Current Biology

Supplemental Information

Genomic Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool

Ayça Omrak, Torsten Günther, Cristina Valdiosera, Emma M. Svensson, Helena Malmström, Henrike Kiesewetter, William Aylward, Jan Storå, Mattias Jakobsson, and Anders Götherström

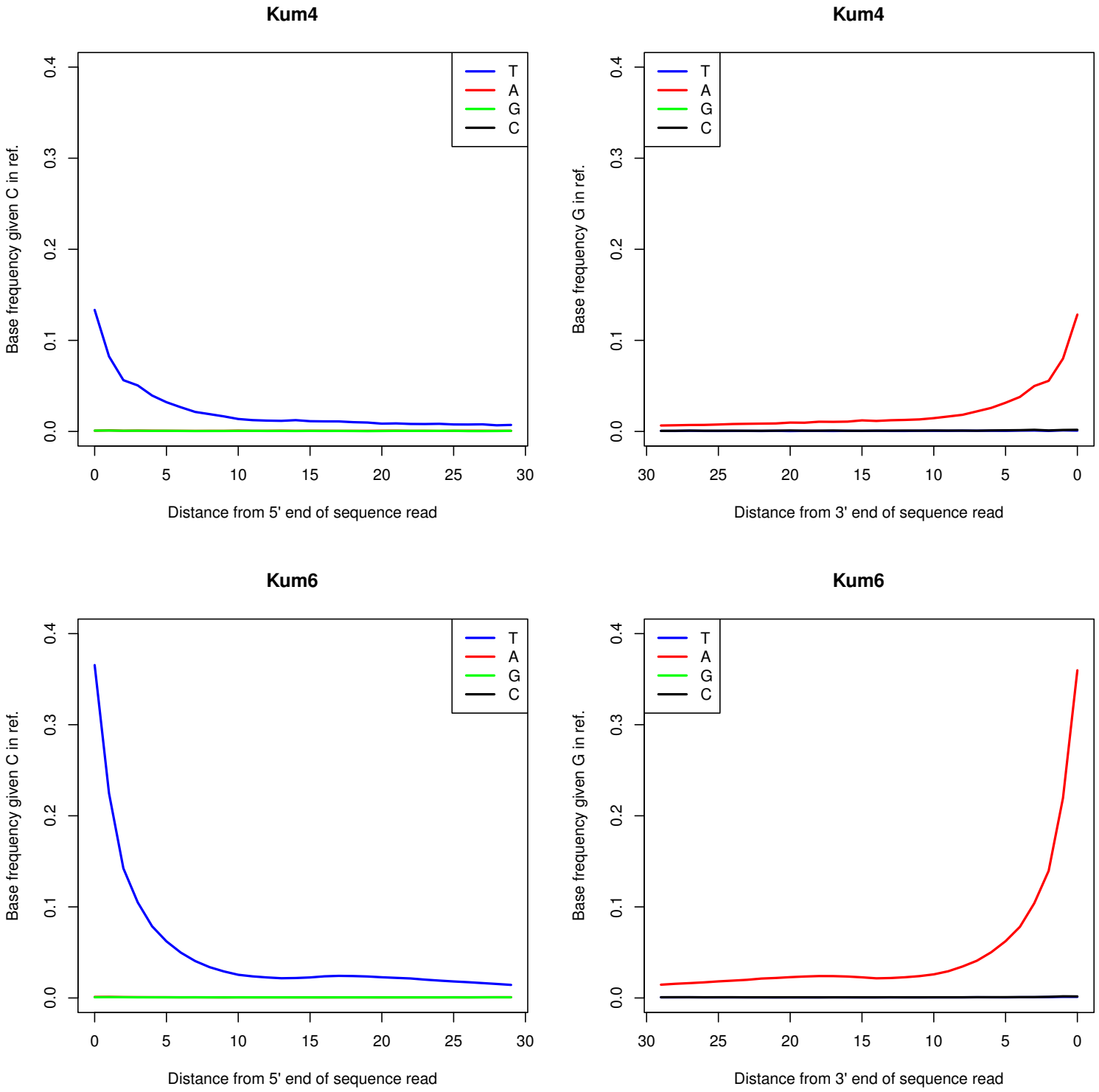

Figure S1. Authentication. *Deamination patterns for Kum4 and Kum6 (Related to Figure 1A)*

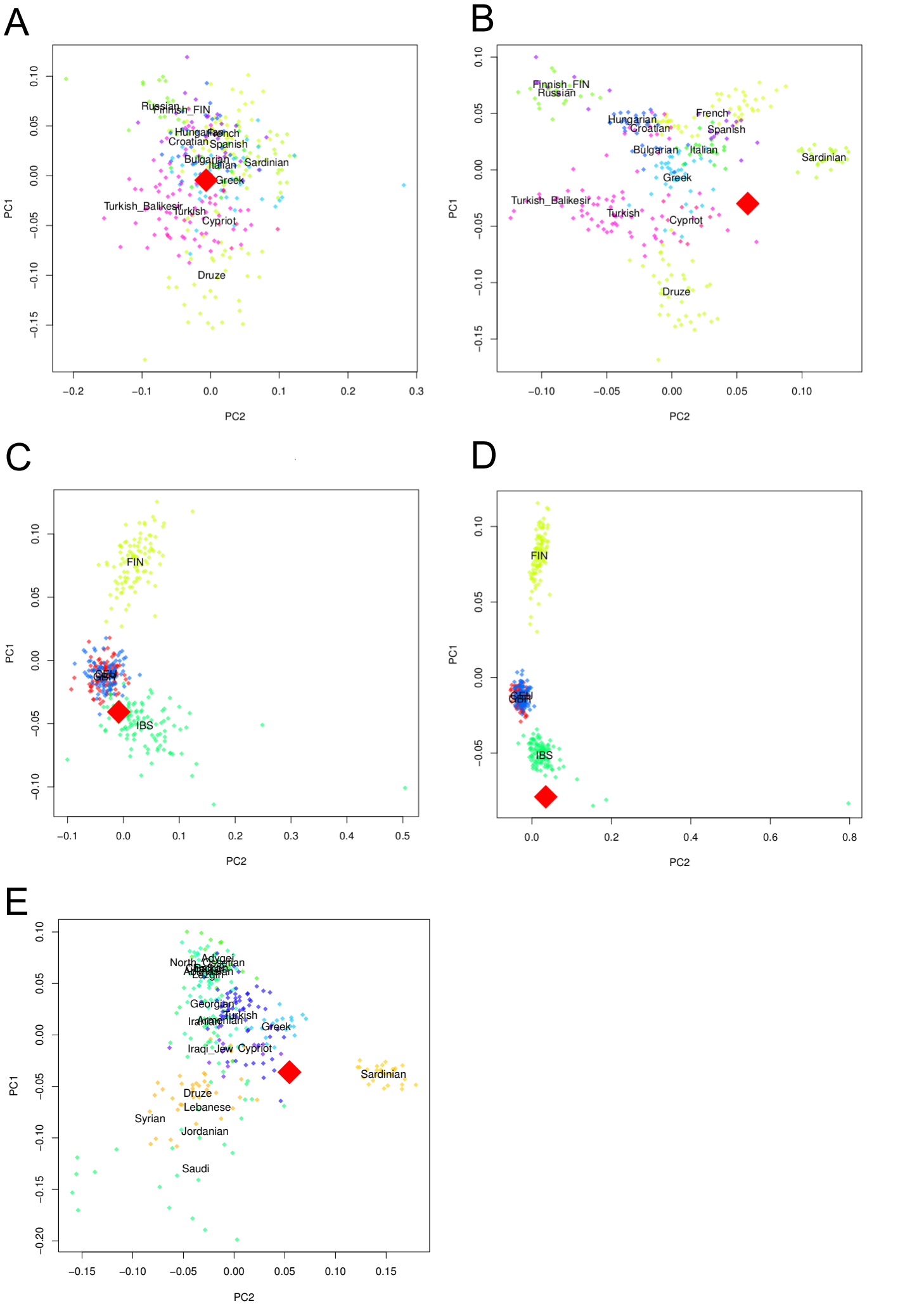

Figure S2. Individual Principal Component Analyses for Kum4 and Kum6. *(a) Kum4 individiual PCA with European Human Origins populations (b) Kum6 individual PCA with European Human Origins populations, (c) Kum4 individual PCA with European 1000 genome populations, (d) Kum6 individual PCA with European 1000 genomes populations, (e) Kum6 individual PCA with South Eastern European and Western Asian populations from the Human Origins data. Sardinians were added as a European farmer-like reference population. (Related to Figure 1B)*

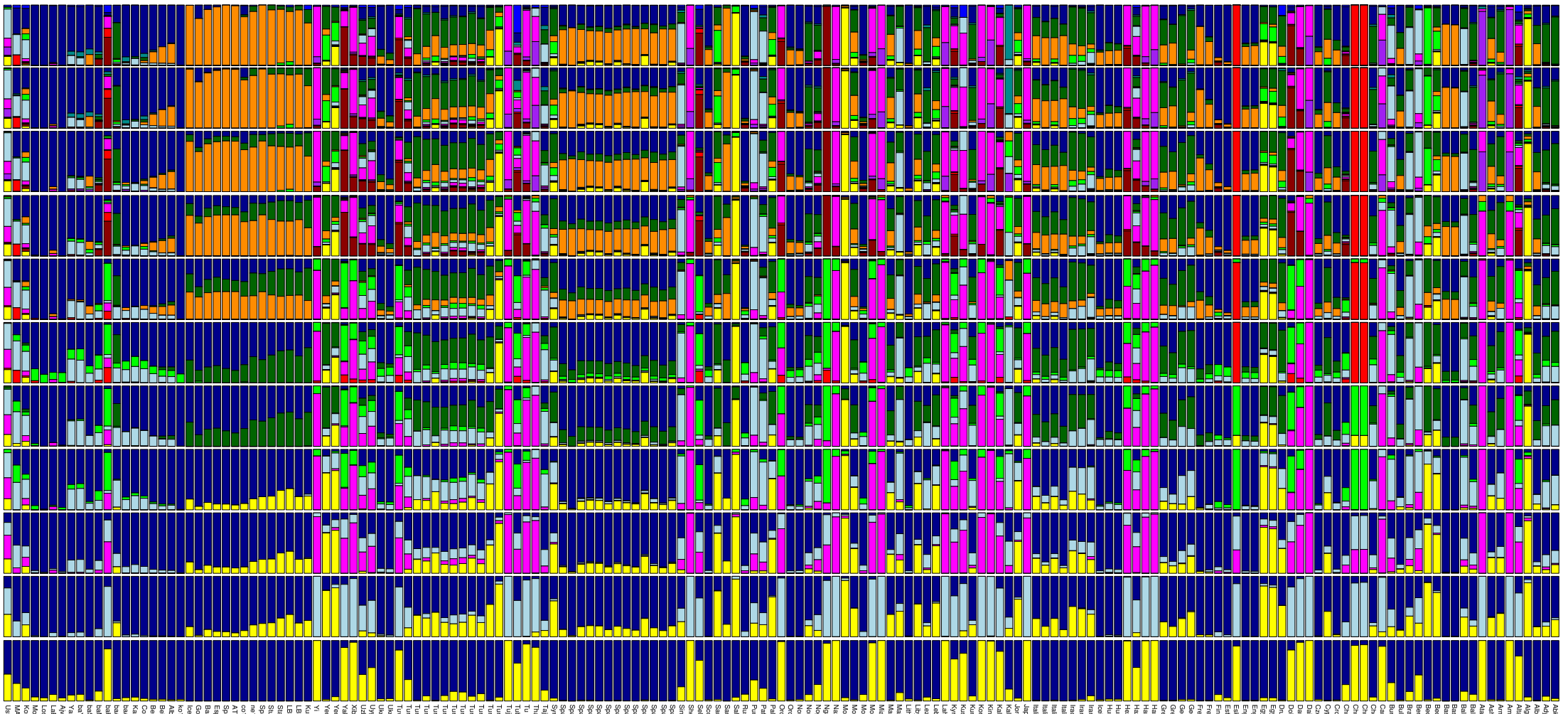

Figure S3. Ancestry proportions inferred from model-based clustering. *Admixture plots of population Q values for K=2 to K=12. (Related to Figure 3)*

Table S1: Samples (Related to Figure 1A)

(a) Overview of all skeletal finds of individuals screened for this study with sex and age-at-death estimation.

Lab Name	Layer	Trench	Excavation Year	Grave (find no. of skeleton)	Osteology (sex, age)
Kum4	Kumtepe IB2	F28	1994	1 (F28.317)	female, 18-23 years
Kum5	Kumtepe IA2	F28	1994	5 (F28.639)	probably male, 20-30 years
Kum6	Kumtepe IA1	F28	1994-1995	6 (F28.992 and F28.1044)	female, 20-30 years
Kum7	Kumtepe IA1	F28	1995	7 (F28.996 and F28.1045)	probably female, 15-17 years
Kum8	Kumtepe IA2	F28	1994	8 (F28.875)	sex?, 13-16 years
Kum9	Kumtepe IA1	F28	1995	9 (F28.1000 and F28.1046)	female, 16-18 years
Kum10	Kumtepe IA1	F28	1995	10 (F28.1021)	male, 35-45 years
Kum12	Kumtepe IA1	F28	1995	12 (F28.1041)	male, 30-40 years

(b) Radiocarbon (C14) dates from human bones from Kumtepe (layer IA1) [S1] (layer IA1) (Calibration curve: IntCal13 [S2] , calibration method: OxCal 4.2.4 [S3]).

Context	Find no.	BP	Std dev	Delta c13	Cal BC(1sigma)	Cal BC (2sigma)
Grave 6	F28.1044	5880	41	-19,81	4795-4710 (68,2%)	4846-4618 (95,4%)
Grave 7	F28.1045	5643	29	-20,1	4520-4445 (68,2%)	4543-4372 (95,4%)
Grave 9	F28.1046	5893	38	-20,17	4800-4715 (68,2%)	4848-4687 (95,4%)

Table S2: Summary statistics for all sequencing runs (Related to Figure 1A)

Sequencing Run	Total	Total Unique	Endogenous (%)	Clonality
Kum6				
D69_TTGAAGT_L004_merged_001_SN344_0250_AD2538ACXX	4 267 856	7 536	0.177	19.77
D69_ACCAACT_L004_merged_001_SN344_0250_AD2538ACXX	16 618 193	26 458	0.159	20.11
D69_ACCAACT_L004_merged_001_SN7001335_0080_BC2KVJACXX	37 951 577	42 033	0.111	21.09
D69_TTGAAGT_L004_merged_001_SN7001335_0080_BC2KVJACXX	9 734 163	10 987	0.113	20.33
D69_GGATCAA_L004_merged_D00458_9999_AC3MD9ACXX	23 009 284	74 380	0.323	98.12
D69S_GCTCGAA_L004_merged_D00458_9999_AC3MD9ACXX	18 117 966	11 511	0.064	21.47
D69_GAATCTC_L004_L005_merged_SN866_0285_BC42E0ACXX	957 325	1 629	0.170	26.76
D69_GGATCAA_L004_merged_D00458_9999_AC3MD9ACXX	23 009 284	2 519	0.011	44.34
D69_TTGAAGT_L005_L005_merged_SN866_0285_BC42E0ACXX	31 126	84	0.270	42.86
D69_CCTAGGT_L008_merged_D00457_0036_BC3R92ACXX	19 393 167	22 672	0.117	24.75
D69a-2_TTGAAGT_L005__merged_D00118_0147_AC4A0BACXX	10 812 021	5 826	0.054	90.49
D69s-2_ACTATCA_L005__merged_D00118_0147_AC4A0BACXX	14 208 130	10 406	0.073	34.00
D69s-3_TTGGATC_L005__merged_D00118_0147_AC4A0BACXX	14 540 177	10 616	0.073	35.58
D71_GCAAGAT_L004.merged_D00458_9999_AC3MD9ACXX	19 155 296	189 148	0.987	19.60
D71_GGATCAA_L008_merged_D00457_0036_BC3R92ACXX	20 948 448	1 721 121	8.216	21.65
D71a-2_CGACCTG_L005__merged_D00118_0147_AC4A0BACXX	14 375 339	132 709	0.923	10.87
D71z-2_GAATCTC_L005__merged_D00118_0147_AC4A0BACXX	30 125 608	112 705	0.374	89.66
D71a-1_GCAAGAT_L007_merged_D00118_0154_BC4VF9ACXX	40 459 798	382 976	0.013	41.45
D71a-2_CGACCTG_L007_merged_D00118_0154_BC4VF9ACXX	54 363 422	565 981	0.009	12.15
D71z-1_GGATCAA_L007_merged_D00118_0154_BC4VF9ACXX	47 182 140	3 807 480	0.072	14.69
Kum4				
KF2a_TAATGCG_L007_merged_D00118_0143_BC431CACXX	19 010 453	112 173	0.590	7.08
KF2a_TAATGCG_L007_merged_D00118_0154_BC4VF9ACXX	48 805 708	274 322	0.005	12.82

Table S3: General statistics (Related to Figure 1A)

(a) Contamination estimates for Kum6.

	Method 1 [S4]						**Method 2 [S5]**
Individual	Point Estimate	Informative Sites	Consensus Alleles	Total Alleles	Lower CI	Upper CI	Proportion Authentic
Kum6	2.04	3	48	49	0	5.99	~ 99%

(b) Mitochondrial haplogroup classification for Kum6.

Individual	Haplogroup	SNPs supporting called hg	Missing hg specific SNP's	Additional SNPs in Kum6 not belonging to called hg
Kum6	H2a	73A, 146T, 195T, 247G, 769G, 825T, 1018G, 2706A, 2758G, 2885T, 3594C, 4104A, 4312C, 4769A, 7146A, 7256C, 8468C, 8655C, 8701A, 9540T, 10398A, 10664C, 10688G, 10873T, 10915T, 11719G, 11914G, 12705C, 13105A, 13276A, 13506C, 13650C, 14766C, 16187C, 16189T, 16223C, 16230A, 16278C, 16311T	152T, 1438A, 16129G	1959T, 4407A, 4678C, 5753A, 5756A, 6527G, 7008A, 8608T, 9008T, 9459T, 10810T, 1627

(c) Genetic sex of the individuals

Sample	Number of alignments to chrX or chrY	Number of alignments to chrY	Ry	SE	95% CI	Assignment
Kum6	271 037	189	0.0007	0.0001	0.0006-0.0008	XX
Kum4	16 665	9	0.0005	0.0002	0.0002-0.0009	XX

Table S4: Ancient samples used for comparative statistics (Related to Figure 1A)

(a) Number of transversion SNPs overlapping with the Human Origins Array for all ancient individuals used in this study.

Individual	Number of SNPs	Reference
Kum4	4,308	This study
Kum6	42,061	This study
Ajv58	335,213	[S6]
Gok2	265,008	[S6]
Iceman	373,182	[S7]
LaBrana1	363,806	[S8]
Stuttgart	378,101	[S9]
Loschbour	377,760	[S9]
Motala12	89,542	[S9]
ne1	375,747	[S10]
co1	220,343	[S10]
I0058	156,721	[S11]
I0100	214,742	[S11]
I0103	210,374	[S11]
I0112	214,042	[S11]
I0118	219,141	[S11]
I0172	172,246	[S11]
I0174	64,301	[S11]
I0176	19,426	[S11]
I0231	218,126	[S11]
I0408	201,879	[S11]
I0412	209,316	[S11]
I0550	38,380	[S11]
I0560	84,840	[S11]
ATP2	217,748	[S12]
MA1	273,695	[S13]
Ust-Ishim	225,993	[S14]
Kostenki14	145,278	[S15]
RISE395	186,151	[S16]
RISE423	79,608	[S16]
RISE497	225,380	[S16]
RISE505	220,016	[S16]
RISE523	201,509	[S16]
RISE552	193,063	[S16]

(b) Number of transversion SNPs overlapping with the 1000 genomes data for all ancient individuals used in this study.

Individual	Number of SNPs
Kum4	16,631
Kum6	173,214
Ajv58	1,485,010
Gok2	931,296
Iceman	1,480,180
LaBrana	1,514,797
LBK	1,792,647
Loschbour	1,811,983
MA1	1,140,461
Motala12	1,528,623
ne1	1,800,002
co1	938,795
Denisovan	1,788,462
Ust-Ishim	1,709,866
Kostenki14	941,997
ATP2	1,599,343
RISE395	1,059,797
RISE423	458,839
RISE497	1,516,411
RISE505	1,473,875
RISE523	1,131,716
RISE552	1,092,023

Table S5: Key results obtained from *D-statistics D*(Outgroup, pop1; pop2, pop3). (Related to Figure 4A)

Kum6 shows higher affinities to the Bronze Age Asian cultures Sintashta, Andronovo and Mezhovskaya than to Mesolithic Europeans.					
Outgroup	**pop1**	**pop2**	**pop3**	**D**	**Z**
Denisovan	Kum6	Sintashta	Loschbour	-0.0455	-3.84
Denisovan	Kum6	Sintashta	LaBrana1	-0.0379	-3.368
Denisovan	Kum6	Sintashta	Motala12	-0.0398	-3.164
Denisovan	Kum6	Andronovo	Loschbour	-0.0417	-3.658
Denisovan	Kum6	Andronovo	Motala12	-0.039	-3.327
Denisovan	Kum6	Andronovo	LaBrana1	-0.0334	-3.068
Denisovan	Kum6	Mezhovskaya	Loschbour	-0.038	-3.16
Denisovan	Kum6	Mezhovskaya	Motala12	-0.0291	-2.393
Denisovan	Kum6	Mezhovskaya	LaBrana1	-0.0274	-2.33
Several Bronze Age Asian Cultures show higher affinities to Kum6 than to early Neolithic Europeans.					
Outgroup	**pop1**	**pop2**	**pop3**	**D**	**Z**
Denisovan	Armenia_BA	Kum6	ATP2	-0.0412	-2.36
Denisovan	Sintashta	Kum6	ATP2	-0.0236	-1.882
Denisovan	Yamnaya_RISE	Kum6	ATP2	-0.0132	-1.108
Denisovan	Karasuk	Kum6	ATP2	0.0001	0.007
Denisovan	Andronovo	Kum6	ATP2	0.0004	0.033
Denisovan	Mezhovskaya	Kum6	ATP2	0.0039	0.328
Denisovan	Karasuk	Kum6	co1	-0.0123	-0.922
Denisovan	Armenia_BA	Kum6	co1	-0.0158	-0.758
Denisovan	Sintashta	Kum6	co1	-0.0083	-0.548
Denisovan	Mezhovskaya	Kum6	co1	-0.0052	-0.35
Denisovan	Yamnaya_RISE	Kum6	co1	0.0035	0.234
Denisovan	Andronovo	Kum6	co1	0.0172	1.37
Denisovan	Andronovo	Kum6	Gok2	-0.0223	-1.703
Denisovan	Armenia_BA	Kum6	Gok2	-0.0317	-1.623
Denisovan	Karasuk	Kum6	Gok2	-0.0199	-1.506
Denisovan	Yamnaya_RISE	Kum6	Gok2	-0.0079	-0.583
Denisovan	Mezhovskaya	Kum6	Gok2	-0.0043	-0.307
Denisovan	Sintashta	Kum6	Gok2	-0.0024	-0.173
Denisovan	Armenia_BA	Kum6	Iceman	-0.0162	-0.92
Denisovan	Karasuk	Kum6	Iceman	0.0047	0.419
Denisovan	Sintashta	Kum6	Iceman	0.027	2.097
Denisovan	Mezhovskaya	Kum6	Iceman	0.0269	2.122
Denisovan	Yamnaya_RISE	Kum6	Iceman	0.0299	2.353
Denisovan	Andronovo	Kum6	Iceman	0.0319	2.861
Denisovan	Armenia_BA	Kum6	ne1	-0.0574	-3.594
Denisovan	Sintashta	Kum6	ne1	-0.0353	-3.016
Denisovan	Karasuk	Kum6	ne1	-0.0235	-2.248
Denisovan	Andronovo	Kum6	ne1	-0.0225	-2.182
Denisovan	Mezhovskaya	Kum6	ne1	-0.0193	-1.719
Denisovan	Yamnaya_RISE	Kum6	ne1	-0.018	-1.499
Denisovan	Armenia_BA	Kum6	Stuttgart	-0.0546	-3.316
Denisovan	Andronovo	Kum6	Stuttgart	-0.0268	-2.477
Denisovan	Mezhovskaya	Kum6	Stuttgart	-0.0276	-2.256
Denisovan	Sintashta	Kum6	Stuttgart	-0.0285	-2.248
Denisovan	Karasuk	Kum6	Stuttgart	-0.0226	-2.015
Denisovan	Yamnaya_RISE	Kum6	Stuttgart	-0.0198	-1.74
Kostenki14 shows higher affinities to Kum6 than to early Neolithic Europeans.					
Outgroup	**pop1**	**pop2**	**pop3**	**D**	**Z**
Denisovan	K14	ATP2	Kum6	0.0126	1.005
Denisovan	K14	co1	Kum6	-0.0047	-0.309
Denisovan	K14	Gok2	Kum6	0.0194	1.239
Denisovan	K14	Iceman	Kum6	0.0017	0.126

Denisovan	K14	ne1	Kum6	0.0306	2.425
Denisovan	K14	Stuttgart	Kum6	0.0336	2.759

Supplemental Experimental Procedures

S1. Archaeological background and Kumtepe

S1.1. Archaeological background

The Neolithic period in Anatolia spans over a period of 6,000 years, starting around 11,000 cal BP and ending around 5,000 cal BP [S17, S18]. The Neolithic sites in Anatolia predate the indications of Neolithization in Europe and it was recognized already a century ago that the farmer-agricultural lifestyle must have spread from here to neighboring regions [S19]. The earliest indications of crop and animal domestication from the Northern Levant and Southeast Anatolia are dated to around 11,700-10,700 cal BP and in Central Anatolia, by 10,500-9,500 BP [S17, S20–S28]. The first indications of plant cultivation and sedentarism appeared in the South Eastern parts of Anatolia almost contemporaneously as in the Levant. Neolithic settlements with early dates are also found in Southern Levant at c.10,500 cal BP [S29] and in Cyprus at c.10,200 cal BP [S30, S31]. Important sites in Anatolia are the large sedentary settlements exhibiting Neolithic economies at Pınarbaşı (10,500 – 10,000 BP), Boncuklu, Aşıklı, and slightly later at Çatalhöyük (9,400 BP) [S32]. Thus, these sites postdate the earliest Neolithic communities by a millennium and are geographically located outside of the region exhibiting the oldest Neolithic communities [S17, S21, S28, S33].

The oldest Neolithic settlements in northwest Anatolia are dated around the beginning of the 9th millennium BP, the final stages of the Pre-Pottery Neolithic[S17, S26, S34, S35]. At Ulucak, close to Izmir, Pre-Pottery Neolithic layers have been dated to c. 9,000 cal BP [S36]. These dates are similar to those of several sites from Greece (at c.8,900-8,400 BP) [S33, S37] which show that the westerly dispersal of Neolithic lifeways was well on the way at this time. However, the initial Neolithic phase in western Anatolia is still not fully understood. It is not until the mid of the 9th millennium BP when larger number of sites with a full “Neolithic package” occur, such as Fikirtepe (c.8,500 cal BP) and other sites around the Sea of Marmara such as Menteşe, Barçın, Ilıpınar [S38, S39]. The 9th Millenium BP saw many sites being established in Western Anatolia and many sites continued to be in use for extended periods after that. The establishment of the settlement at Kumtepe in the Troad around 7,000 BP is probably associated to some level of general reconstruction of the settlements in the area. Changes in the settlement structure and character in the eastern and central parts of Anatolia are probably important for the development also in the western regions of Anatolia and Aegean but the development was complex and dynamic [S17, S28, S40]. Climatic changes causing drier conditions were probably a contributing factor [S41, S42], but this needs further attention [S43, S44].

S1.2. Kumtepe

Kumtepe is a small mound located about two kilometers south of the Dardanelles and 4.5 kilometers northwest of the archaeological site of Troy. It has a dimension of less than 1.4 ha and lies 50 meters from the west arm of the Scamander River [S45, S46]. Initially, four layers were differentiated: Kumtepe IA, IB, IC and II [S47].

Phase IA was settled around 7,000 BP [S45]. This was a time when Neolithic settlements had been found in western Anatolia for almost two millennia [S35]. Geomorphological investigations show that at the time of the first settlement Kumtepe lay directly on the shoreline of a long bay, which was formed by the lower part of the Scamander river and covered by the sea about 7,000-6,000 BP [S48]. Archaeobotanical studies reveal a landscape covered with woodlands dominated by oak trees, but also open areas in the period of Kumtepe IA [S49]. No actual settlement structures were found in sub-phase IA1 and IA2, although, graves, potsherds, flint, obsidian and bone artefacts, shells and animal bones were recovered. A house with a yard was the first remains of a building structure and it was discovered in Kumtepe IA3. IA4 represents a hiatus with natural soil formation [S45].

Kumtepe was abandoned roughly between 6,600-5,500 BP, a period that includes a cooling event in climate (Bond Event no. 4, around 6,000 BP) that may have affected the local environment [S42, S49]. No signs of destruction (e.g fire) mark the end of the first settlement in the IA period. In fact, the absence of valuable and intact objects suggests that the site was

abandoned [S50]. The second settlement phase, Kumtepe IB, is dated to approximately 5,500-4,900 BP [S1]. The architectural remains in the layers of IB and IC have been divided into five building phases: IB1 to IB4 and IC1 [S45]. Archaeobotanical and archaeozoological investigations indicate differences in the subsistence economy between Kumtepe IA and IB [S46]. The subsistence at Kumtepe was from the start largely based on agriculture, animal husbandry, collection of fruits and shells, and to a lower extent on hunting and fishing. The diet of the Kumtepe IA inhabitants consisted mainly of oysters, figs, lentils, bitter vetch, and meat of mostly cattle, but also sheep and fallow deer. In later periods from Kumtepe IB onward cereals and pork were added to the diet and became, together with shells, the main components of the meals. Fishing became more important in the Late Chalcolithic period (Kumtepe IB), while the consumption of figs, fallow deer and beef decreased [S42, S46, S51, S52].

S1.3. The Graves

Four adult skeletons were found during the 1934 test excavation, two female burials lying in a contracted position on the right side on the bedrock in the oldest layer IA, one female skeleton in contracted position was discovered in IB, and one skeleton of a young man in flexed position was found in IC [S47, S53].

A burial ground was discovered in 1994/1995 in layer IA1 which probably extends beyond the excavated area [S54]. Six skeletons were unearthed; five were found in simple pit graves (Graves 6, 7, 9, 10, 12) while one grave (Grave 11) was enclosed by a long slab of limestone on each side with a copper bead and two big potsherds in the fill of the grave. Kum6 was found in one of the pit graves (Grave 6). All bodies were lying in a flexed position on the right side, south-north oriented, with the head pointing south or south-west (Grave 7) facing east. No associated architectural remains were found with the graves. The two first skeletons from Kumtepe IA excavated in 1934 most likely belonged to the same burial ground. Radiocarbon dates for this Kumtepe IA1 graveyard are available from three skeletons (Table S1b).

Four disturbed graves belonging to Kumtepe IA2 were found and the skeletons were in bad conditions. However, it was possible to discern that the dead were buried in a flexed position lying on the right side with the head pointing south. Kum4 was found buried in Kumtepe IB2 layers (Grave 1), a thousand years later from the previous finds, still on the right side in a flexed position and with the head pointing south. The remains of four neonates were also found in the settlement area of Kumtepe IB and IC layers.

Eight skeletal remains, five from IA1, two from IA2 and one from IB1 were attempted for DNA extraction (Table S1a).

S2. DNA Extraction and Library Preparation

Thermal Age [S55] was estimated for Kum6 to assess the biomolecular preservation. Kum6 has a thermal age of 17,7 kyr at 10°C, making it thermally older than all the comparison individuals from different chronologies and environments such as Motala (Motala12 [S9], 5,2kyr at 10°C), Mal'ta (MA-1 [S13], 5,6kyr at 10°C), and Denisova cave (Denisovan [S56], <15kyr at 10°C). This estimation indicates a high level of DNA decay, consistent with the observed poor preservation of DNA in this sample.

Prior to delivery to the clean room facilities dedicated to ancient DNA (aDNA) bones and teeth were decontaminated with chlorine, to further remove contaminants they were UV irradiated (6 J/cm^2 at 254 nm) on both sides, and one millimeter of the surface was removed. The bones were powderized using a dremel drill with a diamond bur. Between 40-120 mg of bone powder was used in each DNA extraction using a variety of silica-based methods [S57–S59]. Each sample was extracted up to three times and at least one negative control was included in each batch of extractions.

Double stranded DNA libraries were prepared using 20µl of extract, with blunt-end ligation as described in [S60] with modifications as in [S12]. The initial nebulization step was omitted since aDNA is already fragmented. Each library was amplified in six replicates, each in a total volume of 25 µl, two negative controls were included in each PCR batch. Each reaction contained, 3 µl DNA library, and the following in final concentrations; 1X AmpliTaqGold Buffer, 2.5mM $MgCl_2$, 250nM of each dNTP, 2.5U AmpliTaqGold (Life Technologies), and 200nM each of the IS4 primer and an Indexed P7primer [S44]. The cycling conditions were 94^o for 10 min followed by 10-14 cycles of 94^o for 30 sec, 60^o for 30 sec, 72^o for 45 sec, and a final extension at 72^o for 10 min. Amplified libraries were pooled and purified with AMPure XP beads

(Agencourt). The libraries were quantified on a 2100 Bioanalyzer using the High Sensitivity Kit (Agilent Technologies). None of the extraction blanks or PCR blanks showed any presence of DNA and were therefore not further sequenced. Between two and 19 libraries were pooled at equimolar concentrations for sequencing on an Illumina HiSeq 2500 at the SNP & SEQ Technology Platform at Science for Life laboratory, Uppsala University. Each pool was sequenced with version 3 chemistry and 100 bp paired-end reads on one or several lanes.

Initial shotgun sequencing results of the material yielded limited amounts of endogenous DNA for six of the individuals and they were thus excluded from further analyses. The remaining two individuals (Kum6, Kum4) yielded variable levels of endogenous DNA in different libraries including those recovered from the same bone fragment (from 0.01% to 8.16%, Table S2)

S3. Sequence Processing and Alignment

We merged read pairs simultaneously as the adapters were removed, requiring an overlap of at least 11 bp and summing up base qualities using MergeReadsFastQ_cc.py [S61]. We collapsed PCR duplicate reads with identical start and end coordinates into consensus sequences using FilterUniqueSAMCons.py [S61]. We mapped the sequence reads to the human genome (build 36 and 37.1) using BWA version 0.5.9 [S62] with parameters -l 16500 -n 0.01 -o 2. We also remapped ancient data from [S6–S10], using the same procedures. We required mismatches to the human reference genome at less than 10% of the positions for each read, and discarded reads of less than 35 bp length.

S3.1 Authentication and contamination estimates

We used PMDtools [S63] for the assessment of nucleotide misincorporations that are characteristic for aDNA [S64] (Figure S1). Mitochondrial DNA contamination estimates for Kum6 were obtained using two different approaches. First, we applied the method by Green et al. [S4] which uses private or nearly private alleles (frequency of less than 5% in 311 modern mitochondrial genomes) in the consensus sequence of Kum6. Sites with coverages of 10 or more and base calls with minimum base quality of 30 were considered informative for the estimation. We also excluded transition polymorphisms if the consensus allele was a C or G to avoid the effect of post-mortem damage. A point estimate of mitochondrial contamination was obtained by adding the counts of the consensus and the alternative bases across all sites, assuming independence among positions. Confidence intervas were obtained by using a binomial approximation.

Second, we used a Bayesian approach which checks whether mitochondrial reads map better to the Kum6 consensus sequence or a set of 311 modern human mtDNAs in order to obtain an estimate of the proportion of authentic sequences [S5]. Contamination estimates for Kum6 are shown in Table S3a.

S3.2 Mitochondrial Haplogroups

We called consensus sequences for the mitochondria of all samples using the mpileup and vcfutils.pl (vcf2fq) tools in the samtools package with default parameters [S65]. Sequence polymorphisms are reported against RSRS [S66] and were phylogenetically analyzed using HaploFind [S67] and PhyloTree Build 16 [S68] to assign sequences to previously annotated haplogroups. The average mitochondrial genome coverage was 21x for Kum6 and 1.5x for Kum4. The Kum6 mitochondrial genome has 39 mutations classifying it as haplogroup H2a (Table S3b). Twelve additional mutations were found in the consensus sequence. One of these is G16274A, which together with T10810C, defines subhaplogroup H2a3. Therefore, Kum6 seems to be an ancestral lineage to H2a3 as it has acquired the defining transition at nucleotide position 16274 but lacks the back mutation at nucleotide position 10810. Nine of the remaining additional mutations are supported by only one or two reads and the majorities of them are unique and not present in PhyloTree. Further, eight of these mutations are C to T or G to A transitions that can likely be attributed to post-mortem damage alterations [S69]. The last additional polymorphism, A6527G, may be a true mutation in Kum6 as it, like most of the 39 haplogroup defining mutations as well as the transitions at nucleotide position 10810 and nucleotide position 16274, is covered by >10 sequence reads.

H is one of the most diverse and common haplogroup in present-day Europe and in the Near East [S70, S71]. The highest frequencies are found in Eastern Europe and Caucasus [S72]. In present-day Turkey, the frequency of haplogroup H is about

25% [S71] and 3.3% of these belong to subhaplogroup H2a [S73]. In ancient European populations, haplogroup H is most commonly associated with Neolithic and subsequent farmer communities [S74] although it has been found in pre-Neolithic hunter-gatherers from the Iberian peninsula as well [S75–S78]. So far, little is known about the mtDNA lineages from the Neolithic Near East with the exception of an ancient mtDNA study on remains from Syria where H was reported as the third most prevalent haplogroup (14.28%) after K (42.80%) and R0 (21.42%) [S79].

S3.3 Genetic Sex Identification

To assess the biological sex of both individuals, we counted the number of reads mapping to the X and Y chromsomes with a quality of at least 30 and compared the ratio to a reference panel [S80]. Both samples could be confidently assigned as females which also confirm the osteological analysis. (Table S3c).

S4. Kum4

This individual yielded low levels of clonality (i.e. low number of PCR duplicates), even though informative amounts of endogenous DNA was obtained in the initial sequencing, the yield from the same library was significantly lower when re-sequenced and it was decided to not continue further sequencing. At this stage we had obtained 0.01x genome coverage and 1.5x mitochondrial coverage. Given the low coverage, mitochondrial DNA contamination estimates were not feasible. However, nucleotide misincorporation patterns, a common type of chemical feature observed in ancient DNA, were detected validating our sequence data (Figure S1). Principal component analyses were performed using one SNP Chip reference panel [S81] (Human Origins data, Figure S2a) and one whole genome data set [S82] (1000 genomes, Figure S2c). With this analysis, we were able to confirm the position of this individual falling within the southern-European variation, following a similar tendency to Kum6 (see Main text and Figure S2b and S2d). We computed outgroup *f3*-statistics in order to test to determine the levels of shared genetic drift between each ancient individual and a set of modern European, Near Eastern, Asian and North African populations (as described in section S6.2). This analysis revealed a closer affinity for Kum4 with Sardinians, consistent with what has been previously observed in early European farmers [S6, S9, S83]. Further, we calculated *D*-statistics for all ancient genomes and Kum4 using the Denisovan genome [S56] as outgroup. We observed that Kum4 showed more affinities with early farmers than with hunter-gatherer individuals. However, comparisons within the farming individuals (Denisovan, Early farmer1; Kum4, Early farmer 2) yielded no significant results due to the low coverage of Kum4. Therefore, our main conclusions are based solely on the analyses of Kum6.

S5. Comparative data

S5.1 SNP genotype data from modern populations

The ancient individuals were merged with the SNP genotype calls of the Human Origins data set [S81]. For each site and ancient individual, we randomly picked one read with minimum mapping quality 30 covering that site and used the respective base as allele if its base quality was 30 or higher. Sites showing indels were excluded. We also excluded all transition sites where the ancient individual carried a T or A in order to avoid potential post-mortem damage. Ancient individuals used in this study and the number of SNPs retained after merging with the Human Origins data set are shown in Table S4a.

S5.2 Sequence data from modern populations

VCF and BAM files for the Yoruban individuals sequenced in phase 3 of the 1000 genomes project were obtained from ftp.1000genomes.ebi.ac.uk. We used vcftools [S84] to extract all SNPs with a minor allele frequency of at least 10% in the Yoruban population and excluded all transition polymorphisms to avoid potential post-mortem damage. The remaining 1,938,919 SNPs were merged with the ancient genomes and Denisovan [S56] (as an outgroup) as described above. The number of SNPs per individual is shown in Table S4b.

S5.3 SNP genotype data from ancient individuals

We also incorporated individuals from a recent SNP capture study [S11] in parts of our analysis. Genotype calls for 354,212 SNPs were downloaded from HYPERLINK "http://genetics.med.harvard.edu/reich/Reich_Lab/Datasets.html",

http://genetics.med.harvard.edu/reich/Reich_Lab/Datasets.html. We extracted the individuals with the highest number of SNP's for the following populations/cultures: Starcevo_EN (I0174), LBKT_EN (I0176), LBK_EN (I0100), Spain_EN (I0412), Spain_MN (I0408), Esperstedt_MN (I0172), Baalberge_MN (I0560), Yamnaya (I0231), Corded_Ware_LN (I0103), Karlsdorf_LN (I0550), BenzigerodeHeimburg_LN (I0058), Bell_Beaker_LN (I0112), and Alberstedt_LN (I0118). Due to the nature of the data and the overlap of sites with the Human Origins array, we treated the data the same way as the genotype data in S5.1: We randomly sampled one allele at overlapping heterozygous sites and coded sites without information as missing data.

S6. Population Genetic Analysis

S6.1 Principal Component Analysis

We performed principal component analysis (PCA) to compare ancient samples with reference to modern populations. Each ancient individual was merged separately with Western Eurasian populations from the Human Origins data set [S9, S81] as described in section S5.1. To account for the non-diploid data of ancient individuals due to the low sequencing coverage, a random allele was selected for each heterozygous modern individual, which made the data set completely homozygous. PCAs were conducted using the *smartpca* tool (version 10210) of the EIGENSOFT package [S81], randomly excluding one SNP of a pair of SNPs with a linkage disequilibrium of $r^2 > 0.7$. In order to present all ancient individuals in one plot (Figure 1B), we used Procrustes analysis [S81, S85] to transform each individual's PC1 and PC2 loadings to the coordinate system of the PCA using all SNPs for the modern populations. The original coordinates of each ancient individuals were set to (0,0), assuming that their influence on the axes of variation is negligible. PC1 and PC2 loadings for the reference individuals were then averaged across all separate analyses. The individual PCA's for both ancient individuals are shown in Figure S2a-e.

We repeated this analysis for Kum4 using 1000 genomes data to gain more power with less sample size but more SNPs. Indeed we recovered 14436 overlapping SNPs as opposed to 4308 SNPs from the previous data set. We wanted to confirm the position of Kum4 and it did indeed group with the Spanish populations in the PCA with 1000 genomes also. (Figure S2c)

S6.2. Outgroup f_3-statistics

In order to measure genetic relationship between two populations - independent of an excess of drift in low-quality samples - an outgroup f_3-statistic[S13, S81]; f3(O; A, B)) was calculated as:

$$f_3(O;A,B)=\frac{\sum (p_O-p_A)(p_O-p_B)-(p_O-p_O^2)/(n_O-1)}{\sum 2\,p_O(1-p_O)}$$

p_O is the allele frequency of the arbitrarily chosen reference allele and n_O is the number of chromosomes in the outgroup population at locus i, with corresponding notations for populations A and B. In the absence of admixture with the outgroup (we use Sub-Saharan Africans to assure that), we expect the value of the statistic to be positive. A positive deviation from zero will be a function of the shared genetic history of two populations A and B.

S6.3 *D*-statistics

To test deviations from a tree population topology of the shape $((A,B);(X,Y))$, we used ADMIXTOOLS to calculate *D*-statistics [S81] :

$$D(A,B;X,Y)=\frac{\sum_{i=1}^{n}\left[\left(p_{iA}-p_{iB}\right)\left(p_{iX}-p_{iY}\right)\right]}{\sum_{i=1}^{n}\left[\left(p_{iA}+p_{iB}-2\,p_{iA}\,p_{iB}\right)\left(p_{iX}+p_{iY}-2\,p_{iX}\,p_{iY}\right)\right]}$$

where p_{iA} is the frequency of one arbitrarily chosen allele in population A at marker i. Numerator and denominator are summed across all sites. Standard errors were then obtained by performing a block jackknife over blocks of 0.5 Mbp. Significant deviations from 0 can be interpreted as deviations from a tree like population history due to gene flow, where positive values suggest excess affinity between A and X and/or Y and B and negative values excess affinity between A and Y and/or B and X. Using an outgroup population (Denisovan in this study) as population A narrows these options down to gene flow between B and X (in case of positive values) or between B and Y (if values are negative). Individuals with a genome coverage of less than 0.5x as well as individuals from a recent SNP capture study [S11] were not included in *D*-statistics due to low numbers of overlapping SNPs with the low coverage Kumtepe individuals, which limited the power to see any clear signals. (Figure 4 and Table S5)

S7. Admixture Graph Inference

TreeMix v1.12 [S86] was used to infer admixture graphs. TreeMix estimates a maximum-likelihood tree from the covariance matrix of allele frequencies and then adds migration edges to account for residual covariance. We included the high coverage Denisovan genome (used as root), Yorubans from the 1000 genomes project [S82] (used for SNP ascertainment), MA1 [S13] (as representative of ancient North Eurasians) and the highest coverage individual from each of the ancient European groups (Kum6, Gok2, Ajv58, NE1, Loschbour, Stuttgart, Motala12, Iceman, LaBraña) [S6–S10]. Overlapping sequence data was available for these individuals at 46,683 transversion SNPs. Correction for low samples sizes was turned off (-noss) since all groups consistent of single individuals; and covariances were estimated using blocks of 500 SNPs.

We tested different numbers of migration edges to account for residual covariance not explained by the tree structure. Each setting was run with ten different random seeds and we report the most common topology per number of migration edges. The general split between farmers and hunter-gatherers was observed in all trees. Kum6 clearly grouped with farmers (Figure 4B), as a sister group to Iceman which is consistent with the results from *D*-statistics and outgroup f_3-statistics (Figure 2). Adding one and two migration edges confirmed known connections between Mesolithic Scandinavia and Paleolithic Siberia [S9] as well as between Scandinavian farmers and hunter-gatherers [S6] Higher numbers of migrations did not show a single most common scenario and migrations often involved outgroups (not shown). This is most likely due to the low number of SNPs in the analysis, as well as low levels of admixture involving Denisovans [S56] and Sub-Saharan Africans [S87] or ascertainment bias.

S8. Model-based clustering

Individuals were clustered using an unsupervised clustering algorithm as implemented in ADMIXTURE [S88]. ADMIXTURE assigns proportions of each individual's genome to one of K (user defined number of clusters) ancestral populations. We included all Eurasian and North African populations from the Human Origins data set [S81]; the ancient individuals with more than 1x coverage, the individual with the highest number of SNPs per early Neolithic group from a recent SNP capture approach [S11], the highest coverage Bronze Age Asian sample [S16] if its coverage was 0.5x or higher and Kum6. In contrast to the PCA (Supplementary Information S6.1), we only use the highest coverage individual per culture from Haak et al. [S11] as well as Allentoft et al. [S16] in order to avoid biases due to sample sizes.

All ancient individuals were merged with the Human Origins data (as described in Supplementary S5.1). Transition SNPs were coded as missing data since they are prone to post-mortem DNA damage. We also picked one random allele for each modern individual at heterozygous sites. Each data set was then thinned for linkage disequilibrium using PLINK 1.07 [S89].

Pairwise LD was calculated in windows of 200 SNPs with a step size of 25 SNPs and an r2 threshold of 0.7. We explored between 2 and 12 clusters using 50 replicates per K. Common signals between the different replicates were identified using the LargeKGreedy mode of CLUMPP [S90]. Clustering on a geographic map was visualized using the GNU R package rworldmap. We display the results for K=9 in the main manuscript (Figure 3) since lower numbers of clusters did not separate European early farmers, hunter-gatherers and modern-day Near Eastern populations into different clusters (Figure S3).

ADMIXTURE assumes perfect genotype calls without any errors. This scenarios is obviously unrealistic for low coverage sequence data so we additionally checked the results of ADMIXTURE using NGSadmix [S91] which works with genotype likelihoods and, thus, accounts for the uncertainty involved in the analysis of genotype calls obtained from low coverage NGS data. We used the same ancient and modern individuals and sites as for the ADMIXTURE analysis. Genotype data published by Haak et al. [S11] was assumed to be without any error, so each individual's genotype got called with a likelihood of 1.0 or a likelihood of one third for all possible genotypes in the case of missing data. We then used ANGSD [S92] to estimate genotype likelihoods for all sequenced ancient individuals using the samtools model, minimum mapping and base qualities of 20 and trimming the last 5bps of each sequenced fragment. NGSadmix was run for 2 to 12 clusters with 25 iterations per K to find the most likely solution. Taking the uncertainty into account made the results noisier and caused some inconsistencies among consecutive values of K. The results for the ancient individuals, however, are qualitatively consistent with the results of ADMIXTURE showing that Kum6 has the highest proportion of West Asian ancestry among all early farmers across different values of K (not shown).

Supplemental References

S1. Kromer, B., Korfmann, M., and Jablonka, P. (2003). Heidelberg radiocarbon dates for Troia I to VIII and Kumtepe. In Troia and the Troad Natural Science in Archaeology., G. A. Wagner, E. Pernicka, and H.-P. Uerpmann, eds. (Springer Berlin Heidelberg), pp. 43–54.

S2. Reimer, P. (2013). IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon *55*, 1869–1887.

S3. Bronk Ramsey, C. (2013). Recent and Planned Developments of the Program OxCal. Radiocarbon *55*. Available at: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/16215 [Accessed December 5, 2015].

S4. Green, R. E., Malaspinas, A.-S., Krause, J., Briggs, A. W., Johnson, P. L. F., Uhler, C., Meyer, M., Good, J. M., Maricic, T., Stenzel, U., et al. (2008). A Complete Neandertal Mitochondrial Genome Sequence Determined by High-Throughput Sequencing. Cell *134*, 416–426.

S5. Fu, Q., Mittnik, A., Johnson, P. L. F., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., et al. (2013). A Revised Timescale for Human Evolution Based on Ancient Mitochondrial Genomes. Current Biology *23*, 553–559.

S6. Skoglund, P., Malmström, H., Omrak, A., Raghavan, M., Valdiosera, C., Günther, T., Hall, P., Tambets, K., Parik, J., Sjögren, K.-G., et al. (2014). Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. Science *344*, 747–750.

S7. Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., et al. (2012). New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat Commun *3*, 698.

S8. 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.

S9. 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.

S10. 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*.

S11. 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.

S12. 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. PNAS *112*, 11917–11922.

S13. Raghavan, M., Skoglund, P., Graf, K. E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford Jr, T. W., Orlando, L., Metspalu, E., et al. (2014). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature *505*, 87–91.

S14. 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.

S15. 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). Genomic structure in Europeans dating back at least 36,200 years. Science *346*, 1113–1118.

S16. 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.

S17. Özdoğan, M. (2011). Archaeological Evidence on the Westward Expansion of Farming Communities from Eastern Anatolia to the Aegean and the Balkans. Current Anthropology *52*, 415–430.

S18. Yakar, J. (2011). Anatolian chronology and terminology. In The Oxford Handbook of Ancient Anatolia: (10,000-323 BCE) (OUP USA), pp. 56–93.

S19. Childe, G. V. (1925). The Dawn of European Civilization (London: Kegan Paul).

S20. Arbuckle, B. S. (2013). The late adoption of cattle and pig husbandry in Neolithic Central Turkey. Journal of Archaeological Science *40*, 1805–1815.

S21. Baird, D. (2012). The Late Epipaleolithic, Neolithic, and Chalcolithic of the Anatolian Plateau, 13,000–4000 BC. In A Companion to the Archaeology of the Ancient Near East, D. T. Potts, ed. (Wiley-Blackwell), pp. 431–465.

S22. Bartl, K. (2012). The Northern Levant. In A Companion to the Archaeology of the Ancient Near East, D. T. Potts, ed. (Wiley-Blackwell), pp. 373–395.

S23. Çakırlar, C. (2012). The evolution of animal husbandry in Neolithic central-west Anatolia: the zooarchaeological record from Ulucak Höyük (c. 7040–5660 cal. BC, Izmir, Turkey). Anatolian Studies *62*, 1–33.

S24. Conolly, J., Colledge, S., Dobney, K., Vigne, J.-D., Peters, J., Stopp, B., Manning, K., and Shennan, S. (2011). Meta-analysis of zooarchaeological data from SW Asia and SE Europe provides insight into the origins and spread of animal husbandry. Journal of Archaeological Science *38*, 538–545.

S25. Düring, B. S. (2010). The Prehistory of Asia Minor: From Complex Hunter-Gatherers to Early Urban Societies

(Cambridge University Press).

S26. Özdoğan, M. (1997). The Beginning of Neolithic Economies in Southeastern Europe: an Anatolian Perspective. Journal of European Archaeology *5*, 1–33.

S27. Özdoğan, M. (1997). Anatolia from the Last Glacial Maximum to the Holocene Climatic Optimum: cultural transformations and the impact of the environmental setting. Paléorient *23/2*, 25–38.

S28. Özdoğan, M. (2014). Anatolia: From the Pre-Pottery Neolithic to the End of the Early Bronze Age (10,500–2000 bce). In The Cambridge World Prehistory 3 Volume Set, C. Renfrew and P. Bahn, eds. (Cambridge: Cambridge University Press), pp. 1508–1544.

S29. Twiss, K. C. (2007). The Neolithic of the southern Levant. Evol. Anthropol. *16*, 24–35.

S30. Şevketoğlu, M. (2006). Cypro-Anatolian relations on the 9th Millenium BC: Akanthou/Tatlisu Rescue Excavation. Anatolia *30*, 119–136.

S31. Vigne, J.-D., Briois, F., Zazzo, A., Willcox, G., Cucchi, T., Thiébault, S., Carrère, I., Franel, Y., Touquet, R., Martin, C., et al. (2012). First wave of cultivators spread to Cyprus at least 10,600 y ago. PNAS *109*, 8445–8449.

S32. Hodder, Ian (2011). Çatalhöyük: A Prehistoric Settlement on the Konya Plain. In The Oxford Handbook of Ancient Anatolia: (10,000-323 BCE) (OUP USA), pp. 934–949.

S33. Brami, M. N. (2014). A graphical simulation of the 2,000-year lag in Neolithic occupation between Central Anatolia and the Aegean basin. Archaeol Anthropol Sci, 1–9.

S34. Schoop, U. D. (2011). The Chalcolithic on the Plateau. In The Oxford Handbook of Ancient Anatolia: (10,000-323 BCE) (OUP USA), pp. 150–173.

S35. Özdoğan, E. (2014). Current Research and New Evidence for the Neolithization Process in Western Turkey. Eur. J. Archaeol. *18*, 33–59.

S36. Çilingiroglu, Ç., and Çakırlar, C. (2013). Towards configuring the neolithisation of Aegean Turkey. Documenta Praehistorica *40*, 21–29.

S37. Perlès, C. (2001). The Early Neolithic in Greece: The First Farming Communities in Europe (Cambridge University Press).

S38. Roodenberg, J. (2011). Ilipinar: A Neolithic Settlement in the Eastern Marmara Region. In The Oxford Handbook of Ancient Anatolia: (10,000-323 BCE) (OUP USA), pp. 950–967.

S39. Yalçikli, D. (2014). New Evidence for Neolithic Habitation in the Hinterland of the Troad. Oxford Journal of Archaeology *33*, 245–255.

S40. Reingruber, A. (2011). Early Neolithic settlement patterns and exchange networks in the Aegean. Documenta Praehistorica *38*, 291.

S41. Berger, J.-F., and Guilaine, J. (2009). The 8200 cal BP abrupt environmental change and the Neolithic transition: A Mediterranean perspective. Quaternary International *200*, 31–49.

S42. Riehl, S., and Marinova, E. (2007). Mid-Holocene vegetation change in the Troad (W Anatolia): man-made or natural? Veget Hist Archaeobot *17*, 297–312.

S43. Asouti, E. (2009). The relationship between Early Holocene climate change and Neolithic settlement in Central Anatolia, Turkey: current issues and prospects for future research. Documenta Praehistorica, 1–5.

S44. Lillie M., Budd C., Alpaslan-Roodenberg S., Karul N., and Pinhasi R. (2012). Musings on Early Farming Communities in Northwest Anatolia; and other Flights of Fancy. Interdisciplinaria Archaeologica. Natural Sciences in Archaeology *III(1)*, 11–22.

S45. Bertram, J.-K., and Karul, N. (2014). Anmerkungen zur Stratigraphie des Kumtepe. Die Ergebnisse der Grabungen in den Jahren 1994 und 1995. In Landschaftsgeschichte der Troas. Bioarchäologische Forschungen (Dr. Rudolf Habelt GmbH), pp. 1059–1084. Available at: https://lirias.kuleuven.be/handle/123456789/463295 [Accessed July 1, 2015].

S46. Riehl, S. (1999). Bronze Age environment and economy in the Troad: the archaeobotany of Kumtepe and Troy. Mo Vince-Verlag.

S47. Sperling, J. W. (1976). Kum Tepe in the Troad: Trial Excavation, 1934. Hesperia: The Journal of the American School of Classical Studies at Athens *45*, 305–364.

S48. Kayan, I. (2014). Geoarchaeological Research at Troia and its Environs. In Troia 1987-2012: Grabungen und Forschungen I. Forschungsgeschichte, Methoden und Landschaft, E. Pernicka, C. Rose, and P. Jablonka, eds. (Bonn: Dr. Rudolf Habelt GmbH), pp. 694–727.

S49. Riehl, S., Marinova, E., and Uerpmann, H.-P. (2014). Landschaftsgeschichte der Troas. Bioarchäologische Forschungen. In Troia 1987-2012: Grabungen und Forschungen I. Forschungsgeschichte, Methoden und Landschaft, E. Pernicka, C. Rose, and P. Jablonka, eds. (Bonn: Dr. Rudolf Habelt GmbH), pp. 732–769.

S50. Gabriel, U. (2014). Die Keramik der troadischen Fundorte Kumtepe IA, Beşik-Sivritepe und Çıplak Köyü im Kontext ihrer überregionalen Vergleichsfunde. In Troia 1987-2012: Grabungen und Forschungen I. Forschungsgeschichte, Methoden und Landschaft, E. Pernicka, C. Rose, and P. Jablonka, eds. (Bonn: Dr. Rudolf Habelt GmbH), pp. 990–1058.

S51. Gündem, C. Y. (2009). Animal Based Economy in Troia and the Troas During The Maritime Troy Culture (c. 3000-2200 BC.) and a General Summary for West Anatolia.

S52. Uerpmann, M. (2006). Von Adler bis Zahnprassen - Der Beitrag der Archäozoologie zur Erforschung Troias. In Troia. Archäologie eines Siedlungshügels und seiner Landschaft, M. O. Korfmann, ed. (Zabern Verlag Mainz), pp. 355–360.

S53. Angel, J. L., and Blegen, C. W. (1951). Troy: The Human Remains. (Princeton University Press).

S54. Korfmann, M., Girgin, Ç., Morçöl, Ç., and Kılıç, S. (1995). Kumtepe 1993: Report on the Rescue Excavation. In Studia Troica (Zabern Verlag Mainz), pp. 237–261.

S55. Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C., and Collins, M. J. (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. Journal of Human Evolution *45*, 203–217.

S56. Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J. G., Jay, F., Prüfer, K., Filippo, C. de, et al. (2012). A High-Coverage Genome Sequence from an Archaic Denisovan Individual. Science *338*, 222–226.

S57. Svensson, E. M., Anderung, C., Baubliene, J., Persson, P., Malmström, H., Smith, C., Vretemark, M., Daugnora, L., and Götherström, A. (2007). Tracing genetic change over time using nuclear SNPs in ancient and modern cattle. Animal Genetics *38*, 378–383.

S58. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., et al. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. PNAS *110*, 15758–15763.

S59. Rohland, N., and Hofreiter, M. (2007). Ancient DNA extraction from bones and teeth. Nat. Protocols *2*, 1756–1762.

S60. Meyer, M., and Kircher, M. (2010). Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture

and Sequencing. Cold Spring Harb Protoc *2010*, pdb.prot5448.

S61. Kircher, M. (2012). Analysis of High-Throughput Ancient DNA Sequencing Data. In Ancient DNA Methods in Molecular Biology., B. Shapiro and M. Hofreiter, eds. (Humana Press), pp. 197–228.

S62. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

S63. Skoglund, P., Northoff, B. H., Shunkov, M. V., Derevianko, A. P., Pääbo, S., Krause, J., and Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. PNAS *111*, 2229–2234.

S64. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., and Pääbo, S. (2012). Temporal Patterns of Nucleotide Misincorporations and DNA Fragmentation in Ancient DNA. PLoS ONE *7*, e34131.

S65. 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.

S66. Behar, D. M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N. M., Kivisild, T., Torroni, A., and Villems, R. (2012). A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from its Root. The American Journal of Human Genetics *90*, 675–684.

S67. Vianello, D., Sevini, F., Castellani, G., Lomartire, L., Capri, M., and Franceschi, C. (2013). HAPLOFIND: A New Method for High-Throughput mtDNA Haplogroup Assignment. Human Mutation *34*, 1189–1194.

S68. van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. *30*, E386–E394.

S69. Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M., et al. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. PNAS *104*, 14616–14621.

S70. Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., et al. (2000). Tracing European Founder Lineages in the Near Eastern mtDNA Pool. The American Journal of Human Genetics *67*, 1251–1276.

S71. Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., et al. (2004). The Molecular Dissection of mtDNA Haplogroup H Confirms That the Franco-Cantabrian Glacial Refuge Was a Major Source for the European Gene Pool. The American Journal of Human Genetics *75*, 910–918.

S72. Pereira, L., Richards, M., Goios, A., Alonso, A., Albarrán, C., Garcia, O., Behar, D. M., Gölge, M., Hatina, J., Al-Gazali, L., et al. (2005). High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome Res *15*, 19–24.

S73. Roostalu, U., Kutuev, I., Loogväli, E.-L., Metspalu, E., Tambets, K., Reidla, M., Khusnutdinova, E. K., Usanga, E., Kivisild, T., and Villems, R. (2007). Origin and Expansion of Haplogroup H, the Dominant Human Mitochondrial DNA Lineage in West Eurasia: The Near Eastern and Caucasian Perspective. Mol Biol Evol *24*, 436–448.

S74. Brandt, G., Haak, W., Adler, C. J., Roth, C., Szécsényi-Nagy, A., Karimnia, S., Möller-Rieker, S., Meller, H., Ganslmeier, R., Friederich, S., et al. (2013). Ancient DNA Reveals Key Stages in the Formation of Central European Mitochondrial Genetic Diversity. Science *342*, 257–261.

S75. Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A., Cabrera, V. M., and de la Rúa, C. (2012). Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. PLoS ONE *7*, e34417.

S76. Gamba, C., Fernández, E., Tirado, M., Deguilloux, M. F., Pemonge, M. H., Utrilla, P., Edo, M., Molist, M., Rasteiro, R., Chikhi, L., et al. (2012). Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. Molecular Ecology *21*, 45–56.

S77. Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Duranthon, F., Guilaine, J., Crubézy, E., and Ludes, B. (2011). Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. PNAS *108*, 9788–9791.

S78. Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., Bosch, A., Guilaine, J., Crubézy, E., and Ludes, B. (2011). Ancient DNA suggests the leading role played by men in the Neolithic dissemination. PNAS *108*, 18255–18259.

S79. Fernández, E., Pérez-Pérez, A., Gamba, C., Prats, E., Cuesta, P., Anfruns, J., Molist, M., Arroyo-Pardo, E., and Turbón, D. (2014). Ancient DNA Analysis of 8000 B.C. Near Eastern Farmers Supports an Early Neolithic Pioneer Maritime Colonization of Mainland Europe through Cyprus and the Aegean Islands. PLoS Genet *10*, e1004401.

S80. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. Journal of Archaeological Science *40*, 4477–4482.

S81. 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.

S82. McVean, G. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

S83. 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.

S84. 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.

S85. Wang, C., Szpiech, Z. A., Degnan, J. H., Jakobsson, M., Pemberton, T., Hardy, J. A., Singleton, A. B., and Rosenberg, N. A. (2010). Comparing Spatial Maps of Human Population-Genetic Variation Using Procrustes Analysis. Statistical Applications in Genetics and Molecular Biology *9(1)*, Article 13.

S86. Pickrell, J. K., and Pritchard, J. K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. PLoS Genet *8*, e1002967.

S87. Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P. H., de Filippo, C., et al. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. Nature *505*, 43–49.

S88. Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

S89. 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. The American Journal of Human Genetics *81*, 559–575.

S90. Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics *23*, 1801–1806.

S91. Skotte, L., Korneliussen, T. S., and Albrechtsen, A. (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data. Genetics *195*, 693–702.