LETTER

doi:10.1038/nature24476

Parallel palaeogenomic transects reveal complex genetic history of early European farmers

Mark Lipson[1]*, Anna Szécsényi-Nagy[2]*, Swapan Mallick[1,3], Annamária Pósa[2], Balázs Stégmár[2], Victoria Keerl[4], Nadin Rohland[1], Kristin Stewardson[1,5], Matthew Ferry[1,5], Megan Michel[1,5], Jonas Oppenheimer[1,5], Nasreen Broomandkhoshbacht[1,5], Eadaoin Harney[1,5], Susanne Nordenfelt[1], Bastien Llamas[6], Balázs Gusztáv Mende[2], Kitti Köhler[2], Krisztián Oross[2], Mária Bondár[2], Tibor Marton[2], Anett Osztás[2], János Jakucs[2], Tibor Paluch[7], Ferenc Horváth[7], Piroska Csengeri[8], Judit Koós[8], Katalin Sebők[9], Alexandra Anders[9], Pál Raczky[9], Judit Regenye[10], Judit P. Barna[11], Szilvia Fábián[12], Gábor Serlegi[2], Zoltán Toldi[13], Emese Gyöngyvér Nagy[14], János Dani[14], Erika Molnár[15], György Pálfi[15], László Márk[16,17,18,19], Béla Melegh[18,20], Zsolt Bánfai[18,20], László Domboróczki[21], Javier Fernández-Eraso[22], José Antonio Mujika-Alustiza[22], Carmen Alonso Fernández[23], Javier Jiménez Echevarría[23], Ruth Bollongino[4], Jörg Orschiedt[24,25], Kerstin Schierhold[26], Harald Meller[27], Alan Cooper[6,28], Joachim Burger[4], Eszter Bánffy[2,29], Kurt W. Alt[30,31,32], Carles Lalueza-Fox[33], Wolfgang Haak[6,34] & David Reich[1,3,5]

Ancient DNA studies have established that Neolithic European populations were descended from Anatolian migrants[1–8] who received a limited amount of admixture from resident hunter-gatherers[3–5,9]. Many open questions remain, however, about the spatial and temporal dynamics of population interactions and admixture during the Neolithic period. Here we investigate the population dynamics of Neolithization across Europe using a high-resolution genome-wide ancient DNA dataset with a total of 180 samples, of which 130 are newly reported here, from the Neolithic and Chalcolithic periods of Hungary (6000–2900 BC, $n = 100$), Germany (5500–3000 BC, $n = 42$) and Spain (5500–2200 BC, $n = 38$). We find that genetic diversity was shaped predominantly by local processes, with varied sources and proportions of hunter-gatherer ancestry among the three regions and through time. Admixture between groups with different ancestry profiles was pervasive and resulted in observable population transformation across almost all cultural transitions. Our results shed new light on the ways in which gene flow reshaped European populations throughout the Neolithic period and demonstrate the potential of time-series-based sampling and modelling approaches to elucidate multiple dimensions of historical population interactions.

The population dynamics of the Neolithization process are of great importance for understanding European prehistory[10–13]. The first quantitative model of the Neolithic transition to integrate archaeological and genetic data was the demic diffusion hypothesis[10], which posited that growing population densities among Near Eastern farmers led to a range expansion that spread agriculture to Europe. Ancient DNA analysis has validated major migrations from populations related to Neolithic Anatolians as driving the introduction of farming in Europe[1–8], but the demic diffusion model does not account for the complexities of the interactions between farmers and hunter-gatherers in Europe throughout the Neolithic period[11–16]. For example, ancient DNA analyses have shown that farmers traversed large portions of Europe with limited initial admixture from hunter-gatherers[3,5,7,8] and, furthermore, that farmers and hunter-gatherers lived in close proximity in some locations long after the arrival of agriculture[15,16]. However, genetic data have not been used systematically to model population interactions and transformations during the course of the Neolithic period. Key open questions include whether migrating farmers mixed with hunter-gatherers at each stage of the expansion (and, if so, how soon after arriving this occurred) and whether the previously observed increase in hunter-gatherer ancestry among farmers in several parts of Europe by the Middle Neolithic period[5–9] represented a continuous versus discrete process and a continent-wide phenomenon versus a collection of parallel, local events.

We compiled a high-resolution dataset of 180 Neolithic and Chalcolithic European genomes (pre-dating the arrival of steppe ancestry in the third millennium BC (ref. 5)) from what are now Hungary, Germany and Spain, of which 130 individuals are newly reported here, 45 with new direct radiocarbon dates (Table 1, Fig. 1a, b, Extended Data Tables 1, 2, Supplementary Tables 1, 2 and Supplementary Information sections 1–3). We enriched for DNA fragments covering a set of approximately 1.23 million single-nucleotide polymorphism (SNP) targets[7] and called one allele at random per site, obtaining mostly high-quality data, with at least 100,000 SNPs hit at least once (average coverage around 0.1 or higher) for 90 of the 130 samples (Methods). Most (90) of our new samples comprise an approximately 3,000-year transect of the prehistory of the Carpathian Basin (Supplementary Information section 1), from both the eastern (Great Hungarian Plain or Alföld) and western (Transdanubia) regions of present-day

[1]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences, Budapest 1097, Hungary. [3]Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [4]Institute of Organismic and Molecular Evolution, Johannes Gutenberg University Mainz, Mainz 55128, Germany. [5]Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. [6]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia. [7]Móra Ferenc Museum, Szeged 6720, Hungary. [8]Herman Ottó Museum, Miskolc 3529, Hungary. [9]Institute of Archaeological Sciences, Eötvös Loránd University, Budapest 1088, Hungary. [10]Laczkó Dezso˝ Museum, Veszprém 8200, Hungary. [11]Balaton Museum, Keszthely 8360, Hungary. [12]Department of Archaeological Excavations and Artefact Processing, Hungarian National Museum, Budapest 1088, Hungary. [13]Jósa András Museum, Nyíregyháza 4400, Hungary. [14]Déri Museum, Debrecen 4026, Hungary. [15]Department of Biological Anthropology, Szeged University, Szeged 6726, Hungary. [16]Department of Biochemistry and Medical Chemistry, University of Pécs, Pécs 7624, Hungary. [17]Imaging Center for Life and Material Sciences, University of Pécs, Pécs 7624, Hungary. [18]Szentágothai Research Center, University of Pécs, Pécs 7624, Hungary. [19]PTE-MTA Human Reproduction Research Group, Pécs 7624, Hungary. [20]Department of Medical Genetics and Szentágothai Research Center, University of Pécs, Pécs 7624, Hungary. [21]Dobó István Castle Museum, Eger 3300, Hungary. [22]Department of Geography, Prehistory, and Archaeology, University of the Basque Country, Investigation Group IT622-13, Vitoria-Gasteiz 01006, Spain. [23]CRONOS SC, Burgos 09007, Spain. [24]Department of Prehistoric Archaeology, Free University of Berlin, Berlin 14195, Germany. [25]Curt-Engelhorn-Centre Archaeometry gGmbH, Mannheim 68159, Germany. [26]Commission for Westphalian Antiquities, Westphalia-Lippe Regional Association, 48157 Münster, Germany. [27]State Office for Heritage Management and Archaeology Saxony-Anhalt and State Heritage Museum, Halle 06114, Germany. [28]Environment Institute, University of Adelaide, Adelaide, South Australia 5005, Australia. [29]Romano-Germanic Commission, German Archaeological Institute, Frankfurt am Main 60325, Germany. [30]Center of Natural and Cultural History of Man, Danube Private University, Krems-Stein 3500, Austria. [31]Department of Biomedical Engineering, University of Basel, Allschwil 4123, Switzerland. [32]Institute for Integrative Prehistory and Archaeological Science, University of Basel, Basel 4055, Switzerland. [33]Institute of Evolutionary Biology (CSIC-UPF), Barcelona 08003, Spain. [34]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena 07745, Germany.
*These authors contributed equally to this work.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

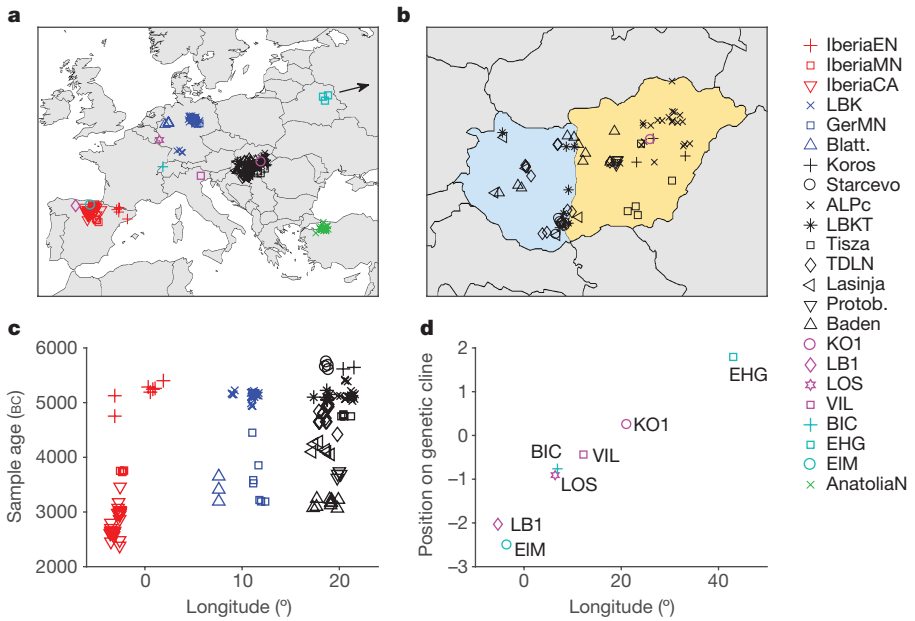


Figure 1 | Spatial and temporal contexts of European Neolithic samples. **a**, **b**, Locations of samples used for analyses, with close-up of Hungary (yellow shading for Alföld and blue for Transdanubia). **c**, Sample ages arranged by longitude. **d**, Hunter-gatherer genetic cline (derived from multidimensional scaling analysis; Supplementary Information section 5) as a function of longitude. The four primary WHG individuals are shown together with 'BIC' (Bichon, around 11700 BC from Switzerland[30]), 'EHG' (eastern hunter-gatherers, 7000–5000 BC from Russia[5,7]) and 'ElM' (El Mirón, around 17000 BC from Spain[20]). Random jitter is added to separate overlapping positions in **a**–**c**. GerMN, Germany Middle Neolithic; Blatt., Blätterhöhle; Protob., Protoboleráz. Map image data from Esri and DeLorme.

Hungary. For our primary analyses, we retained 104 samples from 15 population groupings (Table 1 and Methods), which we merged with 50 Neolithic individuals from the literature[4,5,7,17,18]. We co-analysed these samples with 25 Neolithic individuals (around 6500–6000 BC) from northwestern Anatolia[7] to represent the ancestors of the first European farmers (FEF; Supplementary Information section 4) and four primary European hunter-gatherer individuals[4,7,17,19,20] (WHG, western hunter-gatherers; Table 1).

A principal component analysis of our samples shows that, as expected, all of the Neolithic individuals fall along a cline of admixture between FEF and WHG (Extended Data Fig. 1). Y-chromosome diversity also indicates contributions from ancestral Anatolian farmer and local hunter-gatherer populations, dominated by haplogroups G and I (the latter being especially common in Iberia; Supplementary Information section 3). The European populations are consistent with a common origin in Anatolia (Supplementary Information section 4), reflected by the low differentiation among Early Neolithic groups in the principal component analysis. Over the course of the Neolithic period, we observe a trend of increasing hunter-gatherer ancestry in each region, although at a slower rate in Hungary than in Germany and Spain, and with limited intra-population heterogeneity (Fig. 2a and Supplementary Information section 6). We also find that this hunter-gatherer ancestry is more similar to the eastern WHG individuals (KO1 and VIL; for definitions see Table 1) farther east and more similar to the western WHG individuals (LB1 and LOS) farther west (Fig. 2b). Although this pattern does not demonstrate directly where mixture between hunter-gatherers and farmers took place, it suggests, given that European hunter-gatherers display a strong correlation between genetic and geographic structure (Fig. 1d), that hunter-gatherer ancestry in farmers was to a substantial extent derived from populations that lived in relatively close proximity.

To analyse admixed hunter-gatherer ancestry more formally, we modelled Neolithic farmers in an admixture graph framework. We started with a 'scaffold' model (Extended Data Fig. 2) consisting of Neolithic Anatolians, the four reference WHG individuals and two outgroups (Mbuti and Kostenki 14 (refs 20, 21)), with significant signals of admixture in LB1 and KO1 (Supplementary Information sections 5, 6). We then added each Neolithic population to this model in turn, fitting them as a mixture of FEF and either one or two hunter-gatherer ancestry components. To check for robustness, we repeated our analyses using transversions or outgroup-ascertained SNPs only, with in-solution capture data for LOS, and with additional or alternative hunter-gatherers in the model (Extended Data Table 3 and Supplementary Information section 6), and in all cases the results were qualitatively consistent. We find that almost all ancient groups from Hungary have ancestry significantly closest to one of the more eastern WHG individuals (KO1 or VIL); the samples from present-day

Table 1 | Neolithic population groups and western hunter-gatherer individuals in the study

Population	Location	Samples*	Approximate ages (BC)
Körös EN	Hungary (eastern)	6/5/3†	6000–5500
Starčevo EN	Hungary (western)	5/4/4	6000–5500
ALPc MN	Hungary (eastern)	25/20/22	5500–5000
LBKT MN	Hungary (western)	8/7/7	5500–5000
Vinča MN	Hungary (western)	6/6/0	5500–5000
Tisza LN	Hungary (eastern)	6/6/5	5000–4500
TDLN	Hungary (western)	15/14/14	5000–4500
Tiszapolgár CA	Hungary (eastern)	5/5/0	4500–4000
Lasinja CA	Hungary (western)	7/7/6	4300–3900
Protoboleráz CA	Hungary (eastern)	4/4/4	3800–3600
Baden CA	Hungary	13/12/10	3600–2850
LBK EN	Germany	30/15/29	5500–4850
Germany MN	Germany	8/4/7	4600–3000
Blätterhöhle MN	Germany	4/4/4†	4100–3000
Iberia EN	Spain	7/2/7	5500–4500
Iberia MN	Spain	4/0/4	3900–3600
Iberia CA	Spain	27/15/27	3000–2200
KO1 HG	Hungary (eastern)	1/0/1	5700
LB1 HG	Spain	1/0/1	5900
LOS HG	Luxembourg	1/0/1	6100
VIL HG	Italy (northeastern)	1/0/1	12000

EN/MN/LN, Early/Middle/Late Neolithic; ALPc, Alföld Linear Pottery culture; CA, Chalcolithic; HG, hunter-gatherer (LB1, La Braña 1; LOS, Loschbour; VIL, Villabruna); LBK, Linearbandkeramik; LBKT, Linearbandkeramik in Transdanubia; TDLN, Transdanubian Late Neolithic.
*Total number of samples/new in this study/used in the final analyses.
†Includes one hunter-gatherer individual treated separately.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

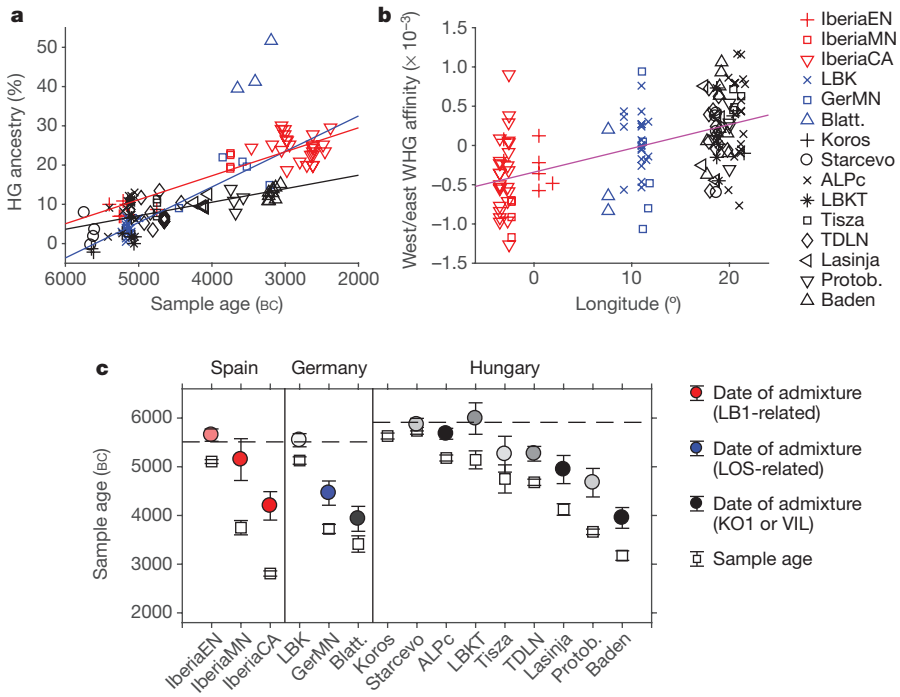


Figure 2 | Admixture parameters for test individuals and populations. **a**, Estimated individual hunter-gatherer ancestry versus sample age, with best-fitting regression lines for each region (excluding Blätterhöhle). Standard errors are around 2% for hunter-gatherer ancestry and 100 years for dates (Methods and Extended Data Tables 1, 2). **b**, Relative affinity of hunter-gatherer ancestry, measured as f_4 (LB1 and LOS, KO1 and VIL; Anatolia, *X*), where *X* indicates any of the European Neolithic individuals (positive, more similar to eastern WHG; negative, more similar to western WHG; standard errors, approximately 5×10^{-4}), with best-fitting regression line ($|Z| > 3$ for aggregate differences among the three regions). **c**, Population-level mean sample ages and dates of admixture, plus or minus two standard errors. Coloured fill indicates the inferred primary hunter-gatherer ancestry component, with darker shades corresponding to higher confidence (all admixed populations, except LBK and Tisza, were significant at $P < 0.05$; see Extended Data Table 3 and Supplementary Information section 6). Dashed lines denote the approximate date of arrival of farming in each region.

Germany have the greatest affinity to LOS; and all three Iberian groups have LB1-related ancestry (Fig. 2c and Extended Data Table 3). This pattern implies that admixture into European farmers occurred multiple times from local hunter-gatherer populations. Moreover, combining the proportions and sources of hunter-gatherer ancestry, populations from the three regions are distinguishable at all stages of the Neolithic period. Therefore, any further long-range migrations that may have occurred after the initial spread of agriculture in the studied regions (and before large incursions from the steppe) were not substantial enough to homogenize the ancestry of farming populations.

Additional insights about population interactions can be gained by studying the dates of admixture events. We used ALDER[22] to estimate dates of admixture for Neolithic individuals based on the recombination-induced breakdown of contiguous blocks of FEF and WHG ancestry over time (Extended Data Tables 1, 2, 4 and Extended Data Fig. 3). The ALDER algorithm is not able to accommodate large amounts of missing data, so we developed a strategy for running it with the relatively low coverage of ancient DNA (Supplementary Information section 7). The dates that we obtain (Fig. 2c) are based on a model of a single wave of admixture, which means that if the true history for a population includes multiples waves or continuous admixture, we will obtain an intermediate value. In particular, for later populations, this history could include mixture with previously admixed groups (either farmers with markedly different hunter-gatherer ancestry proportions or hunter-gatherers with farmer ancestry).

For our most complete time series, from Hungary, we infer admixture dates throughout the Neolithic period that are on average mostly 18–30 generations old (500–840 years), indicating ongoing population transformation and admixture (Fig. 2c and Extended Data Table 4). This pattern is accompanied by a gradual increase in hunter-gatherer ancestry over time, although never reaching the levels that we observe in Middle Neolithic Germany or Iberia (Fig. 2a). Although most of the Early Neolithic samples from Hungary do not have significantly more hunter-gatherer ancestry than Neolithic Anatolians (Fig. 2a and Extended Data Tables 1, 2), one Starčevo individual, BAM17b, is inferred to have $7.8 \pm 1.7\%$ (mean $\pm$ s.e.m.) hunter-gatherer ancestry and a very recent ALDER date of 4.5 ± 1.9 (mean $\pm$ s.e.m.) generations (5865 $\pm$ 65 BC (mean $\pm$s.e.m.); 1.9 ± 0.9 generations using a group-level estimate; Extended Data Table 4), consistent with having one or two hunter-gatherer ancestors in the past few generations. Additionally, one newly sampled Körös individual, TIDO2a, is similar to KO1 in having around 80% WHG and 20% FEF ancestry and an ALDER date of 16.1 ± 3.8 generations, reinforcing the distinctive heterogeneity of the Tiszaszőlős site, the origin of both TIDO2a and KO1. We also infer an average admixture date of 5675 $\pm$ 55 BC for the ALPc Middle Neolithic, again suggesting that in Hungary, interaction between Anatolian migrants and local hunter-gatherers began in the Early Neolithic (compare with refs 14, 23–25). The largest differences between Alföld and Transdanubia are observed in the Middle Neolithic, with substantially more hunter-gatherer ancestry in ALPc than LBKT (Fig. 2 and Extended Data Table 3) and, overall, we observe slight trends towards more hunter-gatherer ancestry to the north and east (Extended Data Fig. 4), as expected based on the greater archaeological evidence of hunter-gatherer settlement and interactions[23]. By the Late Neolithic and Chalcolithic periods, however, and especially in the Baden period (when the region became culturally unified[26]), our results are broadly similar across both halves of present-day Hungary.

From Germany, we analysed 29 individuals from the Early Neolithic Linearbandkeramik (LBK) culture and 11 individuals from the Middle Neolithic period, four of which came from the Blätterhöhle site, which has been shown to have featured a combination of farmer and hunter-gatherer occupation to a relatively late date[15]. The average date of admixture for LBK (5545 $\pm$ 65 BC) is more recent than the dates for Early and Middle Neolithic populations from Hungary and the total

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

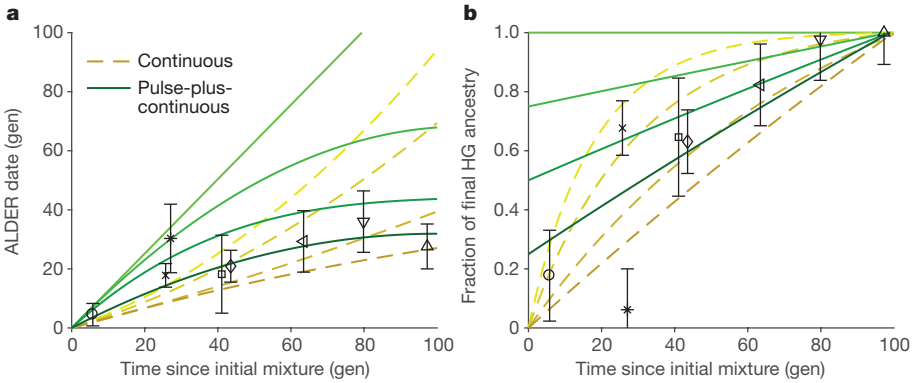


Figure 3 | Hungary time series and simulated data. a, Dates of admixture. **b**, Hunter-gatherer ancestry proportions, normalized to the total of the most recent (rightmost) population. Symbols are as in Figs 1, 2, indicating population-level mean ± 2 s.e.m. Yellow dashed lines represent continuous admixture simulations: from top to bottom, diminishing 5% per generation, diminishing 3%, diminishing 1% and uniform. Green solid lines represent pulse-plus-continuous admixture simulations: from top to bottom, all hunter-gatherer ancestry in a pulse at time zero; three-quarters of final hunter-gatherer ancestry in an initial pulse, followed by uniform continuous gene flow; half in initial pulse and half continuous; and one-quarter in initial pulse. Gen, generation.

hunter-gatherer ancestry proportion in LBK (around 4–5%) is intermediate between LBKT and ALPc. This ancestry is most closely related to a combination of KO1 and LOS, although the assignment of the hunter-gatherer source(s) is not statistically significant (Fig. 2c and Extended Data Table 3). These results are consistent with genetic and archaeological evidence for LBK origins from the early LBKT (ref. 25), followed by additional, Central European WHG admixture after about 5500 BC. Our 'Germany Middle Neolithic' grouping shows increased hunter-gatherer ancestry (around 17%, most closely related to LOS) and a more recent average date of admixture, reflecting gene flow from hunter-gatherers after the LBK period. We successfully sequenced a total of 17 bones from Blätterhöhle cave dating to the Middle Neolithic, most of which had distinct individual labels in ref. 15. Surprisingly, the genome-wide data indicated that they corresponded to only four unique individuals from the cave (Supplementary Information section 8), and we merged data from each sample to represent these four individuals. In accordance with previous results[15], we find that the three farmer individuals (classified based on stable isotopes) had 40–50% hunter-gatherer ancestry, whereas Bla8, who had signatures associated with a hunter-gatherer–fisher lifestyle, was closer genetically to hunter-gatherers, but was also admixed, with around 27% ancestry from farmers. Our results thus provide evidence of asymmetric gene flow between farmers and hunter-gatherers at Blätterhöhle centred around the relatively late date of about 4000 BC (ALDER dates of 10–25 generations).

In Iberia, we again see widespread evidence of local hunter-gatherer admixture, with confidently inferred LB1-related ancestry in all three population groups (Early and Middle Neolithic and Chalcolithic). For Iberia Early Neolithic individuals, we infer an average admixture date of 5650 ± 65 BC, which increases to 5860 ± 110 BC when considering only the five oldest individuals (of which the earliest, CB13 (ref. 18) has an estimate of 5890 ± 105 BC). Given that farming is thought to have begun in Iberia around 5500 BC (ref. 27), these dates suggest the presence of at least a small proportion of hunter-gatherer ancestry in earlier Cardial Neolithic populations acquired along their migration route (although our admixture graph analysis only confidently detected an LB1-related component). The later Iberians have large proportions of hunter-gatherer ancestry, approximately 23% for Middle Neolithic (from the site of La Mina, in north-central Iberia) and 27% for Chalcolithic populations, and also relatively old ALDER dates (approximately 50 generations, or 1,400 years), indicating that most of the admixture occurred well before their respective sample dates. Both populations show evidence of ancestry related to a different WHG individual in addition to LB1 (Extended Data Table 3), suggesting a non-local source for at least some of the hunter-gatherer ancestry gained between the Early and Middle Neolithic periods.

Synthesizing our time series data, we compared the observed ALDER dates and hunter-gatherer ancestry proportions of Neolithic populations to parameters estimated via simulation under different temporal admixture scenarios (Fig. 3, Extended Data Fig. 5 and Supplementary Information section 9). We assumed dates of 5900 BC (Hungary) or 5500 BC (Germany and Spain) for the onset of mixture. Although none of the scenarios match the data perfectly, a good fit for Hungary is provided by a model (bottom solid green curve in both panels of Fig. 3) of an initial admixture pulse (approximately a quarter of the total hunter-gatherer ancestry observed by the end of the time series) followed by continuous gene flow. By contrast, scenarios such as a single admixture pulse or continuous mixture decreasing by 5% or more per generation provide too much hunter-gatherer ancestry at early dates. The series for Alföld and Transdanubia should be considered to be separate, but their parameters follow mostly similar trajectories, with the exception of the Middle Neolithic period, during which LBKT has a relatively old admixture date (albeit with large uncertainty) and ALPc a relatively high hunter-gatherer ancestry proportion (possibly influenced by the bias of sampling in favour of the central and northern parts of the Alföld). Considering the other regions, even after normalizing for the different total hunter-gatherer ancestry proportions, we observe a high degree of local distinctiveness, including the older ALDER dates of Iberia Middle Neolithic and Chalcolithic populations and the markedly higher hunter-gatherer ancestry in Blätterhöhle (Extended Data Fig. 5). Although the simulated data are generated under a model of gene flow from an unadmixed hunter-gatherer source population into a series of farmer populations in a single line of descent, observed admixture could also be influenced by flow in the other direction (from farmers to hunter-gatherers) or could reflect immigration of new farmer populations (either via their own previous hunter-gatherer admixture or new admixture between farming populations with different proportions of hunter-gatherer ancestry). On the basis of archaeological evidence, for example, new hunter-gatherer ancestry could have been introduced into Transdanubia during the Late Neolithic period via gene flow from the northern Balkan Vinča or Sopot cultures to Transdanubia[14,28,29].

Our results provide greatly increased detail to our understanding of population interactions and admixture during the European Neolithic period. In each of our three study regions, the arrival of farmers prompted admixture with local hunter-gatherers, which unfolded over many centuries: almost all sampled populations have more hunter-gatherer ancestry and more recent dates of admixture than their local predecessors, suggesting recurrent changes in genetic composition and substantial hunter-gatherer gene flow beyond initial contact. These transformations left distinct signatures in each region, implying that they resulted from a complex web of local interactions rather than from a uniform demographic phenomenon. Our transect

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

of Hungary, in particular, with representative samples from many archaeological cultures across the region and throughout the Neolithic and Chalcolithic periods, illustrates the power of dense ancient DNA time series. Future work with continually improving datasets and statistical models should yield many more insights about historical population transformations in space and time.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 March; accepted 6 October 2017.
Published online 8 November 2017.

1. Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326,** 137–140 (2009).
2. Haak, W. *et al.* Ancient DNA from European Early Neolithic farmers reveals their Near Eastern affinities. *PLoS Biol*. **8,** e1000536 (2010).
3. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336,** 466–469 (2012).
4. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513,** 409–413 (2014).
5. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522,** 207–211 (2015).
6. Günther, T. *et al.* Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl Acad. Sci. USA* **112,** 11917–11922 (2015).
7. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528,** 499–503 (2015).
8. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl Acad. Sci. USA* **113,** 6886–6891 (2016).
9. Brandt, G. *et al.* Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science* **342,** 257–261 (2013).
10. Ammerman, A. J. & Cavalli-Sforza, L. L. *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton, 1984).
11. Price, T. D. (ed.) in *Europe's First Farmers* 301–318 (Cambridge, 2000).
12. Zvelebil, M. The agricultural transition and the origins of Neolithic society in Europe. *Documenta Praehistorica* **28,** 1–26 (2001).
13. Richards, M. The Neolithic invasion of Europe. *Annu. Rev. Anthropol.* **32,** 135–162 (2003).
14. Tringham, R. in *Europe's First Farmers* (ed. Price, T. D.) 19–56 (Cambridge, 2000).
15. Bollongino, R. *et al.* 2000 years of parallel societies in Stone Age central Europe. *Science* **342,** 479–481 (2013).
16. Skoglund, P. *et al.* Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344,** 747–750 (2014).
17. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5,** 5257 (2014).
18. Olalde, I. *et al.* A common genetic origin for early farmers from Mediterranean Cardial and central European LBK cultures. *Mol. Biol. Evol.* **32,** 3132–3142 (2015).
19. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507,** 225–228 (2014).
20. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534,** 200–205 (2016).
21. Seguin-Orlando, A. *et al.* Genomic structure in Europeans dating back at least 36,200 years. *Science* **346,** 1113–1118 (2014).
22. Loh, P.-R. *et al.* Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* **193,** 1233–1254 (2013).
23. Bánffy, E. Eastern, central and western Hungary — variations of Neolithisation models. *Documenta Praehistorica* **33,** 125–142 (2006).
24. Domboróczki, L., Kaczanowska, M. & Kozłowski, J. The Neolithic settlement at Tiszaszőlős-Domaháza-puszta and the question of the northern spread of the Körös Culture. *Atti Soc. Preist. Protost. Friuli-VG* **17,** 101–155 (2010).
25. Szécsényi-Nagy, A. *et al.* Tracing the genetic origin of Europe's first farmers reveals insights into their social organization. *Proc. R. Soc. Lond. B* **282,** 20150339 (2015).
26. Raczky, P. in *The Copper Age Cemetery of Budakalász* (eds Bondár, M. & Raczky, P.) 475–485 (Pytheas, 2009).
27. Martins, H. *et al.* Radiocarbon dating the beginning of the Neolithic in Iberia: new results, new problems. *J. Medit. Arch.* **28,** 105–131 (2015).
28. Jakucs, J. *et al.* Between the Vinča and Linearbandkeramik worlds: the diversity of practices and identities in the 54th–53rd centuries cal BC in southwest Hungary and beyond. *J. World Prehist.* **29,** 267–336 (2016).
29. Oross, K. *et al.* Midlife changes: the Sopot burial ground at Alsónyék. *Bericht der Römisch-Germanischen Kommission* **94,** 151–178 (2016).
30. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6,** 8912 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank I. Lazaridis, P.-R. Loh, I. Mathieson, I. Olalde, E. Palkopoulou, N. Patterson and P. Skoglund for helpful comments and suggestions; J. Krause for providing the Stuttgart sample for which we generated a new library in this study; A. Whittle and A. Bayliss from The Times of Their Lives project for providing the radiocarbon date for sample VEJ5a; and B. Havasi (Balaton Museum), G. V. Székely (Katona József Museum), C. Farkas (Dobó István Museum), B. Nagy (Herman Ottó Museum), I. Pap, A. Kustár, T. Hajdu (Hungarian Natural History Museum), J. Ódor (Wosinsky Mór Museum), E. Nagy (Janus Pannonius Museum), P. Rácz (King St Stephen Museum), L. Szathmáry (Debrecen University), N. Kalicz, V. Voicsek, O. Vajda-Kiss, V. Majerik and I. Kővári for assistance with samples. This work was supported by the Australian Research Council (grant DP130102158 to B.L. and W.H.), Hungarian National Research, Development and Innovation Office (K 119540 to B.M.), German Research Foundation (AI 287/7-1, 10-1 and 14-1 to K.W.A.), FEDER and Ministry of Economy and Competitiveness of Spain (BFU2015-64699-P to C.L.-F.), National Science Foundation (HOMINID grant BCS-1032255 to D.R.), National Institutes of Health (NIGMS grant GM100233 to D.R.), and Howard Hughes Medical Institute (D.R.).

Author Contributions A.S.-N., J.B., E.B., K.W.A., C.L.-F., W.H. and D.R. designed and supervised the study. B.G.M., K.K., K.O., M.B., T.M., A.O., J.J., T.P., F.H., P.C., J.K., K.Se., A.A., P.R., J.R., J.P.B., S.F., G.S., Z.T., E.G.N., J.D., E.M., G.P., L.M., B.M., Z.B., L.D., J.F.-E., J.A.M.-A., C.A.F., J.J.E., R.B., J.Or., K.Sc., H.M., A.C., J.B., E.B., K.W.A., C.L.-F. and W.H. provided samples and assembled archaeological and anthropological information. A.S.-N., A.P., B.S., V.K., N.R., K.St., M.F., M.M., J.Op., N.B., E.H., S.N. and B.L. performed laboratory work. M.L., A.S.-N., S.M. and D.R. analysed genetic data. M.L., A.S.-N. and D.R. wrote the manuscript with input from all coauthors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to M.L. (mlipson@genetics.med.harvard.edu), A.S.-N. (szecsenyi-nagy.anna@btk.mta.hu) or D.R. (reich@genetics.med.harvard.edu).

Reviewer Information *Nature* thanks P. Bellwood and the other anonymous reviewer(s) for their contribution to the peer review of this work.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Experimental procedures. Teeth and petrous bone samples from Hungary were taken under sterile conditions in the Hungarian Museums and anthropological collections. Samples, other than Gorzsa, were documented, cleaned and ground into powder either in the Anthropological Department of the Johannes Gutenberg University of Mainz during the course of the German Research Foundation project AL 287-10-1, or in the Laboratory of Archaeogenetics of the Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences in Budapest, following published protocols[25]. DNA was extracted in Budapest using 0.08–0.11 g powder according to published methods[31], using High Pure Viral NA Large Volume Kit columns (Roche)[32,33]. DNA extractions were tested by PCR, amplifying the 16,117–16,233-bp fragment of the mitochondrial genome, and visualized on a 2% agarose gel. DNA libraries were prepared from clean and successful extraction batches using UDG-half and UDG-minus treatment methods[5,34]. We included milling (hydroxylapatite blanks to control for cleanness) and extraction negative controls in every batch. Barcode adaptor-ligated libraries were amplified with TwistAmp Basic (Twist DX Ltd), purified with Agencourt AMPure XP (Beckman Coulter) and checked on a 3% agarose gel[5]. The DNA concentration of each library was measured on a Qubit 2.0 fluorometer. Promising libraries after initial quality-control analysis were shipped to Harvard Medical School, where further processing took place. All other samples were prepared similarly in dedicated clean rooms at Harvard Medical School and the University of Adelaide in accordance with published methods[5,7,33]. For samples LHUE2010.11 (one library) and MIR202-037-n105, we used magnetic-bead cleanups instead of MinElute column cleanups between enzymatic reactions with magnetic-bead cleanups and SPRI-bead cleanup instead of the final PCR cleanup[35,36].

We initially screened the libraries via in-solution hybridization to a set of probes targeting mitochondrial DNA (mtDNA)[37] plus roughly 3,000 nuclear SNP targets, using a protocol described previously[5,33] with amplified baits synthesized by CustomArray Inc. Libraries with good screening results—limited evidence of contamination, reasonable damage profiles and substantial coverage on targeted segments—were enriched for a genome-wide set of approximately 1.2 million SNPs[7,33] and sequenced to greater depth. Raw sequence data were processed by trimming barcodes and adapters, merging read pairs with at least 15 base pairs of overlapping sequence and mapping to the human reference genome (version hg19). Reads were filtered for mapping and base quality, duplicate molecules were removed and two terminal bases were clipped to eliminate damage (five for UDG-minus libraries)[5]. All libraries had a rate of at least 4.8% C-to-T substitutions in the final base of screening sequencing reads (Supplementary Table 1), consistent with damage patterns expected for authentic ancient DNA[34,38]. Pseudo-haploid genotypes at each SNP were called by choosing one allele at random from among mapped reads. Sex determinations for each individual were made by manually examining the fractions of reads mapping to the X and Y chromosomes and imposing thresholds for males and females (with any indeterminate samples labelled as unknown).

mtDNA sequences were reassembled in Geneious R10 to rCRS[39] and RSRS[40] and alleles were called if the majority nucleotide had a frequency of at least 0.7 (minimum 3 reads). The assembly and the resulting list of base calls were double-checked against http://phylotree.org/ (mtDNA tree build 17; 18 February 2016). Haplotype calls are given in Extended Data Tables 1, 2 and Supplementary Table 2. On the Y chromosome, 15,100 SNPs were targeted and sequenced and the detected derived and ancestral alleles were compared to the ISOGG Y-tree (https://isogg.org/) version 12.34, updated on 5 February 2017. Haplogroup definitions are detailed in Supplementary Information section 3.

We merged libraries from the same individual (for those with more than one) and then combined our new samples with genome-wide data from the literature (ancient individuals as described and as listed in Extended Data Table 1, 2 and present-day individuals from the SGDP[41]) using all autosomal SNPs (around 1.15 million) from our target set. For two replications of our admixture graph analyses, we restricted either to the subset of transversions (around 280,000 SNPs) or to the subset from panels 4 and 5 of the Affymetrix Human Origins array (ascertained as heterozygous in a San or Yoruba individual; around 260,000 SNPs). For the principal component analysis (PCA) (Extended Data Fig. 1), we merged with a large set of present-day samples[33] and used all autosomal Human Origins SNPs (around 593,000).

To test for possible contamination, we used contamMix[42] and ANGSD[43] to estimate rates of apparent heterozygosity in haploid genome regions (mtDNA and the X chromosome in males, respectively). Any samples with >5% mtDNA mismatching or >2% X chromosome contamination were excluded from further analyses, with the exception of Bla5 (Supplementary Information section 8). We also removed samples identified as clear outliers in PCA, or with significant population genetic differences between all sequencing data and genotypes called only from sequences displaying ancient DNA damage signatures. A total of 19 samples were excluded on the basis of one of these criteria. For individual-level *f*-statistic analyses (Fig. 2a, b), we restricted our analysis to samples with a maximum level of uncertainty, defined as a standard error of at most 7×10^{-4} for the statistic f_4 (Mbuti, WHG; Anatolia, *X*). This threshold (corresponding to an average coverage of approximately 0.05, or around 60,000 SNPs hit at least once) was met by 89 out of 112 samples passing quality control (and 49 out of 50 samples from the literature). We did not impose such a threshold for ALDER analyses, but because low coverage results in a weaker signal, only one of the 23 high-uncertainty individuals in our primary dataset provided an ALDER date (as compared to 89 of the 130 low-uncertainty individuals).

Population assignments. In most cases, population groupings were used that correspond to archaeological culture assignments based on chronology, geography, and material culture traits. Occasionally, we merged populations that appeared similar genetically in order to increase power: we pooled samples from all phases and groups of the eastern Hungarian Middle Neolithic period into a single ALPc population; merged six Sopot with eight Lengyel individuals for the western Hungarian Transdanubian Late Neolithic; combined one Hunyadihalom (Middle Chalcolithic period from the Danube–Tisza interfluve in central Hungary) with Lasinja; pooled four LBK samples from Stuttgart with the majority from further to the northeast (primarily Halberstadt); and merged several cultures of the German Middle Neolithic period into a single group. Other populations vary in their degrees of date and site heterogeneity; our Iberia Middle Neolithic samples are the most homogeneous group, and the Iberia Early Neolithic and Chalcolithic populations are among the most heterogeneous (Extended Data Tables 1, 2 and Supplementary Table 1). For our main analyses, we excluded the Vinča and Tiszapolgár populations because of a lack of sufficient high-quality data.

The designations Early, Middle, Late Neolithic and Chalcolithic have different meanings in different areas. For our study regions, each term generally refers to an earlier period in Hungary than in Germany and Spain (for example, the ALPc and LBKT Middle Neolithic populations in Hungary are roughly contemporaneous with the LBK and Iberia Early Neolithic populations). In order to maintain agreement with the archaeological literature, we use the established definitions, with the appropriate word of caution that they should be treated separately in each region.

Sample dates. We report 52 newly obtained accelerator mass spectrometry radiocarbon dates for Neolithic individuals (45 direct, 7 indirect), focusing on representative high-quality samples from each site and any samples with chronological uncertainty. We combined these with 58 radiocarbon dates from the literature[4,5,7,17,18,25,28,29,44,45]. We report the 95.4% calibrated confidence intervals from OxCal[46] version 4.2 with the IntCal13 calibration curve[47] in Extended Data Tables 1, 2. For use in ALDER analyses (Supplementary Information section 7), we use the mean and standard deviation of the calibrated date distributions (although the distributions are non-normal, we find that on average the mean plus or minus two standard deviations contains more than 95.4% of the probability density). For samples without direct radiocarbon dates, but with dates from other samples or materials at the same site, we form conservative 95.4% confidence intervals by taking the minimum and maximum bounds of any of the calibrated confidence intervals from the site. Finally, for the remaining samples, we use plausible date ranges based on archaeological context; we assume independence across individuals, but as a result take a conservative approach and treat the assigned range as ±1 s.e.m. (for example, an estimated range of 4800–4500 BC becomes 4650 ± 150 BC).

Population genetic analyses. We performed PCA by computing components for present-day populations and then projecting ancient individuals using the 'lsqproject' and 'shrinkmode' options in smartpca[48]. Admixture graphs were tested and *f*-statistics were computed using ADMIXTOOLS[49]. To obtain calendar dates of admixture, we combine the ALDER results (in generations in the past) with the ages of the Neolithic individuals, assuming an average generation time of 28 years[50,51]. All analytical procedures are described in detail in Supplementary Information sections 4–9.

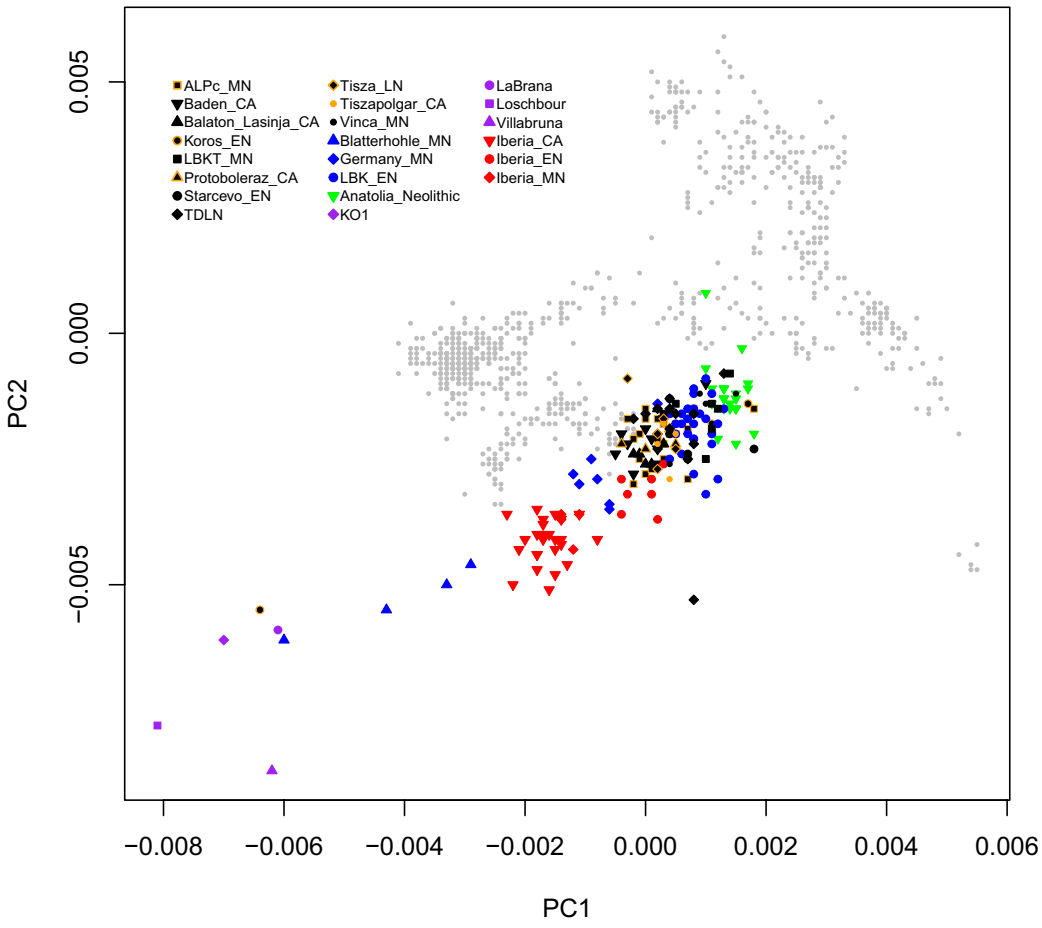
Data availability. The aligned sequences are available through the European Nucleotide Archive under accession number PRJEB22629. Genotype datasets used in analysis are available at https://reich.hms.harvard.edu/datasets.

31. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110,** 15758–15763 (2013).
32. Korlević, P. *et al.* Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques* **59,** 87–93 (2015).
33. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536,** 419–424 (2016).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

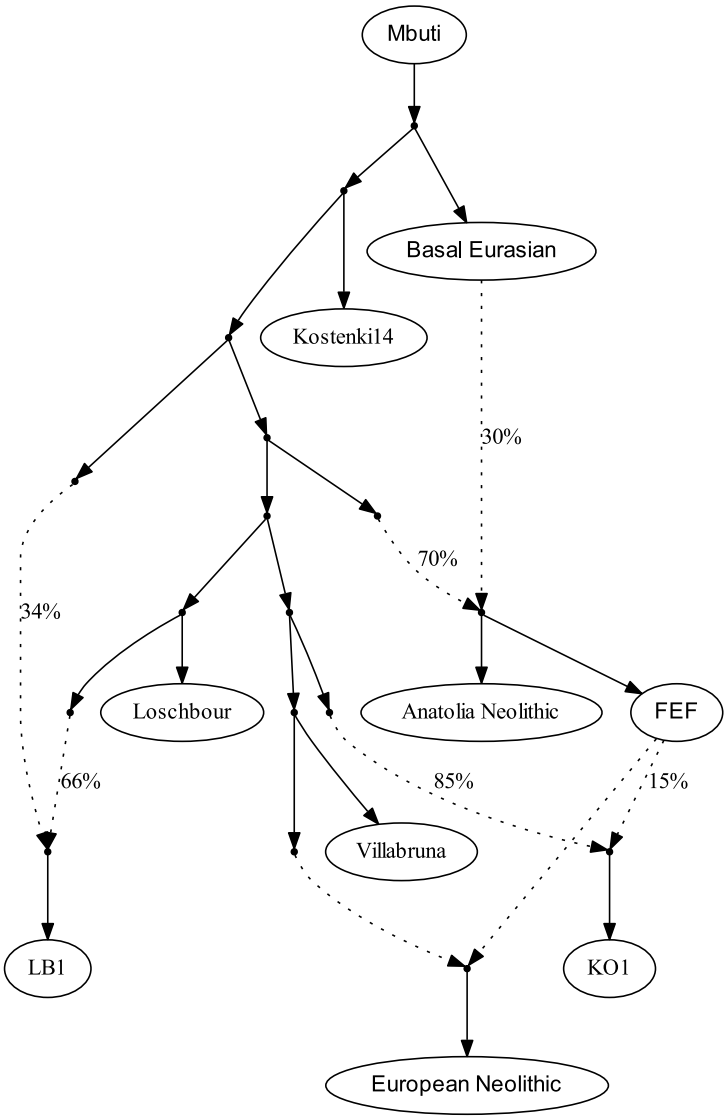
34. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370,** 20130624 (2015).
35. DeAngelis, M. M., Wang, D. G. & Hawkins, T. L. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res.* **23,** 4742–4743 (1995).
36. Rohland, N. & Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22,** 939–946 (2012).
37. Meyer, M. *et al.* A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505,** 403–406 (2014).
38. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Pääbo, S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE* **7,** e34131 (2012).
39. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23,** 147 (1999).
40. Behar, D. M. *et al.* A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90,** 675–684 (2012).
41. Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538,** 201–206 (2016).
42. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110,** 2223–2227 (2013).
43. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15,** 356 (2014).
44. Domboróczki, L. in *The First Neolithic Sites in Central/South-East European Transect. Volume III: The Körös Culture in Eastern Hungary* (eds Anders, A. & Siklósi, Z.) 107–111 (Oxford, 2012).
45. Oross, K. *et al.* The early days of Neolithic Alsónyék: the Starčevo occupation. *Bericht der Römisch-Germanischen Kommission* **94,** 93–121 (2016).
46. Ramsey, C. B. & Lee, S. Recent and planned developments of the program OxCal. *Radiocarbon* **55,** 720–730 (2013).
47. Reimer, P. J. *et al.* Intcal13 and marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55,** 1869–1887 (2013).
48. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2,** e190 (2006).
49. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192,** 1065–1093 (2012).
50. Fenner, J. N. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am. J. Phys. Anthropol.* **128,** 415–423 (2005).
51. Moorjani, P. *et al.* A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc. Natl Acad. Sci. USA* **113,** 5652–5657 (2016).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



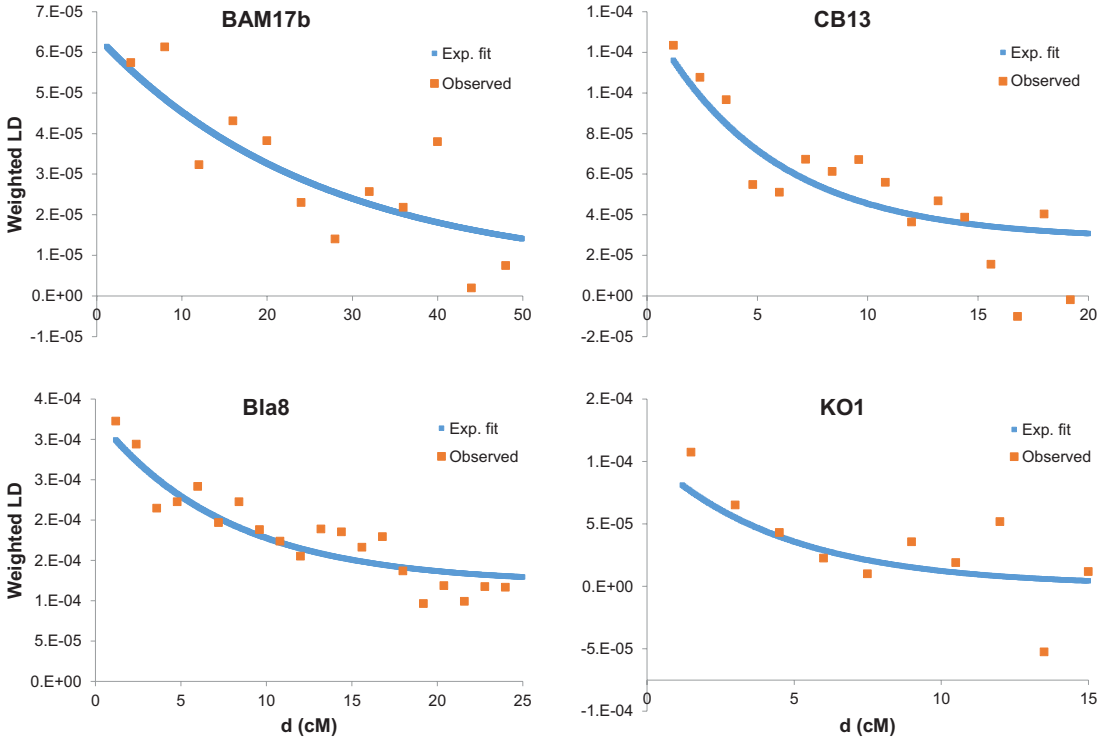
Extended Data Figure 1 | First two principal components from the PCA. We computed the principal components (PCs) for a set of 782 present-day western Eurasian individuals genotyped on the Affymetrix Human Origins array (background grey points) and then projected ancient individuals onto these axes. A close-up omitting the present-day Bedouin population is shown.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



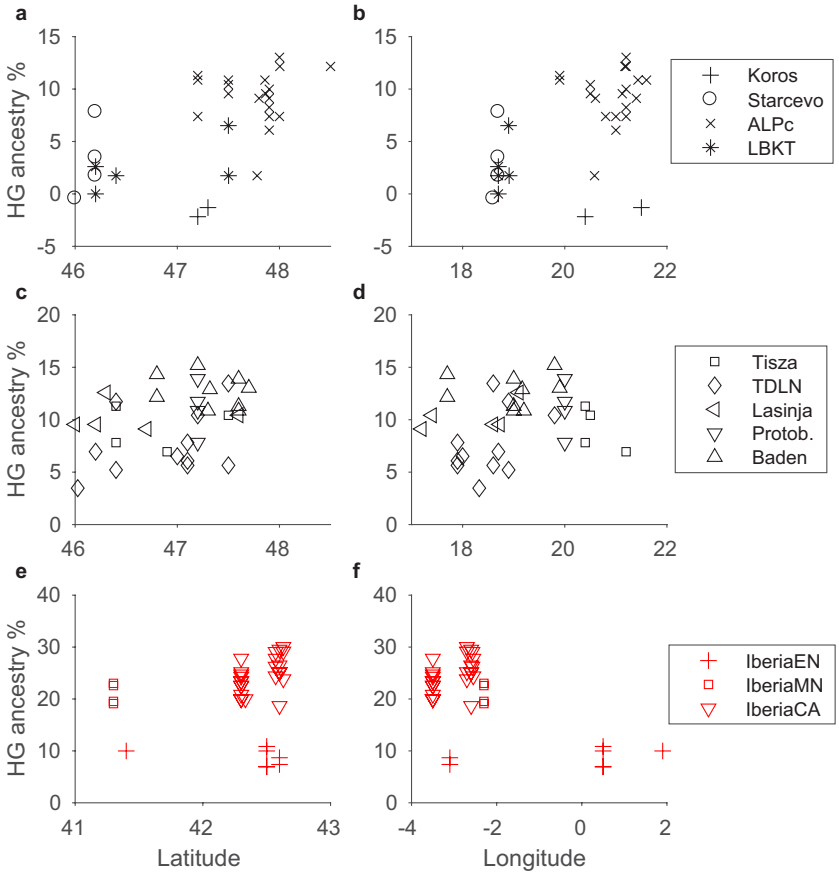
Extended Data Figure 2 | Scaffold admixture graph used for modelling the European Neolithic populations. Dotted lines denote admixture events. Neolithic Anatolians, LB1 and KO1 are modelled as admixed, with basal Eurasian ancestry, deeper European hunter-gatherer ancestry and FEF ancestry, respectively. European test populations were fitted as a mixture of FEF and ancestry related to one or two of the four WHG individuals (here VIL-related as an example). See Supplementary Information section 6 for details.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



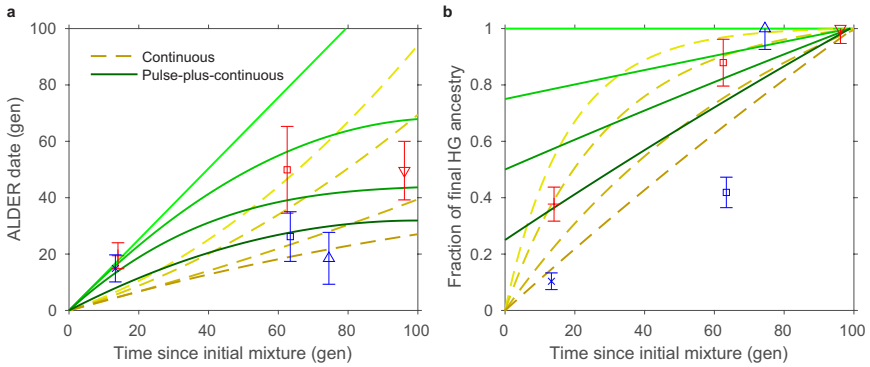
Extended Data Figure 3 | Examples of ALDER weighted linkage disequilibrium decay curves. Weighted linkage disequilibrium (LD) curves are shown as a function of genetic distance *d*, using Neolithic Anatolians and WHG as references, for four individuals: BAM17b (Starčevo Early Neolithic), CB13 (Iberia Early Neolithic), Bla8 (Blätterhöhle hunter-gatherer) and KO1. The results shown here use helper individuals M11-363 (Neolithic Anatolian), L11-322 (Neolithic Anatolian), BIC and LB1, respectively, and have fitted dates (blue curves) of 3.8 ± 1.2, 18.3 ± 6.0, 13.1 ± 2.7 and 21.6 ± 8.8 generations (compared to final individual-level dates of 4.5 ± 1.9, 17.5 ± 3.5, 12.1 ± 2.9 and 21.0 ± 7.0 generations; see Supplementary Information section 7). Note that the *x*-axis scales are different for the four plots.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



Extended Data Figure 4 | Hunter-gatherer ancestry as a function of latitude and longitude for Neolithic individuals. a, **b**, Early and Middle Neolithic Hungary. **c**, **d**, Late Neolithic and Chalcolithic Hungary. **e**, **f**, Iberia. HG, hunter-gatherer; Protob., Protoboleráz.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



Extended Data Figure 5 | Germany and Iberia time series and simulated data. a, Dates of admixture. **b**, Hunter-gatherer ancestry proportions, normalized to the total of the most recent (rightmost) population. Symbols are as in Figs 1, 2 and indicate population-level mean ± 2 s.e.m. Yellow dashed lines represent continuous admixture simulations: from top to bottom, diminishing 5% per generation, diminishing 3%, diminishing 1% and uniform. Green solid lines represent pulse-plus-continuous admixture simulations: from top to bottom, all hunter-gatherer ancestry in a pulse at time zero; three-quarters of final hunter-gatherer ancestry in an initial pulse followed by uniform continuous gene flow; half in initial pulse and half continuous; and one-quarter in initial pulse.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 1 | Information for the Neolithic individuals from Hungary

ID	Population	Site	Lat.	Long.	Date	Sex	Mt Hap	Y Hap	Cov.	HG%	*ALD*	Ref.
GEN68	Körös	Törökszentmiklós road 4 site 3	47.2	20.4	5706–5541	F	K1a	..	6.16	-2.16±1.5	0±0.0	
HUNG276, KO2	Körös	Berettyóújfalu-Morotva-liget	47.3	21.5	5713–5566	F	K1a	..	0.91	-1.49±1.6	0±0.0	[7,17]
TIDO2a	Körös	Tiszaszőlős-Domaháza	47.6	20.7	5736–5547	M	K1	I2a2	0.45	79.3±2.1	16±3.8	
BAM17b	Starčevo	Alsónyék-Bátaszék, Mérnöki telep	46.2	18.7	5832–5667	M	T1a2	H2	1.47	7.76±1.7	4.5±1.9	
BAM25	Starčevo	Alsónyék-Bátaszék, Mérnöki telep	46.2	18.7	5702–5536	M	N1a1a1	H2	0.22	1.62±1.9	0±0.0	[5,7]
BAM4a	Starčevo	Alsónyék-Bátaszék, Mérnöki telep	46.2	18.7	5641–5547	M	K1a4	G2a2a1	0.20	3.39±2.0	0±0.0	
LGCS1a	Starčevo	Lánycsók	46.0	18.6	*5800–5500*	M	W5	G2a2b2b1a	0.77	-0.63±1.6	0±0.0	
BAL25b	LBKT	Bátaszék-Lajvér	46.2	18.7	5208–4948	M	K1b1a	G2a2a1	2.77	0.06±1.5	0±0.0	
BOVO1b	LBKT	Bölcske-Gyűrűsvölgy	46.7	19.0	*5300–4900*	F	H	..	0.01	10.9±6.3	0±0.0	
BUD4a	LBKT	Budakeszi-Szőlőskert	47.5	18.9	*5300–4900*	M	T1a	G2a2b2a	0.17	6.72±2.3	36±6.1	
BUD9a	LBKT	Budakeszi-Szőlőskert	47.5	18.9	*5300–4900*	F	U2	..	1.10	1.87±1.6	13±5.3	
GEN18	LBKT	Alsónyék, site 11	46.2	18.7	5309–5074	M	T2c1	G2a2b2b1	1.48	2.66±1.5	35±12	
KON3	LBKT	Enese elkerülő, Kóny, Proletár-dülő, M85, site 2	47.6	17.4	*5300–4900*	F	T2b	..	0.03	2.79±4.0	0±0.0	
SZEH4	LBKT	Szemely-Hegyes	46.4	18.7	5207–4944	F	N1a1a1a3	..	0.07	1.88±3.0	0±0.0	[5,7]
CEG07b	ALPc	Cegléd, site 4/1	47.2	19.9	*5300–4900*	M	J2b1	G2a2b2a	0.30	11.4±1.9	0±0.0	
CEG08b	ALPc	Cegléd, site 4/1	47.2	19.9	*5300–4900*	F	J1c1	..	0.19	11.0±2.2	23±3.0	
EBSA2a	ALPc	Ebes-Sajtgyár	47.5	21.5	*5300–4900*	F	K1a	..	0.05	16.2±3.1	0±0.0	
EBVO5a	ALPc	Ebes-Zsongvölgy	47.5	21.5	*5300–4900*	M	V1a	CT	0.04	9.25±3.3	0±0.0	
HAJE10a	ALPc	Hajdúnánás-Eszlári út	47.9	21.4	5221–5000	M	J2b1	I	0.29	10.8±1.8	0±0.0	
HAJE7a	ALPc	Hajdúnánás-Eszlári út	47.9	21.4	5302–5057	M	K1a	I2c	1.57	9.15±1.7	6.2±5.7	
HELI11a	ALPc	Hejőkürt-Lidl	47.9	21.0	5209–4912	M	N1a1a1	I2a2a1b	0.99	6.01±1.8	14±2.0	
HELI2a	ALPc	Hejőkürt-Lidl	47.9	21.0	*5300–4900*	M	U8b1b	I	0.09	7.39±2.6	4.4±1.7	
HUNG302, NE2	ALPc	Debrecen Tocopart Erdoalja	47.5	21.6	5291–5056	F	H	..	4.88	11.0±1.7	0±0.0	[7,17]
HUNG372, NE5	ALPc	Kompolt-Kígyósér	47.2	20.8	5295–4950	M	J1c1	C1a2	4.25	7.48±1.6	0±0.0	[7,17]
HUNG86, NE3	ALPc	Garadna-Elkerülő út site 2	48.5	21.2	5281–5026	F	X2b-T226C	..	3.32	12.1±1.7	18±3.2	[7,17]
MEMO24b	ALPc	Mezőkövesd-Mocsolyás	47.8	20.6	*5500–5300*	M	U8b1b	CT	0.04	11.7±3.3	26±12	
MEMO2b	ALPc	Mezőkövesd-Mocsolyás	47.8	20.6	*5500–5300*	F	K1a1	..	2.28	8.99±1.7	24±5.2	
MEMO7a	ALPc	Mezőkövesd-Mocsolyás	47.8	20.6	5481–5361	F	HV	..	0.26	1.64±1.9	13±6.1	
PF325, NE1	ALPc	Polgár-Ferenci-hát	47.9	21.2	5306–5071	F	U5b2c	..	1.52	8.12±1.8	11±3.9	[7,17]
PF839/1198, NE4	ALPc	Polgár-Ferenci-hát	47.9	21.2	5211–5011	F	J1c5	..	3.49	9.95±1.7	25±10	[7,17]
POPI5a	ALPc	Polgár-Piócás	47.9	21.1	*5300–4900*	M	K1a1	I2a2a	0.31	9.75±2.0	11±3.7	
PULE1.18a	ALPc	Pusztataskony-Ledence	47.5	20.5	*5300–4900*	F	T2c1d1	..	0.29	10.6±1.8	0±0.0	
PULE1.23a	ALPc	Pusztataskony-Ledence	47.5	20.5	*5300–4900*	F	H1e	..	0.17	9.52±2.2	11±3.4	
TISO13a	ALPc	Tiszadob-Ókenéz	48.0	21.2	5208–4942	M	J1c2	I2a2a	1.21	12.9±1.7	22±7.6	
TISO1b	ALPc	Tiszadob-Ókenéz	48.0	21.2	*5300–4900*	M	H7	I2a2a1b1	0.11	7.24±2.4	0±0.0	
TISO3a	ALPc	Tiszadob-Ókenéz	48.0	21.2	*5300–4900*	F	U5b2b1a	..	0.27	12.1±2.1	8.4±5.2	
SEKU10a	Vinča	Szederkény-Kukorica-dülő	45.6	18.3	5320–5080	M	K2a	G2a2b2a1a	0.24	2.28±1.9	0±0.0	
SEKU6a	Vinča	Szederkény-Kukorica-dülő	45.6	18.3	5321–5081	F	H26	..	1.15	9.16±1.7	9.0±9.4	
VEGI17a	Vinča	Versend-Gilencsa	45.6	18.3	*5400–5000*	F	U2	..	0.01	-6.14±5.6	0±0.0	
VEGI3a	Vinča	Versend-Gilencsa	45.6	18.3	*5400–5000*	M	T2b	H2	0.41	0.53±1.8	0±0.0	
Gorzsa18	Tisza	Hódmezővásárhely-Gorzsa	46.4	20.4	*5000–4500*	M	U5b2c	I2a1	6.87	7.77±1.6	13±4.3	
Gorzsa4	Tisza	Hódmezővásárhely-Gorzsa	46.4	20.4	*5000–4500*	F	T1a	..	0.06	11.3±3.0	22±11	
KOKE3a	Tisza	Hódmezővásárhely-Kökénydomb Vörös tanya	46.4	20.2	*5000–4500*	M	K1b1	I	0.06	13.7±3.2	0±0.0	
PULE1.24	Tisza	Pusztataskony-Ledence	47.5	20.5	*5000–4500*	F	K1a4	..	0.40	10.4±1.9	18±7.2	
VSM3a	Tisza	Vésztő-Mágor	46.9	21.2	*5000–4500*	M	H26	G2a	0.09	6.92±2.6	0±0.0	
ALE14a	TDLN	Alsónyék-Elkerülő site 2	46.2	18.7	5030–4848	M	U8b1b	G2a2b	0.05	-1.11±3.2	0±0.0	
ALE4a	TDLN	Alsónyék-Elkerülő site 2	46.2	18.7	5016–4838	M	T2c1	F	0.03	10.6±3.6	0±0.0	
BAL3a	TDLN	Bátaszék-Lajvér	46.2	18.7	*4800–4500*	M	T2f	H2	0.91	6.89±1.7	22±9.0	
CSAT19a	TDLN	Csabdi-Télizöldes	47.5	18.6	*4800–4500*	M	H	H	0.52	5.82±1.8	34±9.6	
CSAT25a	TDLN	Csabdi-Télizöldes	47.5	18.6	4826–4602	M	T2b	I2c	0.43	13.5±1.9	26±8.1	
FAGA1a	TDLN	Fajsz-Garadomb	46.4	18.9	*5100–4750*	M	HV0a	I	0.09	5.08±2.4	0±0.0	
FAGA2a	TDLN	Fajsz-Garadomb	46.4	18.9	5195–4842	F	H	..	0.49	11.9±1.8	14±4.1	
FEB3a	TDLN	Felsőörs-Bárókert	47.0	18.0	*4800–4500*	M	H44	J2a	0.16	6.31±2.1	0±0.0	
HUNG347, NE7	TDLN	Apc-Berekalja	47.2	19.8	4491–4357	M	N1a1a1a	I	4.85	10.6±1.6	19±3.1	[7,17]
SZEH5a	TDLN	Szemely-Hegyes	46.0	18.3	4904–4709	M	K1b1a	G	0.01	10.8±6.5	0±0.0	
SZEH7b	TDLN	Szemely-Hegyes	46.0	18.3	4930–4715	F	K1a	..	0.52	3.44±1.7	0±0.0	
VEJ12a	TDLN	Veszprém Jutasi út	47.1	17.9	*4800–4500*	M	U8b1a2b	H	0.10	6.17±2.3	0±0.0	
VEJ2a	TDLN	Veszprém Jutasi út	47.1	17.9	*4800–4500*	M	T2b	C1a2	0.34	5.63±1.8	0±0.0	
VEJ5a	TDLN	Veszprém Jutasi út	47.1	17.9	4936–4742	M	J1c2	G2a2a1	0.62	7.78±1.8	15±2.9	
GEN67	Tiszapolgár	Törökszentmiklós road 4 site 3	47.2	20.4	4444–4257	M	H1	I2a2a1b	2.28	13.0±1.7	50±15	
PULE1.10a	Tiszapolgár	Pusztataskony-Ledence	47.5	20.5	*4500–4000*	M	T2c1	I2a	0.28	9.03±2.0	0±0.0	
PULE1.13a	Tiszapolgár	Pusztataskony-Ledence	47.5	20.5	*4500–4000*	M	T2c1	G2a2b2a1a1c1a	0.38	10.3±1.9	0±0.0	
PULE1.9a	Tiszapolgár	Pusztataskony-Ledence	47.5	20.5	*4500–4000*	M	H26	G2a2b	0.11	11.6±2.4	0±0.0	
GEN100	Lasinja	Alsónyék, site 11	46.2	18.7	*4300–3900*	F	T2b	..	1.81	9.51±1.6	45±11	
GEN49	Lasinja	Nemesnádudvar-Papföld	46.3	19.1	4228–3963	M	T2b23	C1a2a	0.97	12.8±1.8	27±6.8	
KEFP2a	Lasinja	Keszthely-Fenékpuszta	46.7	17.2	*4300–3900*	F	J2b1a	..	0.74	9.12±1.7	21±5.4	
KON2a	Lasinja	Enese elkerülő, Kóny, Proletár-dülő, M85, site 2	47.6	17.4	4333–4072	F	K2a	..	2.13	10.3±1.7	21±6.4	
M6-116.12a	Lasinja	Lánycsók, Csata-alja	46.0	18.6	4232–4046	F	T2f8a	..	0.64	9.68±1.7	29±11	
VEJ9a	Lasinja	Veszprém Jutasi út	47.1	17.9	4339–4237	M	H40	C1a2	0.05	8.83±3.2	0±0.0	
GEN60	Protoboleráz	Abony, Turjányos-dűlő	47.2	20.0	3909–3651	M	H	G2a2b2a	1.88	14.0±1.6	37±8.8	
GEN61	Protoboleráz	Abony, Turjányos-dűlő	47.2	20.0	*3800–3600*	M	J1c	I2c	0.76	10.8±1.7	65±13	
GEN62	Protoboleráz	Abony, Turjányos-dűlő	47.2	20.0	3762–3636	F	N1a1a1a3	..	4.81	8.00±1.6	37±9.6	
GEN63	Protoboleráz	Abony, Turjányos-dűlő	47.2	20.0	3658–3384	M	U5a1c1	I2c	1.92	11.9±1.7	34±8.1	
GEN12a	Baden	Budakalász-Luppa csárda	47.6	19.0	3340–2945	M	H26a	G2a2b2a1a1b1	1.98	13.8±1.6	34±7.2	
GEN13a	Baden	Budakalász-Luppa csárda	47.6	19.0	3332–2929	M	HV	G2a2b2a1a	2.65	11.3±1.6	27±6.6	
GEN15a	Baden	Budakalász-Luppa csárda	47.6	19.0	3367–3103	M	J2a1a1	G2a2b2a1a1c1a	1.66	10.8±1.7	22±9.3	
GEN16a	Baden	Alsónémedi	47.3	19.2	3346–2945	F	T2b	..	4.30	12.9±1.6	38±16	
GEN17a	Baden	Alsónémedi	47.3	19.2	3359–3098	M	U5b3f	G2a2a	0.82	10.7±1.7	21±6.4	
GEN21	Baden	Balatonlelle-Felső-Gamász	46.8	17.7	*3600–2850*	M	K1a	I2a1	0.67	12.3±1.7	0±0.0	
GEN22	Baden	Balatonlelle-Felső-Gamász	46.8	17.7	3332–2929	M	U5a1	I2a1a1a	2.31	14.5±1.7	25±6.6	
GEN55	Baden	Vámosgyörk	47.7	19.9	*3600–2850*	F	T2c1d1	..	0.81	13.1±1.7	22±6.6	
HUNG353, CO1	Baden	Apc-Berekalja	47.2	19.8	3315–2923	F	H	..	4.56	15.1±1.7	0±0.0	[7,17]
Vors1	Baden	Vörs	46.7	17.3	*3300–2850*	F	T2f	..	0.03	4.47±4.2	0±0.0	

Lat., latitude; Long., longitude; Mt Hap, mtDNA haplogroup; Y hap, Y-chromosome haplogroup; Cov., mean coverage per SNP; HG%, inferred percentage of hunter-gatherer ancestry (mean ± s.e.m.); ALD, inferred date of admixture (generations in the past; mean ± s.e.m.; zero implies no date obtained); Ref., reference for published data; if blank, newly published sample in this study (asterisk denotes a published individual with new sequencing data added). Radiocarbon dates are set as roman text, whereas dates estimated from archaeological context are in italics. Further information can be found in Supplementary Table 1.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 2 | Information for the Neolithic individuals from Germany and Spain

ID	Population	Site	Lat.	Long.	Date	Sex	Mt Hap	Y Hap	Cov.	HG%	*ALD*	Ref.
HAL03a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5295–5057	F	T2b	..	0.01	-5.13±6.8	0±0.0	
HAL07a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5212–4992	F	N1a1a1	..	0.05	1.72±3.2	0±0.0	
HAL15a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5199–4857	M	N1a1a1a3	G2	0.02	5.26±5.0	0±0.0	
HAL17b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	V1	..	0.02	9.21±4.2	0±0.0	
HAL18a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	K2a	..	0.02	0.27±4.6	0±0.0	
HAL19	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	K1a2	..	0.86	7.10±1.7	16±7.6	[7]*
HAL2	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5211–4963	M	N1a1a1a2	G2a2a1	0.76	1.91±1.7	11±2.4	[5,7]*
HAL20b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	M	K1a2	G2a2a	0.06	2.53±3.1	0±0.0	
HAL21a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	M	T2b	G2a2a	0.01	-4.41±5.8	0±0.0	
HAL22b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	T2b	..	0.02	-7.71±4.7	0±0.0	
HAL24	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5201–4850	M	X2d1	G2a2a1	0.42	6.39±1.8	0±0.0	[5,7]*
HAL25	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5210–5002	M	K1a	G2a2a1	0.49	2.58±1.7	18±6.6	[5,7]*
HAL27a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	M	N1a1a3	G2a2a	0.05	3.84±3.0	0±0.0	
HAL31a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5295–5057	F	K1	..	0.12	4.54±2.3	11±3.1	
HAL32b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	H26	..	0.23	3.34±2.0	23±4.4	
HAL34	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5219–5021	F	N1a1a1	..	0.25	5.63±2.0	9.2±5.0	[5,7]*
HAL35b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	J1c	..	0.10	3.93±2.4	0±0.0	
HAL38a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	V1	..	0.29	1.10±1.9	0±0.0	
HAL39b	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5210–5002	M	H1e	G2a2a1	0.08	3.96±2.6	0±0.0	
HAL4	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5202–4852	F	N1a1a1a	..	6.92	6.55±1.6	18±5.9	[5,7]*
HAL40a	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	*5500–4850*	F	T2b	..	0.17	2.50±2.1	0±0.0	
HAL5	LBK	Halberstadt-Sonntagsfeld	51.9	11.0	5211–4991	F	T2c1	..	2.23	2.98±1.6	15±5.4	[5,7]*
KAR16A	LBK	Karsdorf	51.3	11.7	*5500–4850*	M	H46b	T1a	0.09	0.28±2.6	13±5.1	[7]
KAR6	LBK	Karsdorf	51.3	11.7	5217–5041	M	H1/H1au1b	CT	0.10	5.78±2.5	0±0.0	[5,7]
LBK1976	LBK	Viesenhäuser Hof	48.8	9.2	*5500–4850*	F	T2e	..	0.44	3.46±1.7	18±4.4	[5,7]
LBK1992	LBK	Viesenhäuser Hof	48.8	9.2	*5500–4850*	F	T2b	..	2.66	5.68±1.6	12±4.3	[5,7]
LBK2155	LBK	Viesenhäuser Hof	48.8	9.2	*5500–4850*	F	T2b	..	3.63	4.84±1.5	13±4.4	[5,7]
Stuttgart	LBK	Viesenhäuser Hof	48.8	9.2	5310–5076	F	T2c1d1	..	9.65	3.00±1.6	22±8.1	[4]*
UWS4	LBK	Unterwiederstedt	51.7	11.5	5223–5021	F	J1c17	..	18.6	5.70±1.6	13±14	[5,7]
ESP30	GermanyMN	Esperstedt	51.4	11.7	3970–3710	M	H1e1a	I	0.09	22.0±2.7	0±0.0	[5,7]
HAL13a	GermanyMN	Halberstadt-Sonntagsfeld	51.9	11.0	*4600–4300*	F	V1a	..	0.11	9.04±2.4	13±4.3	
QLB15D	GermanyMN	Quedlinburg	51.8	11.1	3654–3527	M	HV	R1b1a	0.16	20.9±2.2	36±8.7	[5,7]
QLB18A	GermanyMN	Quedlinburg	51.8	11.1	3640–3376	F	T2e1	..	0.41	19.6±1.8	23±4.9	[5,7]
SALZ3B	GermanyMN	Salzmünde-Schiebzig	51.5	11.8	*3400–3025*	M	U3a1	G2a2a1	0.09	14.9±2.7	0±0.0	[7]
SALZ57A	GermanyMN	Salzmünde-Schiebzig	51.5	11.8	3345–3097	F	H3	..	0.02	25.0±4.4	0±0.0	
SALZ77A	GermanyMN	Salzmünde-Schiebzig	51.5	11.8	*3400–3025*	M	H3	IJK (x J)	0.02	21.3±5.0	0±0.0	
Bla16	Blätterhöhle	Blätterhöhle Cave	51.4	7.6	3958–3344	M	U5b2a2	R1b1a	0.80	39.5±1.9	15±5.8	
Bla28	Blätterhöhle	Blätterhöhle Cave	51.4	7.6	3337–3024	M	J1c1b1	R1b1a	0.10	51.9±2.7	11±4.5	
Bla5	Blätterhöhle	Blätterhöhle Cave	51.4	7.6	3704–3117	F	H5	..	5.07	41.2±1.9	24±4.7	
Bla8	Blätterhöhle	Blätterhöhle Cave	51.4	7.6	4038–3532	M	U5b2b2	I2a1	4.58	72.6±2.0	12±2.9	
CB13	Iberia EN	Cova Bonica	41.4	1.9	5469–5327	F	K1a2a	..	0.98	9.97±1.7	17±3.5	[18]
E-06-Ind1	Iberia EN	El Prado de Pancorbo	42.6	-3.1	4827–4692	F	K1a4a1	..	0.47	8.72±1.8	17±2.3	
E-14-Ind2	Iberia EN	El Prado de Pancorbo	42.6	-3.1	5216–5031	F	H1	..	0.38	7.52±1.8	19±2.8	
Troc1	Iberia EN	Els Trocs	42.5	0.5	5311–5218	F	J1c3	..	0.69	7.15±1.7	12±9.1	[5,7]
Troc3	Iberia EN	Els Trocs	42.5	0.5	5294–5066	M	T2c1d/T2c1d2	R1b1a	1.31	9.91±1.8	49±22	[5,7]
Troc5	Iberia EN	Els Trocs	42.5	0.5	5310–5078	M	N1a1a1	I2a1b1	13.8	6.83±1.6	6.8±2.8	[5,7]
Troc7	Iberia EN	Els Trocs	42.5	0.5	5303–5075	F	V	..	1.57	11.0±1.7	18±4.8	[5,7]
Mina18	Iberia MN	La Mina	41.3	-2.3	3893–3661	F	U5b1	..	13.6	22.8±1.7	42±18	[5,7]
Mina3	Iberia MN	La Mina	41.3	-2.3	*3900–3600*	M	K1a1b1	H2	0.38	19.5±1.9	80±20	[5,7]
Mina4	Iberia MN	La Mina	41.3	-2.3	*3900–3600*	M	H1	I2a2a1b2	3.95	22.6±1.9	25±6.2	[5,7]
Mina6	Iberia MN	La Mina	41.3	-2.3	*3900–3600*	F	K1b1a1	..	1.36	18.9±1.7	46±8.2	[5,7]
1.-K11	Iberia CA	La Chabola de la Hechicera	42.6	-2.6	3263–2903	M	X2b	I2a2	0.18	27.8±2.1	68±28	
3.-K11	Iberia CA	La Chabola de la Hechicera	42.6	-2.6	3627–3363	F	J2a1a1	..	0.12	24.4±2.4	27±11	
5.-K18	Iberia CA	La Chabola de la Hechicera	42.6	-2.6	3090–2894	M	J1c1	I2a2	0.10	18.5±2.5	43±11	
ES.1/4	Iberia CA	El Sotillo	42.6	-2.6	2571–2347	M	H3	I2	0.07	25.4±2.8	0±0.0	
ES-6G-110	Iberia CA	El Sotillo	42.6	-2.6	2916–2714	M	H3	I2a2a	0.05	25.4±3.2	0±0.0	
Inventario0/4	Iberia CA	El Sotillo	42.6	-2.6	2481–2212	M	X2b	I2a2a	0.12	29.6±2.5	56±23	
LHUE11J.5	Iberia CA	Alto de la Huesera	42.6	-2.6	3092–2877	F	U5b1	..	1.19	26.7±1.9	40±9.7	
LHUE2010.10	Iberia CA	Alto de la Huesera	42.6	-2.6	3014–2891	F	J1c1	..	0.11	25.2±2.5	64±13	
LHUE2010.11	Iberia CA	Alto de la Huesera	42.6	-2.6	3092–2918	M	V	G2a2a	5.36	28.9±1.8	38±12	
LHUE2014.11J	Iberia CA	Alto de la Huesera	42.6	-2.6	3100–2850	F	U5b2b	..	0.06	26.3±3.0	0±0.0	
LY.II.A.10.15066	Iberia CA	Las Yurdinas II	42.6	-2.7	3350–2750	M	U5b2b3a	I2a2a2a	1.93	30.0±1.8	0±0.0	
LY.II.A.10.15067	Iberia CA	Las Yurdinas II	42.6	-2.7	3350–2750	F	J2a1a1	..	0.30	23.8±2.0	0±0.0	
LY.II.A.10.15068	Iberia CA	Las Yurdinas II	42.6	-2.7	3350–2750	F	K1a4a1	..	0.39	29.2±1.9	26±10	
LY.II.A.10.15069	Iberia CA	Las Yurdinas II	42.6	-2.7	3354–2943	F	J1c3	..	4.24	25.1±1.7	28±15	
MIR1	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	F	K1a	..	0.24	24.2±2.1	0±0.0	[7]
MIR13	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	F	H3c3	..	0.10	27.8±2.4	0±0.0	[7]
MIR14	Iberia CA	El Mirador Cave	42.3	-3.5	2568–2346	M	H3	I2a2a	0.94	23.3±1.8	57±15	[7]
MIR17	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	F	J1c1	..	0.22	23.6±2.2	0±0.0	[7]
MIR18	Iberia CA	El Mirador Cave	42.3	-3.5	2865–2575	F	H1t	..	1.58	20.0±1.6	0±0.0	[7]
MIR19	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	M	H3	I	0.06	21.8±3.1	0±0.0	[7]
MIR2	Iberia CA	El Mirador Cave	42.3	-3.5	2857–2496	F	K1b1a1	..	0.98	22.6±1.7	56±8.9	[7]
MIR202-037-n105	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	M	K1a	I2a2a	5.73	19.9±1.7	0±0.0	
MIR21	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	M	H3	I	0.11	24.7±2.4	55±17	[7]
MIR22	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	F	K1a2a	..	2.79	22.6±1.7	62±10	[7]
MIR24	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	M	J2b1a3	G2a2b2b	0.06	20.0±3.0	0±0.0	[7]
MIR25	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2346	M	U3a1	I2a1a1	0.73	25.3±1.7	34±15	[7]
MIR5, MIR6	Iberia CA	El Mirador Cave	42.3	-3.5	2900–2679	M	X2b	I2a2a2a	10.4	20.7±1.7	0±0.0	[7]

Lat., latitude; Long., longitude; Mt Hap, mtDNA haplogroup; Y hap, Y-chromosome haplogroup; Cov., mean coverage per SNP; HG%, inferred percentage of hunter-gatherer ancestry (mean ± s.e.m.); ALD, inferred date of admixture (generations in the past; mean ± s.e.m.; zero implies no date obtained); Ref., reference for published data; if blank, newly published sample in this study (asterisk denotes a published individual with new sequencing data added). Radiocarbon dates are set as roman text, whereas dates estimated from archaeological context are in italics. Further information can be found in Supplementary Table 1.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 3 | Admixture graph results for Neolithic populations

	Main scaffold		Alternative scaffold	
Population	**HG ancestry**	**WHG affinity**	**HG ancestry**	**WHG affinity**
Körös EN	0.0 ± 1.2%		0.0 ± 1.2%	
Starčevo EN	2.3 ± 1.0%	KO1/VIL*	2.3 ± 1.0%	VIL
ALPc MN	8.8 ± 0.6%	KO1* + VIL	9.5 ± 0.6%	KO1* + VIL
LBKT MN	0.8 ± 0.9%	VIL*	0.5 ± 0.9%	VIL
Tisza LN	8.4 ± 1.3%	KO1/VIL	9.8 ± 1.3%	KO1/VIL + EHG
TDLN	8.2 ± 0.7%	KO1/VIL*	8.4 ± 0.7%	KO1*
Lasinja CA	10.7 ± 0.9%	KO1/VIL*	10.6 ± 0.9%	KO1/VIL*
Protoboleráz CA	12.7 ± 0.9%	KO1/VIL*	12.5 ± 0.9%	KO1/VIL
Baden CA	13.0 ± 0.7%	KO1/VIL*	13.4 ± 0.7%	KO1*
LBK EN	4.2 ± 0.6%	KO1 + LOS	5.0 ± 0.6%	KO1*
Germany MN	17.0 ± 1.1%	LOS*	18.3 ± 1.1%	LOS + KO1
Blätterhöhle MN	40.6 ± 1.5%	KO1/VIL* + LOS	42.6 ± 1.5%	KO1* + LOS
Iberia EN	10.0 ± 0.8%	LB1*	10.4 ± 0.8%	LB1*
Iberia MN	23.3 ± 1.1%	LB1* + LOS	24.8 ± 1.1%	LB1* + LOS
Iberia CA	26.5 ± 0.7%	LB1* + LOS/KO1/VIL*	27.5 ± 0.7%	LB1* + VIL*

Hunter-gatherer ancestry in Neolithic populations as inferred from admixture graph analyses. The inferred ancestry proportions for the best-fitting FEF + WHG model are shown, along with the WHG individual(s) inferred to be related to the hunter-gatherer sources, with an asterisk (*) denoting statistical significance (*$P < 0.05$) (Methods). The two sets of results are for the primary scaffold model (Extended Data Fig. 2) and an alternative admixture graph scaffold that includes EHG (Supplementary Information section 6). Plus signs indicate two components, whereas slashes indicate single components with one of two or three possibilities.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Extended Data Table 4 | Mean dates of admixture for Neolithic populations

Population	Individual-based	Group-based	Average sample date (BCE)
Körös EN			5631 ± 31
Starčevo EN	4.5 ± 1.9	1.9 ± 0.9	5738 ± 35
ALPc MN	17.8 ± 2.0	16.4 ± 2.6	5180 ± 31
LBKT MN	30.3 ± 5.8	31.5 ± 10.9	5142 ± 93
Tisza LN	18.2 ± 6.6	12.6 ± 3.1	4750 ± 145
TDLN	20.9 ± 2.7	19.1 ± 3.8	4681 ± 32
Lasinja CA	29.3 ± 5.2	23.0 ± 4.1	4123 ± 59
Protoboleráz CA	44.3 ± 6.4	19.8 ± 5.4	3674 ± 35
Baden CA	27.6 ± 3.8	26.2 ± 6.9	3176 ± 49
LBK EN	14.9 ± 2.4	15.4 ± 3.6	5128 ± 38
Germany MN	26.2 ± 4.4	55.0 ± 41.2	3724 ± 46
Blätterhöhle MN	18.5 ± 4.6	23.1 ± 6.2	3414 ± 84
Iberia EN	19.4 ± 2.3	17.5 ± 5.9	5107 ± 20
Iberia MN	49.9 ± 7.7	40.0 ± 6.9	3749 ± 74
Iberia CA	49.6 ± 5.2	56.5 ± 7.9	2808 ± 27

Dates of admixture (in generations in the past) as inferred from ALDER through two different methods. Left, the mean individual-level dates used in our main analyses. Right, direct estimates for population groups. By default, for group-level estimates, we used all individuals that yielded a date in our standard ALDER procedure, but because of missing data, for some populations we used a subset of individuals (typically those with highest coverage): Starčevo (BAM17b, BAM4a and LGCS1a; we note that in this case only BAM17b had an ALDER signal individually), ALPc (HAJE7a, HELI11a, MEMO2b, NE1, NE3, NE4 and TISO13a), Tisza (Gorzsa18 and PULE1.24), Baden (GEN12a, GEN13a, GEN15a, GEN17a, GEN22 and GEN55), LBK (HAL19, HAL2, HAL4, HAL5, LBK1992 and Stuttgart) and Iberia Chalcolithic (LHUE11J.5, LHUE2010.11, LY.II.A.10.15066, LY.II.A.10.15069, MIR14, MIR2 and MIR22) populations. For the group-level estimate for the Iberia Middle Neolithic population, we use a fitting start point of 0.8 centimorgans instead of the program-inferred minimum of 0.6, because of a noticeably lower standard error. For our main analyses, we omit the outlier Protoboleráz individual GEN61, yielding an average date of 36.0 ± 5.2 generations, to help to capture uncertainty due to the disagreement between the individual-level and group-level estimates shown here. Mean sample dates (except for Körös) are based on the same weighting as the individual-level average dates of admixture for compatibility (Supplementary Information section 7).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

nature research

Corresponding author(s): David Reich

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Sample sizes were not predetermined; as many ancient samples as possible were included in the analyses.

2. Data exclusions

Describe any data exclusions.

Some samples were omitted for data-quality reasons; otherwise data were only excluded in one or two places for the sake of defining population groups (as described in the text)

3. Replication

Describe whether the experimental findings were reliably reproduced.

N/A

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

N/A

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

All samples were processed in the same manner regardless of their origin

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Only existing population genetics software tools were used (Admixtools package and ALDER)

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	N/A

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	N/A
b. Describe the method of cell line authentication used.	N/A
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	N/A