Current Biology, Volume *27*

Supplemental Information

The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers

Eppie R. Jones, Gunita Zarina, Vyacheslav Moiseyev, Emma Lightfoot, Philip R. Nigst, Andrea Manica, Ron Pinhasi, and Daniel G. Bradley

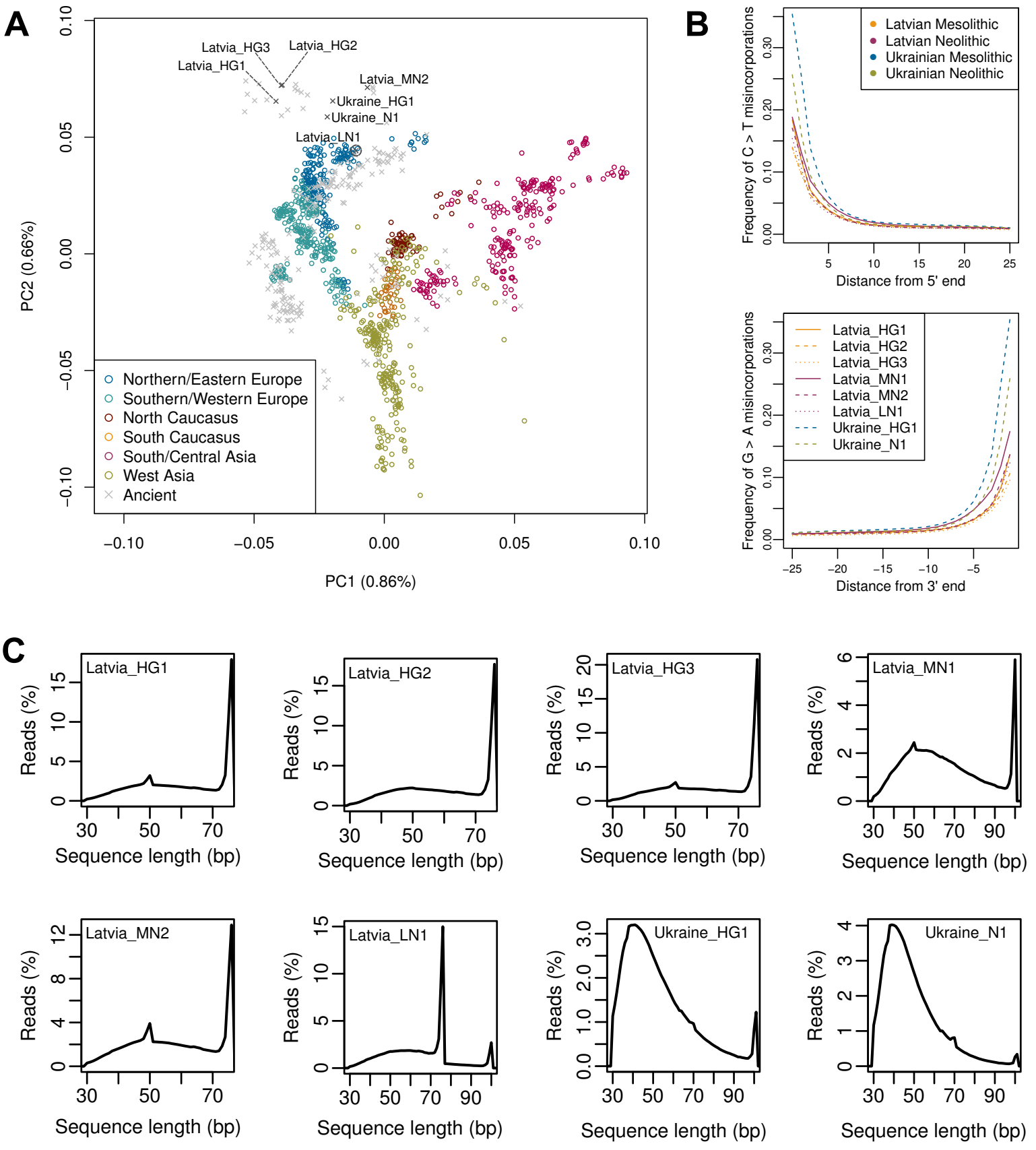

Figure S1. PCA and patterns of molecular damage. Related to Figure 2. A. PCA with modern individuals shown in colour and ancient individuals shown in grey. Datapoints are the same as in Figure 2A. B. Damage patterns for ancient samples. Plots show mismatch frequency relative to the reference genome as a function of read position for Latvian and Ukrainian samples. C. Sequence length distribution plots for ancient samples. Some samples have blips in their distributions as they were sequenced with more than one cycling format.

Figure S2. ADMIXTURE analysis showing modern individuals only. Related to Figure 2. 2-20 clusters (K) are shown.

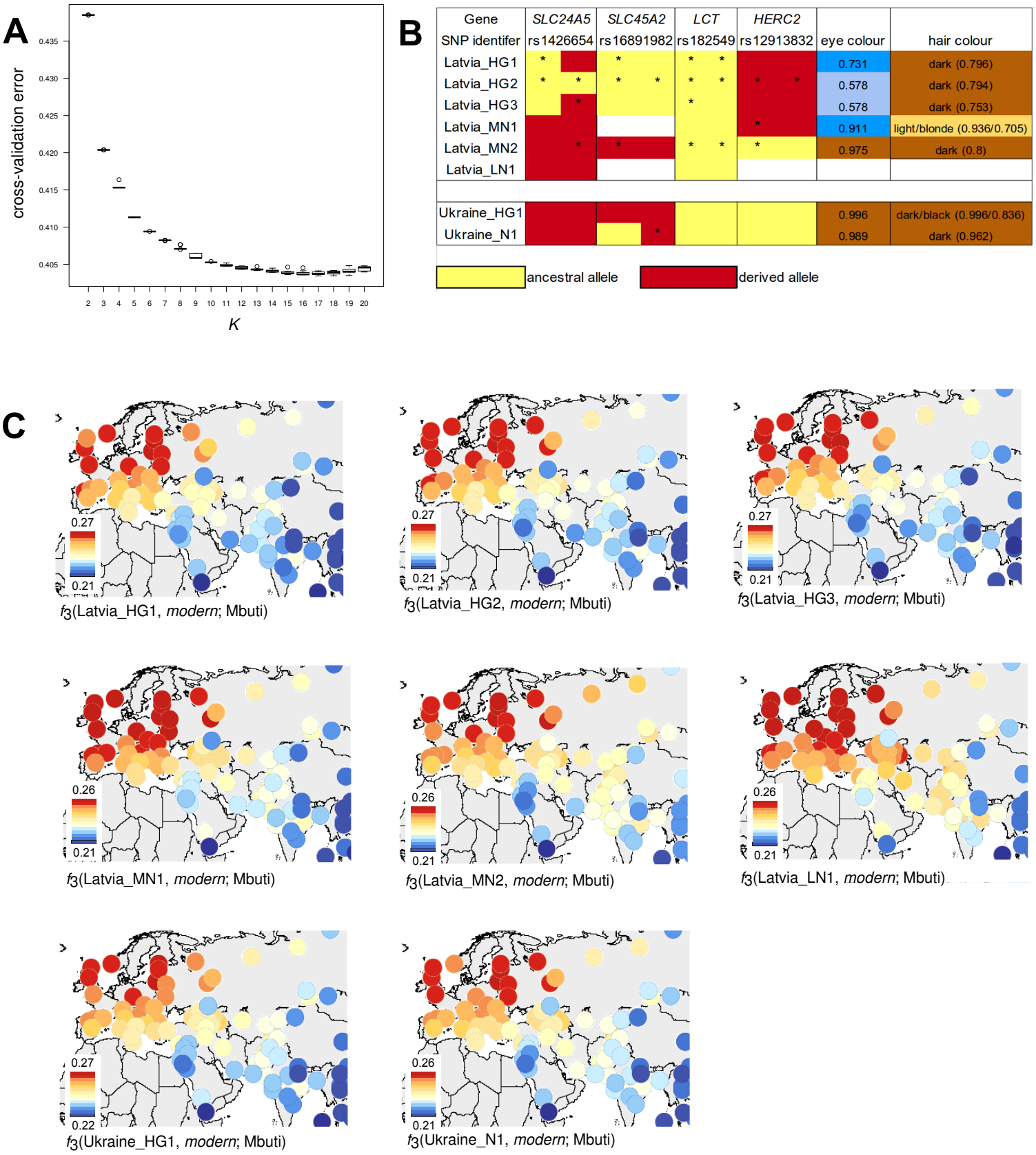

Figure S3. Cross-validation error for ADMIXTURE analysis, imputed genotypes and outgroup f_3-statistics. Related to Figure 2. A. ADMIXTURE analysis cross validation error as a function of the number of clusters (*K*). 10 replicates were performed for each value of *K*. The minimal error was found at *K*=17, but the error already started plateauing from roughly *K*=10, implying little improvement from this value onwards. B. Selected imputed genotypes (probability ≥ 0.7) for ancient Latvian and Ukrainian samples along with predicted hair and eye colour. The derived alleles in the *SLC24A5* and *SLC45A2* genes are associated with skin de-pigmenation [S1,S2] and *HERC2* with light iris colour [S3,S4]. The derived allele of the *LCT* gene is associated with the ability to digest lactose into adulthood [S5]. Hair and eye colour were predicted using the Hirisplex model [S6]. Stars show imputed genotypes which are supported by observed alleles. Two stars show that the genotype is supported by at least 2 reads, one star shows that the genotype is supported by one read. C. Outgroup f_3-statistics for Latvian and Ukrainian ancient samples. Statistics show that all samples share the most affinity to modern populations from Northern and Eastern Europe.

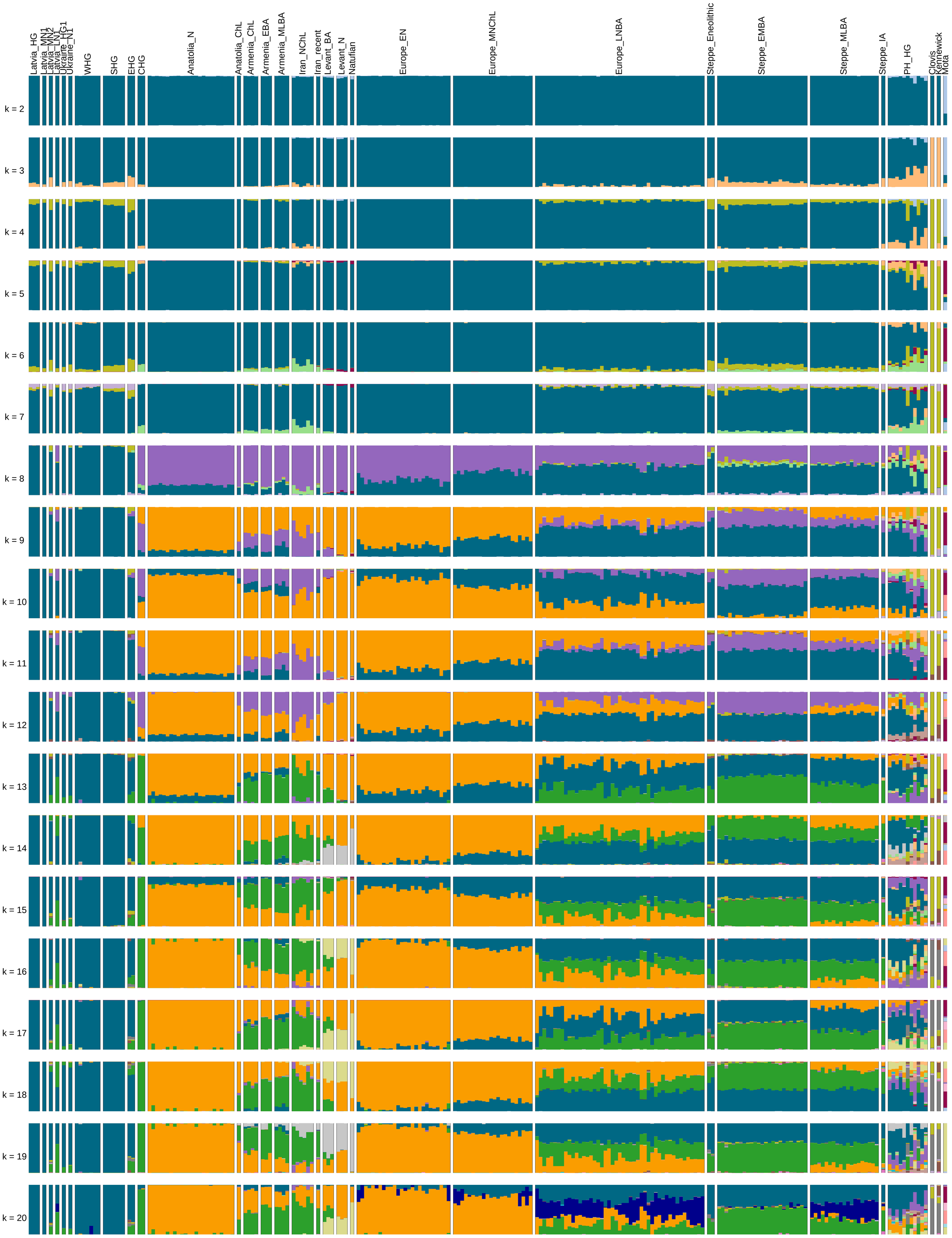

Figure S4. ADMIXTURE analysis showing ancient individuals only. Related to Figure 2. 2-20 clusters (K) are shown.

Table S1. Imputed genotypes probabilities for the SNP panel used in the Hirisplex prediction system.

				Latvia_HG1		Latvia_HG2		Latvia_HG3		Latvia_MN1		Latvia_MN2		Latvia_LN1		Ukraine_HG1		Ukraine_N1	
Gene	**SNP identifier**	**allele 0**	**allele 1**	**geno**	**P**	**geno**	**P**	**geno**	**P**	**geno**	**P**	**geno**	**P**	**geno**	**P**	**geno**	**P**	**geno**	**P**
TYR	rs1042602	C	A	0/0	0.932	0/0	1	0/0	1	0/0	0.821	0/0	0.997	0/0	0.989	0/0	0.992	0/0	1
MC1R	rs1110400	T	C	0/0	1	0/0	1	0/0	1	0/0	1	0/0	1	0/0	1	0/0	0.999	0/0	1
MC1R	rs11547464	G	A	0/0	0.999	0/0	0.999	0/0	0.98	0/0	0.982	0/0	1	0/0	0.983	0/0	0.994	0/0	0.985
IRF4	rs12203592	C	T	0/1	0.529	0/1	0.616	0/1	0.531	0/0	0.797	0/0	0.993	0/0	0.975	0/0	0.853	0/0	0.646
KITLG	rs12821256	T	C	0/0	0.997	0/0	1	0/0	0.999	0/0	0.984	0/0	1	0/0	0.969	0/0	0.992	0/0	0.934
SLC24A4	rs12896399	G	T	0/0	1	0/1	0.984	0/1	1	0/1	0.995	0/0	1	0/0	0.664	0/1	0.859	0/0	0.778
HERC2	rs12913832	A	G	1/1	0.848	1/1	0.995	1/1	0.773	1/1	0.761	0/0	1	0/1	0.523	0/0	0.742	0/0	0.969
TYR	rs1393350	G	A	0/0	1	0/0	1	0/0	1	0/0	0.997	0/0	1	0/0	1	0/0	1	0/0	0.988
SLC45A2	rs16891982	C	G	0/0	0.999	0/0	0.999	0/0	0.883	0/1	0.643	1/1	0.997	0/0	0.528	0/0	0.964	0/1	1
OCA2	rs1800407	C	T	0/0	1	0/0	1	0/0	1	0/0	0.998	0/0	1	0/0	0.993	0/1	0.528	0/0	0.998
MC1R	rs1805005	G	T	0/0	0.963	0/0	0.976	0/0	0.814	0/0	0.794	0/0	0.997	0/0	0.875	0/0	0.869	0/0	0.676
MC1R	rs1805006	C	A	0/0	1	0/0	1	0/0	1	0/0	0.998	0/0	1	0/0	1	0/0	0.972	0/0	1
MC1R	rs1805007	C	T	0/0	0.993	0/0	0.999	0/0	0.972	0/0	0.974	0/0	1	0/1	0.53	0/0	0.966	0/0	0.979
MC1R	rs1805008	C	T	0/0	1	0/0	1	0/0	1	0/0	0.99	0/0	0.997	0/0	0.999	0/0	0.985	0/0	0.997
MCIR	rs1805009	G	C	0/0	1	0/0	1	0/0	1	0/0	0.986	0/0	0.995	0/0	0.998	0/0	0.987	0/0	0.987
MC1R	rs2228479	G	A	0/0	1	0/0	1	0/0	1	0/0	0.974	0/0	0.999	0/0	1	0/0	0.996	0/0	0.991
PIGU/ASIP	rs2378249	G	A	1/1	1	1/1	1	1/1	1	1/1	0.999	1/1	1	1/1	0.997	1/1	0.999	0/1	0.995
SLC24A4	rs2402130	G	A	1/1	1	1/1	0.999	1/1	1	1/1	0.795	1/1	1	0/1	0.98	0/1	0.871	1/1	0.981
SLC45A2	rs28777	C	A	0/0	0.997	0/0	0.951	0/0	0.842	1/1	0.538	1/1	1	0/1	0.626	0/0	0.957	0/1	0.998
EXOC2	rs4959270	C	A	0/1	0.999	0/1	1	1/1	0.998	0/0	0.563	0/0	1	0/1	0.99	0/1	0.702	1/1	0.951
TYRP1	rs683	C	A	0/1	1	0/1	1	1/1	1	0/1	0.999	1/1	1	1/1	0.989	0/0	1	0/1	0.997
MC1R	rs86insA	C	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC1R	rs885479	G	A	0/0	1	0/0	1	0/0	1	0/0	0.962	0/1	0.703	0/0	0.969	0/0	0.929	0/0	0.891
MC1R	Y152OCH	C	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

geno, genotype; P, probability

Sample details are shown in Table 1.

Table S2. Key *D*-statistics of the form *D*(*A*,*B*; *X*,*Y*) for Ukrainian samples.

A	B	X	Y	D	Z-score	Loci
The Ukrainian samples form a clade to the exclusion of other tested samples.						
Mbuti	***Loschbour (WHG)***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0453*	*-2.298*	*133784*
Mbuti	***Latvia_LN1***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0703*	*-1.526*	*25338*
Mbuti	***BR2***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0264*	*-1.357*	*133824*
Mbuti	***Kotias (CHG)***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0258*	*-1.328*	*133744*
Mbuti	***Bichon (WHG)***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0238*	*-1.144*	*132164*
Mbuti	*Sardinian*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0205*	*-0.997*	*132316*
Mbuti	*Dinka*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0206*	*-0.926*	*133142*
Mbuti	*Han*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0173*	*-0.807*	*132983*
Mbuti	***Latvia_HG***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0125*	*-0.717*	*133260*
Mbuti	***Stuttgart***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0140*	*-0.701*	*133719*
Mbuti	*Karitiana*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0122*	*-0.567*	*132697*
Mbuti	*Mandenka*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0113*	*-0.534*	*132062*
Mbuti	*French*	***Ukraine_HG1***	***Ukraine_N1***	*-0.0107*	*-0.513*	*133120*
Mbuti	***Latvia_MN1***	***Ukraine_HG1***	***Ukraine_N1***	*-0.0160*	*-0.370*	*21119*
Mbuti	*Yoruba*	***Ukraine_HG1***	***Ukraine_N1***	*0.0005*	*0.022*	*133213*
Mbuti	*Dai*	***Ukraine_HG1***	***Ukraine_N1***	*0.0028*	*0.121*	*132618*
Mbuti	***NE1***	***Ukraine_HG1***	***Ukraine_N1***	*0.0033*	*0.164*	*133836*
Mbuti	***Karelia (EHG)***	***Ukraine_HG1***	***Ukraine_N1***	*0.0062*	*0.290*	*103705*
Mbuti	***Latvia_MN2***	***Ukraine_HG1***	***Ukraine_N1***	*0.0085*	*0.355*	*98909*
Mbuti	*San*	***Ukraine_HG1***	***Ukraine_N1***	*0.0112*	*0.506*	*133136*
Mbuti	***MA1***	***Ukraine_HG1***	***Ukraine_N1***	*0.0156*	*0.636*	*84801*
Mbuti	*Papuan*	***Ukraine_HG1***	***Ukraine_N1***	*0.0265*	*1.257*	*132572*
There may have been increased connectivity with northern Eurasia from the Mesolithic to the Neolithic in Ukraine.						
Mbuti	**AfontovaGora3**	**WHG**	**Ukraine_HG1**	0.0846	2.411	4087
Mbuti	**AfontovaGora3**	**WHG**	**Ukraine_N1**	0.1462	4.461	4310
Mbuti	**MA1**	**WHG**	**Ukraine_HG1**	0.0032	0.158	11652
Mbuti	**MA1**	**WHG**	**Ukraine_N1**	0.0708	3.369	11892
Mbuti	**EHG**	**WHG**	**Ukraine_HG1**	0.0237	1.417	15359
Mbuti	**EHG**	**WHG**	**Ukraine_N1**	0.0421	2.546	15634
Mbuti	**CHG**	**WHG**	**Ukraine_HG1**	0.0143	0.915	16477
Mbuti	**CHG**	**WHG**	**Ukraine_N1**	0.0140	0.886	16724
Mbuti	***MA1***	***Loschbour (WHG)***	***Ukraine_HG1***	*0.0114*	*0.936*	*447952*
Mbuti	***MA1***	***Loschbour (WHG)***	***Ukraine_N1***	*0.0510*	*4.179*	*439711*
Mbuti	***MA1***	***Bichon (WHG)***	***Ukraine_HG1***	*0.0062*	*0.520*	*444959*
Mbuti	***MA1***	***Bichon (WHG)***	***Ukraine_N1***	*0.0509*	*4.102*	*437042*
Mbuti	***Karelia (EHG)***	***Loschbour (WHG)***	***Ukraine_HG1***	*0.0269*	*2.451*	*546828*
Mbuti	***Karelia (EHG)***	***Loschbour (WHG)***	***Ukraine_N1***	*0.0491*	*4.426*	*531401*
Mbuti	***Karelia (EHG)***	***Bichon (WHG)***	***Ukraine_HG1***	*0.0348*	*3.069*	*543214*
Mbuti	***Karelia (EHG)***	***Bichon (WHG)***	***Ukraine_N1***	*0.0560*	*4.998*	*528197*
Mbuti	***Kotias (CHG)***	***Loschbour (WHG)***	***Ukraine_HG1***	*0.0016*	*0.155*	*695299*
Mbuti	***Kotias (CHG)***	***Loschbour (WHG)***	***Ukraine_N1***	*0.004*	*0.381*	*675471*
Mbuti	***Kotias (CHG)***	***Bichon (WHG)***	***Ukraine_HG1***	*0.0054*	*0.494*	*688140*
Mbuti	***Kotias (CHG)***	***Bichon (WHG)***	***Ukraine_N1***	*0.0057*	*0.551*	*669022*

Latvia_HG, Latvian hunter-gatherers; WHG, western hunter-gatherers; EHG, eastern hunter-gatherers; CHG, Caucasus hunter-gatherers.

Tests performed using the whole genome panel are italicized, otherwise tests were performed using the Human Origin transversion SNP panel. Ancient samples are highlighted in bold. Samples include in each ancient group can be found in Supplemental Dataset 1. Key *D*-statistics for Latvian samples are shown in Table 2.

Supplemental Experimental Procedures

Archaeological information on sites, samples and their context

The Neolithic transition in the Baltic and Dnieper Rapids, Ukraine

The first evidence for the adoption of some elements of a food producing economy in the Baltic began ~6,400 cal BP, and continued over the following 5,000 years with the gradual establishment of farming communities [S7,S8]. Stable isotope analyses at Zvejnieki indicates a considerable variability across time from the late Mesolithic to the end of the Middle Neolithic, with a diet that is rich in terrestrial/freshwater proteins and of high trophic level species [S9]. There is in general a shift from a diet which relies mainly on freshwater fish during the late Mesolithic and Early Neolithic periods, to one which relies more on terrestrial animals and plants, during the Late Neolithic [S9].

Similarly at the Dnieper Rapids, Ukraine, the appearance of pottery in several of the cemeteries dates to between 7,500-7,000 cal BP but the first appearance of domesticates occurs only after 7,000 cal BP. Stable isotope analysis from cemetery populations in this region suggests variation in either the exploitation or availability of certain freshwater resources across the Mesolithic to Neolithic periods [S7].

Zvejnieki, Latvia

The site of Zvejnieki is situated in Northern Latvia, on the north-eastern bank of Lake Burtnieks. The shores of the lake have been quarried for gravel since the early 1960s and this activity led to the discovery of several prehistoric burials, some of which had traces of red ochre. This was followed by test excavations in 1964. Continuation of fieldwork during the 1960s and 1970s, covering a total of 4200 m^2, revealed 317 burials and a rich archaeological assemblage which included flint spearheads, arrows, bone harpoons, bone pendants, amber ornaments and pottery. The Zvejnieki archaeological complex consists of several groups of burials. Close to the cemetery, two settlement phases are known: the Mesolithic settlement Zvejnieki II and Neolithic settlement Zvejnieki I [S10,S11]. Six individuals from this site were selected for genetic analysis. The following information on these burials was taken from [S10–S13].

Mesolithic samples

Latvia_HG1: Burial 313. Adult female. Fragmentary skeletal material. ^{14}C date: LuS 8220, 7525 ±60 BP (8,417-8,199 cal BP).

Latvia_HG2: Burial 93. Adult male. Buried in extended supine position with head facing southwest. Grave goods included 23 teeth pendants, one beaver bone and three bird bones. Ochre layer surrounded the skeleton. ^{14}C date: Hela-1212, 6840 ±55 BP (7,791-7,586 cal BP).

Latvia_HG3: Burial 121. Adult. Described as female based on morphology but genetically determined to be male. Buried in extended supine position with head facing south. Animal teeth pendants were scattered around the burial as well as on the breast, shoulders and along the legs. Grave goods included a perforated animal phalange, two bird bones, a stone object (possibly representing an animal) and a flint chip. ^{14}C date: Ua-19883, 6145 ±80 BP (7,252-6,802 cal BP).

Neolithic samples

Latvia_MN1: Burial 124. Adult male. Buried in supine position with head facing southwest. Fragmentary skeletal material. Seven teeth pendants found in the vicinity of the head and breast. ^{14}C date: Ua-3639, 5280 ±55 BP (6,201-5,926 cal BP).

Latvia_MN2: Burial 221. Adult. Fragmentary skeletal material. Described as male based on morphology but genetically determined to be female. Part of a collective burial which contained five adults and one child. Buried in extended supine position. Heavily strewn with ochre. Grave goods included rich amber artefacts, two flint arrowheads, five flint chips, a two-headed water-bird figurine made of antler, an ornamented clay figurine and two bone fishing hook fragments. ^{14}C date: Ua-19813, 5180 ±65 BP (6,179-5,750 cal BP).

Latvia_LN1: Burial 137. Adult. Fragmentary bone material. Described as male based on morphology but genetically determined to be female. Body buried on left side, head facing east. Grave goods included a stone awl, a chisel initially thought to have been made from a deer bone but later determined to be from a domesticated goat, and potsherds. ^{14}C date: Ua-19811,4280 ±60 BP (5,039-4,626 cal BP).

Vasilyevka 3 and Vovnigi 2, Ukraine

Human remains from the cemeteries of Vasilyevka 3 (Mesolithic) and Vovnigi 2 (Neolithic) were originally studied by Konduktorova and Gokhman [S14,S15]. One individual from each graveyard was selected for genetic analysis.

Vasilyevka 3

The cemetery of Vasilyevka 3 is located on the third loess terrace of the left bank of the Dnieper river in the vicinity of Vasilyevka village (Dnipropetrovsk region). This is the biggest Mesolithic cemetery in the area. The cemetery was excavated by Telegin in 1955. 45 skeletons were found in a total of 100 m^2. It was suggested that approximately the same number of burials were destroyed by a 4 m wide gully which cut through the central part of the cemetery.

Based on the positions of the skeletons, two main burial groups were defined. The first group included 34 individuals buried in a crouched position on their right or left side, the second group included 7 individuals who were buried in supine position. The crouched burials are characteristic of Mesolithic burials from Eastern Ukraine. Most graves contained a single individual however there were also several double and triple burials.

In contrast to other Mesolithic cemeteries from the Dnieper area, for example Vasilyevka 1 and Voloshsky, where most of burials had east or south-east orientation, in Vasilyevka 3, different orientation for both main groups were found. It is possible that the heterogeneity in Vasilyevka 3 burial practices point to the introduction of a new type of funerary rite. Burials in supine position prevailed towards the end of the Mesolithic and into the Neolithic period [S16]. Artefacts found in the Vasilievka 3 burials included flint microliths, sometimes installed in bone weapons, and perforated shells.

The sample taken for genetic analyses (Skeleton 37; 14C date: Poz-83447, 9530 ±80 BP, 11,143-10,591 cal BP) was from a triple burial. All skeletons in the burial were determined to be male based on morphology. They were found in a crouched position on their left side. One skeleton in the burial had a microlith arrow point stuck in his left tenth rib. Several flint pieces were also found in his chest area. Because of this, and the very robust characteristics of all skeletons in the burial, it has been proposed that they were warriors who died and were buried simultaneously [S15]. It is believed to be one of the earliest graves in the cemetery as it was the deepest and another burial was found above it (skeleton 11). Radiocarbon dating of two crouched and one extended burial show that this graveyard was in use between 12,350 and 11,150 cal BC [S17–19].

Vovnigi 2

The Neolithic cemetery of Vovnigi 2 was excavated by Bodjanskiy and Rudinskiy from 1949 to 1952. The graveyard is located on the right bank of the Dnieper River in the very center of the village of Vovnigi (Dnipropetrovsk region, Ukraine). In the excavated area covering ~100 m^2, a total of 131 burials were excavated. It has been proposed that the large number of burials at Vovnigi, compared to earlier Neolithic cemeteries in the region, resulted from a population increase towards the end of the Neolithic [S20].

All burials were arranged in three rows with 70 burials forming the central group. Such specific cemetery organization is characteristic of the Dnieper-Donets Neolithic culture. Most skeletons were in western or north-western orientation. They were buried in supine position with their hands positioned over their pelvis or in some cases their arms were extended along their body. The skeletal remains used in our genetic analyses ((skeleton 2; 14C date: Poz-83446, 5590 ±50 BP, 6,469-6,293 cal BP) came from a single burial located within the main core of burials. This individual was buried in extended supine position in north-south orientation.

Overlapping burials were typical, especially in the central group of the cemetery. Many skeletons were covered with a thick layer of ochre. Remarkably, in all cases, burials with ochre overlay burials without it. It has therefore been proposed that the use of ochre in burial practices was developed during the later stages of the use of the cemetery [S21]. Artefacts in the burials ("grave goods") were not numerous, consisting of small flint blades and flakes, microliths, pearl beads and shell fragments (belonging mostly to the genus *Unio*). Some of the shell fragments had artificial holes which point to their use as pendants. The presence of deer teeth and a few

pottery sherds in several burials was reported. It has been suggested that the small fragments of pottery found in graves most probably came from broken pots used during funeral feasts. Most of these pottery fragments show traces of fire. The fragments were similar to pottery found at Dnieper-Donets Neolithic settlements. The Dnieper-Donets culture was spread across the western Azov Sea area, the Dnieper and the Crimean Steppes during Neolithic. Although pottery appeared in settlements of the Ukrainian Steppe zone during Early Neolithic times (8,250 cal BP), its use in funerary rituals began only 300 years later [S22]. The Vovnigi 2 site has been dated by Jacobs to 7,430 to 6,700 cal BP [S7] with the date of our sample extending the use of the cemetery by several centuries.

DNA extraction, library preparation, sequencing and alignment

In dedicated ancient DNA facilities, DNA from ancient Latvian and Ukrainian samples was extracted from (powdered) petrous bones using a silica column based protocol [S23] as described by Gamba et al., 2014 [S24]. Next-generation sequencing libraries were constructed and amplified with AccuPrime Pfx Supermix (Life Technology) following Gamba et al., 2014 ([S24]; a protocol adapted from Meyer and Kircher 2010 [S25]). Libraries were first screened to assess their human DNA content on an Illumina MiSeq platform at TrinSeq, Dublin using 50 base pair (bp) single-end sequencing. Selected libraries were further sequencing using 70 bp single-end sequencing (MiSeq platform at TrinSeq), 75 bp single-end sequencing (NextSeq platform at University of Potsdam) or 100 bp single-end sequencing (HiSeq 2,000 platform at Beijing Genomics Institute or at Macrogen).

Reads needed to be within 1 bp of their given index in order to be used in the alignment. Adapter sequences were removed from the 3' end of reads using Cutadapt (version 1.3) [S26], requiring an overlap of at least 1 bp between the read and the adapter sequence. Reads were aligned to the GRCh37 build of the human genome with the mitochondrial sequence replaced by the revised Cambridge reference sequence (NCBI accession number NC_012920.1) using Burrows-Wheeler Aligner (BWA) version 0.7 [S27]. Sequences from the same sample, but from different sequencing runs, were merged using Picard MergeSamFiles (http://picard.sourceforge.net/) and duplicate reads were removed using SAMtools version 0.1.19 [S28]. A minimum length threshold of 30 bp was imposed and Genome Analysis Toolkit (GATK) DepthofCoverage [S29] was used to calculate the average depth of coverage. Indels were realigned using GATK RealignerTargetCreator and IndelRealigner [S29]. Sequences with a mapping quality of ≥ 30 were retained using SAMtools [S28] and 2 bp were soft-clipped from the start and end of every read.

Authenticity of data

The authenticity of data was first assessed by looking for short average sequence length as described in Gallego Llorente *et al.,* 2015 [S30]. MapDamage 2.0 [S31], with default parameters, was then used to look for and evaluate patterns of molecular damage which are typical of ancient DNA. Contamination in ancient mitochondrial sequences was estimated by assessing the number of non-consensus bases (minimum base quality threshold of 20) at haplogroup defining positions as a function of the total coverage for each of these sites [S24,S32]. X chromosome contamination levels were also evaluated for male samples with greater than 0.5-fold coverage (a cut-off recommended by Allentoft *et al.,* [S33]), using ANGSD [S34] with parameters recommended on the ANGSD website and imposing a minimum base quality threshold of 20.

SNP calling and merging with modern dataset and other ancient samples

Genotypes for modern individuals from the Human Origins dataset [S35] were merged with additional, publicly available modern genotypes (from Assyrian, French, Iranian, Italian, Lebanese, Libyan, Moroccan and Romanian populations) outlined in Lazaridis *et al.,* 2016 [S36]. Overlapping SNP positions between both datasets were retained and genotypes on sex chromosomes were removed. Using PLINK [S37], this modern dataset was merged with ancient genotypes provided in the Lazaridis *et al.,* 2016 [S36] dataset which included ancient samples described in [S24,S30,S33,S35,S38–S49]. For our Latvian and Ukrainian ancient samples, we called genotypes at positions which overlapped with this dataset using GATK Pileup [S29]. Bases were required to have a minimum quality of 20 and only alleles found in the Human Origins dataset were considered. For SNP positions with more than one base call, one allele was randomly chosen with a probability equal to the frequency of the base at that position. This allele was duplicated to form a homozygous diploid genotype which was used to represent the individual at that SNP position [S50]. This method of SNP calling was also used to call

genotypes in ancient samples from Cassidy *et al.,* 2015 [S51]. These data as well as ancient data from Fu *et al.,* 2016 [S52] were merged with the modern data and other ancient samples using PLINK [S37]. For all populations genetic analyses (PCA, ADMIXTURE, *D*-statistics and f_3-statistics), transition sites were removed from the dataset in order to reduce the impact of cytosine deamination on results and only ancient samples with ≥ 15,000 called SNPs were included.

D-statistics were also performed on a panel of whole genome sequences which contained both modern [S38] and ancient genomes (samples from this study along with Bichon, Kotias [S45], NE1, BR2 [S24], Mal'ta [S38], Loschbour, Stuttgart [S35] and Karelia [S52]). Ancient genomes were aligned, filtered and pseudo-diploid genotypes called as described above, imposing a minimum base quality threshold of 30. Modern genotypes were also made pseudo-diploid by choosing one allele at random and duplicating it. Only autosomal transversion SNPs found in the Phase 1 of the 1,000 Genomes Project were used in these analyses.

Population genetic analyses

PCA was performed using EIGENSOFT 5.0.1 smartpca [S53] to project ancient data onto the first two principal components defined by a Eurasian subset of our modern dataset. This was carried out with the lsqproject option on and the outlier removal option off. One SNP from each pair in linkage disequilibrium with $r^2>0.2$ was removed. ADMIXTURE version 1.23 [S54] was used to perform a clustering analysis on the dataset described above. Single-nucleotide polymorphisms in linkage disequilibrium were thinned using PLINK (v1.07) [S37] with parameters --indep-pairwise 200 25 0.5 [S43]. Clusters (*K*) (2–20) were explored using 10 runs with fivefold cross-validation at each *K* using different random seeds. The minimal cross validation error was found at *K*=17 (see Figure S6; results for *K*2-20 are shown in Figure S5 and Figure S7 for modern and ancient individuals respectively). *D*-statistics [S55] and f_3-statistics [S56,S57] were computed using the qpDstat and 3PopTest programs respectively from the ADMIXTOOLS package [S57].

Uniparental haplogroups

For mitochondrial and Y chromosome haplogroup determination, sequences were rescaled using mapDamage 2.0 [S31] to reduce the impact of terminal sequence deamination on results. ANGSD [S34] was used to create mitochondrial consensus sequences setting the minimum depth of coverage to 3 and the minimum base quality to 20. Haplogroups were then determined using the HAPLOFIND [S58] web-tool which classifies mitochondrial sequences according to previously annotated haplogroups. YFitter [S59] was used to determine the Y chromosomal haplogroups of our male samples. Genotypes at positions along the Y-chromosome which overlapped with the dbSNP (build 137) database were called in male samples using GATK [S29]. VCFtools [S60] was used to convert the resulting VCF files to PLINK format files [S37] which were subsequently converted to qcall format using the tped2qcall.py script provided in the YFitter package. Yfitter was run with default parameters using the build 37 Y chromosome tree outlined in Karafet *et al.,* 2008 [S61].

Genotypes associated with particular phenotypes

For selected markers, genotypes with alleles present in Phase 1 of 1,000 genomes dataset [S62] and with base quality ≥ 20, were called using GATK Unified genotyper [S29]. As many diagnostic markers were not sequenced or had low coverage, we used imputation to infer genotypes at these positions. Imputation was performed as described in Gamba *et al.,* 2014 [S24], imputing at least 1 Mb upstream and downstream of the SNP position where possible and using 10 iterations to estimate genotypes at ungenotyped markers. The Hirisplex prediction model [S6] was used to explore hair and eye colour with a genotype probability threshold of ≥ 0.7 imposed.

Supplemental References

S1. Lamason, R.L., Mohideen, M.-A.P.K., Mest, J.R., Wong, A.C., Norton, H.L., Aros, M.C., Jurynec, M.J., Mao, X., Humphreville, V.R., Humbert, J.E., *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science *310*, 1782–1786.

S2. Sturm, R.A. (2009). Molecular genetics of human pigmentation diversity. Hum. Mol. Genet. *18*, R9–17.

S3. Sturm, R.A., Duffy, D.L., Zhao, Z.Z., Leite, F.P.N., Stark, M.S., Hayward, N.K., Martin, N.G., and Montgomery, G.W. (2008). A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. Am. J. Hum. Genet. *82*, 424–431.

S4. Eiberg, H., Troelsen, J., Nielsen, M., Mikkelsen, A., Mengel-From, J., Kjaer, K.W., and Hansen, L. (2008). Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Hum. Genet. *123*, 177–187.

S5. Enattah, N.S., Sahi, T., Savilahti, E., Terwilliger, J.D., Peltonen, L., and Järvelä, I. (2002). Identification of a variant associated with adult-type hypolactasia. Nat. Genet. *30*, 233–237.

S6. Walsh, S., Liu, F., Wollstein, A., Kovatsi, L., Ralf, A., Kosiniak-Kamysz, A., Branicki, W., and Kayser, M. (2013). The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci. Int. Genet. *7*, 98–115.

S7. Lillie, M., and Budd, C. (2011). The Mesolithic-Neolithic Transition in Eastern Europe: Integrating Stable Isotope Studies of Diet with Palaeopathology to Identify Subsistence Strategies and Economy. In Human Bioarchaeology of the Transition to Agriculture, Jay T. Stock and Ron Pinhasi, eds. (John Wiley & Sons, Ltd.), pp. 43–62.

S8. Zvelebil, M. (2006). Mobility, contact, and exchange in the Baltic Sea basin 6000–2000 BC. Journal of Anthropological Archaeology *25*, 178–192.

S9. Eriksson, G. (2006). Stable isotope analysis of human and faunal remains from Zvejnieki. In Back to the Origin New Research in the Mesolithic-Neolithic Zvejnieki Cemetery & Environment, I. Larsson and L. Zagorska, eds., pp. 183–216.

S10. Zagorskis, F.A. (1987). Zvejnieku akmens laikmeta kapulauks (Zinatne).

S11. Zagorskis, F.A. (2004). Zvejnieki (Northern Latvia) Stone Age cemetery (Oxford, England: Archaeopress).

S12. Denisova, R.J. (1975.). Antropologiya drevnih baltov. Riga: Zinātne.

S13. Larsson, L., and Zagorska, I. eds. (2006). Back to the Origin : New Research in the Mesolithic-Neolithic Zvejnieki Cemetery and Environment, North Latvia (Almqvist & Wiksell International).

S14. Konduktorova, T.S. (1973). Anthropology of Ukrainian populations in Mesolithic, Neolithic and Bronze Ages. Moscow: Nauka. p. 127 (In Russian).

S15. Gokhman, I.I. (1966). The populations of Ukraine in Mesolithic and Neolithic times. Moscow: Nauk: 196 (In Russian).

S16. Telegin, D.L. (1989). The Mesolithic of South-West of the USSR (Ukraine and Moldaviya). Mezolit SSSR. Moscow: Nauka., 106–124 (In Russian).

S17. Jacobs, K. (1993). Human Postcranial Variation in the Ukrainian Mesolithic-Neolithic. Curr. Anthropol. *34*, 311–324.

S18. Kośko, A., Domańska, L., and Jacobs, K. (1998). The Ukrainian Steppe as a region of intercultural contacts between Atlantic and Mediterranean zones of European Mesolithic. Baltic-Pontic-Studies *5*, 102–119.

S19. Lillie, M., Richards, M.P., and Jacobs, K. (2003). Stable isotope analysis of 21 individuals from the Epipalaeolithic cemetery of Vasilyevka III, Dnieper Rapids region, Ukraine. J. Archaeol. Sci. *30*, 743–752.

S20. Rudinskiy, M.Y. (1955). The Vovnigi Late Neolithic cemeteries. Kratkiye soobscheniya Instituta

Arkheologii *4*, 147–151 (In Russian).

S21. Bodjanskiy, A.V., and Kotova, N.S. (1994). Vovnigskiy 2 posdneneoliticheskiy mogilnik (The Vovnigi 2 Late Neolithic Cemetery). Archeologichni pamjaki ta istorija sta rodavnogo naselennja Ukrainy (Archaeological sites and history of ancient population in Ukraine). Vezsha, Lutsk: 98–110. (In Russian).

S22. Kotova, N., and Spitsyna, L. (2010). Ceramics in the burial rites of the Neolithic-Early Bronze Age in the Ukrainian steppe. Documenta Praehistorica XXXVII, 179–184.

S23. Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., and Saunders, S.R. (1998). Technical note: improved DNA extraction from ancient bones using silica-based spin columns. Am. J. Phys. Anthropol. *105*, 539–543.

S24. Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., *et al.* (2014). Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. *5*, 5257.

S25. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, db.prot5448.

S26. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10–12.

S27. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

S28. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

S29. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

S30. Gallego Llorente, M., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, J.W., Curtis, M.C., Stock, J.T., Coltorti, M., Pieruccini, P., *et al.* (2015). Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. Science *350*, 820–822.

S31. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.

S32. Sánchez-Quinto, F., Schroeder, H., Ramirez, O., Ávila-Arcos, M.C., Pybus, M., Olalde, I., Velazquez, A.M.V., Marcos, M.E.P., Encinas, J.M.V., Bertranpetit, J., *et al.* (2012). Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. Curr. Biol. *22*, 1494–1499.

S33. Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., *et al.* (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.

S34. Korneliussen, T., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.

S35. Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., *et al.* (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

S36. Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., *et al.* (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.

S37. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., *et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

S38. Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford, T.W., Jr, Orlando, L., Metspalu, E., *et al.* (2013). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature *505*, 87–91.

S39. Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., Santpere, G., Chiang, C.W.K., DeGiorgio, M., Prado-Martinez, J., Rodríguez, J.A., Rasmussen, S., Quilez, J., *et al.* (2014). Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature *507*, 225–228.

S40. Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Jr, Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., *et al.* (2014). The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature *506*, 225–229.

S41. Seguin-Orlando, A., Korneliussen, T.S., Sikora, M., Malaspinas, A.-S., Manica, A., Moltke, I., Albrechtsen, A., Ko, A., Margaryan, A., Moiseyev, V., *et al.* (2014). Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. Science *346*, 1113–1118.

S42. Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S.M., Bondarev, A.A., Johnson, P.L.F., Aximu-Petri, A., Prüfer, K., de Filippo, C., *et al.* (2014). Genome sequence of a 45,000-year-old modern human from western Siberia. Nature *514*, 445–449.

S43. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., *et al.* (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211.

S44. Günther, T., Valdiosera, C., Malmström, H., Ureña, I., Rodriguez-Varela, R., Sverrisdóttir, Ó.O., Daskalaki, E.A., Skoglund, P., Naidoo, T., Svensson, E.M., *et al.* (2015). Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. Proc. Natl. Acad. Sci. U. S. A. *112*, 11917–11922.

S45. Jones, E.R., Gonzalez-Fortes, G., Connell, S., Siska, V., Eriksson, A., Martiniano, R., McLaughlin, R.L., Gallego Llorente, M., Cassidy, L.M., Gamba, C., *et al.* (2015). Upper Palaeolithic genomes reveal deep roots of modern Eurasians. Nat. Commun. *6*, 8912.

S46. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., *et al.* (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503.

S47. Olalde, I., Schroeder, H., Sandoval-Velasco, M., Vinner, L., Lobón, I., Ramirez, O., Civit, S., García Borja, P., Salazar-García, D.C., Talamo, S., *et al.* (2015). A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. Mol. Biol. Evol. *32*, 3132–3142.

S48. Rasmussen, M., Sikora, M., Albrechtsen, A., Korneliussen, T.S., Moreno-Mayar, J.V., Poznik, G.D., Zollikofer, C.P.E., Ponce de León, M.S., Allentoft, M.E., Moltke, I., *et al.* (2015). The ancestry and affiliations of Kennewick Man. Nature *523*, 455–458.

S49. Fu, Q., Hajdinjak, M., Moldovan, O.T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N., Rohland, N., Lazaridis, I., Nickel, B., *et al.* (2015). An early modern human from Romania with a recent Neanderthal ancestor. Nature *524*, 216–219.

S50. Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., and Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science *336*, 466–469.

S51. Cassidy, L.M., Martiniano, R., Murphy, E.M., Teasdale, M.D., Mallory, J., Hartwell, B., and Bradley, D.G. (2016). Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. Proc. Natl. Acad. Sci. U. S. A. *113*, 368–373.

S52. Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., *et al.* (2016). The genetic history of Ice Age Europe. Nature *534*, 200–205.

S53. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190.

S54. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

S55. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., *et al.* (2010). A draft sequence of the Neandertal genome. Science *328*, 710–722.

S56. Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009). Reconstructing Indian population history. Nature *461*, 489–494.

S57. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

S58. Vianello, D., Sevini, F., Castellani, G., Lomartire, L., Capri, M., and Franceschi, C. (2013). HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. Hum. Mutat. *34*, 1189–1194.

S59. Jostins, L., Xu, Y., McCarthy, S., Ayub, Q., Durbin, R., Barrett, J., and Tyler-Smith, C. (2014). YFitter: Maximum likelihood assignment of Y chromosome haplogroups from low-coverage sequence data. arXiv [q-bio.PE]. Available at: http://arxiv.org/abs/1407.7988.

S60. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., *et al.* (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

S61. Karafet, T.M., Mendez, F.L., Meilerman, M.B., Underhill, P.A., Zegura, S.L., and Hammer, M.F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res. *18*, 830–838.

S62. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.