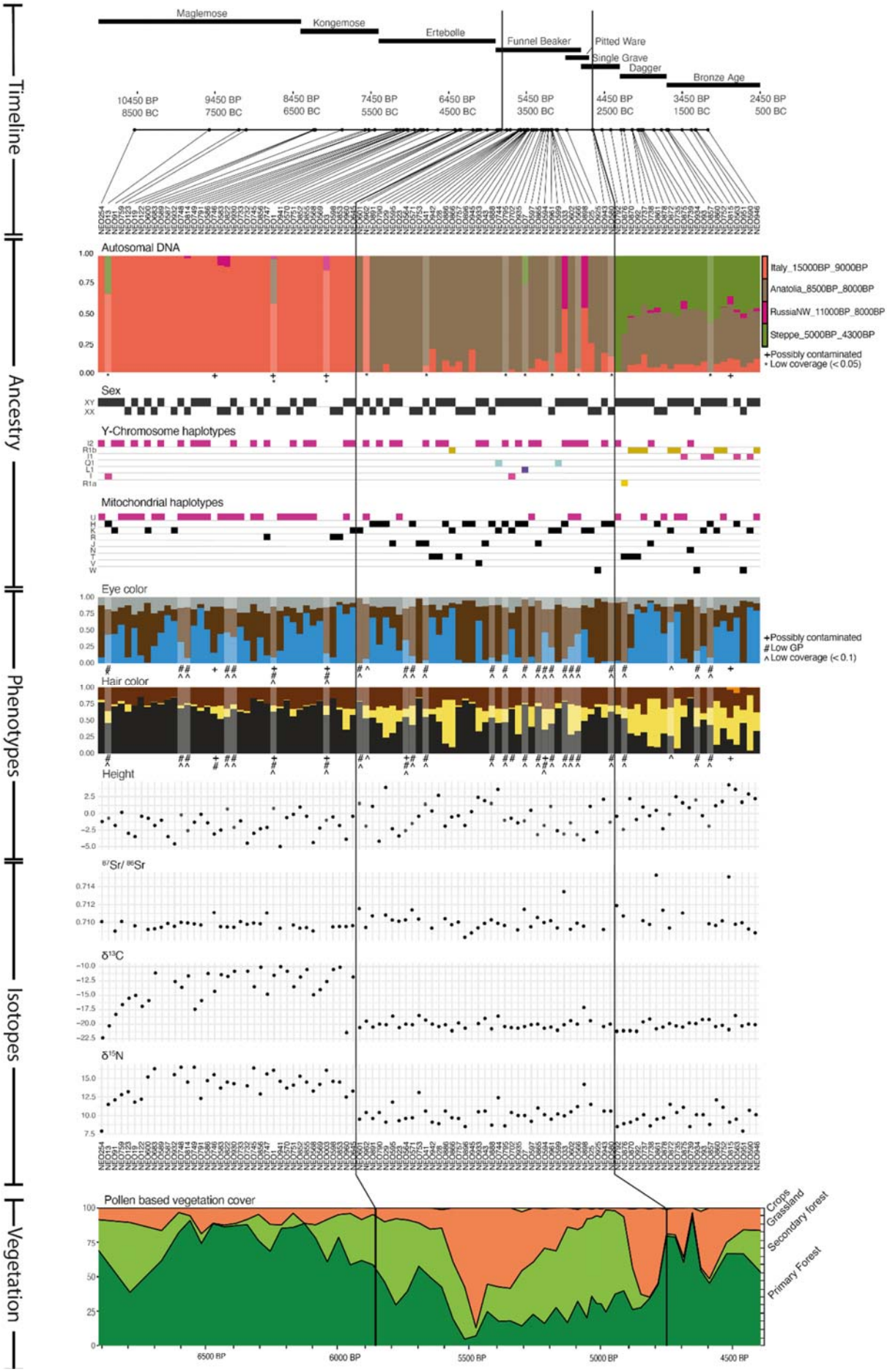
Further, we find evidence for regional stratification in early Neolithic farmer ancestries in subsequent Neolithic groups. Specifically, southern European early farmers appear to have provided major genetic ancestry to mid- and late Neolithic groups in Western Europe, while central European early farmer ancestry is mainly observed in subsequent Neolithic groups in eastern Europe and Scandinavia (Extended Data Fig. 7D-F). These results are consistent with distinct migratory routes of expanding farmer populations as previously suggested[8]. For example, similarities in material culture and flint mining activities could suggest that the first farmers in South Scandinavia originated from or had close social relations with the central European Michelsberg Culture[56].

The second continental-wide and CWC-mediated transition from Neolithic farmer ancestry to Steppe-related ancestry was found to differ markedly between geographic regions. The contribution of local Neolithic farmer ancestry to the incoming groups was high in eastern, western and southern Europe, reaching >50% on the Iberian Peninsula ("postNeol" set; Extended Data Fig. 4, 6B, C)[34]. Scandinavia, however, portrays a dramatically different picture, with a near-complete replacement of the local Neolithic farmer population inferred across all sampled individuals (Extended Data Fig. 7B, C). Following the second transition, Neolithic Anatolian-related farmer ancestry remains in Scandinavia, but the source is now different. It can be modelled as deriving almost exclusively from a genetic cluster associated with the Late Neolithic Globular Amphora Culture (GAC) (Poland_5000BP_4700BP; Extended Data Fig. 4). Strikingly, after the Steppe-related ancestry was first introduced into Europe (Steppe_5000BP_4300BP), it expanded together with GAC-related ancestry across all sampled European regions (Extended Data Fig. 7I). This suggests that the spread of steppe-related ancestry throughout Europe was predominantly mediated through groups that were already admixed with GAC-related farmer groups of the eastern European plains. This finding has major implications for understanding the emergence of the CWC. A stylistic connection from GAC ceramics to CWC ceramics has long been suggested, including the use of amphora-shaped vessels and the development of cord decoration patterns[57]. Moreover, shortly prior to 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[58], indicating close interaction between the groups. Previous ancient genomic data from a few individuals suggested that this was limited to cultural influences and not population admixture[59]. However, in the light of our new genetic evidence it appears that this zone, and

possibly other similar zones of contact between GAC and Yamnaya (or other closely-related steppe/forest-steppe groups) 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[cf. 60]. This resulted in regionally diverse situations of interaction and admixture[61,62] but a significant part of the CWC dispersal happened through corridors of cultural and demic transmission which had been established by the GAC during the preceding period[63,64].

Fine-scale structure and multiproxy analysis of Danish transect

We present a detailed and continuous sequence of multiproxy data from Denmark, from the Early Mesolithic Maglemose, via the Kongemose and Late Mesolithic Ertebølle epochs, the Early and Middle Neolithic Funnel Beaker Culture and the Single Grave Culture, to Late Neolithic and Bronze Age individuals (Fig. 4). To integrate multiproxy data from as many skeletons as possible we made use of non-imputed data for the admixture analyses (Supplementary Note S3d) which were not restricted to the >0.1X coverage cut-off used elsewhere. This provided genetic profiles from 100 Danish individuals (Fig 4), spanning c. 7,300 years from the earliest known skeleton in Denmark (the Mesolithic "Koelbjerg Man" (NEO254, 10,648-10,282 cal. BP, 95% probability interval) and formerly known as the "Koelbjerg Woman"[65]), to a Bronze Age skeleton from Hove Å (NEO946) dated to 3322-2967 cal. BP (95%). Two temporal shifts in genomic admixture proportions confirm the major population genetic turnovers (Fig. 4) that was inferred from imputed data (Fig. 3). The multiproxy evidence, however, unveils the dramatic concomitant changes in all investigated phenotypic, environmental and dietary parameters (Fig. 4).


Timeline
Ancestry
Phenotypes
Isotopes
Vegetation
Maglemose
Kongemose
Ertebølle
Funnel Beaker
Pitted Ware
Single Grave
Dagger
Bronze Age
10450 BP 8500 BC
9450 BP 7500 BC
8450 BP 6500 BC
7450 BP 5500 BC
6450 BP 4500 BC
5450 BP 3500 BC
4450 BP 2500 BC
3450 BP 1500 BC
2450 BP 500 BC
Autosomal DNA
Italy_15000BP_9000BP
Anatolia_8500BP_8000BP
RussiaNW_11000BP_8000BP
Steppe_5000BP_4300BP
+Possibly contaminated
• Low coverage (< 0.05)
Sex
XY
XX
Y-Chromosome haplotypes
Mitochondrial haplotypes
Eye color
+Possibly contaminated
Low GP
^ Low coverage (< 0.1)
Hair color
Height
87Sr/86Sr
δ13C
δ15N
Pollen based vegetation cover
Crops
Grassland
Secondary forest
Primary Forest
6500 BP
6000 BP
5500 BP
5000 BP
4500 BP

542

Fig 4. Environmental, dietary, phenotypic and ancestry shifts in Denmark through time. Two dramatic population turnovers are evident from chronologically-sorted multiproxy data representing 100 Danish Stone Age and early Bronze Age skeletons sequenced in this study. The figure shows concomitant changes in several investigated parameters including (from the top) admixture proportions from non-imputed autosomal genome-wide data, Y-chromosomal and mitochondrial haplogroups, genetic phenotype predictions (based on imputed data) as well as $^{87}Sr/^{86}Sr$ and $\delta^{13}C$ and $\delta^{15}N$ isotope data as possible proxies for mobility and diet, respectively. Predicted height values represent differences (in cm) from the average height of the present-day Danish population, based on genotypes at 310 height-associated loci (Supplementary Note 4f). Probabilities for the indicated natural eye and hair colours are based on genotypes at 18 pigmentation-associated loci (Supplementary Note 4f) with grey denoting probability of intermediate eye colour (including grey, green and hazel). Lower panel shows changes in vegetation as predicted from pollen analyses at Lake Højby in Zealand (Supplementary Note 12). Black vertical lines mark the first presence of Anatolian farmer ancestry and Steppe-related ancestry, respectively.

During the Danish Mesolithic, individuals from the Maglemose, Kongemose and Ertebølle cultures displayed a remarkable genetic homogeneity across a 5,000 year transect deriving their ancestry almost exclusively from a southern European source (source: Italy_15000BP_9000BP) that later predominates in western Europe (Fig. 2). These cultural transitions occurred in genetic continuity, apparent in both autosomal and uniparental markers, which rules out demic diffusion and supports the long-held assumption of a continuum of culture and population[e.g. 66–68]. Genetic predictions indicate blue eye pigmentation with high probability in several individuals throughout the duration of the Mesolithic (Supplementary Note 4f), consistent with previous findings [11,20,45]. In contrast, none of the analysed Mesolithic individuals displayed high probability of light hair pigmentation. Height predictions for Mesolithic individuals generally suggest slightly lower or perhaps less variable genetic values than in the succeeding Neolithic period. However, we caution that the relatively large genetic distance to modern individuals included in the GWAS panel make these scores poorly applicable to Mesolithic individuals (Supplementary Note 4c) and are dependent on the choice of GWAS filters used. Unfortunately, only a fraction of the 100 Danish skeletons included were suitable for stature estimation by actual measurement, why these values are not reported.

Stable isotope $\delta^{13}C$ values in collagen inform on the proportion of marine versus terrestrial protein, while $\delta^{15}N$ values reflect the trophic level of protein sources[69,70]. Both the Koelbjerg Man and the second earliest human known from Denmark, (Tømmerupgårds Mose – not part of the present study; see[71]) showed more depleted dietary isotopic values, representing a lifestyle of inland hunter-fisher-gatherers of the early Mesolithic forest. A second group consisted of coastal fisher-hunter-gatherers dating to the late half of the Maglemose epoch onwards (Supplementary Figs. S10.1 and S10.2). During this period global sea-level rise gradually changed the landscape of present-day Denmark from an interior part of the European continent to an archipelago, where all human groups had ample access to coastal resources within their annual territories. Increased $\delta^{13}C$ and $\delta^{15}N$ values imply that from the late Maglemose marine foods gradually increased in importance, to form the major supply of proteins in the final Ertebølle period[71,cf. 72]. Interestingly, rather stable $^{87}Sr/^{86}Sr$ isotope ratios throughout the Mesolithic indicate limited mobility, in agreement with the evidence for genetic continuity reported here and modelled in previous work[73,74 Fig. 3], and/or dietary sources from homogeneous environments.

The arrival of Neolithic farmer-related ancestry at c. 5,900 BP in Denmark resulted in a population replacement with very limited genetic contribution from the local hunter-gatherers. The shift is abrupt and brings changes in all the measured parameters. This is a clear case of demic diffusion, which settles a long-standing debate concerning the neolithisation process in Denmark[15,56,75,76], at least at a broader population level. The continuing use of coastal kitchen middens well into the

Neolithic[77,78] remains, however, an enigma, although this may represent sites where local remnants of Mesolithic groups survived in partly acculturated form, or it could be middens taken over by the newcomers. Concomitant shifts in both autosomal and uniparental genetic markers show that the migration by incoming farmers was not clearly sex-biased but more likely involved nuclear family units. Diet shifted abruptly to terrestrial sources evidenced by $\delta^{13}C$ values around -20 ‰ and $\delta^{15}N$ values around 10 ‰ in line with archaeological evidence that domesticated crops and animals were now providing the main supply of proteins (Supplementary Note 6). Isotope values remained stable at these levels throughout the following periods, although with somewhat greater variation after c. 4,500 BP. However, five Neolithic and Early Bronze Age individuals have $\delta^{13}C$ and $\delta^{15}N$ values indicating intake of high trophic marine food. This is most pronouncedly seen for NEO898 (Svinninge Vejle) who was one of the two aforementioned Danish Neolithic individuals displaying typical Swedish PWC hunter-gatherer ancestry. A higher variability in $^{87}Sr/^{86}Sr$ values can be seen with the start of the Neolithic and this continues in the later periods, which suggests that the Neolithic farmers in Denmark consumed food from more diverse landscapes and/or they were more mobile than the preceding hunter-gatherers (Supplementary Note 11). The Neolithic transition also marks a considerable rise in frequency of major effect alleles associated with light hair pigmentation[79], whereas polygenic score predictions for height are generally low throughout the first millennium of the Neolithic (Funnel Beaker epoch), echoing previous findings based on a smaller set of individuals[45,80].

We do not know how the Mesolithic Ertebølle population disappeared. Some may have been isolated in small geographical pockets of brief existence and/or adapted to a Neolithic lifestyle but without contributing much genetic ancestry to subsequent generations. The most recent individual in our Danish dataset with Mesolithic WHG ancestry is "Dragsholm Man" (NEO962), dated to 5,947-5,664 cal. BP (95%) and archaeologically assigned to the Neolithic Funnel Beaker farming culture based on his grave goods[81,82]. Our data confirms a typical Neolithic diet matching the cultural affinity but contrasting his WHG ancestry. Thus, Dragsholm Man represents a local person of Mesolithic ancestry who lived in the short Mesolithic-Neolithic transition period and adopted a Neolithic culture and diet. A similar case of very late Mesolithic WHG ancestry in Denmark was observed when analysing human DNA obtained from a piece of chewed birch pitch dated to 5,858–5,661 cal. BP (95%)[83].

The earliest example of Anatolian Neolithic ancestry in our Danish dataset is observed in a bog skeleton of a female from Viksø Mose (NEO601) dated to 5,896-5,718 cal. BP (95%) (and hence potentially contemporaneous with Dragsholm Man) whereas the most recent Danish individual showing Anatolian ancestry without any Steppe-related ancestry is NEO943 from Stenderup Hage, dated to 4,818-4,415 cal. BP (95%). Using Bayesian modelling we estimate the duration between the first appearance of Anatolian ancestry to the first appearance of Steppe-related ancestry in Denmark to be between 876 and 1100 years (95% probability interval, Supplementary Note 9) indicating that the typical Neolithic ancestry was dominant for less than 50 generations in Denmark. From this point onwards the steppe-ancestry was introduced, signalling the rise of the late Neolithic Corded Ware derived cultures in Denmark (i.e. Single Grave Culture), followed by the later Neolithic Dagger epoch and Bronze Age cultures. While this introduced a major new component in the Danish gene pool, it was not accompanied by apparent shifts in diet. Our complex trait predictions indicate an increase in "genetic height" occurring concomitant with the introduction of Steppe-related ancestry, which is consistent with Steppe individuals (e.g., Yamnaya) being genetically taller on average[45] and with previous results from other European regions[80,84].

These major population turnovers were accompanied by significant environmental changes, as apparent from high-resolution pollen diagrams from Lake Højby in Northwest Zealand

reconstructed using the Landscape Reconstruction Algorithm (LRA[85] (Supplementary Note 8). While the LRA has previously been applied at low temporal resolution regional scale [e.g. 86,87], and at local scale to Iron Age and later pollen diagrams [e.g. 88,89], this is the first time this quantitative method is applied at local scale to a pollen record spanning the Mesolithic and Neolithic periods in Denmark. Comparison with existing pollen records show that the land cover changes demonstrated here reflect the general vegetation development in eastern Denmark, while the vegetation on the sandier soils of western Jutland maintains a more open character throughout the sequence (Supplementary Note 12). We find that during the Mesolithic (i.e. before c. 6,000 BP) the vegetation was dominated by primary forest trees (*Tilia, Ulmus, Quercus, Fraxinus, Alnus* etc.). The forest composition changed towards more secondary, early successional trees (*Betula and then Corylus*) in the earliest Neolithic, but only a minor change in the relationship between forest and open land is recorded. From c. 5,650 BP deforestation intensified, resulting in a very open grassland-dominated landscape. This open phase was short-lived, and secondary forest expanded from 5,500 to 5,000 BP, until another episode of forest clearance gave rise to an open landscape during the last part of the Funnel Beaker epoch. We thus conclude that the agriculture practice was characterised by repeated clearing of the forest with fire, followed by regrowth. This strategy changed with the onset of the Single Grave Culture, when the forest increased again, but this time dominated by primary forest trees, especially *Tilia* and *Ulmus*. This reflects the development of a more permanent division of the landscape into open grazing areas and forests. In contrast, in western Jutland this phase was characterised by large-scale opening of the landscape, presumably as a result of human impact aimed at creating pastureland[90].

Finally, we investigated the fine-scale genetic structure in southern Scandinavia after the introduction of Steppe-related ancestry using a temporal transect of 38 Late Neolithic and Early Bronze Age Danish and southern Swedish individuals. Although the overall population genomic signatures suggest genetic stability, patterns of pairwise IBD-sharing and Y-chromosome haplogroup distributions indicate at least three distinct ancestry phases during a ~1,000-year time span: i) An early stage between ~4,600 BP and 4,300 BP, where Scandinavians cluster with early CWC individuals from Eastern Europe, rich in Steppe-related ancestry and males with an R1a Y-chromosomal haplotype (Extended Data Fig. 8A, B); ii) an intermediate stage until c. 3,800 BP, where they cluster with central and western Europeans dominated by males with distinct sub-lineages of R1b-L51 (Extended Data Fig. 8C, D; Supplementary Note 3b) and includes Danish individuals from Borreby (NEO735, 737) and Madesø (NEO752) with distinct cranial features (Supplementary Note 6); and iii) a final stage from c. 3,800 BP onwards, where a distinct cluster of Scandinavian individuals dominated by males with I1 Y-haplogroups appears (Extended Data Fig. 8E). Using individuals associated with this cluster (Scandinavia_4000BP_3000BP) as sources in supervised ancestry modelling (see "postBA", Extended Data Fig. 4), we find that it forms the predominant source for later Iron- and Viking Age Scandinavians, as well as ancient European groups outside Scandinavia who have a documented Scandinavian or Germanic association (e.g., Anglo-Saxons, Goths; Extended Data Fig. 4). Y-chromosome haplogroup I1 is one of the dominant haplogroups in present-day Scandinavians,s, and we document its earliest occurrence in a ~4,000-year-old individual from Falköping in southern Sweden (NEO220). The rapid expansion of this haplogroup and associated genome-wide ancestry in the early Nordic Bronze Age indicates a considerable reproductive advantage of individuals associated with this cluster over the preceding groups across large parts of Scandinavia.

Hunter-gatherer resilience east of the Urals

In contrast to the significant number of ancient hunter-gatherer genomes from western Eurasia studied to date, genomic data from hunter-gatherers east of the Urals remain sparse. These regions are characterised by an early introduction of pottery from areas further east and are inhabited by complex hunter-gatherer-fisher societies with permanent and sometimes fortified settlements (Supplementary Note 5; [91]).

Here, we substantially expand the knowledge on ancient Stone Age populations of this region by reporting new genomic data from 38 individuals, 28 of which date to pottery-associated hunter-gatherer contexts [e.g. 92] between 8,300-5,000 BP (Supplementary Table II).The majority of these genomes form a previously only sparsely sampled [48,93] "Neolithic Steppe" cline spanning the Siberian Forest Steppe zones of the Irtysh, Ishim, Ob, and Yenisei River basins to the Lake Baikal region (Fig. 1C; Extended Data Fig. 2A, 3E). Supervised admixture modelling (using the "deep" set of ancestry sources) revealed contributions from three major sources in these hunter gatherers from east of Urals: early West Siberian hunter-gatherer ancestry (SteppeC_8300BP_7000BP) dominated in the western Forest Steppe; Northeast Asian hunter-gatherer ancestry (Amur_7500BP) was highest at Lake Baikal; and Paleosiberian ancestry (SiberiaNE_9800BP) was observed in a cline of decreasing proportions from northern Lake Baikal westwards across the Forest Steppe (Extended Data Fig. 4, 9). [93]

We used these Neolithic hunter-gatherer clusters ("postNeol" ancestry source set, Extended Data Fig. 4) as putative source groups in more proximal admixture modelling to investigate the spatiotemporal dynamics of ancestry compositions across the Steppe and Lake Baikal after the Neolithic period. We replicate previously reported evidence for a genetic shift towards higher Forest Steppe hunter-gatherer ancestry (SteppeCE_7000BP_3600BP) in late Neolithic and early Bronze Age individuals (LNBA) at Lake Baikal [93,94]. However, ancestry related to this cluster is already observed at ~7,000 BP in herein-reported Neolithic hunter-gatherer individuals both at Lake Baikal (NEO199, NEO200), and along the Angara river to the north (NEO843). Both male individuals at Lake Baikal belonged to Y-chromosome haplogroup Q1, characteristic of the later LNBA groups in the same region. (Extended Data Fig. 3, 6A). Together with an estimated date of admixture of ~6,000 BP for the LNBA groups, these results suggest gene flow between hunter-gatherers of Lake Baikal and the south Siberian forest steppe regions already during the early Neolithic. This is consistent with archaeological interpretations of contact. In this region, bifacially flaked tools first appeared near Baikal [95] from where the technique spread far to the west. We find its reminiscences in Late Neolithic archaeological complexes (Shiderty 3, Borly, Sharbakty 1, Ust-Narym, etc.) in Northern and Eastern Kazakhstan, around 6,500-6,000 BP [96,97]. Our herein-reported genomes 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 hunter-gatherer ancestry in the Okunevo[93,94], suggesting that they instead originate from a three-way mixture of two different genetic clusters of Siberian forest steppe hunter-gatherers and Steppe-related ancestry (Extended Data Fig. 4D). We date the admixture with Steppe-related ancestry to ~4,600 BP, consistent with gene flow from peoples of the Afanasievo culture that existed near Altai and Minusinsk Basin during the early eastwards' expansion of Yamnaya-related groups [20,94].

From around 3,700 BP, individuals across the Steppe and Lake Baikal regions display markedly different ancestry profiles (Fig. 3; Extended Data Fig. 4D, 9). We document a sharp increase in non-local ancestries, with only limited ancestry contributions from local hunter-gatherers. The early stages of this transition are characterised by influx of Yamnaya-related ancestry, which decays over

time from its peak of ~70% in the earliest individuals. Similar to the dynamics in western Eurasia, Yamnaya-related ancestry is here correlated with late Neolithic GAC-related farmer ancestry (Poland_5000BP_4700BP; Extended Data Fig. 9G), recapitulating the previously documented eastward expansion of admixed Western Steppe pastoralists from the Sintashta and Andronovo complexes during the Bronze Age[20,48,98]. However, GAC-related ancestry is notably absent in individuals of the Okunevo culture, providing further support for two distinct eastward migrations of Western Steppe pastoralists during the early (Yamnaya) and later (Sintashta, Andronovo) Bronze Age. The later stages of the transition are characterised by increasing Central Asian (Turkmenistan_7000 BP_5000BP) and Northeast Asian-related (Amur_7500BP) ancestry components (Extended Data Fig. 9G). Together, these results show that deeply structured hunter-gatherer 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, including a large-scale introduction of domesticated horse lineages concomitant with new equestrian equipment and spoke-wheeled chariotry [20,48,98,99].

Genetic legacy of Stone Age Europeans

To investigate the distribution of Stone Age and Early Bronze Age ancestry components in modern populations, we used ChromoPainter [100] to "paint" the chromosomes of individuals in the UK Biobank (https://www.ukbiobank.ac.uk) using a panel of 10 ancient donor populations (Supplementary Note 3h). Painting was done following the pipeline of Margaryan et al. [101] based on GLOBETROTTER [102], and admixture proportions were estimated using Non-Negative Least squares. Haplotypes in the modern genomes are assigned to the genetically closest ancient population as measured by meiosis events, which favours more recent matches in time. Therefore, ancestry proportions assigned to the oldest groups (e.g. WHG) should be interpreted as an excess of this ancestry, which cannot be explained by simply travelling through more recent ancient populations up to present times.

First, we selected non-British individuals from the UK Biobank if their country of birth was European, African, or Asian. Because many of these individuals are admixed or British, we set up a pipeline (Supplementary Note 3g) to select individuals of a typical ancestral background for each country. This resulted in 24,511 individuals from 126 countries, who were then chromosome painted to assess the average admixture proportions for each ancestry per country.

The various hunter-gatherer ancestries are not homogeneously distributed amongst modern populations (Fig. 5). WHG-related ancestry is highest in present-day individuals from the Baltic States, Belarus, Poland, and Russia; EHG-related ancestry is highest in Mongolia, Finland, Estonia and Central Asia; and CHG-related ancestry is maximised in countries east of the Caucasus, in Pakistan, India, Afghanistan and Iran, in accordance with previous results [103]. The CHG-related ancestry likely reflects both Caucasus hunter-gatherer and Iranian Neolithic signals, explaining the relatively high levels in south Asia [104]. Consistent with expectations [105,106], Neolithic Anatolian-related farmer ancestry is concentrated around the Mediterranean basin, with high levels in southern Europe, the Near East, and North Africa, including the Horn of Africa, but is less frequent in Northern Europe. This is in direct contrast to the Steppe-related ancestry, which is found in high levels in northern Europe, peaking in Ireland, Iceland, Norway, and Sweden, but decreases further south. There is also evidence for its spread into southern Asia. Overall, these results refine global patterns of spatial distributions of ancient ancestries amongst modern populations.

788 The availability of a large number of modern genomes (n=408,884) from self-identified "white"
789 British individuals who share similar PCA backgrounds [107] allowed us to further examine the
790 distribution of ancient ancestries at high resolution in Britain (Supplementary Note 3h). Although
791 regional ancestry distributions differ by only a few percent, we find clear evidence of geographical
792 heterogeneity across the United Kingdom as visualised by assigning individuals to their birth
793 county and averaging ancestry proportions per county (Fig. 5, inset boxes). The proportion of
794 Neolithic farmer ancestry is highest in southern and eastern England today and lower in Scotland,
795 Wales, and Cornwall. Steppe-related ancestry is inversely distributed, peaking in the Outer
796 Hebrides and Ireland, a pattern only previously described for Scotland [108]. This regional pattern was
797 already evident in the Pre-Roman Iron Age and persists to the present day even though immigrating
798 Anglo-Saxons had relatively less Neolithic farmer ancestry than the Iron-Age population of
799 southwest Briton (Extended Data Fig. 4). Although this Neolithic farmer/steppe-related dichotomy
800 mirrors the modern 'Anglo-Saxon'/'Celtic' ethnic divide, its origins are older, resulting from
801 continuous migration from a continental population relatively enhanced in Neolithic farmer
802 ancestry, starting as early as the Late Bronze Age [109]. By measuring haplotypes from these
803 ancestries in modern individuals, we are able to show that these patterns differentiate Wales and
804 Cornwall as well as Scotland from England. We also found higher levels of WHG-related ancestry
805 in central and Northern England. These results demonstrate clear ancestry differences within an
806 'ethnic group' (white British) traditionally considered relatively homogenous, which highlights the
807 need to account for subtle population structure when using resources such as the UK Biobank
808 genomes.
809

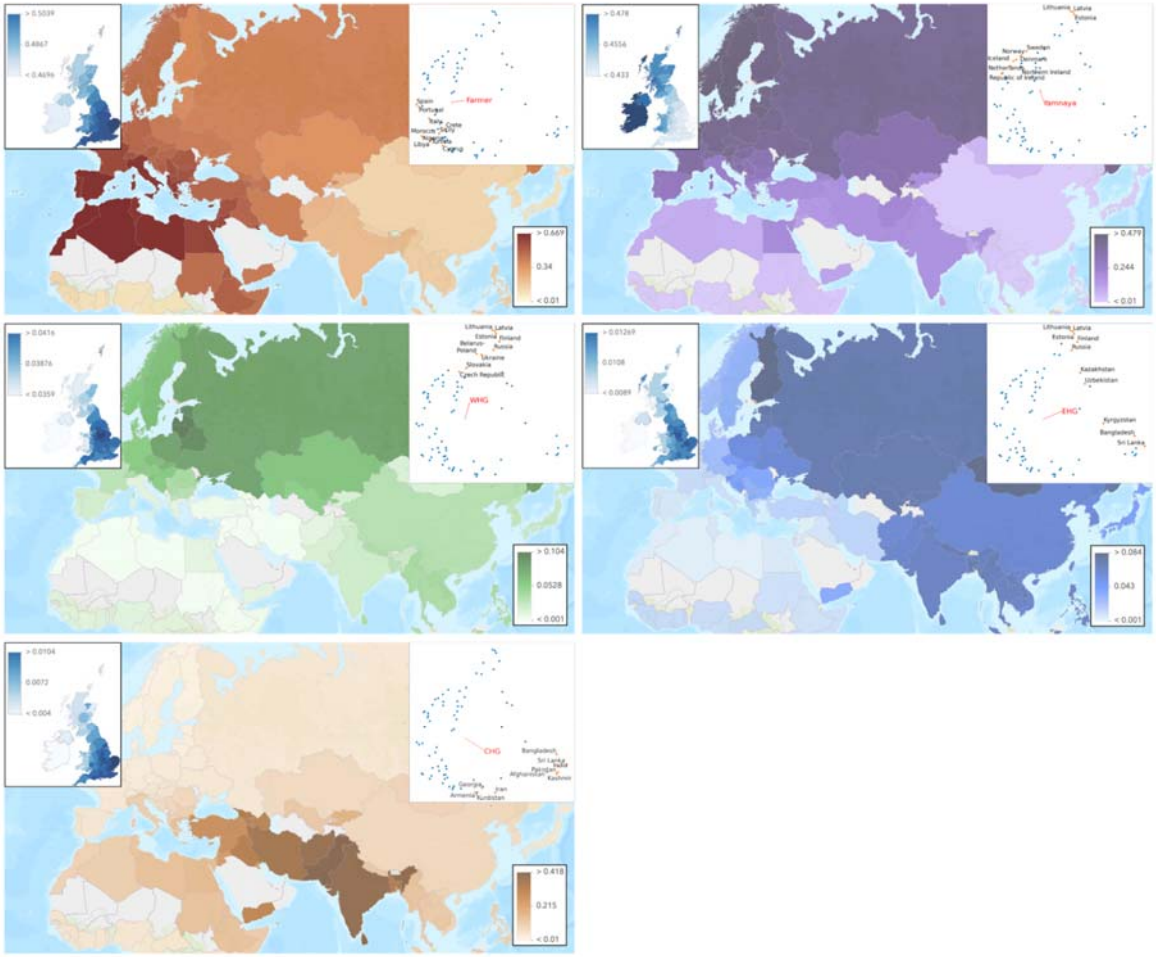


Fig 5. The genetic legacy of Stone Age ancestry in modern populations.
From top left clockwise: Neolithic Farmer, Yamnaya, Caucasus hunter-gatherer, Eastern hunter-gatherer, Western hunter-gatherer. Panels show average admixture proportion in modern individuals per country estimated using NNLS (large maps), average per county within the UK (top left insert), and PCA (PC2 vs PC1) of admixture proportions, with the top 10 highest countries by admixture fraction labelled and PCA loadings for that ancestry.

Sociocultural insights

We used patterns of pairwise IBD sharing between individuals and runs of homozygosity (ROH) within individuals (measured as the fraction of the genome within a run of homozygosity f(ROH)) to examine our data for temporal shifts in relatedness within genetic clusters. Both measures show clear trends of a reduction of within-cluster relatedness over time, in both western and eastern Eurasia (Fig. 6). This pattern is consistent with a scenario of increasing effective population sizes during this period [110]. 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 changes substantially during the Neolithic transition (~9,000 to ~6,000 BP), where clusters of Neolithic farmer-associated individuals show overall reduced IBD sharing and f(ROH) compared to clusters of HG-associated individuals (Fig. 6A,C). In the east, genetic relatedness remains high until ~4,000 BP, consistent with a much longer persistence of smaller localised hunter-gatherer groups (Fig. 6B,D).

Next, we examined the data for evidence of recent parental relatedness, by identifying individuals harbouring a large fraction of their genomes (> 50cM) in long (>20cM) ROH segments [111]. We only detect 39 such individuals out of a total sample of 1,540 imputed ancient genomes (Fig. 6E), in line with recent results indicating that close kin mating was not common in human prehistory [41,103,111,112]. With the exception of eight ancient American individuals from the San Nicolas Islands in California [113], no obviously discernible spatiotemporal or cultural clustering was observed among the individuals with recent parental relatedness. Interestingly, an ~1,700-year-old Sarmatian individual from Temyaysovo (tem003) [114] was found homozygous for almost the entirety of chromosome 2, but without evidence of ROHs elsewhere in the genome, suggesting an ancient case of uniparental disomy. Among several noteworthy familial relationships (see Supplementary Fig. S3c.2), we report a Mesolithic father/son burial at Ertebølle (NEO568/NEO569), as well as a Mesolithic mother/daughter burial at Dragsholm (NEO732/NEO733).

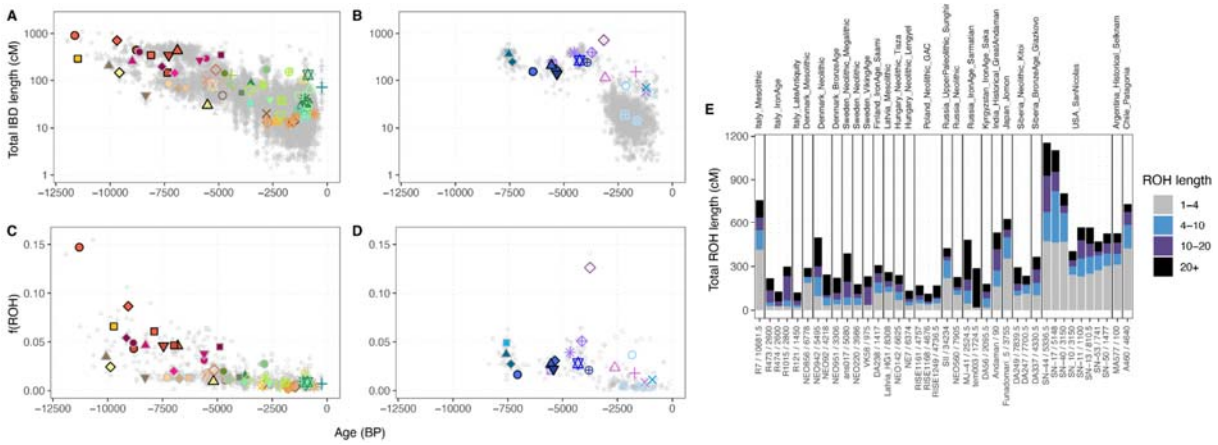


Fig 6. Patterns of co-ancestry. (A)-(D) Panels show within-cluster genetic relatedness over time, measured either as the total length of genomic segments shared IBD between individuals (A, B) or the proportion of individual genomes found in a run of homozygosity f(ROH) (C,D). Results for both measures are shown separately for individuals from western (A, C) or eastern Eurasia (B, D). Small grey dots indicate estimates for individual pairs (A, B) or individuals (C, D), with larger coloured symbols indicating median values within genetic clusters. (E) Distribution of ROH lengths for 39 individuals with evidence for recent parental relatedness (>50 cM total in ROHs > 20 cM).

Pathogenic structural variants in ancient vs. modern-day humans

Rare, recurrent copy-number variants (CNVs) are known to cause neurodevelopmental disorders and are associated with a range of psychiatric and physical traits with variable expressivity and incomplete penetrance[115,116]. To understand the prevalence of pathogenic structural variants over time we examined 50 genomic regions susceptible to recurrent CNV, known to be the most prevalent drivers of human developmental pathologies[117]. The analysis included 1442 ancient imputed genomes passing quality control for CNV analysis (Supplementary Note 4i) and 1093 modern human genomes for comparison [118,119]. We identified CNVs in ancient individuals at ten loci using a read-depth based approach and digital Comparative Genomic Hybridization [120] (Supplementary Table S4i.1; Supplementary Figs. S4i.1-S41.20). Although most of the observed CNVs (including duplications at 15q11.2 and *CHRNA7*, and CNVs spanning parts of the TAR locus and 22q11.2 distal) have not been unambiguously associated with disease in large studies, the identified CNVs include deletions and duplications that have been associated with developmental delay, dysmorphic features, and neuropsychiatric abnormalities such as autism (most notably at 1q21.1, 3q29, 16p12.1 and the DiGeorge/VCFS locus, but also deletions at 15q11.2 and

duplications at 16p13.11). The individual harbouring the 16p13.1 deletion, RISE586 [20], a 4,000 BP woman aged 20-30 from the Únětice culture (modern day Czech Republic), had almost complete skeletal remains, which allowed us to test for the presence of various skeletal abnormalities associated with the 16p13.11 microdeletion [121]. RISE586 exhibited a hypoplastic tooth, spondylolysis of the L5 vertebrae, incomplete coalescence of the S1 sacral bone, among other minor skeletal phenotypes. The skeletal phenotypes observed in this individual are relatively common (~10%) in European populations and are not specific to 16p13.1 thus do not indicate strong penetrance of this mutation in RISE586 [122–125]. However, these results do highlight our ability to link putatively pathogenic genotypes to phenotypes in ancient individuals. Overall, the carrier frequency in the ancient individuals is similar to that reported in the UK Biobank genomes (1.25% vs 1.6% at 15q11.2 and *CHRNA7* combined, and 0.8% vs 1.1% across the remaining loci combined) [126]. These results suggest that large, recurrent CNVs that can lead to several pathologies were present at similar frequencies in the ancient and modern populations included in this study.

Ancestry-stratified patterns of natural selection in the last 13,000 years

The Neolithic transition led to a fundamental change in lifestyle, diet and exposure to pathogens that imposed drastically new selection pressures on human populations. To detect genetic candidate targets of selection, we used a set of 1,015 imputed ancient genomes from West Eurasia that were fitted to a four-way admixture model of demographic history in this region (Supplementary Note 3i) and identified phenotype-associated variants with evidence for directional selection over the last 13,000 years, with a special focus on the Neolithic transition (Supplementary Note 4a). We adapted CLUES [127] to model time-series data (Supplementary Note 4a) and used it to infer allele frequency trajectories and selection coefficients for 33,323 quality-controlled phenotype-associated variants ascertained from the GWAS Catalogue [128]. An equal number of putatively neutral, frequency-paired variants were used as a control set. To control for possible confounders, we built a causal model to distinguish direct effects of age on allele frequency from indirect effects mediated by read depth, read length, and/or error rates (Supplementary Note 4b), and developed a mapping bias test used to evaluate systematic differences between data from ancient and present-day populations (Supplementary Note 4a). Because admixture between groups with differing allele frequencies can confound interpretation of allele frequency changes through time, we also applied a novel chromosome painting technique, based on inference of a sample's nearest neighbours in the marginal trees of a tree sequence (Supplementary Note 3i). This allowed us to accurately assign ancestral path labels to haplotypes found in both ancient and present-day individuals. By conditioning on these haplotype path labels, we could infer selection trajectories while controlling for changes in admixture proportions through time (Supplementary Note 4a).

Our analysis identified no genome-wide significant ($p < 5e-8$) selective sweeps when using genomes from present-day individuals alone (1000 Genomes Project populations GBR, FIN and TSI), although trait-associated variants were enriched for signatures of selection compared to the control group ($p < 2.2e-16$, Wilcoxon signed-rank test). In contrast, when using imputed aDNA genotype probabilities, we identified 11 genome-wide significant selective sweeps in the GWAS variants, and none in the control group, consistent with selection acting on trait-associated variants (Supplementary Note 4a, Supplementary Figs. S4a.4 to S4a.14). However, when conditioned on one of our four ancestral histories—genomic regions arriving in present day genomes through Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) or Anatolian farmers (ANA)—we identified 21 genome-wide significant selection peaks (including the 11 from the pan-ancestry analysis) (Fig. 7). This suggests that admixture between ancestral populations has masked evidence of selection at many trait associated loci in modern populations.

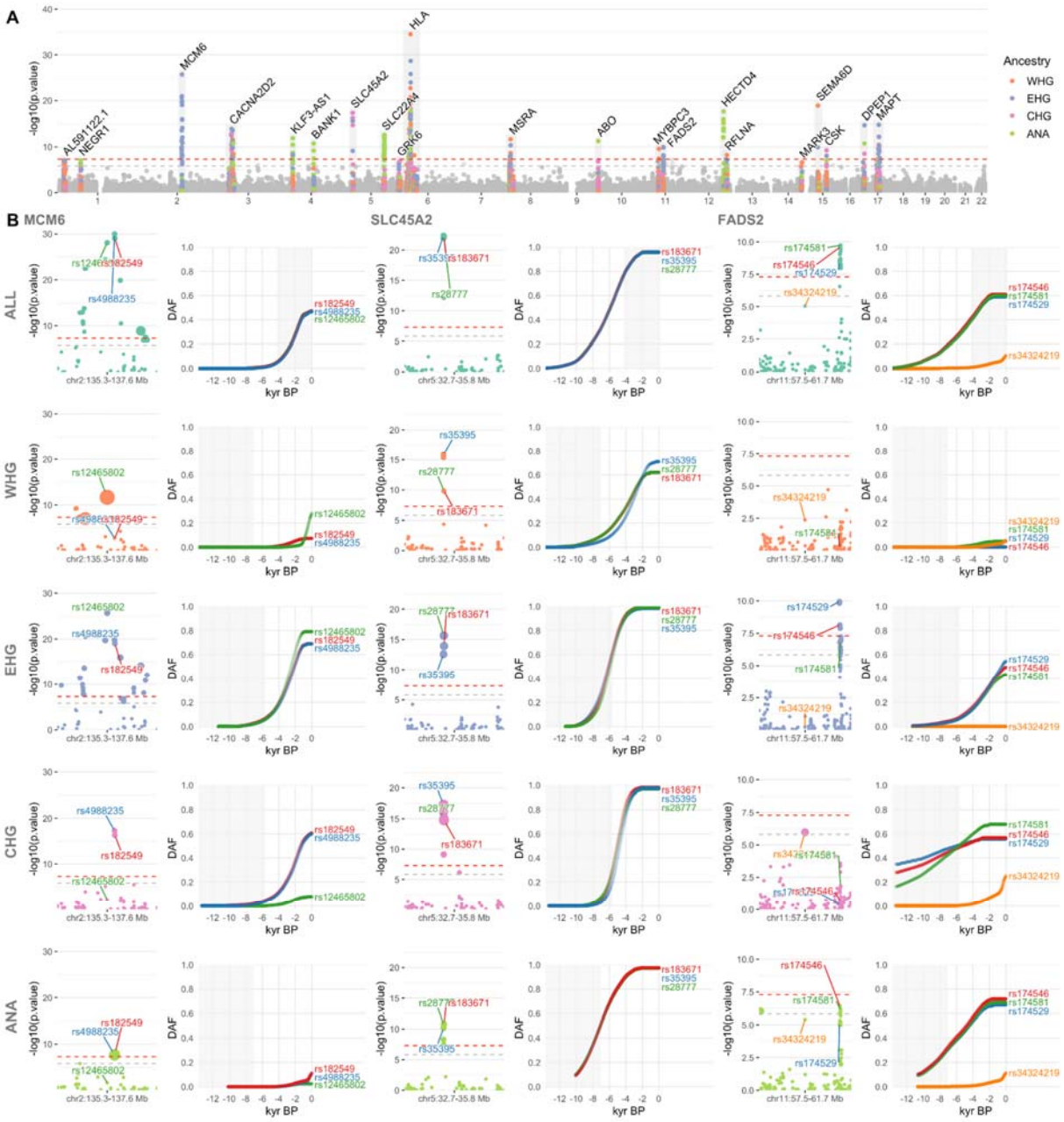


Fig 7. Genome-wide selection scan for trait associated variants. A) Manhattan plot of p-values from selection scan with CLUES, based on a time-series of imputed aDNA genotype probabilities. Twenty-one genome-wide significant selection peaks highlighted in grey and labelled with the most significant gene within each locus. Within each sweep, SNPs are positioned on the y-axis and coloured by their most significant marginal ancestry. Outside of the sweeps, SNPs show p-values from the pan-ancestry analysis and are coloured grey. Red dotted lines indicate genome-wide significance ($p < 5e-8$), while the grey dotted line shows the Bonferroni significance threshold, corrected for the number of tests ($p < 1.35e-6$). B) Detailed plots for three genome-wide significant sweep loci: (i) MCM6, lactase persistence; (ii) SLC45A2, skin pigmentation; and (iii) FADS2, lipid metabolism. Rows show results for the pan-ancestry analysis (ALL) plus the four marginal ancestries: Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) and Anatolian farmers (ANA). The first column of each loci shows zoomed Manhattan plots of the p-values for each ancestry (significant SNPs sized by their selection coefficients), and column two shows allele

trajectories for the top SNPs across all ancestries (grey shading for the marginal ancestries indicates approximate temporal extent of the pre-admixture population).

Selection on diet-associated loci

We find strong changes in selection associated with lactose digestion after the introduction of farming, but prior to the expansion of the Yamnaya pastoralists into Europe around 5,000 years ago [20,21], settling controversies regarding the timing of this selection [129–132]. The strongest overall signal of selection in the pan-ancestry analysis is observed at the *MCM6 / LCT* locus (rs4988235; p=9.86e-31; s=0.020), where the derived allele results in lactase persistence [133,134] (Supplementary Note 4a). The trajectory inferred from the pan-ancestry analysis indicates that the lactase persistence allele began increasing in frequency only c. 7,000 years ago, and has continued to increase up to present times (Fig. 7). Our ancestry-stratified analysis shows, however, that selection at the MCM6/LCT locus is much more complex than previously thought. In the pan-ancestry analysis, this sweep is led by the lactase persistence SNP rs4988235, whereas in the ancestry-stratified analysis, this signal is primarily driven by sweeps in two of the ancestral backgrounds (EHG and CHG), each of which differ in their most significant SNPs (Fig. 7). Conversely, in the WHG background, we find no evidence for selection at rs4988235, but strong selection at rs12465802 within the last c. 2,000 years. Overall, our results suggest that there were multiple, asynchronous selective sweeps in this genomic region in recent human history, and possibly targeting different loci.

We also find strong selection in the FADS gene cluster — *FADS1* (rs174546; p=2.65e-10; s=0.013) and *FADS2* (rs174581; p=1.87e-10; s=0.013) — which are associated with fatty acid metabolism and known to respond to changes in diet from a more/less vegetarian to a more/less carnivorous diet [135–140]. In contrast to previous results [138–140], we find that much of the selection associated with a more vegetarian diet occurred in Neolithic populations before they arrived in Europe, but then continued during the Neolithic (Fig. 7). The strong signal of selection in this region in the pan-ancestry analysis is driven primarily by a sweep occurring on the EHG, CHG and ANA haplotypic backgrounds (Fig. 7). Interestingly, we find no evidence for selection at this locus in the WHG background, and most of the allele frequency rise in the EHG background occurs after their admixture with CHG (around 8ka, [141]), within whom the selected alleles were already close to present-day frequencies. This suggests that the selected alleles may already have existed at substantial frequencies in early farmer populations in the Middle East and among Caucasus Hunter gatherers (associated with the ANA and CHG and backgrounds, respectively) and were subject to continued selection as eastern groups moved northwards and westwards during the late Neolithic and Bronze Age periods.

When specifically comparing selection signatures differentiating ancient hunter-gatherer and farmer populations [142], we also observe a large number of regions associated with lipid and sugar metabolism, and various metabolic disorders (Supplementary Note 4e). These include, for example, a region in chromosome 22 containing *PATZ1*, which regulates the expression of *FADS1*, and *MORC2*, which plays an important role in cellular lipid metabolism [143–145]. Another region in chromosome 3 overlaps with *GPR15,* which is both related to immune tolerance and to intestinal homeostasis [146–148]. Finally, in chromosome 18, we recover a selection candidate region spanning *SMAD7,* which is associated with inflammatory bowel diseases such as Crohn's disease [149–151]. Taken together these results suggest that the transition to agriculture imposed a substantial amount of selection for humans to adapt to our new diet and that some diseases observed today in modern societies can likely be understood as a consequence of this selection.

Selection on immunity-associated variants

In addition to diet-related selection, we observe selection in several loci associated with immunity/defence functions and with autoimmune disease (Supplementary Note 4a). Some of these selection events occurred earlier than previously claimed and are likely associated with the transition to agriculture and may help explain the high prevalence of autoimmune diseases today. Most notably, we detect a 33 megabase (Mb) wide selection sweep signal in chromosome 6 (chr6:19.1–50.9 Mb), spanning the human leukocyte antigen (HLA) region (Supplementary Note 4a). The selection trajectories of the variants within this locus support multiple independent sweeps, occurring at different times and with differing intensities. The strongest signal of selection at this locus in the pan-ancestry analysis is at an intergenic variant, located between *HLA-A* and *HLA-W* (rs7747253; p=8.86e-17; s=-0.018), associated with heel bone mineral density [152], the derived allele of which rapidly reduced in frequency, beginning c. 8,000 years ago (Extended Data Fig. 10). In contrast, the signal of selection at *C2* (rs9267677; p= 9.82e-14; s= 0.04463), also found within this sweep, and associated with educational attainment [153], shows a gradual increase in frequency beginning c. 4,000 years ago, before rising more rapidly c. 1,000 years ago. This highlights the complex temporal dynamics of selection at the HLA locus, which not only plays a role in the regulation of the immune system, but also has association with many other non-immune-related phenotypes. The high pleiotropy in this region makes it difficult to determine which selection pressures may have driven these increases in frequencies at different periods of time. However, profound shifts in lifestyle in Eurasian populations during the Holocene, including a change in diet and closer contact with domestic animals, combined with higher mobility and increasing population sizes, are likely drivers for strong selection on loci involved in immune response.

We also identified selection signals at the *SLC22A4* (rs35260072; p=1.15e-10; s=0.018) locus, associated with increased itch intensity from mosquito bites [154], and find that the derived variant has been steadily rising in frequency since c. 9,000 years ago (Extended Data Fig. 11). However, in the same *SLC22A4* candidate region as rs35260072, we find that the frequency of the previously reported SNP rs1050152 plateaued c. 1,500 years ago, contrary to previous reports suggesting a recent rise in frequency [45]. Similarly, we detect selection at the *HECTD4* (rs11066188; p=3.02e-16; s=0.020) and *ATXN2* (rs653178; p=1.92e-15; s=0.019) locus, associated with celiac disease and rheumatoid arthritis [155], which has been rising in frequency for c. 9,000 years (Extended Data Fig. 12), also contrary to previous reports of a more recent rise in frequency [45]. Thus, several disease-associated loci previously thought to be the result of recent adaptation may have been subject to selection for a longer period of time.

Selection on the 17q12.13 locus

We further detect signs of strong selection in a 12 Mb sweep in chromosome 17 (chr17:36.1–48.1 Mb), spanning a locus on 17q21.3 implicated in neurodegenerative and developmental disorders (Supplementary Note 4a). The locus includes an inversion and other structural polymorphisms with indications of a recent positive selection sweep in some human populations [156,157]. Specifically, partial duplications of the *KANSL1* gene likely occurred independently on the inverted (H2) and non-inverted (H1) haplotypes (Fig. 8B) and both are found in high frequencies (15-25%) among current European and Middle Eastern populations but are much rarer in Sub-Saharan African and East Asian populations. We used both SNP genotypes and WGS read depth information to determine inversion (H1/H2) and *KANSL1* duplication (d) status in the ancient individuals studied here (see Supplementary Note 4g).

The H2 haplotype is observed in two of three previously published genomes[158] of Anatolian aceramic Neolithic individuals (Bon001 and Bon004) from around 10,000 BP, but data were insufficient to identify *KANSL1* duplications. The oldest evidence for *KANSL1* duplications is observed in an Iranian early Neolithic individual (AH1 from 9,900 BP[2]) followed by two Georgian Mesolithic individuals (NEO281 from 9,724 BP and KK1 [6] from 9,720 BP) all of whom are heterozygous for the inversion and carry the inverted duplication. The *KANSL1* duplications are also detected in two Russian Neolithic individuals: NEO560 from 7,919 BP (H1d) and NEO212 from 7,390 BP (H2d). With both H1d and H2d having spread to large parts of Europe with Anatolian Neolithic Farmer ancestry, their frequency seems unchanged in most of Europe as Steppe-related ancestry becomes dominant in large parts of the subcontinent (Extended Data Fig. 8D). The fact that both H1d and H2d are found in apparently high frequencies in both early Anatolian Farmers and the earliest Yamnaya/Steppe-related ancestry groups suggests that any selective sweep acting on the H1d and H2d variants would probably have occurred in populations ancestral to both.

We note that the strongest signal of selection observed in this locus is at *MAPT* (rs4792897; p=4.65e-10; s=0.03 (Fig. 8A; Supplementary Note 4a), which codes for the tau protein [159] and is involved in a number of neurodegenerative disorders, including Alzheimer's disease and Parkinson's disease [160–164]. However, the region is also enriched for evidence of reference bias in our imputed dataset—especially around the *KANSL1* gene—due to complex structural polymorphisms (Supplementary Note 4i).

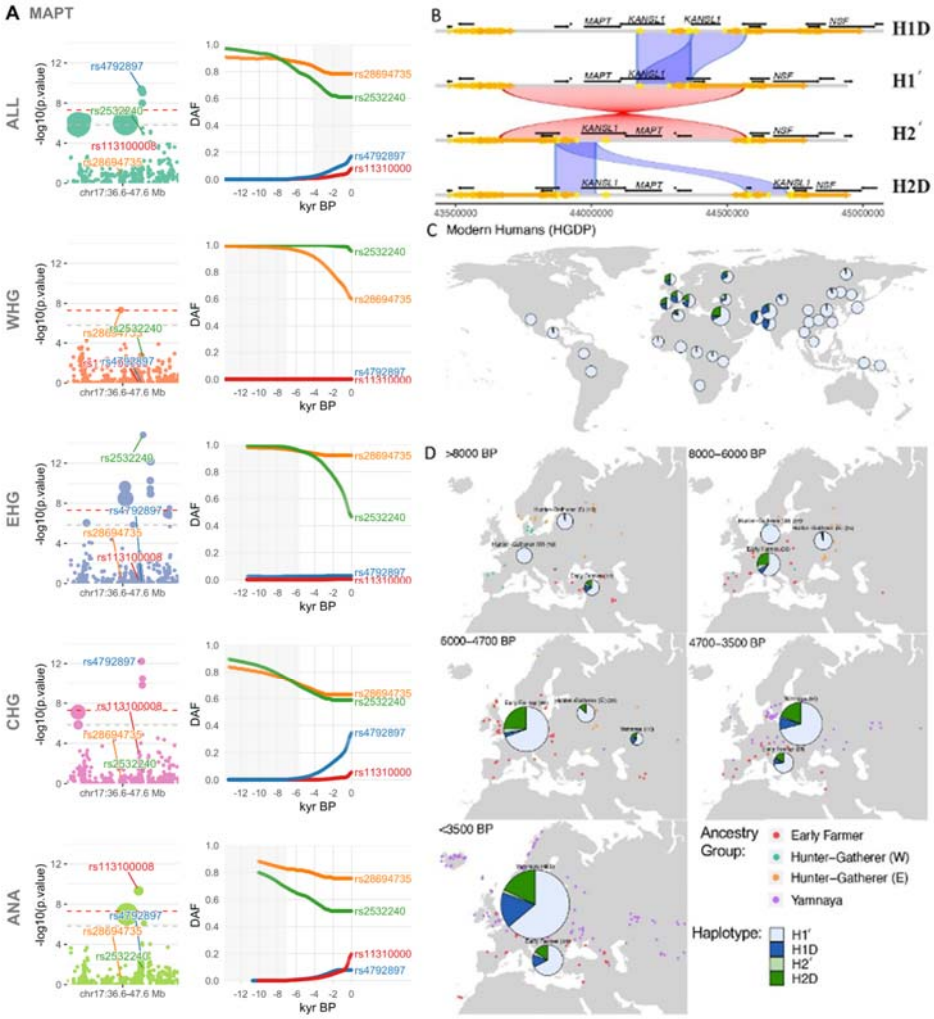


Fig 8. Selection at the *MAPT* / 17q21.31 inversion locus. A) Results for the pan-ancestry analysis (ALL) plus the four marginal ancestries: Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) and Anatolian farmers (ANA). Grey shading for the marginal ancestries indicates approximate temporal extent of the pre-admixture population. B) Haplotypes of the 17q21.31 locus: the ancestral (non-inverted) H1 17q21.31 and the inverted H2 haplotype. Duplications of the *KANSL1* gene have occurred independently on both lineages yielding H1D and H2D haplotypes. C) Frequency of the 17q21.31 inversion and duplication haplotypes across modern-day global populations (Human Genome Diversity Project [119]). D) Change in the frequency of the 17q21.31 inversion haplotype through time.

Selection on pigmentation-associated variants

Our results identify strong selection for lighter skin pigmentation in groups moving northwards and westwards, in agreement with the hypothesis that selection is caused by reduced UV exposure and resulting vitamin D deficiency. We find that the most strongly selected alleles reached near-fixation several thousand years ago, suggesting that this was not associated with recent sexual selection as proposed [165,166] (Supplementary Note 4a).

In the pan-ancestry analysis we detect strong selection at the *SLC45A2* locus (rs35395; p=4.13e-23; s=0.022) locus [167,168], with the selected allele (responsible for lighter skin), increasing in frequency from c. 13,000 years ago, until plateauing c. 2,000 years ago (Fig. 7). The dominating hypothesis is that high melanin levels in the skin are important in equatorial regions owing to its protection

against UV radiation, whereas lighter skin has been selected for at higher latitudes (where UV radiation is less intense) because some UV penetration is required for cutaneous synthesis of vitamin D [169,170]. Our findings confirm pigmentation alleles as major targets of selection during the Holocene [45,171,172] particularly on a small proportion of loci with large effect sizes [168].

Additionally, our results provide unprecedentedly detailed information about the duration and geographic spread of these processes (Fig. 7) suggesting that an allele associated with lighter skin was selected for repeatedly, probably as a consequence of similar environmental pressures occurring at different times in different regions. In the ancestry-stratified analysis, all marginal ancestries show broad agreement at the *SLC45A2* locus (Fig. 7) but differ in the timing of their frequency shifts. The ANA ancestry background shows the earliest evidence for selection, followed by EHG and WHG around c. 10,000 years ago, and CHG c. 2,000 years later. In all ancestry backgrounds except WHG, the selected haplotypes reach near fixation by c. 3,000 years ago, whilst the WHG haplotype background contains the majority of ancestral alleles still segregating in present-day Europeans. This finding suggests that selection on this allele was much weaker in ancient western hunter-gatherer groups during the Holocene compared to elsewhere. We also detect strong selection at the *SLC24A5* (rs1426654; p=6.45e-09; s=0.019) which is also associated with skin pigmentation [167,173]. At this locus, the selected allele increased in frequency even earlier than *SLC45A2* and reached near fixation c. 3,500 years ago (Supplementary Note 4a). Selection on this locus thus seems to have occurred early on in groups that were moving northwards and westwards, and only later in the Western hunter-gatherer background after these groups encountered and admixed with the incoming populations.

Selection among major axes of ancient population variation

Beyond patterns of genetic change at the Mesolithic-Neolithic transition, much genetic variability observed today reflects high genetic differentiation in the hunter-gatherer groups that eventually contributed to modern European genetic diversity [142]. Indeed, a substantial number of loci associated with cardiovascular disease, metabolism and lifestyle diseases trace their genetic variability prior to the Neolithic transition, to ancient differential selection in ancestry groups occupying different parts of the Eurasian continent (Supplementary Note 4d). These may represent selection episodes that preceded the admixture events described above, and led to differentiation between ancient hunter-gatherer groups in the late Pleistocene and early Holocene. One of these overlaps with the *SLC24A3* gene which is a salt sensitivity gene significantly expressed in obese individuals [174,175]. Another spans *ROPN1* and *KALRN*, two genes involved in vascular disorders [176–178]. A further region contains *SLC35F3*, which codes for a thiamine transport and has been associated with hypertension in a Han Chinese cohort [179,180]. Finally, there is a candidate region containing several genes (*CH25H, FAS*) associated with obesity and lipid metabolism [181–183] and another peak with several genes (*ASXL2, RAB10, HADHA, GPR113*) involved in glucose homeostasis and fatty acid metabolism [184–193]. These loci predominantly reflect ancient patterns of extreme differentiation between Eastern and Western Eurasian genomes, and may be candidates for selection after the separation of the Pleistocene populations that occupied different environments across the continent (roughly 45,000 years ago [103]).

Genetic trait reconstruction and the phenotypic legacy of ancient Europeans

When comparing modern European genomes in the UK Biobank to ancient Europeans, we find strong differentiation at certain sets of trait-associated variants, and differential contribution of different ancestry groups to various traits. We reconstructed polygenic scores for phenotypes in ancient individuals, using effect size estimates obtained from GWASs performed using the

>400,000 UK Biobank genomes [107] (http://www.nealelab.is/uk-biobank) and looked for overdispersion among these scores across ancient populations, beyond what would be expected under a null model of genetic drift [194] (Supplementary Note 4c). We stress that polygenic scores and Qx statistic may both be affected by population stratification, so these results should be interpreted with caution [195–198]. The most significantly overdispersed scores are for variants associated with pigmentation, anthropometric differences and disorders related to diet and sugar levels, including diabetes (Fig. 9). We also find psychological trait scores with evidence for overdispersion related to mood instability and irritability, with Western Hunter-gatherers generally showing smaller genetic scores for these traits than Neolithic Farmers. Intriguingly, we find highly inconsistent predictions of height based on polygenic scores in western hunter-gatherer and Siberian groups computed using effect sizes estimated from two different - yet largely overlapping - GWAS cohorts (Supplementary Note 4c), highlighting how sensitive polygenic score predictions are to the choice of cohort, particularly when ancient populations are genetically divergent from the reference GWAS cohort [198]. Taking this into account, we do observe that the Eastern hunter-gatherer and individuals associated with the Yamnaya culture have consistently high genetic values for height, which in turn contribute to stature increases in Bronze Age Europe, relative to the earlier Neolithic populations [45,80,199].

We performed an additional analysis to examine the data for strong alignments between axes of trait-association [200] and ancestry gradients, rather than relying on particular choices for population clusters (Supplementary Note 4e). Along the population structure axis separating ancient East Asian and Siberian genomes from Steppe and Western European genomes (Fig. 1), we find significant correlations with trait-association components related to impedance, body measurements, blood measurements, eye measurement and skin disorders. Along the axis separating Mesolithic hunter-gatherers from Anatolian and Neolithic farmer individuals, we find significant correlations with trait-association components related to skin disorders, diet and lifestyle traits, mental health status, and spirometry-related traits (Fig. 9). Our findings show that these phenotypes were genetically different among ancient groups with very different lifestyles. However, we note that the realised value of these traits is highly dependent on environmental factors and gene-environment interactions, which we do not model in this analysis.

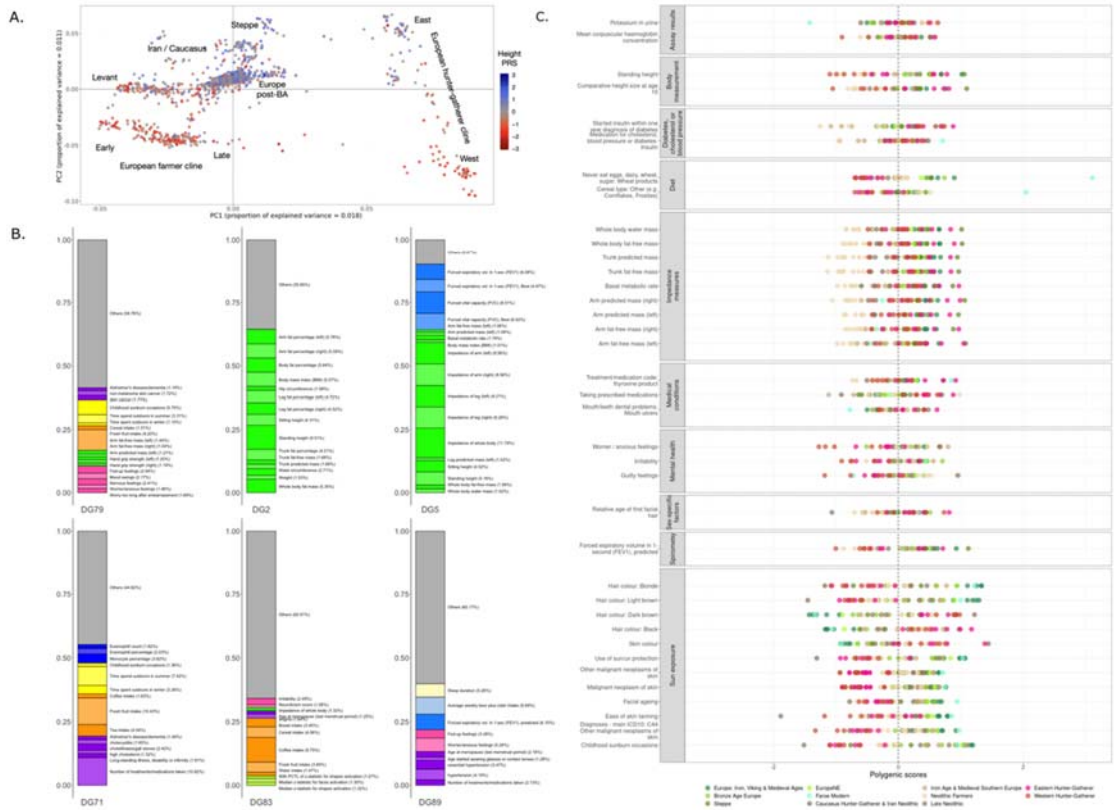


Fig. 9 A. First two principal components of a PCA of ancient Western Eurasian samples, coloured by height polygenic scores. B. Top 6 UK Biobank trait-association components from DeGAs that have the highest correlation ($P < 5e\text{-}3$) with the principal component separating ancient farmers from hunter-gatherer populations (PC1 in Panel A). C. Top UK Biobank traits with highest overdispersion in polygenic scores among ancient populations, out of a total of 320 sets of trait-associated SNPs tested ($Q_X > 79.4$, $P < 0.05/320$).

In addition to the above reconstructions of genetic traits among the ancient individuals, we also estimated the contribution from different ancestral populations (EHG, CHG, WHG, Yamnaya and Anatolian farmer) to variation in polygenic phenotypes in present-day individuals, leveraging the exceptional resolution offered by the UK Biobank genomes [107] to investigate this. We calculated ancestry-specific polygenic risk scores based on the chromosome painting of the >400,000 UKB genomes (Supplementary Note 4h); this allowed us to identify if any of the ancient ancestry components were over-represented in modern UK populations at loci significantly associated with a given trait, and also avoids exporting risk scores over space and time. Working with large numbers of imputed ancient genomes provides high statistical power to use ancient populations as "ancestral sources". We focused on phenotypes whose polygenic scores were significantly over-dispersed in the ancient populations (Supplementary Note 4c), as well as a single high effect variant, ApoE4, known to be a significant risk factor in Alzheimer's Disease ([201,202]). We emphasise that this approach makes no reference to ancient phenotypes but describes how these ancestries contributed to the modern genetic landscape. In light of the ancestry gradients within the British Isles and Eurasia (Fig. 5), these results support the hypothesis that ancestry-mediated geographic variation in disease risks and phenotypes is commonplace. It points to a way forward for disentangling how ancestry contributed to differences in risk of genetic disease – including metabolic and mental health disorders – between present-day populations.

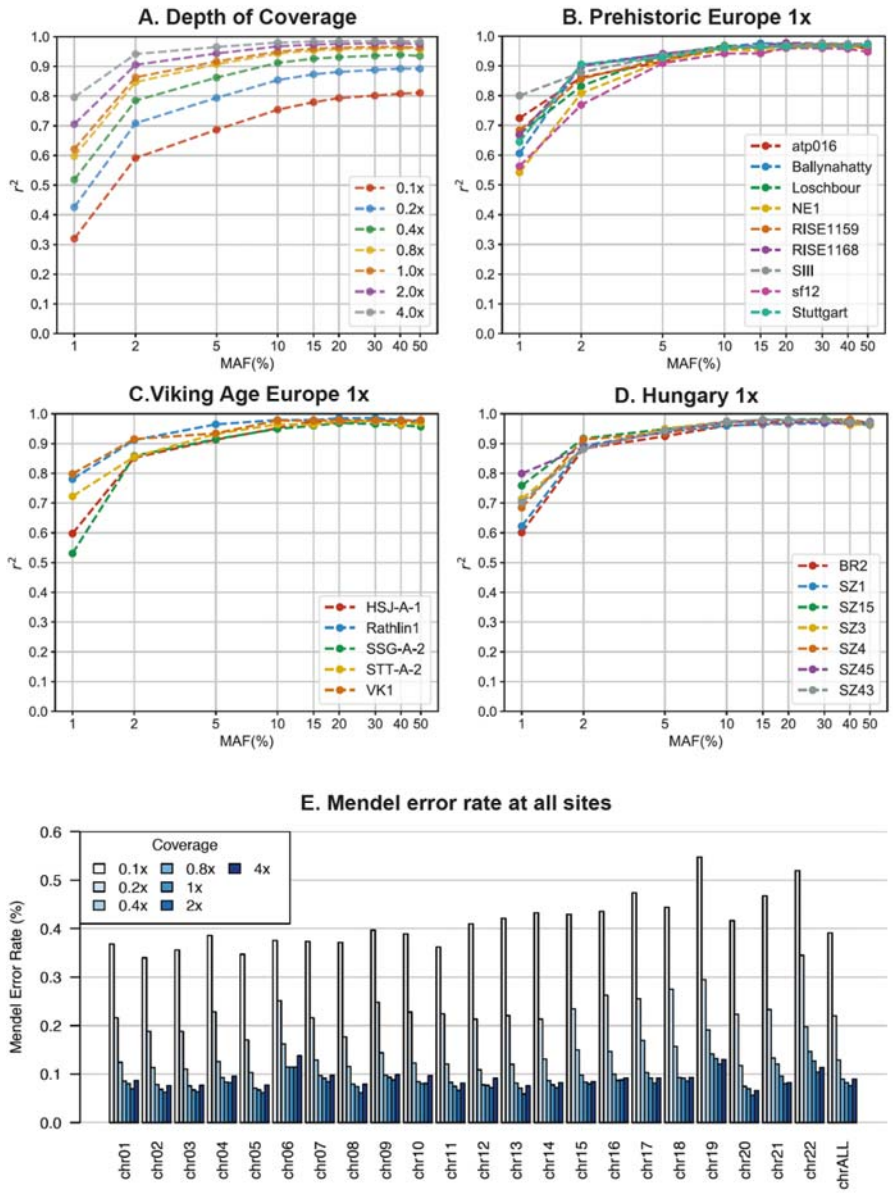
Taken together, these analyses help to settle the famous discussion of selection in Europe relating to height [45,80,203]. The finding that steppe individuals have consistently high genetic values for height (Supplementary Note 4c), is mirrored by the UK Biobank results, which find that the 'Steppe' ancestral components (Yamnaya/EHG) contributed to increased height in present-day populations (Supplementary Note 4h). This shows that the height differences in Europe between north and south may not be due to selection, as claimed in many previous studies, but may be a consequence of differential ancestry.

Likewise, European hunter gatherers are genetically predicted to have dark skin pigmentation and dark brown hair [11,20,21,79,83,168,204,205], and indeed we see that the WHG, EHG and CHG components contributed to these phenotypes in present-day individuals whereas the Yamnaya and Anatolian farmer ancestry contributed to light brown/blonde hair pigmentation (Supplementary Note 4h). Interestingly, loci associated with overdispersed mood-related polygenic phenotypes recorded among the UK Biobank individuals (like increased anxiety, guilty feelings, and irritability) showed an overrepresentation of the Anatolian farmer ancestry component; and the WHG component showed a strikingly high contribution to traits related to diabetes. We also found that the ApoE4 effect allele is preferentially found on a WHG/EHG haplotypic background, suggesting it likely was brought to western Europe by early hunter-gatherers (Supplementary Note 4h). This is in line with the present-day European distribution of this allele, which is highest in north-eastern Europe, where the proportion of these ancestries are higher than in other regions of the continent [206].

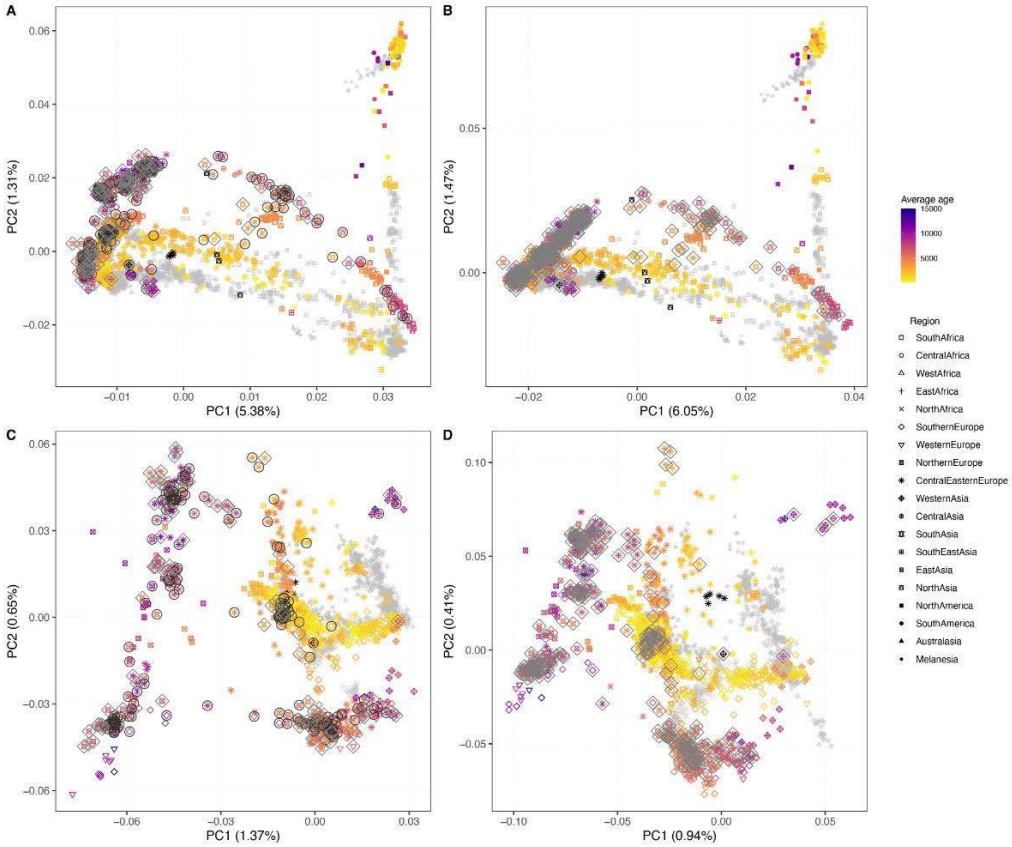
Conclusion

Our study has provided fundamental new insights into one of the most transformative periods of human biological and cultural evolution. We have demonstrated that a clear east-west division known from Stone Age material culture, extending from the Black Sea to the Baltic and persisting across several millennia, was genetically deeply rooted in populations with different ancestries. We showed that the genetic impact of the Neolithic transition was highly distinct, east and west of this boundary. We have identified a hitherto unknown source of ancestry in hunter-gatherers from the Middle Don region contributing ancestry to the Yamnaya pastoralists, and we have documented how the later spread of steppe-related ancestry into Europe was very rapid and mediated through admixture with people from the Globular Amphora Culture. Additionally, we have observed two near-complete population replacements in Denmark within just 1,000 years, concomitantly with major changes in material culture, which rules out cultural diffusion as a main driver and settles generation-long archaeological debates. Our analyses revealed that the ability to detect signatures of natural selection in modern human genomes is drastically limited by conflicting selection pressures in different ancestral populations masking the signals. Developing methods to trace selection in individual ancestry components allowed us to effectively double the number of significant selection peaks, which helped clarify the trajectories of a number of traits related to diet and lifestyle. Our results emphasise how the interplay between major ancient selection and admixture events occurring across Europe and Asia in the Stone and Bronze Ages have profoundly shaped patterns of genetic variation in modern human populations.

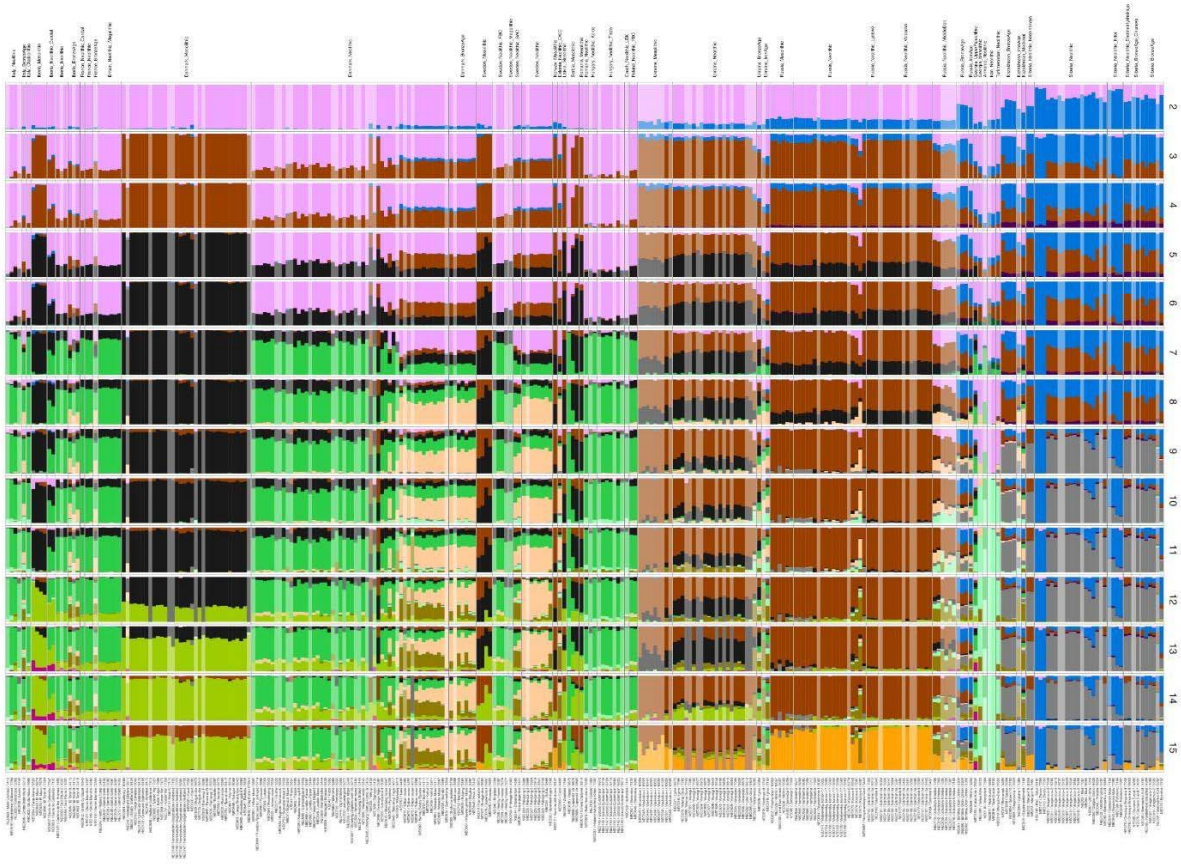
Extended Data Figures



Extended Data Fig.1. Imputation accuracy of aDNA. Panel A shows imputation accuracy across the 42 high-coverage ancient genomes when downsampled to lower depth of coverage values. Panels B-D show imputation accuracy for 1X depth of coverage across 21 high-coverage ancient European genomes. In panels A-D, imputation accuracy is shown as the squared Pearson correlation between imputed and true genotype dosages as a function of minor allele frequency of the target variant sites. Panel E shows imputation accuracy as Mendelian error rates for a re-sequenced Neolithic trio [27] downsampled to a range of low coverage values.

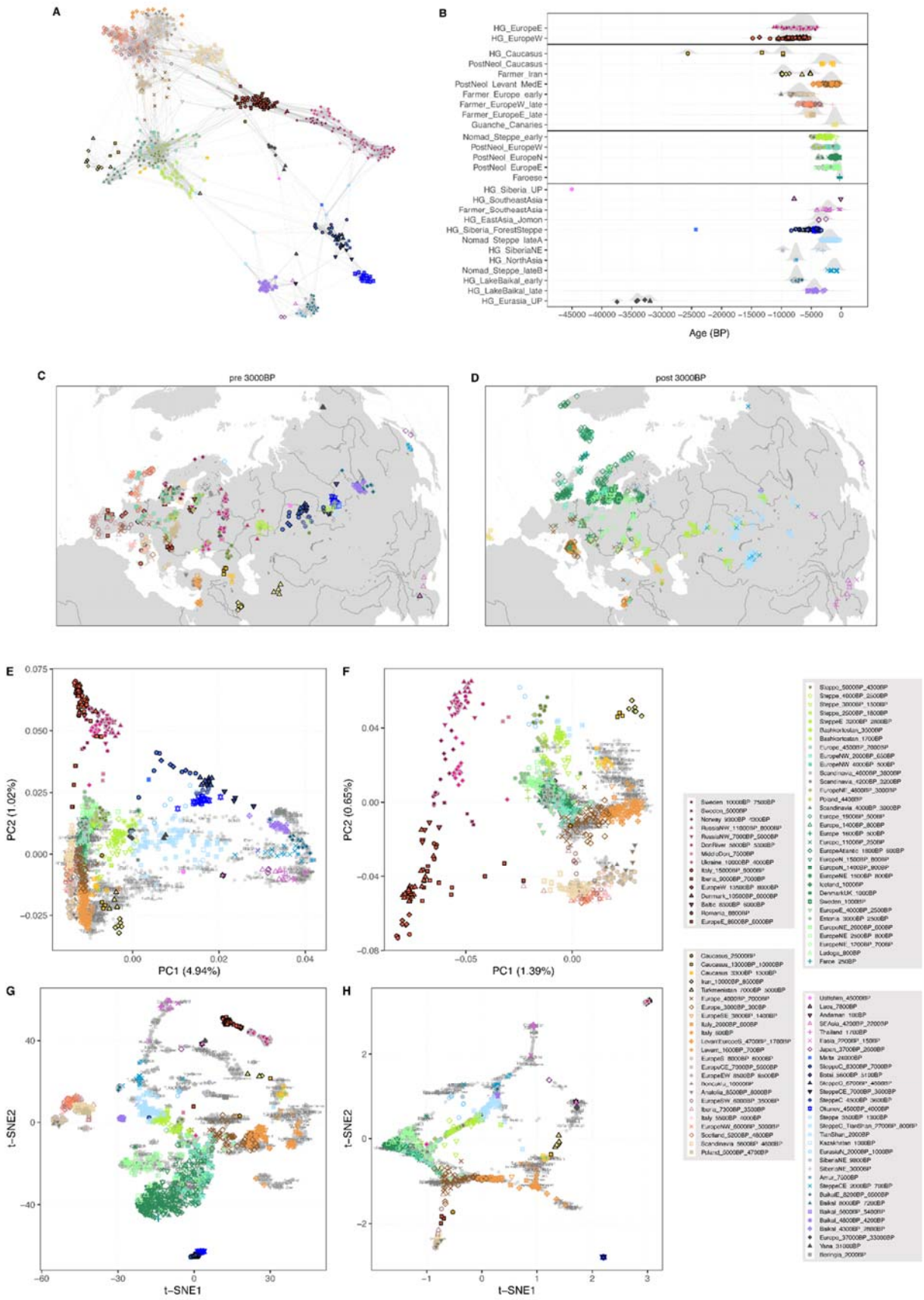

A
B
C
D
PC1 (5.38%)
PC2 (1.31%)
PC1 (6.05%)
PC2 (1.47%)
PC1 (1.37%)
PC2 (0.65%)
PC1 (0.94%)
PC2 (0.41%)
Average age
15000
10000
5000
Region
SouthAfrica
CentralAfrica
WestAfrica
EastAfrica
NorthAfrica
SouthernEurope
WesternEurope
NorthernEurope
CentralEasternEurope
WesternAsia
CentralAsia
SouthAsia
SouthEastAsia
EastAsia
NorthAsia
NorthAmerica
SouthAmerica
Australasia
Melanesia

E


2
3
4
5
6
7
8
9
10
11
12
13
14
15

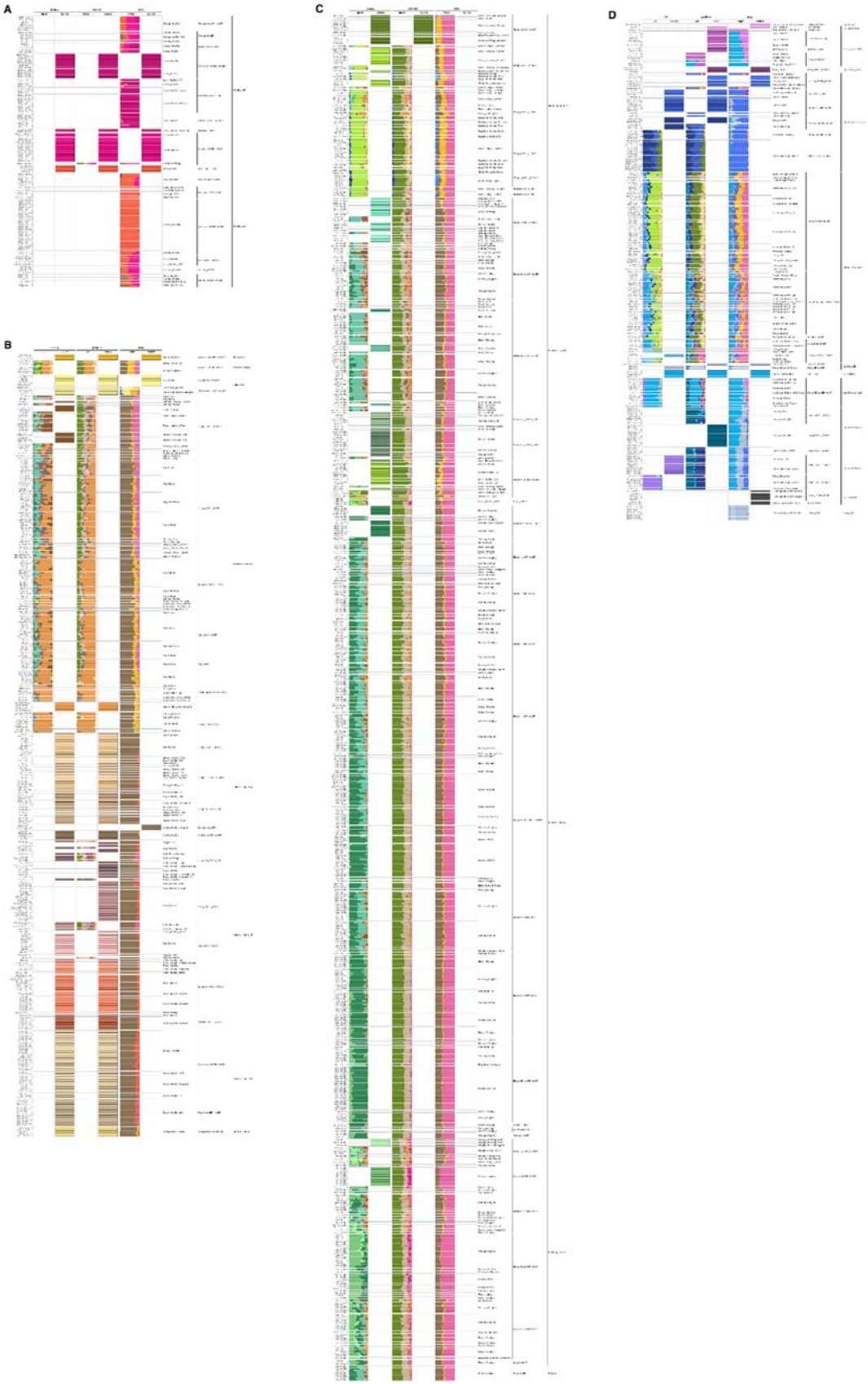
1239

1240 **Extended Data Fig. 2. Genetic structure of the 317 herein-reported ancient genomes.** (A)-(D) Principal component
1241 analysis of 3,316 modern and ancient individuals from Eurasia, Oceania and the Americas (A, B), as well as restricted
1242 to 2,126 individuals from western Eurasia (west of Urals) (C, D). Shown are analyses with principal components
1243 inferred either using both modern and imputed ancient genomes passing all filters, and projecting low coverage ancient
1244 genomes (A, C); or only modern genomes and projecting all ancient genomes (B, D). Ancient genomes sequenced in
1245 this study are indicated either with black circles (imputed genomes) or grey diamonds (projected genomes). (E) Model-
1246 based clustering results using ADMIXTURE for 284 newly reported genomes (excluding close relatives and individuals
1247 flagged for possible contamination). Results shown are based on ADMIXTURE runs from K=2 to K=15 on 1,584
1248 ancient individuals. Low-coverage individuals represented by pseudo-haploid genotypes are indicated with alpha
1249 transparency.
1250


A
B
HG_EuropeE
HG_EuropeW
HG_Caucasus
PostNeol_Caucasus
Farmer_Iran
PostNeol_Levant_MedE
Farmer_Europe_early
Farmer_EuropeW_late
Farmer_EuropeE_late
Guanche_Canaries
Nomad_Steppe_early
PostNeol_EuropeW
PostNeol_EuropeN
PostNeol_EuropeE
Faroese
HG_Siberia_UP
HG_SoutheastAsia
Farmer_SoutheastAsia
HG_EastAsia_Jomon
HG_Siberia_ForestSteppe
Nomad_Steppe_lateA
HG_SiberiaNE
HG_NorthAsia
Nomad_Steppe_lateB
HG_LakeBaikal_early
HG_LakeBaikal_late
HG_Eurasia_UP
−45000 −40000 −35000 −30000 −25000 −20000 −15000 −10000 −5000 0
Age (BP)
C
pre 3000BP
D
post 3000BP
E
PC2 (1.02%)
PC1 (4.94%)
F
PC2 (0.65%)
PC1 (1.39%)
G
t−SNE2
t−SNE1
H
t−SNE2
t−SNE1

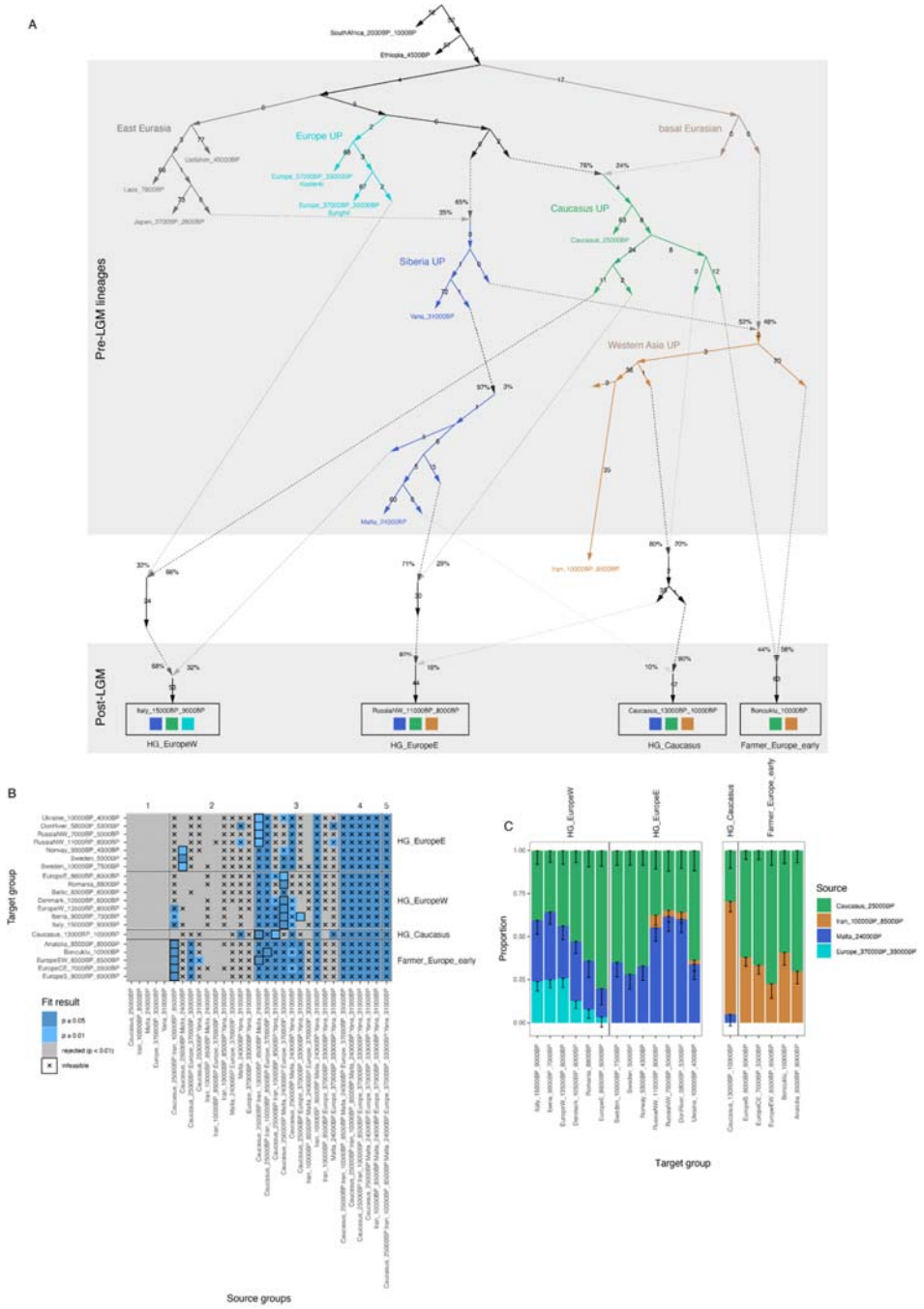
1251

Extended Data Fig. 3. Genetic clustering of ancient individuals. Genetic clusters inferred from pairwise identity-by-descent (IBD) sharing of 1,401 ancient Eurasian individuals, indicated using colored symbols throughout (A) 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. (B) Temporal distribution of clustered individuals, grouped by broad ancestry cluster. (C), (D) Geographical distribution of clustered individuals, shown for individuals predating 3,000 BP (C) and after 3,000 BP (D). (E)-(H) Fine-scale population structure among genetic clusters. Modern individuals are shown in gray, with population labels corresponding to their median coordinates. (E), (F) PCA of 3,119 Eurasian (E) or 2,126 west Eurasian (F) individuals. (G), (H) t-distributed stochastic neighbour embedding (t-SNE) using the first 12 principal components of the all Eurasian panel (E). Shown are embeddings with two different exaggeration factors ρ, emphasising local (G, ρ=1) or global (H, ρ=30) structure.

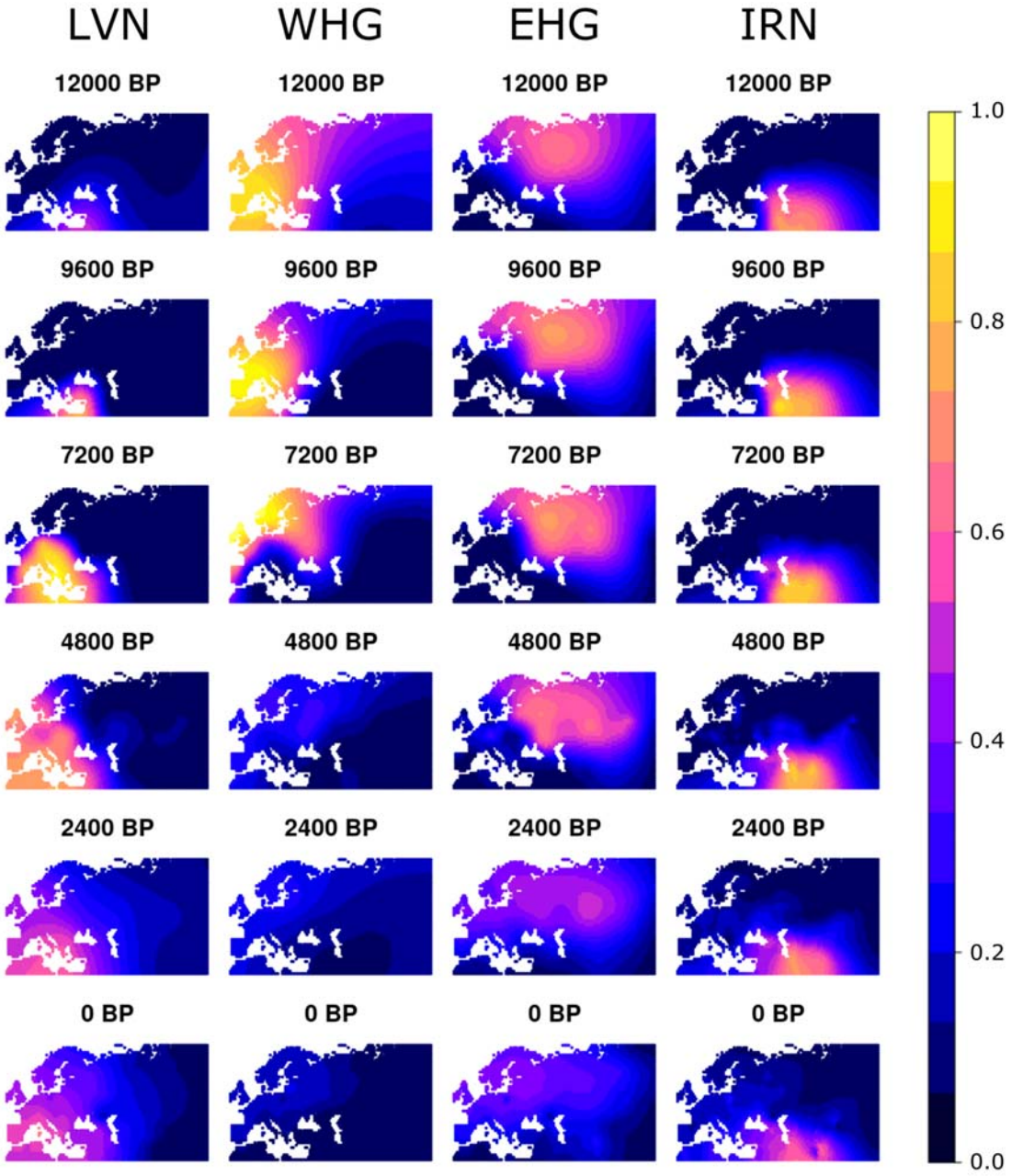


1266

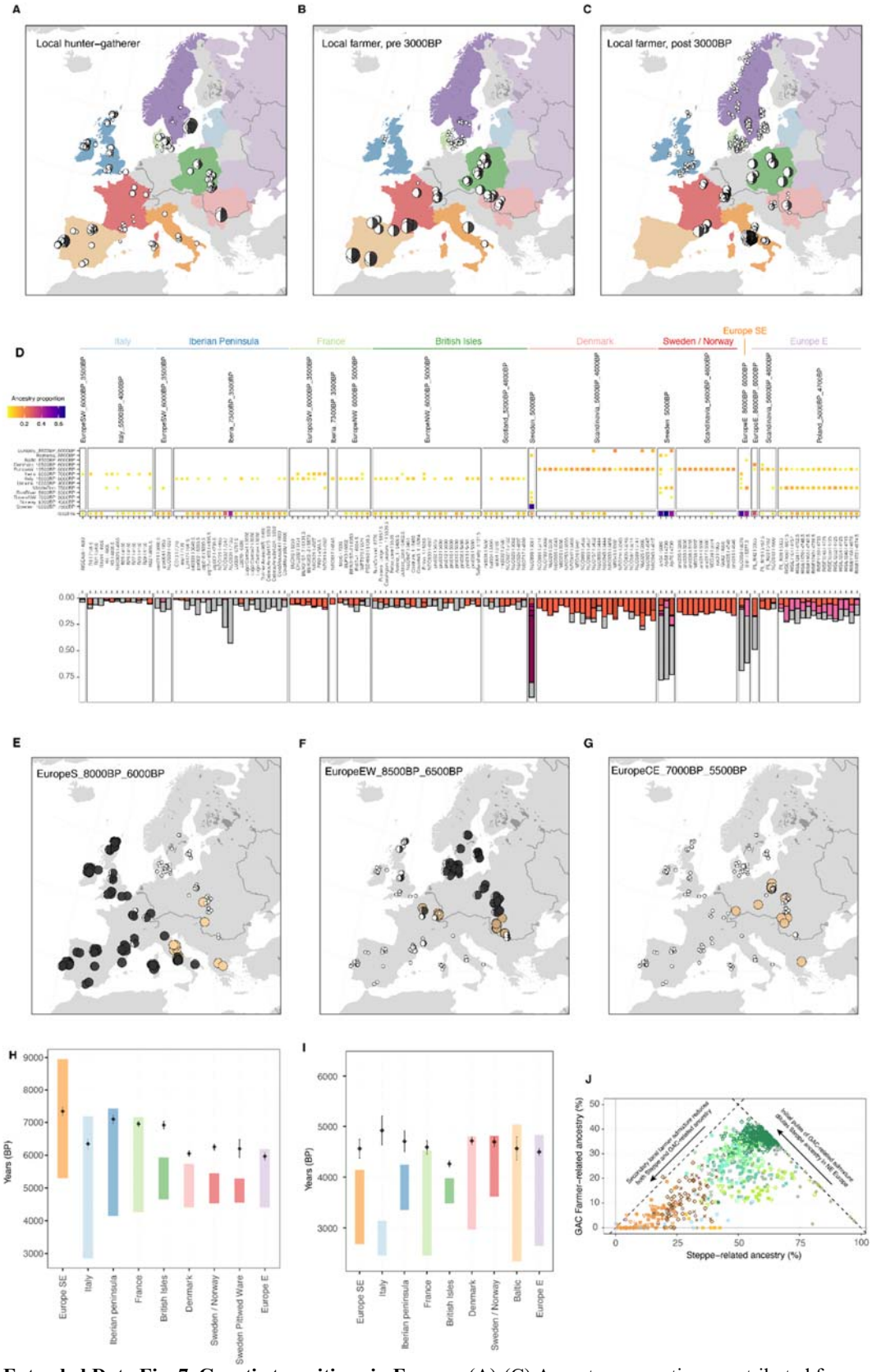
Extended Data Fig. 4. Admixture modelling. Supervised admixture modelling using non-negative least squares on IBD sharing profiles. Panels (A)-(D) show estimated ancestry proportions of four global Eurasian clusters, corresponding to (A) European hunter-gatherers before 4,000 BP; (B) Individuals from Europe and Western Asia from around 10,000 BP until historical times, including Neolithic farmers, Caucasus hunter-gatherers and recent individuals with genetic affinity to the Levant; (C) European individuals after 5,000 BP, as well as pastoralist groups from the Eurasian Steppe; (D) Central, East and North Asian individuals with east Eurasian genetic affinities. Column pairs show results of modelling target individuals (left columns) using three panels of increasingly distal source groups (right columns): "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. Note that some clusters of individuals can be either sources or targets across distinct panels.



Extended Data Fig. 5. Deep Eurasian population structure. (A) Admixture graph fit relating deep Eurasian lineages predating the Last Glacial Maximum (LGM) to later West Eurasian ancestry clusters (worst |Z| = 3.65). (B) Rotating outgroup *qpAdm* analysis showing fit results for modelling post-LGM target groups as mixtures of all possible combinations involving one to five source groups. Colours of the individual matrix cells indicate the fit for a particular model, either rejected at $p<0.01$ (grey), $0.01 \leq p < 0.05$ (light blue) or $p \geq 0.05$ (dark blue). Cells with crosses indicate infeasible models involving negative admixture proportions. (C) Estimated ancestry proportions from *qpAdm* for post-LGM target groups inferred from the model fitting with least number of source groups.

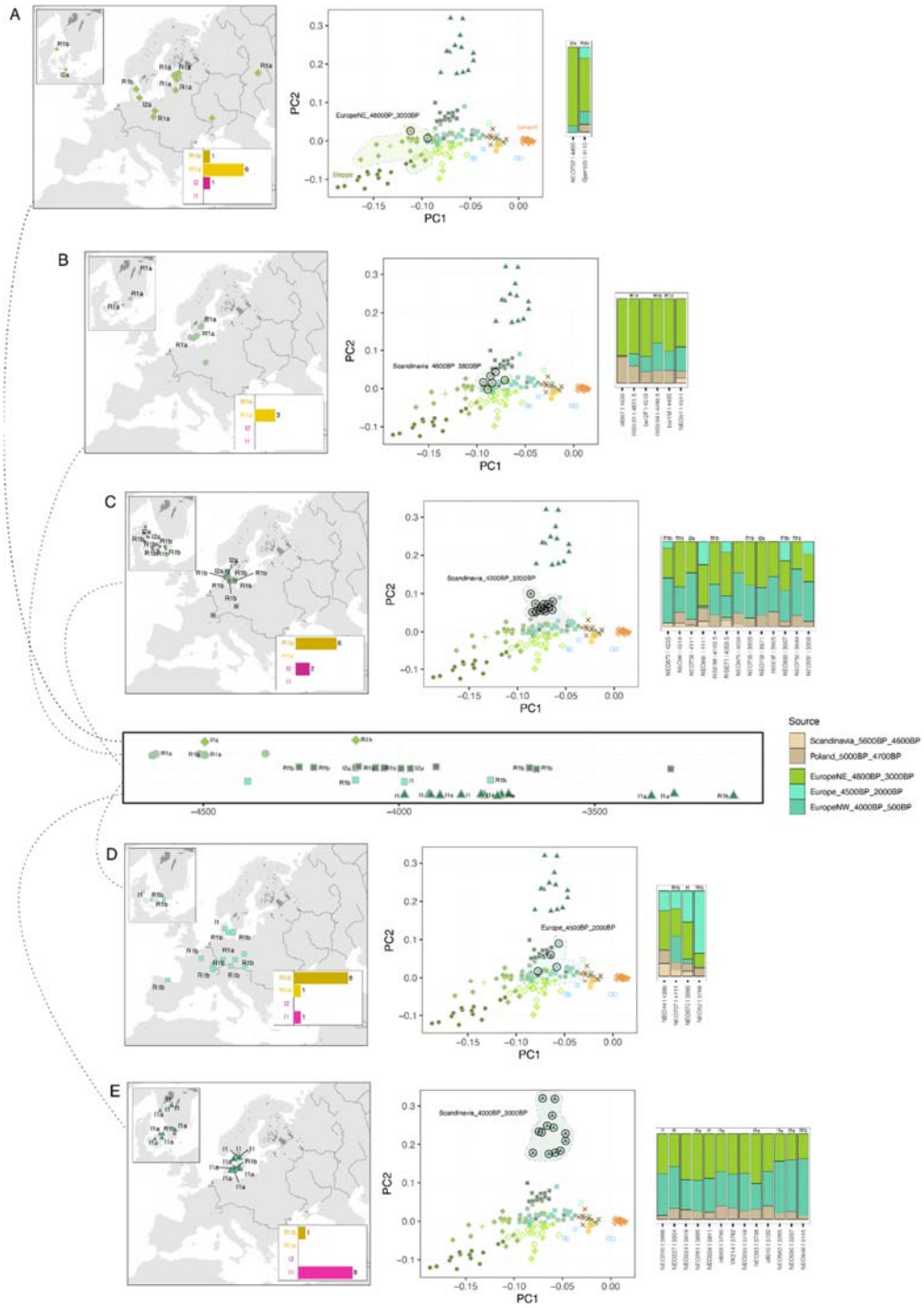


Extended Data Fig 6. Spatiotemporal kriging of four major ancestry clusters over the last 12,000 years of human history. LVN = ancestry maximized in Anatolian farmer populations. WHG = ancestry maximized in western European hunter-gatherers. EHG = ancestry maximized in eastern European hunter-gatherers. IRN = ancestry maximized in Iranian Neolithic individuals and Caucasus hunter-gatherers.

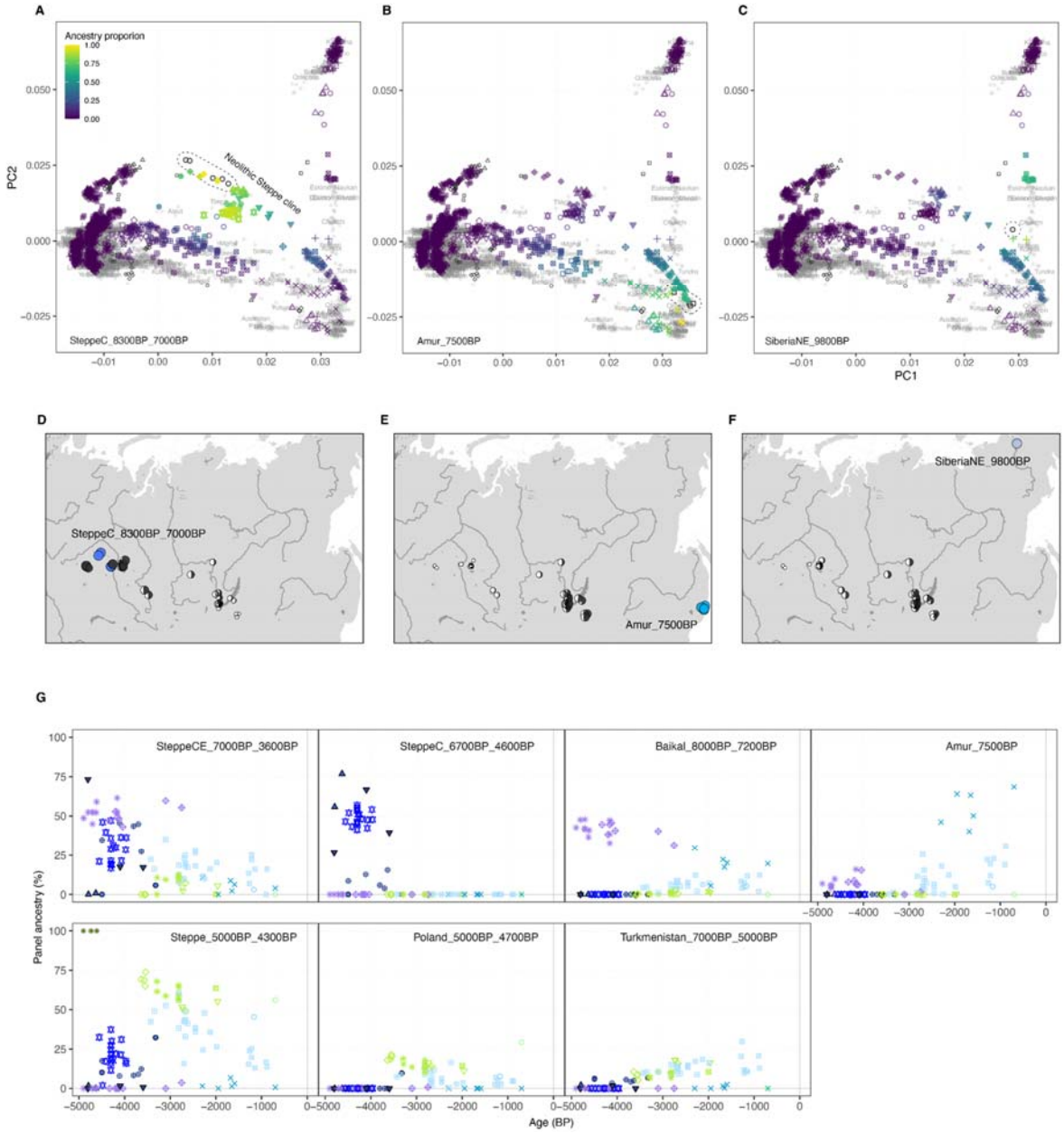


Extended Data Fig. 7. Genetic transitions in Europe. (A)-(C) Ancestry proportions contributed from preceding local groups to later individuals during the two major western Eurasian genetic transitions. (A) contribution to individuals

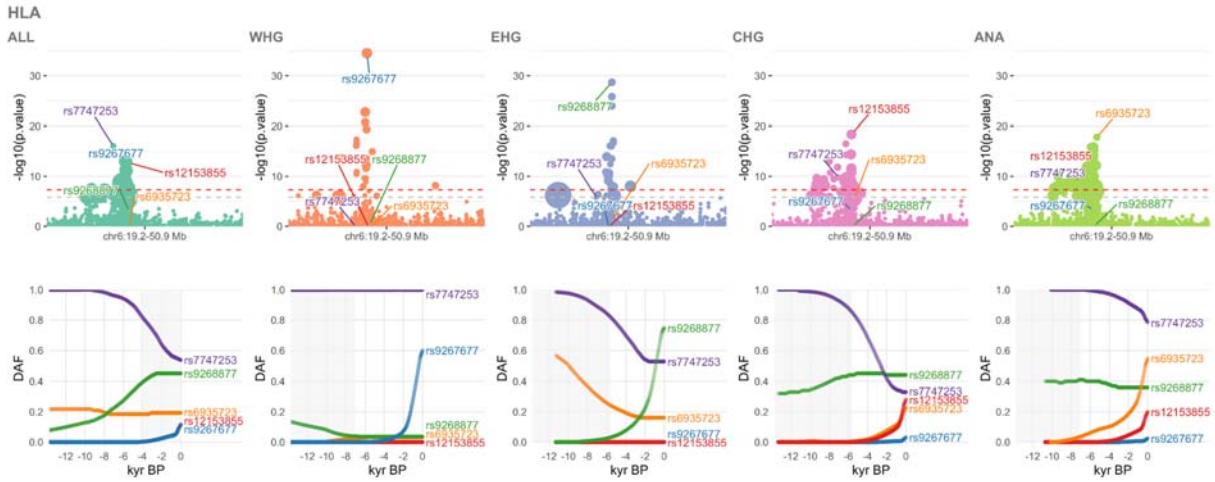
with farmer ancestry from preceding local hunter-gatherer groups; (B,C) contribution to individuals with Steppe-related ancestry from preceding local farmer groups. Coloured areas in all maps indicate the geographic extent of individuals included in respective regions. (D) Composition of hunter-gatherer ancestry proportions from different source groups in individuals with farmer ancestry, shown as heatmap (top) and barplots (bottom). Grey bars represent contributions from local hunter-gatherers (E)-(G) Moon charts showing spatial distribution of estimated ancestry proportions of mid- to late Neolithic farmer individuals from three 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. (H, I) Estimated time of admixture between (H) local hunter-gatherer groups and farmers and (I) eastern European farmers with GAC-related ancestry and Steppe pastoralist groups. 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. (J) Correlation between estimated proportions of Steppe-related and GAC farmer-related ancestries, across west Eurasian target individuals. Symbol shape and colour indicate the genetic cluster of respective individuals.



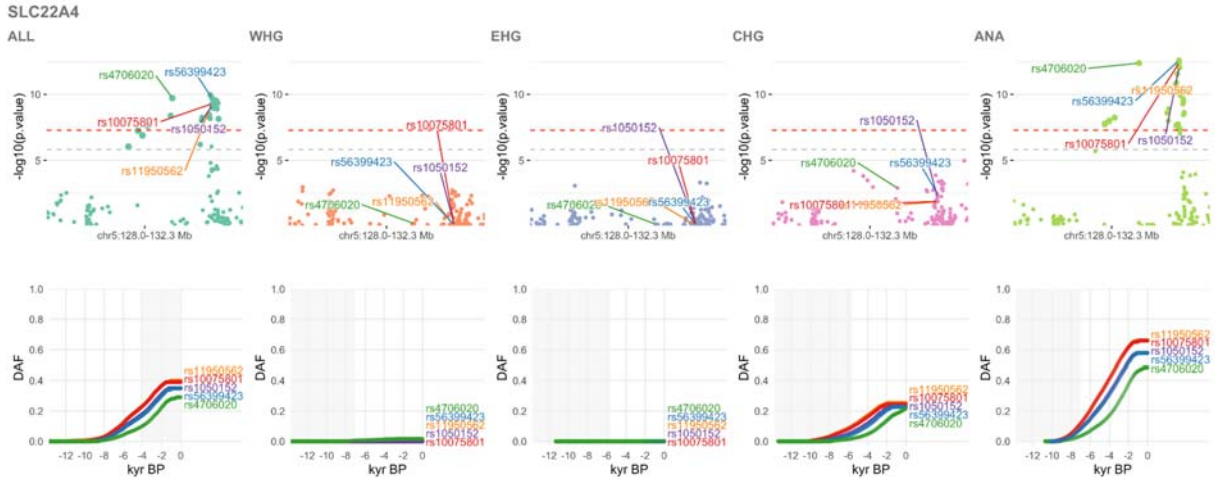
Extended Data Fig. 8. Fine-scale structure in Late Neolithic Scandinavians. (A)-(E) Geographic locations and PCA based on pairwise IBD sharing (middle) of 148 European individuals predating 3,000 BP. Geographic locations are shown for 65 individuals belonging to the five genetic clusters observed in 38 ancient Scandinavians (temporal sequence shown in timeline in centre of plot). Individual assignments and frequency distribution of major Y chromosome haplogroups are indicated in maps and timeline. Plot symbols with black circles indicate the 38 Scandinavian individuals in the PCA panels. Ancestry proportions for the 38 Scandinavian individuals estimated using proximal source groups from outside Scandinavia ("postNeolScand" source set) are shown on the right of the respective cluster results.



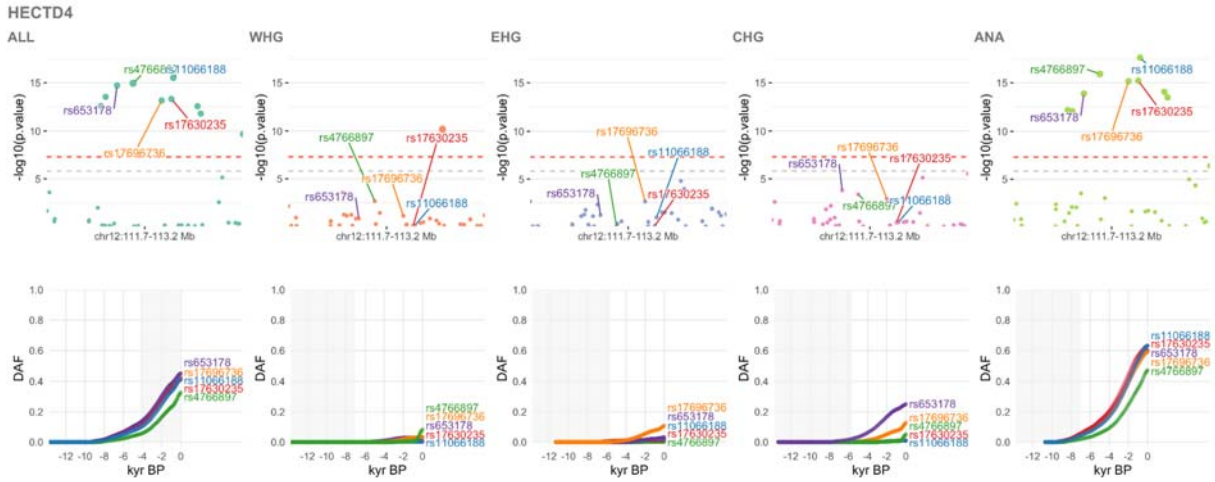
Extended Data Fig. 9. Genetic transformations across the Eurasian Steppe. (A)-(C) Principal component analysis of modern and ancient individuals from Eurasia, Oceania and the Americas, highlighting estimated ancestry proportions from "deep" Siberian ancestry sources (individuals highlighted with dashed line). Present-day individuals are shown in gray, with population labels corresponding to their median coordinates. (D)-(E) Moon charts showing spatial distribution of estimated ancestry proportions of Siberian hunter-gatherers before 5,000 BP from "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. (G) 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.



Extended Data Fig. 10. Selection at the HLA locus. Results for the pan-ancestry analysis (ALL) plus the four marginal ancestries: Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) and Anatolian farmers (ANA). Row one shows zoomed Manhattan plots of the p-values for each ancestry (significant SNPs sized by their selection coefficients), and row two shows allele trajectories for the top SNPs across all ancestries (grey shading for the marginal ancestries indicates approximate temporal extent of the pre-admixture population).



Extended Data Fig. 11. Selective sweep at the *SLC22A4* locus. Results for the pan-ancestry analysis (ALL) plus the four marginal ancestries: Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) and Anatolian farmers (ANA). Row one shows zoomed Manhattan plots of the p-values for each ancestry (significant SNPs sized by their selection coefficients), and row two shows allele trajectories for the top SNPs across all ancestries (grey shading for the marginal ancestries indicates approximate temporal extent of the pre-admixture population).



Extended Data Fig. 12. Selective sweep at the *HECTD4* locus. Results for the pan-ancestry analysis (ALL) plus the four marginal ancestries: Western hunter-gatherers (WHG), Eastern hunter-gatherers (EHG), Caucasus hunter-gatherers (CHG) and Anatolian farmers (ANA). Row one shows zoomed Manhattan plots of the p-values for each ancestry (significant SNPs sized by their selection coefficients), and row two shows allele trajectories for the top SNPs across all ancestries (grey shading for the marginal ancestries indicates approximate temporal extent of the pre-admixture population).

Data availability

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 Accelerator Mass Spectrometry results) are included in the online supplementary materials of this submission.

Code availability

The modified version of CLUES used in this study is available from https://github.com/standard-aaron/clues. The pipeline and conda environment necessary to replicate the analysis of allele frequency trajectories of trait-associated variants in Supplementary Note 4a are available on Github at https://github.com/ekirving/mesoneo_paper. The pipeline to replicate the analyses for Supplementary Note 4c-4e can be found at https://github.com/albarema/neo. All other analyses relied upon available software which has been fully referenced in the manuscript and detailed in the relevant supplementary notes.

References

1. Riehl, S., Zeidi, M. & Conard, N. J. Emergence of agriculture in the foothills of the Zagros Mountains of Iran. *Science* **341**, 65–67 (2013).

2. Broushaki, F. *et al*. Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).

3. Shennan, S. *The First Farmers of Europe: An Evolutionary Perspective*. (Cambridge University Press, 2018).

4. Ellis, E. C., Klein Goldewijk, K., Siebert, S., Lightman, D. & Ramankutty, N. Anthropogenic transformation of the biomes, 1700 to 2000. *Glob. Ecol. Biogeogr.* no–no (2010).
5. Burger, J. R. & Fristoe, T. S. Hunter-gatherer populations inform modern ecology. *Proceedings of the National Academy of Sciences of the United States of America* vol. 115 1137–1139 (2018).
6. 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).
7. 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).
8. 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).
9. Furholt, M. Mobility and Social Change: Understanding the European Neolithic Period after the Archaeogenetic Revolution. *Journal of Archaeological Research* **29**, 481–535 (2021).
10. 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).
11. Brace, S. *et al.* Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat Ecol Evol* **3**, 765–771 (2019).
12. Fowler, C., Harding, J. & Hofmann, D. Introduction. in *The Oxford Handbook of Neolithic Europe* (eds. Fowler, C., Harding, J. & Hofmann, D.) (Oxford University Press, 2015).
13. Kislenko, A. & Tatarintseva, N. The eastern Ural steppe at the end of the Stone Age. *Late prehistoric exploitation of the Eurasian steppe* 183–216 (1999).
14. Mittnik, A. *et al.* The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
15. Fischer, A. & Kristiansen, K. *The Neolithisation of Denmark. 150 years of debate*. (J.R. Collis Publications. Sheffield, 2002).
16. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368 (2017).
17. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
18. Fort, J. Demic and cultural diffusion propagated the Neolithic transition across different regions of Europe. *J. R. Soc. Interface* **12**, (2015).
19. Fort, J., Mercè Pareta, M. & Sørensen, L. Estimating the relative importance of demic and cultural diffusion in the spread of the Neolithic in Scandinavia. *J. R. Soc. Interface* **15**, 20180597 (2018).

20. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).

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

22. Page, A. E. *et al.* Reproductive trade-offs in extant hunter-gatherers suggest adaptive mechanism for the Neolithic expansion. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 4694–4699 (2016).

23. Marciniak, S., Bergey, C., Silva, A. M. & Hałuszko, A. An integrative skeletal and paleogenomic analysis of prehistoric stature variation suggests relatively reduced health for early European farmers. *bioRxiv* (2021).

24. Nielsen, R. Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**, 197–218 (2005).

25. Vitti, J. J., Grossman, S. R. & Sabeti, P. C. Detecting natural selection in genomic data. *Annu. Rev. Genet.* **47**, 97–120 (2013).

26. Rubinacci, 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).

27. Schroeder, H. *et al.* Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 10705–10710 (2019).

28. 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).

29. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).

30. Lazaridis, I., Belfer-Cohen, A., Mallick, S. & Patterson, N. Paleolithic DNA from the Caucasus reveals core of West Eurasian ancestry. *BioRxiv* (2018).

31. Günther, T. *et al.* Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).

32. Klütsch, C. F. C., Manseau, M. & Wilson, P. J. Phylogeographical analysis of mtDNA data indicates postglacial expansion from multiple glacial refugia in woodland caribou (Rangifer tarandus caribou). *PLoS One* **7**, e52661 (2012).

33. Horreo, J. L. & Fitze, P. S. Postglacial Colonization of Northern Europe by Reptiles. in *Origin and Evolution of Biodiversity* (ed. Pontarotti, P.) 197–214 (Springer International Publishing, 2018).

34. Olalde, I. *et al.* The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).

35. Villalba-Mouco, V. *et al.* Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula. *Curr.*

Biol. **29**, 1169–1177.e7 (2019).

36. 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).
37. 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).
38. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
39. Antonio, M. L. *et al.* Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).
40. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 12791–12798 (2020).
41. Cassidy, L. M. *et al.* A dynastic elite in monumental Neolithic society. *Nature* **582**, 384–388 (2020).
42. Racimo, F. *et al.* The spatiotemporal spread of human migrations during the European Holocene. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8989–9000 (2020).
43. Saag, L. *et al.* Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Curr. Biol.* **27**, 2185–2193.e6 (2017).
44. 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.e10 (2021).
45. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
46. 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).
47. 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. U. S. A.* **114**, 897–902 (2017).
48. Narasimhan, V. M. *et al.* The formation of human populations in South and Central Asia. *Science* **365**, (2019).
49. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
50. Malmström, H. *et al.* Ancient mitochondrial DNA from the northern fringe of the Neolithic farming expansion in Europe sheds light on the dispersion process. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20130373 (2015).
51. Coutinho, A. *et al.* The Neolithic Pitted Ware culture foragers were culturally but not genetically influenced by the Battle Axe culture herders. *Am. J. Phys. Anthropol.* **172**, 638–649 (2020).

52. Becker, C. J. Den grubekeramiske kultur i Danmark. Aarbøger for nordisk Oldkyndighed og Historie. 153–274 (1951).

53. Iversen, R. In a World of Worlds: The Pitted Ware Complex in a large scale perspective. *Acta Archaeol.* **81**, 5–43 (2010).

54. Iversen, R., Philippsen, B. & Persson, P. Reconsidering the Pitted Ware chronology. *Praehistorische Zeitschrift* **96**, 44–88 (2021).

55. Klassen, L., Iversen, R. & Rasmussen, L. W. *The Pitted Ware Culture on Djursland Supra-regional: Significance and Contacts in the Middle Neolithic of Southern Scandinavia*. (Aarhus Universitetsforlag, 2020).

56. Sørensen, L. From hunter to farmer in Northern Europe: Migration and adaptation during the Neolithic and Bronze Age. *Acta Archaeologica* **85**, (2014).

57. Furholt, M. Die Złota-Gruppe in Kleinpolen: Ein Beispiel für die Transformation eines Zeichensystems? *Germania: Anzeiger der Römisch-Germanischen Kommission des Deutschen Archäologischen Instituts* **86**, 1–28 (2008).

58. Szmyt, M. In the Reaches Far of Two Worlds: On the Study of Contacts between the Societies of the Globular Amphora and Yamnaya Cultures. in *A Turning of Ages: Jubilee Book Dedicated to Professor Jan Machnik on His 70th Anniversary* (ed. Kadrow, S.) 443–466 (Institute of Archaeology and Ethnology, Polish Academy of Sciences, Cracow Branch, 2000).

59. Tassi, F. *et al.* Genome diversity in the Neolithic Globular Amphorae culture and the spread of Indo-European languages. *Proc. Biol. Sci.* **284**, (2017).

60. Nordqvist, K. & Heyd, V. The Forgotten Child of the Wider Corded Ware Family: Russian Fatyanovo Culture in Context. *Proceedings of the Prehistoric Society* **86**, 65–93 (2020).

61. Papac, L. *et al.* Dynamic changes in genomic and social structures in third millennium BCE central Europe. *Sci Adv* **7**, (2021).

62. Saag, L. *et al.* Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci Adv* **7**, (2021).

63. Johannsen, N. & Laursen, S. Routes and Wheeled Transport in Late 4th–Early 3rd Millennium Funerary Customs of the Jutland Peninsula: Regional Evidence and European Context. **85**, 15–58 (2010).

64. 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).

65. Hansen, J. *et al.* The Maglemosian skeleton from Koelbjerg revisited: Identifying sex and provenance. *Danish*

Journal of Archaeology **6**, 55–66 (2017).

66. Petersen, P. V. Chronological and regional variation in the Late Mesolithic of Eastern Denmark. *Journal of Danish Archaeology* **3**, 7–18 (1984).
67. Fischer, A. People and the sea – settlement and fishing along the Mesolithic coasts. in *The Danish Storebælt Since the Ice Age – Man, Sea and Forest* (eds. Pedersen, L., Fischer, A. & Aaby, B.) 63–77 (A/S Storebælt Fixed Link, 1997).
68. Sørensen, S. A. *The Kongemose Culture*. (University Press of Southern Denmark, 2017).
69. Schoeninger, M. J. & Moore, K. Bone stable isotope studies in archaeology. *Journal of World Prehistory* **6**, 247–296 (1992).
70. Hedges, R. E. M. & Reynard, L. M. Nitrogen isotopes and the trophic level of humans in archaeology. *J. Archaeol. Sci.* **34**, 1240–1251 (2007).
71. Fischer, A. *et al.* Coast-inland mobility and diet in the Danish Mesolithic and Neolithic: evidence from stable isotope values of humans and dogs. *J. Archaeol. Sci.* **34**, 2125–2150 (2007).
72. Tauber, H. 13C evidence for dietary habits of prehistoric man in Denmark. *Nature* **292**, 332–333 (1981).
73. Loog, L. *et al.* Estimating mobility using sparse data: Application to human genetic variation. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12213–12218 (2017).
74. Racimo, F. *et al.* The Spatiotemporal Spread of Human Migrations during the European Holocene. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8989–9000 (2020).
75. Worsaae, J. J. A. Om en ny deling af Steen- og Bronzealderen. Oversigten over det Kongelige Danske Videnskabernes Selskabs Forhandlinger (reprint). in *The Neolithisation of Denmark. 150 years of debate* (eds. Fischer, A. & Kristiansen, K.) 47–56 (J.R. Collis Publications, 1859).
76. Becker, C. J. Mosefundne lerkar fra yngre stenalder (Neolithic pottery in Danish bogs), Copenhagen. *Aarbøger for nordisk Oldkyndighed og Historie* (1947).
77. Andersen, S. H. 'Køkkenmøddinger' (shell middens) in Denmark: A survey. *Proc. Prehist. Soc.* **66**, 361–384 (2000).
78. Gutiérrez-Zugasti, I. *et al.* Shell midden research in Atlantic Europe: State of the art, research problems and perspectives for the future. *Quat. Int.* **239**, 70–85 (2011).
79. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
80. Cox, S. L., Ruff, C. B., Maier, R. M. & Mathieson, I. Genetic contributions to variation in human stature in

prehistoric Europe. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21484–21492 (2019).

81. Brinch Petersen, E. Gravene ved Dragsholm. Fra jægere til bønder for 6000 år siden. *Nationalmuseets arbejdsmark* **1974**, 112–120 (1974).

82. Price, T. D. *et al.* New information on the Stone Age graves at Dragsholm, Denmark. *Acta Archaeol.* **78**, 193–219 (2007).

83. Jensen, T. Z. T. *et al.* A 5700 year-old human genome and oral microbiome from chewed birch pitch. *Nat. Commun.* **10**, 5520 (2019).

84. Grasgruber, P., Sebera, M., Hrazdíra, E., Cacek, J. & Kalina, T. Major correlates of male height: A study of 105 countries. *Econ. Hum. Biol.* **21**, 172–195 (2016).

85. Sugita, S. Theory of quantitative reconstruction of vegetation II: all you need is LOVE. *Holocene* **17**, 243–257 (2007).

86. Nielsen, A. B. *et al.* Quantitative reconstructions of changes in regional openness in north-central Europe reveal new insights into old questions. *Quat. Sci. Rev.* **47**, 131–147 (2012).

87. Githumbi, E. *et al.* Pollen-based maps of past regional vegetation cover in Europe over twelve millennia - evaluation and potential. *Frontiers in Ecology and Evolution* **(Provisionally accepted)**, (2022).

88. Nielsen, A. B. & Odgaard, B. V. Quantitative landscape dynamics in Denmark through the last three millennia based on the Landscape Reconstruction Algorithm approach. *Veg. Hist. Archaeobot.* **19**, 375–387 (2010).

89. Søe, N. E., Odgaard, B. V., Nielsen, A. B., Olsen, J. & Kristiansen, S. M. Late Holocene landscape development around a Roman Iron Age mass grave, Alken Enge, Denmark. *Veg. Hist. Archaeobot.* **26**, 277–292 (2017).

90. Kristiansen, K., Melheim, L., Bech, J.-H., Mortensen, M. F. & Frei, K. M. Thy at the Crossroads: A Local Bronze Age Community's Role in a Macro-Economic System. in *Contrasts of the Nordic Bronze Age* 269–282 (Brepols Publishers, 2020).

91. Borzunov, V. A. The neolithic fortified settlements of the Western Siberia and Trans-Urals. *Rossijskaâ arheologiâ* 20–34 (2013).

92. Piezonka, H. *et al.* The emergence of hunter-gatherer pottery in the Urals and West Siberia: New dating and stable isotope evidence. *J. Archaeol. Sci.* **116**, 105100 (2020).

93. Damgaard, P. de B. *et al.* The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, (2018).

94. Yu, H. *et al.* Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia. *Cell* **181**, 1232–1245.e20 (2020).

95. Okladnikov, A. P. Neolit i bronzovyi vek Pribaikaliya [Neolithic and Bronze Age of the Baikal region]. Materialy i issledovaniya po arkheologii SSSR; №. 18 [Materials and researches on Archeology of the USSR; Vol. 18]. Moscow, Leningrad, AS USSR Publ., 1950, Part 1 and 2, 411 p. (1950).

96. Merz, V. On the origin of the complexes of the Baikal type in the Eneolithic of Kazakhstan. in *Paleodemography and migration processes in Western Siberia in antiquity and the Middle Ages* (ed. Kiryushin, Y. F.) 39–42 (Altai State University, 1994).

97. Merz, V. Periodization of the Holocene complexes of Northern and Central Kazakhstan based on the materials of the multilayer site Shiderty 3. (Kemerovo State University, 2008).

98. Damgaard, P. de B. *et al.* 137 ancient human genomes from across the Eurasian steppes. *Nature* **557**, 369–374 (2018).

99. Librado, P. *et al.* The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* **598**, 634–640 (2021).

100. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).

101. Margaryan, A. *et al.* Population genomics of the Viking world. *Nature* **585**, 390–396 (2020).

102. Hellenthal, G. *et al.* A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).

103. Sikora, M. *et al.* The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188 (2019).

104. Shinde, V. *et al.* An Ancient Harappan Genome Lacks Ancestry from Steppe Pastoralists or Iranian Farmers. *Cell* **179**, 729–735.e10 (2019).

105. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6886–6891 (2016).

106. Feldman, M. *et al.* Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).

107. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

108. Galinsky, K. J., Loh, P.-R., Mallick, S., Patterson, N. J. & Price, A. L. Population Structure of UK Biobank and Ancient Eurasians Reveals Adaptation at Genes Influencing Blood Pressure. *Am. J. Hum. Genet.* **99**, 1130–1139 (2016).

109. Patterson, N. *et al.* Large-scale migration into Britain during the Middle to Late Bronze Age. *Nature* (2021) doi:10.1038/s41586-021-04287-4.

110. 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).

111. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).

112. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).

113. Scheib, C. L. *et al.* Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024–1027 (2018).

114. Krzewińska, M. *et al.* Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci Adv* **4**, eaat4457 (2018).

115. Girirajan, S., Campbell, C. D. & Eichler, E. E. Human copy number variation and complex genetic disease. *Annu. Rev. Genet.* **45**, 203–226 (2011).

116. Weise, A. *et al.* Microdeletion and microduplication syndromes. *J. Histochem. Cytochem.* **60**, 346–358 (2012).

117. Girirajan, S. *et al.* Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N. Engl. J. Med.* **367**, 1321–1331 (2012).

118. Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).

119. Bergström, A. *et al.* Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, (2020).

120. Sudmant, P. H. *et al.* Diversity of human copy number variation and multicopy genes. *Science* **330**, 641–646 (2010).

121. Nagamani, S. C. S. *et al.* Phenotypic manifestations of copy number variation in chromosome 16p13.11. *Eur. J. Hum. Genet.* **19**, 280–286 (2011).

122. Cremin, B., Goodman, H., Spranger, J. & Beighton, P. Wormian bones in osteogenesis imperfecta and other disorders. *Skeletal Radiol.* **8**, 35–38 (1982).

123. Foreman, P. *et al.* L5 spondylolysis/spondylolisthesis: a comprehensive review with an anatomic focus. *Childs. Nerv. Syst.* **29**, 209–216 (2012).

124. Vleeming, A. *et al.* The sacroiliac joint: an overview of its anatomy, function and potential clinical implications. *J. Anat.* **221**, 537–567 (2012).

125. Salanitri, S. & Seow, W. K. Developmental enamel defects in the primary dentition: aetiology and clinical

management. *Aust. Dent. J.* **58**, 133–40; quiz 266 (2013).

126. Crawford, K. *et al.* Medical consequences of pathogenic CNVs in adults: analysis of the UK Biobank. *J. Med. Genet.* **56**, 131–138 (2019).

127. Stern, A. J., Wilton, P. R. & Nielsen, R. An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data. *PLoS Genet.* **15**, e1008384 (2019).

128. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).

129. Enattah, N. S. *et al.* Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *Am. J. Hum. Genet.* **82**, 57–72 (2008).

130. Itan, Y., Powell, A., Beaumont, M. A., Burger, J. & Thomas, M. G. The origins of lactase persistence in Europe. *PLoS Comput. Biol.* **5**, e1000491 (2009).

131. Ségurel, L. & Bon, C. On the Evolution of Lactase Persistence in Humans. *Annu. Rev. Genomics Hum. Genet.* **18**, 297–319 (2017).

132. Segurel, L. *et al.* Why and when was lactase persistence selected for? Insights from Central Asian herders and ancient DNA. *PLoS Biol.* **18**, e3000742 (2020).

133. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* **30**, 233–237 (2002).

134. Bersaglieri, T. *et al.* Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120 (2004).

135. Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013).

136. Gallois, A. *et al.* A comprehensive study of metabolite genetics reveals strong pleiotropy and heterogeneity across time and context. *Nat. Commun.* **10**, 4788 (2019).

137. Ligthart, S. *et al.* Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. *BMC Genomics* **17**, 443 (2016).

138. Buckley, M. T. *et al.* Selection in Europeans on Fatty Acid Desaturases Associated with Dietary Changes. *Mol. Biol. Evol.* **34**, 1307–1318 (2017).

139. Ye, K., Gao, F., Wang, D., Bar-Yosef, O. & Keinan, A. Dietary adaptation of FADS genes in Europe varied across time and geography. *Nat Ecol Evol* **1**, 167 (2017).

140. Mathieson, S. & Mathieson, I. FADS1 and the Timing of Human Adaptation to Agriculture. *Mol. Biol. Evol.* **35**, 2957–2970 (2018).

141. Lazaridis, I. The evolutionary history of human populations in Europe. *Curr. Opin. Genet. Dev.* **53**, 21–27 (2018).

142. Luu, K., Bazin, E. & Blum, M. G. B. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* **17**, 67–77 (2017).

143. Fairn, G. D. & McMaster, C. R. Emerging roles of the oxysterol-binding protein family in metabolism, transport, and signaling. *Cell. Mol. Life Sci.* **65**, 228–236 (2008).

144. Lehto, M. & Olkkonen, V. M. The OSBP-related proteins: a novel protein family involved in vesicle transport, cellular lipid metabolism, and cell signalling. *Biochim. Biophys. Acta* **1631**, 1–11 (2003).

145. Sánchez-Solana, B., Li, D.-Q. & Kumar, R. Cytosolic functions of MORC2 in lipogenesis and adipogenesis. *Biochim. Biophys. Acta* **1843**, 316–326 (2014).

146. Zhong, X. *et al.* The zinc-finger protein ZFYVE1 modulates TLR3-mediated signaling by facilitating TLR3 ligand binding. *Cell. Mol. Immunol.* **17**, 741–752 (2020).

147. Kim, S. V. *et al.* GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. *Science* **340**, 1456–1459 (2013).

148. Nguyen, L. P. *et al.* Role and species-specific expression of colon T cell homing receptor GPR15 in colitis. *Nature Immunology* vol. 16 207–213 (2015).

149. Monteleone, G., Boirivant, M., Pallone, F. & MacDonald, T. T. TGF-β1 and Smad7 in the regulation of IBD. *Mucosal Immunology* vol. 1 S50–S53 (2008).

150. Kennedy, B. W. C. Mongersen, an Oral SMAD7 Antisense Oligonucleotide, and Crohn's Disease. *The New England journal of medicine* vol. 372 2461 (2015).

151. Garo, L. P. *et al.* Smad7 Controls Immunoregulatory PDL2/1-PD1 Signaling in Intestinal Inflammation and Autoimmunity. *Cell Rep.* **28**, 3353–3366.e5 (2019).

152. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* **51**, 258–266 (2018).

153. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).

154. Jones, A. V. *et al.* GWAS of self-reported mosquito bite size, itch intensity and attractiveness to mosquitoes implicates immune-related predisposition loci. *Hum. Mol. Genet.* **26**, 1391–1406 (2017).

155. Gutierrez-Achury, J. *et al.* Functional implications of disease-specific variants in loci jointly associated with coeliac disease and rheumatoid arthritis. *Hum. Mol. Genet.* **25**, 180–190 (2016).

156. Stefansson, H. *et al.* A common inversion under selection in Europeans. *Nat. Genet.* **37**, 129–137 (2005).

157. Steinberg, K. M. *et al.* Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat. Genet.* **44**, 872–880 (2012).

158. Kılınç, G. M. *et al.* The Demographic Development of the First Farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).

159. Andreadis, A., Brown, W. M. & Kosik, K. S. Structure and novel exons of the human tau gene. *Biochemistry* **31**, 10626–10633 (1992).

160. Alonso, A. C., Grundke-Iqbal, I. & Iqbal, K. Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. *Nat. Med.* **2**, 783–787 (1996).

161. Desikan, R. S. *et al.* Genetic overlap between Alzheimer's disease and Parkinson's disease at the MAPT locus. *Mol. Psychiatry* **20**, 1588–1595 (2015).

162. Satake, W. *et al.* Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* **41**, 1303–1307 (2009).

163. Simón-Sánchez, J. *et al.* Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41**, 1308–1312 (2009).

164. Shulman, J. M. & De Jager, P. L. Evidence for a common pathway linking neurodegenerative diseases. *Nature genetics* vol. 41 1261–1262 (2009).

165. Aoki, K. Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited. *Ann. Hum. Biol.* **29**, 589–608 (2002).

166. Frost, P. The puzzle of European hair, eye, and skin color. *Advances in Anthropology* **2014**, (2014).

167. Lona-Durazo, F. *et al.* Meta-analysis of GWA studies provides new insights on the genetic architecture of skin pigmentation in recently admixed populations. *BMC Genet.* **20**, 59 (2019).

168. Ju, D. & Mathieson, I. The evolution of skin pigmentation-associated variation in West Eurasia. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

169. Jablonski, N. G. & Chaplin, G. The evolution of human skin coloration. *J. Hum. Evol.* **39**, 57–106 (2000).

170. Engelsen, O. The relationship between ultraviolet radiation exposure and vitamin D status. *Nutrients* **2**, 482–495 (2010).

171. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A Map of Recent Positive Selection in the Human Genome. *PLoS Biol.* **4**, e72 (2006).

172. Sabeti, P. C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).

173. Martin, A. R. *et al.* An Unexpectedly Complex Architecture for Skin Pigmentation in Africans. *Cell* **171**, 1340–1353.e14 (2017).

174. Wu, H. *et al.* Transcriptome Sequencing to Detect the Potential Role of Long Noncoding RNAs in Salt-Sensitive Hypertensive Rats. *Biomed Res. Int.* **2019**, 2816959 (2019).

175. Logsdon, B. A., Hoffman, G. E. & Mezey, J. G. Mouse obesity network reconstruction with a variational Bayes algorithm to employ aggressive false positive control. *BMC Bioinformatics* **13**, 53 (2012).

176. Wang, L. *et al.* Peakwide Mapping on Chromosome 3q13 Identifies the Kalirin Gene as a Novel Candidate Gene for Coronary Artery Disease. *The American Journal of Human Genetics* vol. 80 650–663 (2007).

177. Ikram, M. A., Seshadri, S. & Bis, J. C. Genomewide Association Studies of Stroke. *Journal of Vascular Surgery* vol. 50 467 (2009).

178. Krug, T. *et al.* Kalirin: a novel genetic risk factor for ischemic stroke. *Hum. Genet.* **127**, 513–523 (2010).

179. Zang, X.-L. *et al.* Association of a SNP in SLC35F3 Gene with the Risk of Hypertension in a Chinese Han Population. *Frontiers in Genetics* vol. 7 (2016).

180. Zhang, K. *et al.* Genetic implication of a novel thiamine transporter in human hypertension. *J. Am. Coll. Cardiol.* **63**, 1542–1555 (2014).

181. Russo, L. *et al.* Cholesterol 25-hydroxylase (CH25H) as a promoter of adipose tissue inflammation in obesity and diabetes. *Mol Metab* **39**, 100983 (2020).

182. Zhao, J., Chen, J., Li, M., Chen, M. & Sun, C. Multifaceted Functions of CH25H and 25HC to Modulate the Lipid Metabolism, Immune Responses, and Broadly Antiviral Activities. *Viruses* **12**, (2020).

183. Demir, A., Kahraman, R., Candan, G. & Ergen, A. The role of FAS gene variants in inflammatory bowel disease. *Turk. J. Gastroenterol.* **31**, 356–361 (2020).

184. Izawa, T. *et al.* ASXL2 Regulates Glucose, Lipid, and Skeletal Homeostasis. *Cell Rep.* **11**, 1625–1637 (2015).

185. Zou, W. *et al.* Myeloid-specific Asxl2 deletion limits diet-induced obesity by regulating energy expenditure. *J. Clin. Invest.* **130**, 2644–2656 (2020).

186. Park, U.-H., Yoon, S. K., Park, T., Kim, E.-J. & Um, S.-J. Additional Sex Comb-like (ASXL) Proteins 1 and 2 Play Opposite Roles in Adipogenesis via Reciprocal Regulation of Peroxisome Proliferator-activated Receptor γ. *Journal of Biological Chemistry* vol. 286 1354–1363 (2011).

187. Ponsuksili, S. *et al.* Epigenome-wide skeletal muscle DNA methylation profiles at the background of distinct metabolic types and ryanodine receptor variation in pigs. *BMC Genomics* **20**, 492 (2019).

188. Samad, M. B. *et al.* [6]-Gingerol, from Zingiber officinale, potentiates GLP-1 mediated glucose-stimulated insulin

secretion pathway in pancreatic β-cells and increases RAB8/RAB10-regulated membrane presentation of GLUT4 transporters in skeletal muscle to improve hyperglycemia in Leprdb/db type 2 diabetic mice. *BMC Complementary and Alternative Medicine* vol. 17 (2017).

189. Vazirani, R. P. *et al.* Disruption of Adipose Rab10-Dependent Insulin Signaling Causes Hepatic Insulin Resistance. *Diabetes* **65**, 1577–1589 (2016).

190. Hsieh, P. *et al.* Exome Sequencing Provides Evidence of Polygenic Adaptation to a Fat-Rich Animal Diet in Indigenous Siberian Populations. *Mol. Biol. Evol.* **34**, 2913–2926 (2017).

191. Thapa, D. *et al.* The protein acetylase GCN5L1 modulates hepatic fatty acid oxidation activity via acetylation of the mitochondrial β-oxidation enzyme HADHA. *J. Biol. Chem.* **293**, 17676–17684 (2018).

192. Baloni, P. *et al.* Genome-scale metabolic model of the rat liver predicts effects of diet restriction. *Sci. Rep.* **9**, 9807 (2019).

193. Ong, H. S. & Yim, H. C. H. Microbial Factors in Inflammatory Diseases and Cancers. *Regulation of Inflammatory Signaling in Health and Disease* 153–174 (2017) doi:10.1007/978-981-10-5987-2_7.

194. Berg, J. J. & Coop, G. A population genetic signal of polygenic adaptation. *PLoS Genet.* **10**, e1004412 (2014).

195. Berg, J. J. *et al.* Reduced signal for polygenic adaptation of height in UK Biobank. *Elife* **8**, (2019).

196. Sohail, M. *et al.* Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife* **8**, (2019).

197. Irving-Pease, E. K., Muktupavela, R., Dannemann, M. & Racimo, F. Quantitative Human Paleogenetics: What can Ancient DNA Tell us About Complex Trait Evolution? *Front. Genet.* **12**, 703541 (2021).

198. Refoyo-Martínez, A. *et al.* How robust are cross-population signatures of polygenic adaptation in humans? *bioRxiv* 2020.07.13.200030 (2021) doi:10.1101/2020.07.13.200030.

199. Ruff, C. B. *Skeletal Variation and Adaptation in Europeans: Upper Paleolithic to the Twentieth Century*. (John Wiley & Sons, 2017).

200. Tanigawa, Y. *et al.* Components of genetic associations across 2,138 phenotypes in the UK Biobank highlight adipocyte biology. *Nat. Commun.* **10**, 4064 (2019).

201. Corder, E. H. *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921–923 (1993).

202. Strittmatter, W. J. *et al.* Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1977–1981 (1993).

203. Rosenstock, E. *et al.* Human stature in the Near East and Europe ca. 10,000–1000 BC: its spatiotemporal

development in a Bayesian errors-in-variables model. *Archaeol. Anthropol. Sci.* **11**, 5657–5690 (2019).

204. Wilde, S. *et al.* Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4832–4837 (2014).

205. González-Fortes, G. *et al.* Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Curr. Biol.* **27**, 1801–1810.e10 (2017).

206. Belloy, M. E., Napolioni, V. & Greicius, M. D. A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. *Neuron* **101**, 820–838 (2019).

Acknowledgements

This publication, the culmination of a research effort lasting over a decade, is dedicated to the memory of Pia Bennike, who was part of a small core team that initiated the project. Sadly Pia passed away in 2017 but this study, and many others, testifies to her tremendous efforts and knowledge on Danish prehistoric skeletal material.

We are deeply indebted to former and present staff members at the National Museum, the Anthropological Laboratory, the regional museums and citizen scientists of Denmark, who for many generations carefully collected, recorded and curated the prehistoric skeletal remains that form a key component of this study. We are equally thankful to curators of the many other institutions across major parts of Eurasia, who to the benefit of following generations curated human skeletal remains and gave us access and permission to sample this precious material. We also thank all the former and current staff at the Lundbeck Foundation GeoGenetics Centre and the GeoGenetics Sequencing Core, and to colleagues across the many institutions detailed below. We are particularly grateful to Line Olsen as project manager for the Lundbeck Foundation GeoGenetics Centre project, and to Pernille Selmer Olsen for assisting with sample processing. We thank UK Biobank Ltd. for access to the UK Biobank genomic resource. We are thankful to Illumina Inc. for collaboration and to L. Speidel for assistance in running *Relate*. EW thanks St. John's College, Cambridge, for providing a stimulating environment of discussion and learning.

The Lundbeck Foundation GeoGenetics Centre is supported by the the Lundbeck Foundation (R302-2018-2155, R155-2013-16338), the Novo Nordisk Foundation (NNF18SA0035006), the Wellcome Trust (UNS69906), Carlsberg Foundation (CF18-0024), the Danish National Research Foundation (44113220) and the University of Copenhagen (KU2016 programme). 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 KK. 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); AP is funded by Wellcome grant WT214300, B.S.d.M and O.D. by the Swiss National Science Foundation (SFNS PP00P3_176977) and European Research Council (ERC 679330); M.N. by the Human Frontier Science Program Postdoctoral Fellowship (LT000143/2019-L4); 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 (gNNF20OC0062491); N.N.J. by Aarhus University Research Foundation; 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.V.L. and I.V.S. by the Science Committee, Ministry of Education and Science of the Republic of Kazakhstan (AP08856317); B.G.R. and MGM 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 trough 30 thousand Years' and '2008 prot. 2008B4J2HS_001 Origin and diffusion of farming in central-southern Italy: a molecular approach'); 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/2019/061) 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). L.Y. acknowledges funding by the Science Committee of the Armenian Ministry of Education and Science (Project 21AG-1F025), L.V. by 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. by the Russian Ministry of Science and Higher Education (075-15-2020-910); 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); R.D. by the Wellcome Trust (WT214300); R.N. by the National Institute of General Medical Sciences (NIH grant R01GM138634); F. Racimo by a Villum Fonden Young Investigator Grant (no. 00025300). T.W. and V.A. are supported by the Lundbeck Foundation iPSYCH initiative (R248-2017-2003).

Author Information

These authors contributed equally: Morten E. Allentoft, Martin Sikora, Alba Refoyo-Martínez, Evan K. Irving-Pease, Anders Fischer, William Barrie & Andrés Ingason

These authors equally supervised research: Thorfinn Korneliussen, Richard Durbin, Rasmus Nielsen, Olivier Delaneau, Thomas Werge, Fernando Racimo, Kristian Kristiansen & Eske Willerslev

Affiliations

Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
Morten E. Allentoft, Martin Sikora, Alba Refoyo-Martínez, Evan K. Irving-Pease, Andrés Ingason, Jesper Stenderup, Fabrice Demeter, Maria Novosolov, Ruairidh Macleod, Rasmus A. Henriksen, Tharsika Vimala, Hugh McColl, Lasse Vinner, Gabriele Scorrano, Abigail Ramsøe, Anders Rosengren, Anthony Ruter, Anne Birgitte Gotfredsen, Charleen Gaunitz, Fulya Eylem Yediay, Isin Altinkaya, Lei Zhao, Peter de Barros Damgaard, Kurt H. Kjær, Thorfinn Korneliussen, Rasmus Nielsen, Thomas Werge, Fernando Racimo, Kristian Kristiansen & Eske Willerslev

Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Perth, Australia

Morten E. Allentoft

GeoGenetics Group, Department of Zoology, University of Cambridge, Cambridge, UK
William Barrie, Alice Pearson, Ruairidh Macleod & Eske Willerslev

Department of Historical Studies, University of Gothenburg, Gothenburg, Sweden
Anders Fischer, Karl-Göran Sjögren, Bettina Schulz Paulsson, Malou Blank & Kristian Kristiansen

Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Roskilde, Denmark
Andrés Ingason, Vivek Appadurai,

The National Museum of Denmark, Ny Vestergade 10, Copenhagen, Denmark
Lasse Sørensen, Poul Otto Nielsen, Morten Fischer Mortensen, Peter Rasmussen, Peter Vang Petersen,

Sealand Archaeology, Gl. Roesnaesvej 27, 4400 Kalundborg, Denmark
Anders Fischer

Department of Genetics, University of Cambridge, Cambridge, UK
Alice Pearson & Richard Durbin

Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK
Richard Durbin & Eske Willerslev

Department of Computational Biology, University of Lausanne, Switzerland
Bárbara Sousa da Mota & Olivier Delaneau

Swiss Institute of Bioinformatics, University of Lausanne, Switzerland
Bárbara Sousa da Mota

Department of Integrative Biology, University of California, Berkeley, USA
Alma S. Halgren, Peter H. Sudmant & Rasmus Nielsen

Center for Computational Biology, University of California, Berkeley, USA
Andrew Vaughn, Aaron J. Stern, Peter H. Sudmant & Rasmus Nielsen

Laboratory of Biological Anthropology, Department of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark
Marie Louise Schjellerup Jørkov & Niels Lynnerup

Muséum National d'Histoire Naturelle, Paris, France
Fabrice Demeter

Research department of Genetics, Evolution and Environment, University College London, London, UK
Ruairidh Macleod

Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
Ashot Margaryan, Theis Zetner Trolle Jensen, Hannes Schroeder, Enrico Cappellini

Centre for Evolutionary Hologenomics, University of Copenhagen, Copenhagen, Denmark
Ashot Margaryan

Anthropology Department, University of Utah, USA
Melissa Illardo

Department of Geology, Lund University, Lund, Sweden
Anne Birgitte Nielsen

Department of Archaeology and Ancient History, Lund University, Lund, Sweden
Lars Larsson

Tårnby Gymnasium og HF, Kastrup, Denmark
Mikkel Ulfeldt Hede

Department of Health Technology, Section of Bioinformatics, Technical University of Denmark, Kongens Lyngby, Denmark
Gabriel Renaud

Department of Archaeology and Heritage Studies, Aarhus University, Aarhus, Denmark
Niels N. Johannsen, Rikke Maring

Institute of Biological Psychiatry, Mental Health Centre Sct Hans, Copenhagen University Hospital, Denmark
Andrew J. Schork, Anders Rosengren & Thomas Werge

Department of Clinical Medicine, University of Copenhagen, 2200 Copenhagen N, Denmark
Thomas Werge

Neurogenomics Division, The Translational Genomics Research Institute (TGEN), Phoenix, AZ, USA
Andrew J. Schork

Terra Ltd., Letchik Zlobin St. 20, Voronezh, 394055, Russian Federation
Andrey Skorobogatov, Ruslan Turin

Department of Archaeology, University of Exeter, Exeter, UK
Alan K. Outram & Catriona J. McKenzie

Institute of Archaeology and Ethnography, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation
Aleksey A. Timoshchenko, Dmitri V. Pozdnyakov, Liudmila N. Mylnikova, Marina S. Nesterova & Vyacheslav I. Molodin

Institute of Ethnology and Anthropology, Russian Academy of Sciences, Moscow, Russian Federation
Elizaveta V. Veselovskaya, Sergey Vasilyev

Centre for Egyptological studies, Russian Academy of Sciences, Moscow, Russia
Sergey Vasilyev

Research Institute and Museum of Anthropology, Lomonosov Moscow State University, Mokhovaya str. 11, Moscow, Russian Federation
Alexandra Buzhilova, Natalia Berezina, Svetlana Borutskaya

Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
Alfredo Coppa, Dušan Borić

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint Petersburg, Russian Federation
Alisa Zubova & Vyacheslav Moiseyev

CIAS, Department of Life Science, University of Coimbra, Coimbra, Portugal
Ana Maria Silva

UNIARQ, University of Lisbon, Lisbon, Portugal
Ana Maria Silva & João Zilhão

Kostanay Regional University A. Baitursynov, Kostanay, Kazakhstan
Andrey V. Logvin, Irina V. Shevnina

Vesthimmerlands Museum, Søndergade 44, Aars, Denmark
Bjarne Henning Nielsen

Museum Nordsjælland, Frederiksgade 9, 3400 Hillerød
Per Lotz, Søren Anker Sørensen, Thomas Jørgensen

Museum Vestsjælland, Klosterstræde 18, 4300 Holbæk, Denmark
Per Lotz

Vendsyssel Historiske Museum, DK-9800 Hjørring, Denmark
Per Lysdahl, Sidsel Wåhlin

Moesgaard Museum, Moesgård Allé 15, Højbjerg, Denmark
Søren H. Andersen

Grupo EvoAdapta, Departamento de Ciencias Históricas, Universidad de Cantabria, Santander, Spain
Borja González-Rabanal

Museu de Ciències Naturals de Barcelona, Barcelona, Spain
Carles Lalueza-Fox

Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain
Carles Lalueza-Fox

ICREA, University of Barcelona, Barcelona, Spain
João Zilhão

Departamento de Prehistoria y Arqueología Department, Universidad Autónoma de Madrid, Madrid, Spain
Concepción Blasco, Corina Liesau, Patricia Ríos

Department of Biology, University of Rome "Tor Vergata", Rome, Italy
Cristina Martinez-Labarga, Olga Rickards

Department of History, Humanities and Society, University of Rome "Tor Vergata", Rome, Italy
Mario Federico Rolfo

Instituto Internacional de Investigaciones Prehistóricas de Cantabria, Universidad de Cantabria, Santander, Spain
David Cuenca-Solana, Igor Gutiérrez Zugasti & Manuel González-Morales

Centre de Recherche en Archéologie, Archeosciences, Histoire (CReAAH), UMR-6566 CNRS, Rennes, France
David Cuenca-Solana

Georgian National Museum, Tbilisi, Georgia
David O. Lordkipanidze

Tbilisi State University, Tbilisi, Georgia
David O. Lordkipanidze

IPND, Tyumen Scientific Centre, Siberian Branch of the Russian Academy of Sciences, Tyumen, Russian Federation
Dmitri Enshin, Olga Poshekhonova, Svetlana N. Skochina

Departament de Prehistòria, Arqueologia i Història Antiga, Universitat de València, València, Spain
Domingo C. Salazar-García, J.Emili Aura Tortosa

Department of Geological Sciences, University of Cape Town, Cape Town, South Africa
Domingo C. Salazar-García

Department of Anthropology, New York University, New York, USA.
Dušan Borić

Institute of Humanities, Ivanovo State University, Ivanovo, Russian Federation
Elena Kostyleva

Saryarka Archaeological Institute, Buketov Karaganda University, Karaganda, Kazakhstan

Emma R. Usmanova

South Ural State University, Chelyabinsk, Russia
Emma R. Usmanova

The Saxo Institute, University of Copenhagen, Copenhagen, Denmark
Erik Brinch Petersen, Rune Iversen

Museum Østjylland, Stemannsgade 2, Randers, Denmark
Esben Kannegaard, Lutz Klassen

Soprintendenza Archeologia Belle Arti e Paesaggio per la Città Metropolitana di Bari, Via Pier l'Eremita, 25, 70122, Bari, Italy
Francesca Radina

Soprintendenza per i Beni Archeologici delle Marche, Via Birarelli 18, 60100, Ancona, Italy
Mara Silvestrini

Soprintendenza Archeologia, Belle Arti e Paesaggio per la provincia di Cosenza, Cosenza, Italy
Paola Aurino

UMR 5199 PACEA, CNRS, Université de Bordeaux, 33615 Pessac, France
Henri Duday, Patrice Courtaud

Institute of Archaeology, National Academy of Sciences of Ukraine, Kyiv, Ukraine
Inna Potekhina

National University of Kyiv-Mohyla Academy, Kyiv, Ukraine
Inna Potekhina

Collège de France, 75231 Paris cedex 05, France
Jean Guilaine

Odense City Museums, Overgade 48, Odense, Denmark
Jesper Hansen

Museum Sydøstdanmark, Algade 97, 4760 Vordingborg, Denmark
Kristoffer Buck Pedersen

Institute of Archaeology and Ethnology, Polish Academy of Sciences, Kraków, Poland
Krzysztof Tunia, Piotr Włodarczak

CNRS UMR 5608, Toulouse Jean Jaurès University, Maison de la Recherche, 5 Allées Antonio Machado, 31058 Toulouse, Cedex 9, France
Laure Metz & Ludovic Slimak

ARGEA Consultores SL, C. de San Crispín, Madrid, Spain

Jorge Vega, Roberto Menduiña

Institute of Molecular Biology, National Academy of Sciences, Yerevan, Armenia
Levon Yepiskoposyan

Russian-Armenian University, Yerevan, Armenia
Levon Yepiskoposyan

Institute of Archaeology and Ethnography, National Academy of Sciences, Yerevan, Armenia
Ruben Badalyan

HistorieUdvikler, Gl. Roesnaesvej 27, DK-4400 Kalundborg, Denmark
Lisbeth Pedersen

Department of history and cultural heritage, University of Siena, Siena, Italy
Lucia Sarti, Mauro Calattini

Centre d'Anthropobiologie et de Génomique de Toulouse, CNRS UMR 5288, Université Paul Sabatier, Toulouse, France
Ludovic Orlando

Västergötlands museum, Stadsträdgården, Skara, Sweden
Maria Vretemark

Cabinet of Anthropology, Tomsk State University, Tomsk, Russian Federation
Marina P. Rykun

Institute for Eastern Research, Adam Mickiewicz University in Poznań, Poznań, Poland
Marzena H. Szmyt

Institute of Archaeology, Jagiellonian University, Ul. Gołębia 11, 31-007, Kraków, Poland
Marcin Przybyła

Zoological Institute of Russian Academy of Sciences, Universitetskaya nab. 1, 199034, St. Petersburg, Russian Federation
Mikhail Sablin

Department of Anthropology, Czech National Museum, Prague, Czech Republic
Miluše Dobisíková

Department of Health and Nature, University of Greenland, Greenland
Morten Meldgaard

The Viking Ship Museum, Vindeboder 12, Roskilde, Denmark
Morten Johansen, Otto Christian Uldum

Archaeology Institute, University of Highlands and Islands, Scotland, UK

Nick Card

Laboratory for Experimental Traceology, Institute for the History of Material Culture of the Russian Academy of Sciences, Dvortsovaya nab., 18, 191186, St. Petersburg, Russian Federation
Olga V. Lozovskaya

Scientific Research Center "Baikal region", Irkutsk State University; 1, K. Marx st., Irkutsk, 664003, Russian Federation
Nikolai A. Saveliev

Paleoecology Laboratory, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russian Federation
Pavel Kosintsev

Department of History of the Institute of Humanities, Ural Federal University, Ekaterinburg, Russian Federation
Pavel Kosintsev

Centre for the Study of Early Agricultural Societies, Department of Cross-Cultural and Regional Studies, University of Copenhagen, 2300 Copenhagen, Denmark
Peder Mortensen

Museum of Cultural History, University of Oslo, P.O. Box 6762. St. Olavs Plass NO-0130 Oslo, Norway
Per Persson

ArchaeoScience, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
Pernille Bangsgaard

Department of History, University of Santiago de Compostela, Spain
Pilar Prieto Martinez

Lipetsk Regional Scientific Public Organisation "Archaeological Research", Lipetsk, Russian Federation
Roman V. Smolyaninov

Laboratory for Archaeological Chemistry, Department of Anthropology, University of Wisconsin-Madison, Madison, USA
T. Douglas Price

Nizhny Tagil State Socio-Pedagogical Institute, Nizhny Tagil, Russia
Yuri B. Serikov

Institute for History of Medicine, First Faculty of Medicine, Charles University, Prague, Czech Republic
Vaclav Smrcka

Centre for Archaeological Research Toraighyrov University, Pavlodar, Kazakhstan
Victor Merz

Malmö Museer, Malmöhusvägen 6, Malmö, Sweden
Yvonne Magnusson

Institute of Statistical Sciences, School of Mathematics, University of Bristol, Bristol, UK
Daniel J. Lawson

Novo Nordisk Foundation Centre for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen N, Denmark
Simon Rasmussen

MARUM, University of Bremen, Bremen, Germany
Eske Willerslev

Contributions

M.E.A., M.S., A.R.-M., E.K.I.-P., A.F., W.B., and A.I. contributed equally to this work. 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., C.L.-F., R.N., T.W., K.K. and E.W. conceptualised 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.K. and E.W. acquired funding for research. 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.V.L., B.H.N., B.G.R, C.B., C.L., C.M-L., D.V., D.C.-S., D.L., D.N., D.C.S.-G., D.B., E.K., E.V.V., E.R.U., E. Kannegaard, F. Radina, H.D., I.G.Z., I.P., I.V.S., J.G., J.H., J.E.A.T., J.Z., J.V., K.B.P., K.T., L.N., L.L., L.M., L.Y., L.P., L. Sarti, L. Slimak, L.K., M.G.M., M. Silvestrini, M.V., M.S.N., M.P.R., M.H.S., M.P., M.C., M. Sablin, N.C., O.P., O.R., O.V.L., P.A., P.K., P.C., P. Ríos, P. Lotz, P. Lysdahl, P.P., P.B., P.d.B.D., P.V.P., P.P.M., P.W., R.V.S., R. Maring, R. Menduiña, R.B., R.T., S.V., S.W., S.B., S.N.S., S.A.S., S.H.A., T.D.P., T.J., Y.B.S., V.I.M., V.S., V.M, Y.M. 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., L.V., A.J. Stern, C.G., F.E.Y, D.J.L., T.S.K., R.D., R.N., O.D., F. Racimo, K.K. and E.W. were involved in developing and applying methodology. J.S., C.G. and L.V. led the DNA laboratory work research component. K.G.S. 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.S.H., R.A.H., T.V., H.M., A.M., A.V., A.B.N., P. Rasmussen, G.R., A. Ramsøe, A.S., 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., D.J.L., P.H.S., T.S.K., and F. Racimo 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.S.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.S., A. Rosengren, A.K.O., A.B., A.C., A.G., A.V.L., A.B.G., C.J.M., D.C.S.-G., E. Kostyleva, E.R.U., E. Kannegaard, I.G.Z., I.P., I.V.S., J.G., J.H., J.E.A.T., L.Z, L.Y., L.P., L.K., M.B., M.G.M., M.V., M.P.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.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.G.S., A.S.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.S., A.J. Schork, A. Ruter, A.K.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., 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., A.F., K.G.S., F.D., R. Macleod, H.M. and T.V. led this). All co-authors read, commented on, and agreed upon the submitted manuscript.

Corresponding authors

Correspondence to Morten E. Allentoft (morten.allentoft@curtin.edu.au), Martin Sikora (martin.sikora@sund.ku.dk), Eske Willerslev (ew482@cam.ac.uk).

Ethics declarations

Competing interests

The authors declare no competing interests.