

Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques

Torsten Günther^{a,1}, Cristina Valdiosera^{a,b,c,1}, Helena Malmström^a, Irene Ureña^{c,d}, Ricardo Rodríguez-Varela^{c,d}, Ódny Osk Sverrisdóttir^a, Evangelia A. Daskalaki^{a,e}, Pontus Skoglund^{a,e,f}, Thijessen Naidoo^a, Emma M. Svensson^{a,g}, José María Bermúdez de Castro^h, Eudald Carbonellⁱ, Michael Dunn^j, Jan Storå^e, Eneko Iriarte^k, Juan Luis Arsuaga^{c,d}, José-Miguel Carretero^{c,k}, Anders Götherström^e, and Mattias Jakobsson^{a,l,2}

^aDepartment of Evolutionary Biology, Uppsala University, 75236 Uppsala, Sweden; ^bDepartment of Archaeology and History, La Trobe University, Melbourne, VIC 3086, Australia; ^cCentro Mixto, Universidad Complutense de Madrid–Instituto de Salud Carlos III de Evolución y Comportamiento Humano, 28029 Madrid, Spain; ^dDepartamento de Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, 28040 Madrid, Spain; ^eDepartment of Archaeology and Classical Studies, Stockholm University, 10691 Stockholm, Sweden; ^fDepartment of Genetics, Harvard Medical School, Boston, MA 02115; ^gDepartment of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden; ^hCentro Nacional de Investigación sobre la Evolución Humana, 09002 Burgos, Spain; ⁱInstitut Català de Paleoeologia Humana i Evolució Social, CEIP Marcel·lí Domingo, 43007 Tarragona, Spain; ^jDepartment of Linguistics and Philology, Uppsala University, 75238 Uppsala, Sweden; ^kLaboratorio de Evolución Humana, Departamento de Ciencias Históricas y Geografía, Universidad de Burgos, 09001 Burgos, Spain; and ^lScience for Life laboratory, Uppsala University, 75123 Uppsala, Sweden

Edited by Eske Willerslev, University of Copenhagen, Copenhagen, Denmark, and approved July 29, 2015 (received for review May 21, 2015)

The consequences of the Neolithic transition in Europe—one of the most important cultural changes in human prehistory—is a subject of great interest. However, its effect on prehistoric and modern-day people in Iberia, the westernmost frontier of the European continent, remains unresolved. We present, to our knowledge, the first genome-wide sequence data from eight human remains, dated to between 5,500 and 3,500 years before present, excavated in the El Portalón cave at Sierra de Atapuerca, Spain. We show that these individuals emerged from the same ancestral gene pool as early farmers in other parts of Europe, suggesting that migration was the dominant mode of transferring farming practices throughout western Eurasia. In contrast to central and northern early European farmers, the Chalcolithic El Portalón individuals additionally mixed with local southwestern hunter-gatherers. The proportion of hunter-gatherer-related admixture into early farmers also increased over the course of two millennia. The Chalcolithic El Portalón individuals showed greatest genetic affinity to modern-day Basques, who have long been considered linguistic and genetic isolates linked to the Mesolithic whereas all other European early farmers show greater genetic similarity to modern-day Sardinians. These genetic links suggest that Basques and their language may be linked with the spread of agriculture during the Neolithic. Furthermore, all modern-day Iberian groups except the Basques display distinct admixture with Caucasus/Central Asian and North African groups, possibly related to historical migration events. The El Portalón genomes uncover important pieces of the demographic history of Iberia and Europe and reveal how prehistoric groups relate to modern-day people.

Ancient DNA | human prehistory | population genomics

The first practice of farming started ~11,000 years before present (yBP) in the Near East, before spreading west and north across Europe and transforming the way of life from small, mobile hunter-gatherer (HG) groups to larger sedentary, agrarian societies. Genomic studies of Stone-Age human remains from northern and central Europe have shown that the Neolithic transition was driven by migration, followed by subsequent admixture with HG groups (1–3). More than 3,000 km away from the Near Eastern origin of farming practices, the Iberian Peninsula represents the western geographic extreme of the Neolithic transition, which reached the Spanish Mediterranean coast ~7,650–7,550 yBP, the Iberian interior 7,350–7,250 yBP (4), and the Northern Cantabrian coast around 7,300–6,800 yBP (5). Investigations of mitochondrial DNA (mtDNA) haplogroup variation from Neolithic and Chalcolithic Iberian samples have been inconclusive, with some suggesting modern-day Near Eastern affinities (6) and others suggesting complex admixture between hunter-gatherers and farmers (7). The

processes and consequences of the Neolithic transition in the Iberian Peninsula are, thus, contentious. The relationship of early Iberian farmers to their Central European counterparts and HG groups in Iberia, as well as their contribution to the present-day genomic variation in Iberia, is crucial to understanding the demographic history of the region. Of particular interest is the connection of Basques to different ancient groups in Iberia, to shed light on their long-debated origins (6–13).

We investigated the remains of eight individuals from the Chalcolithic and Bronze Age periods excavated from the cave of the El Portalón de Cueva Mayor, of the Sierra de Atapuerca (Fig. 1A and *SI Appendix*, Fig. S1)—a site with a remarkably rich archaeological record, with human occupation from the Paleolithic to the historical period (14) (*SI Appendix*, section S1). The

Significance

The transition from a foraging subsistence strategy to a sedentary farming society is arguably the greatest innovation in human history. Some modern-day groups—specifically the Basques—have been argued to be a remnant population that connect back to the Paleolithic. We present, to our knowledge, the first genome-wide sequence data from eight individuals associated with archaeological remains from farming cultures in the El Portalón cave (Atapuerca, Spain). These individuals emerged from the same group of people as other Early European farmers, and they mixed with local hunter-gatherers on their way to Iberia. The El Portalón individuals showed the greatest genetic affinity to Basques, which suggests that Basques and their language may be linked with the spread of agriculture across Europe.

Author contributions: T.G., C.V., A.G., and M.J. designed research; C.V. organized and coordinated the experiments; C.V., Ó.O.S., and E.M.S. extracted DNA; C.V. and E.M.S. built libraries; C.V., H.M., I.U., R.R.-V., E.A.D., and E.M.S. conducted post PCR work; T.G., C.V., H.M., P.S., T.N., and M.J. analyzed genetic data with the following contributions: processing the data and population genomic analyses (T.G.), mtDNA (T.G. and H.M.), Y-chromosome (T.G. and T.N.); M.D., J.S., E.I., J.L.A., and J.-M.C. analyzed archaeological, linguistic, and anthropological data; C.V., I.U., R.R.-V., E.I., J.L.A., and J.-M.C. conducted field work; C.V., J.M.B.d.C., E.C., E.I., J.L.A., and J.-M.C. provided samples; and T.G., C.V., and M.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the European Nucleotide Archive (accession no. [PRJEB9783](https://www.ebi.ac.uk/ena/record/PRJEB9783)).

¹T.G. and C.V. contributed equally to this work.

²To whom correspondence should be addressed. Email: mattias.jakobsson@ebc.uu.se.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1509851112/-DCSupplemental.

Table 1. Sequence information for the eight ancient individuals in this study

Individual	Genome coverage	mt coverage	mt haplogroup	Y haplogroup	Biological sex	Age (C14 cal yBP)	Contamination estimate (95% CI) (41)
ATP2	4.08	341.2	U5b3	H2	XY	4,849–4,628	0.1 (0–0.3)
ATP3	0.03	14.4	K1a2b	—	XY	5,466–5,312	0 (0–7.2)
ATP7	0.04	16.8	J1c1b1	—	XX	5,295–4,894	9.2 (3.8–14.6)
ATP9	0.41	35.4	U5b1b	—	XX	3,700–3,568	1.4 (0–3.3)
ATP16	1.11	68.7	X2c	—	XX	5,211–4,866	1.4 (0.2–2.7)
ATP17	0.03	3.6	H3	—	XY	4,957–4,821	—
ATP20	0.01	3.6	U5a1c	—	XX	4,239–4,000	—
ATP12-1420	1.21	230.8	H3c	I2a2a	XY	4,960–4,829	0.5 (0–1.1)

CI, confidence interval; —, not determined.

phenotypes of the El Portalón individuals are in line with previous studies (*SI Appendix, section S6*) (1, 2, 10). The El Portalón mtDNA-haplogroup composition differs from a Chalcolithic sample of the El Mirador cave, also located in the Atapuerca cave system (6), which can be a consequence of somewhat limited sample sizes for this comparison. However, if this observation is supported by more extensive data from the two caves, it suggests stratification at a small geographic scale and complex population composition among early Iberian farmers, although detailed interpretations are limited for these single-locus comparisons.

Principal component analysis (PCA) of the El Portalón farmers together with relevant published Chalcolithic, Neolithic, and Mesolithic European genome-wide sequence data (1, 2, 10, 12, 19) (Fig. 1*A* and *B*) and a large set of modern-day Europeans (1) (Fig. 1*C* and *SI Appendix, section S7*) revealed a distinct separation of hunter-gatherers and early farmers irrespective of their geographic location (Fig. 1*C*). The closest extant populations to hunter-gatherers from Iberia, Scandinavia, and Central Europe are Northern Europeans; however, the hunter-gatherers fell outside the range of modern-day European genetic variation (Fig. 1*C*). In contrast, early farmers from Iberia, Scandinavia, and Central Europe grouped with modern-day Southern Europeans, consistent with outgroup f_3 statistics (*SI Appendix, Fig. S8*). These results

demonstrate that early European farmers, including those in Iberia, emerged from a common group of people (*SI Appendix, section S11* and *Dataset S3*). This observation indicates that farming was brought to Iberia via migration, similar to the process in Scandinavia (2, 3) and Central Europe (1).

We fitted an admixture graph model for ancient individuals with the greatest genome coverage (20) (using the Denisovan genome as an outgroup and ascertaining genetic variants in Yorubans) (21) (*SI Appendix, section S9*), which confirmed the known connection between Scandinavian hunter-gatherers (Motala12) and a Paleolithic Siberian (MA1) (1, 2). Chalcolithic farmers (Iberian ATP2 and the Tyrolean Iceman) and Scandinavian Neolithic farmers (Gok2) traced a substantial amount of their genetic ancestry to European HG groups, in contrast to the earliest farmers of Central Europe (NE1 and Stuttgart) (Fig. 2*A*), and this increase in HG admixture across Europe was significant as a function of time (Fig. 2*B*) ($R^2 = 0.69$, $P = 0.001$). The best fitting source for the HG admixture into the El Portalón individuals was the common ancestor of the nearby La Braña Mesolithic individual and a Mesolithic individual from Luxembourg (Loschbour) (Fig. 2*A*) whereas contemporary farmers from Central Europe (Iceman) and Scandinavia received their (best-fit) HG admixture from Scandinavian hunter-gatherers (Fig. 2*A*). These inferred admixture events

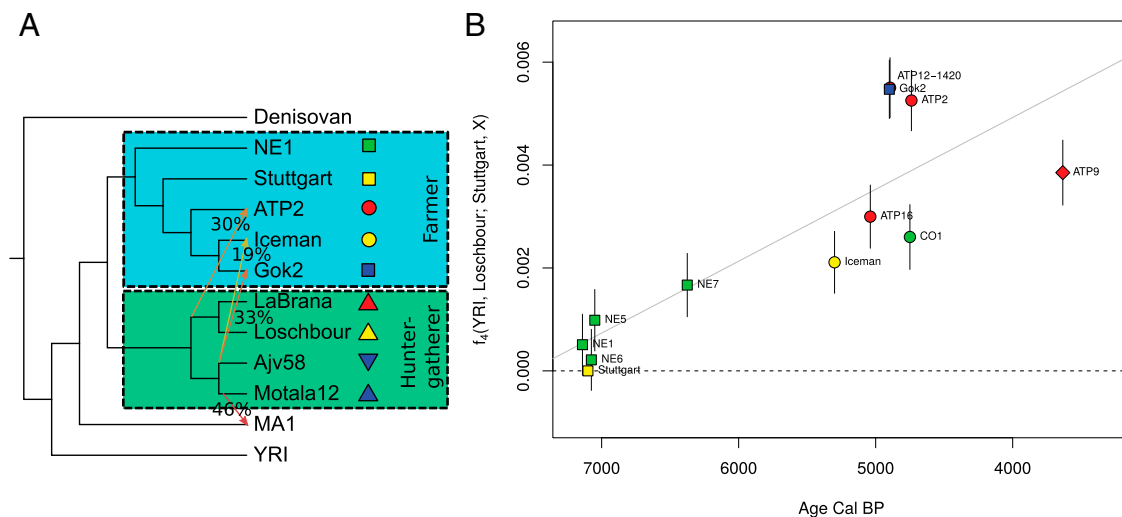


Fig. 2. Admixture among ancient groups based on genome sequence data. (A) Admixture graph with four migration edges for the individual with the highest sequencing coverage for each geographic site, region, or time period. To improve clarity, branch lengths are not drawn to scale and represent only the hierarchical clustering (see *SI Appendix, Fig. S10* for the topology drawn to scale). (B) Genetic affinity of the different early farmer individuals (f_4 -statistic) to a Central European Mesolithic individual (Loschbour, highest coverage Mesolithic genome), as a function of the dating of the early farmer individuals. A central European LBK individual is used as a baseline early Neolithic farmer (highest coverage early Neolithic individual, but the choice of reference individual does not affect the qualitative result) (*SI Appendix, section S11*). The Bronze Age, Chalcolithic, and Scandinavian farmers show greater levels of admixture with HG groups than the temporally older Neolithic Central European farmers (Pearson correlation $R^2 = 0.69$, $P = 0.001$).

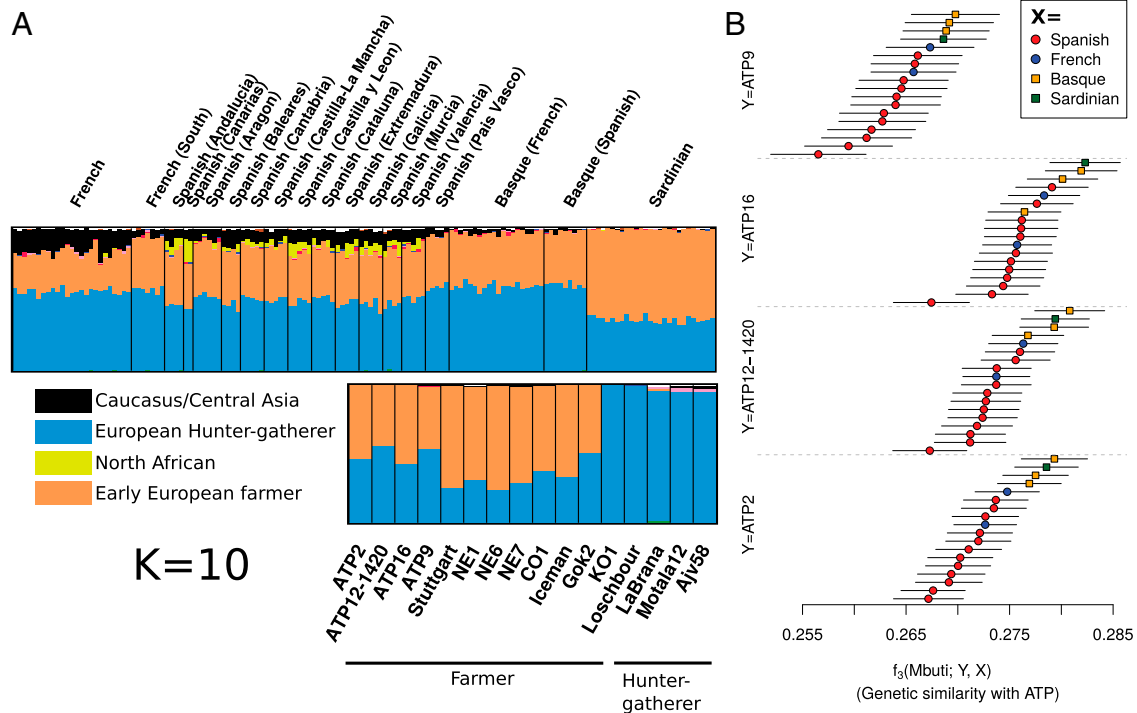


Fig. 3. Population structure of ancient and modern-day individuals. (A) Admixture fractions among modern-day individuals from Eurasia and North Africa together with 16 ancient individuals. Only ancient and modern-day individuals from Southwestern Europe are shown (see Dataset S1 for the complete plot with all individuals). Admixture components are labeled based on the populations/geographic regions in which they are modal. (B) Genetic similarity (measured using outgroup f_3 statistics) between Southwestern European groups and four El Portalón individuals. Error bars show ± 1 SE. Higher values represent greater genetic similarity between the El Portalón farmers and the modern-day Southwestern populations in the legend.

demonstrate that different farmer populations had different HG groups as the best proxy for the source of admixture (D-tests showed similar results of multiple admixture events in different parts of Europe) (SI Appendix, section S11 and Dataset S4). These analyses showed that, whereas early farmers—who were likely more numerous than the hunter-gatherers (SI Appendix, Fig. S9)—spread across Europe, they assimilated HG populations, a process that continued locally for several millennia (2, 17).

To further investigate the relationship between the El Portalón farmers and modern-day individuals, we inferred admixture fractions (22) among a large set of modern-day individuals from Eurasia and North Africa (Fig. 3A and Datasets S1 and S5). All modern-day Iberian groups displayed ancestry from early farmers and hunter-gatherers and also showed admixture from North Africa (Fig. 3A, yellow component) (23) and the Caucasus/Central Asia (Fig. 3A, dark purple component), potentially related to the observed migration during the Bronze Age (24, 25) or the later Roman Empire ruling of Iberia. Basques (including French Basques) were an exception; they display ancestry from early farmers and hunter-gatherers, similar to other modern-day Iberian groups, but little or no admixture from North Africa and the Caucasus/Central Asia (1, 23) (Fig. 3A and SI Appendix, section S10). Interestingly, among all European groups, Basques and Sardinians displayed strong genetic affinity to the El Portalón farmers (Fig. 3B and SI Appendix, Fig. S8). However, all other early farmers were closer to Sardinians (SI Appendix, Figs. S11 and S12), and Basques were closer to El Portalón individuals (or equally close for Gok2) compared with all other early farmers (SI Appendix, Fig. S13). To further test the scenario of Basques being the genetically most similar group to the El Portalón farmers, we computed D-statistics for different population topologies. All topologies where Basques were an outgroup to the highest coverage El Portalón individual (ATP2) and

another modern-day Spanish population—D(Mbuti, Basques; other Spanish, ATP2)—were rejected [36 tests, false discovery rate (FDR) < 0.01] (SI Appendix, section S11) whereas all topologies using Basques as an ingroup with ATP2, and another Spanish population as an outgroup, were consistent with the data. Test results for the other ATP individuals showed qualitatively similar patterns (SI Appendix, section S11 and Datasets S6 and S7). Our data suggest that modern-day Basques traced their genetic ancestry to early Iberian farmers.

This high similarity of modern Basques to El Portalón individuals was surprising because Basques have been posited as a remnant isolated population with a close relationship to the Mesolithic inhabitants of the region, based on classical genetic markers (26) and mtDNA haplogroup data (8), although the level of continuity has been unclear (12, 13). The Basque language (Euskara) is a linguistic isolate, with no proven relationships with any languages now spoken in Europe or elsewhere (11), and it has commonly been concluded that the Basque language is a relict of the ancient, preagricultural linguistic diversity of Europe, with roots as far back as the Paleolithic (SI Appendix, section S12) (9). Our data, suggesting that Basques trace their genetic ancestry to early Iberian farmers, challenges this assumption. The alternative interpretations of the linguistic history of Europe are, however, unclear. The remaining languages of Western Europe belong to the Indo-European family (27). The origin of the Indo-European language family is itself controversial (28), with most debate polarized between proponents of the steppe hypothesis, that Indo-European was introduced into Europe from the East during the Bronze Age (~4,500 yBP) (29), and the Anatolian hypothesis, that Indo-European language dispersed from Anatolia during the Neolithic (30, 31). There was genetic turnover associated with Yamnaya and Corded Ware cultures at ~4,500 yBP, which may

thus be associated with a primary or secondary dispersal of Indo-European languages (25). A possible interpretation of the role of Basque in this scenario would be that it is a descendent of the language (or one of the languages) of the early farmers, and some scholars have posited that the Basque language was related to the pre-Roman language of Sardinia (Paleosardo) (32). The two Southern European population isolates of Sardinians and Basques were genetically associated with the early farmers of Europe that drove the Neolithic transition (1–3), and close contacts between Iberia and Sardinia in the Neolithic are also indicated by archaeological finds (33). However, the possibility remains that the Basque language is a retention of the pre-agricultural linguistic diversity.

Conclusions

In summary, our ancient genomic sequence data from the El Portalón individuals and our analyses suggest the following model of events: The incoming early farmers, who could have spoken a non-Indo-European language, assimilated resident hunter-gatherers, with this admixed group becoming the ancestors of modern-day Iberian groups. Basques remained relatively isolated (compared with other Iberian groups) with marked continuity since the Neolithic/Chalcolithic period, but not since the Mesolithic (contrary to refs. 8, 9, and 26). Later migration into Iberia, possibly during the long reign of the Roman Empire and the 7th to 13th century period of Moorish rule of the peninsula, led to distinct and additional admixture in all Iberian groups but the Basque population (23).

Materials and Methods

Archaeological Samples. Sixteen bone and teeth human remains, representing sixteen individuals from the Chalcolithic and Bronze Age site of El Portalón (Spain) (14) were sampled for ancient DNA analyses. The samples had been excavated between 2000 and 2012, and C14 dates were obtained for each of them using accelerator mass spectrometry (AMS). See *SI Appendix, section S1* for details.

Sequencing. DNA was extracted from bones and teeth (34–36); DNA extracts were converted into blunt-end Illumina libraries (37). All samples were prepared in dedicated ancient DNA (aDNA) facilities at the Evolutionary Biology Center in Uppsala, Sweden. The libraries were sequenced on Illumina's HiSeq platform at the SNP&SEQ Technology Platform SciLife Sequencing Centre in Uppsala. All 16 samples were screened for human DNA, and only individuals with over 1% of human DNA content ($n = 8$) were used for downstream analysis. See *SI Appendix, section S2* and *Dataset S8* for details.

Next Generation Sequencing Data Processing and Authentication. Paired-end reads were merged, and remaining adapters were trimmed (38). The merged and trimmed reads were subsequently mapped to the human reference genome using BWA (39); potential PCR duplicates with identical start and end coordinates were collapsed into consensus sequences. The sequences showed a deamination pattern toward fragment ends, which are characteristic for ancient DNA (16). Contamination was estimated based on

discordant sites in mitochondria and the X chromosome in males (40–42). A detailed description can be found in *SI Appendix, section S3*.

Uniparental Haplogroups. Consensus sequences for the mitochondrial genomes of all samples were called using the *samtools* package (43). We used *haplofind* (44) to assign the mitochondrial genome to known mitochondrial haplogroups. Y haplogroups were assigned based on PhylotreeY (45). We excluded all non-SNP sites, transition sites (to avoid deamination damage), and A/T and G/C SNPs (to avoid strand misidentification). See *SI Appendix, sections S4 and S5* for details.

Modern Reference Data. The ancient samples were merged with the Human Origins genotype data (1, 46), excluding transition sites and sites showing indels. Most of the ancient samples have sequencing depths too low to confidently call diploid genotypes. Therefore, we randomly sampled one allele per individual and SNP site. Only reads and bases with a minimum mapping and base quality of 30 were considered.

To increase power for the comparison of sequenced ancient individuals, we repeated the same procedure for 1.9 million transversion SNPs, which were polymorphic in Yorubans of the 1000 Genomes Project phase 3 data (21). See *SI Appendix, section S7* for a detailed description.

Population Genetic Analysis. PCAs of ancient individuals and modern European populations from the Human Origins dataset were conducted using EIGENSOFT (47). Ancient individuals were projected onto the PC1-PC2 space using Procrustes analysis (48). We calculated D-statistics (46) to check for consistency of the data with different tree topologies and f_3 and f_4 statistics (46, 49) to estimate affinities among populations. See *SI Appendix, section S8* for more details.

TreeMix (20) was used to infer a bifurcating population tree plus admixture events for the highest coverage individuals from each of the ancient European groups using Denisovans (50) as an outgroup and restricting the analysis to transversion SNPs that were polymorphic in Yorubans (21). Correction for low sample sizes was turned off (-noss), and SEs were estimated using blocks of 500 SNPs. We show the admixture graph for $m = 4$ in Fig. 2A, and a detailed description of the methods and results for other migration events can be found in *SI Appendix, section S9*.

Model-based clustering of the ancient individuals together with Eurasian and North African populations from the Human Origins dataset (1) was conducted with ADMIXTURE (22). The genotype data were pruned for linkage disequilibrium, and we tested different numbers of clusters from $K = 2$ to $K = 15$. The results of 50 iterations per number of clusters were combined using CLUMPP (51) and plotted with Distruct (52). We chose to display $K = 10$ in the main paper (Fig. 3A) because it is the lowest value of K to show a clear distinction between hunter-gatherer, early farmer, North African, and Near Eastern components. The plot for all K s is shown in *Dataset S1*, and more details can be found in *SI Appendix, section S10*.

ACKNOWLEDGMENTS. We thank C. Bycroft for discussions and the Portalón Excavation Team for extensive research and field work on the site. Sequencing was conducted at the Uppsala University SNP&SEQ Technology Platform. This project was supported by grants from the Wenner-Gren Foundations (to T.G.), the Knut and Alice Wallenberg Foundation (to M.J.), the Swedish Research Council (to M.J.), the European Research Council (to M.J.) a Transforming Human Societies Research Focus Area and Bridging Fellowship from La Trobe University (to C.V.), and Spanish Ministerio de Economía y Competitividad (CGL2012-38434-C03-01). Computations were performed at the Swedish National Infrastructure for Computing (SNIC-UPPMAX).

- Lazaridis I, et al. (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513(7518):409–413.
- Skoglund P, et al. (2014) Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* 344(6185):747–750.
- Skoglund P, et al. (2012) Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336(6080):466–469.
- García-Martínez de Lagrán I (2015) Recent data and approaches on the Neolithization of the Iberian Peninsula. *Eur J Archaeol* 18(3):429–453.
- Fano MA, Cubas M, Wood R (2014) The first farmers in Cantabrian Spain: Contribution of numerical chronology to understand an historical process. *Quat Int* 364:153–161.
- Gómez-Sánchez D, et al. (2014) Mitochondrial DNA from El Mirador cave (Atapuerca, Spain) reveals the heterogeneity of Chalcolithic populations. *PLoS One* 9(8):e105105.
- Hervella M, et al. (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(4):e34417.
- Behar DM, et al.; Genographic Consortium (2012) The Basque paradigm: Genetic evidence of a maternal continuity in the Franco-Cantabrian region since pre-Neolithic times. *Am J Hum Genet* 90(3):486–493.
- Cavalli-Sforza L (1988) The basque population and ancient migrations in Europe. *MUNIBE (Antropología y Arqueología)* 6 (Suppl):129–137.
- Gamba C, et al. (2014) Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun* 5:5257.
- Kurlansky M (1999) *The Basque History of the World* (Jonathan Cape, London).
- Olalde I, et al. (2014) Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* 507(7491):225–228.
- Sánchez-Quinto F, et al. (2012) Genetic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr Biol* 22(16):1494–1499.
- Carretero JM, et al. (2008) A Late Pleistocene-Early Holocene archaeological sequence of Portalón de Cueva Mayor (Sierra de Atapuerca, Burgos, Spain). *Munibe (Antropología-Arqueología)* 59:67–80.
- Castilla M, Carretero JM, Gracia A, Arsuaga JL (2014) Evidence of rickets and/or scurvy in a complete Chalcolithic child skeleton from the El Portalón site (Sierra de Atapuerca, Spain). *J Anthropol Sci* 92:257–271.
- Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7(3):e34131.
- Brandt G, et al.; Genographic Consortium (2013) Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science* 342(6155):257–261.

18. Malmström H, et al. (2009) Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Curr Biol* 19(20):1758–1762.
19. Keller A, et al. (2012) New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat Commun* 3:698.
20. Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* 8(11):e1002967.
21. Abecasis GR, et al.; 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56–65.
22. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19(9):1655–1664.
23. Botigué LR, et al. (2013) Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc Natl Acad Sci USA* 110(29):11791–11796.
24. Allentoft ME, et al. (2015) Population genomics of Bronze Age Eurasia. *Nature* 522(7555):167–172.
25. Haak W, et al. (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522(7555):207–211.
26. Cavalli-Sforza LL, Menozzi P, Piazza A (1996) *The History and Geography of Human Genes* (Princeton Univ Press, Princeton).
27. Gorrochategui J, Lakarra JA (2013) Why Basque language cannot be, unfortunately, an Indo-European language. *J Indo-Eur Stud* 41(1-2):203–237.
28. Mallory JP (2013) Twenty-first century clouds over Indo-European homelands. *J Lang Relatsh* 9:145–154.
29. Chang W, Cathcart C, Hall D, Garrett A (2015) Ancestry-constrained phylogenetic analysis supports the Indo-European Steppe hypothesis. *Language* 91(1):194–244.
30. Bouckaert R, et al. (2012) Mapping the origins and expansion of the Indo-European language family. *Science* 337(6097):957–960.
31. Gray RD, Atkinson QD (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426(6965):435–439.
32. Blasco Ferrer E (2010) Iberian ordumeles, Paleo-Sardinian Ortimele, Ortarani and Araunele: Cognitive semantics and Substrata Research. *J Indo-Eur Stud* 38(3&4):373–383.
33. Terradas X, et al. (2014) Neolithic diffusion of obsidian in the western Mediterranean: New data from Iberia. *J Archaeol Sci* 41:69–78.
34. Dabney J, et al. (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci USA* 110(39):15758–15763.
35. Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. *Nat Protoc* 2(7):1756–1762.
36. Yang DY, Eng B, Wayne JS, Dудар JC, Saunders SR (1998) Technical note: Improved DNA extraction from ancient bones using silica-based spin columns. *Am J Phys Anthropol* 105(4):539–543.
37. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010(6):t5448.
38. Kircher M (2012) Analysis of high-throughput ancient DNA sequencing data. *Methods Mol Biol* 840:197–228.
39. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
40. Fu Q, et al. (2013) DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci USA* 110(6):2223–2227.
41. Green RE, et al. (2008) A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* 134(3):416–426.
42. Rasmussen M, et al. (2011) An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334(6052):94–98.
43. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
44. Vianello D, et al. (2013) HAPLOFIND: A new method for high-throughput mtDNA haplogroup assignment. *Hum Mutat* 34(9):1189–1194.
45. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30(2):E386–E394.
46. Patterson N, et al. (2012) Ancient admixture in human history. *Genetics* 192(3):1065–1093.
47. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2(12):e190.
48. Wang C, et al. (2010) Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Stat Appl Genet Mol Biol* 9(1):13.
49. Reich D, Thangaraj K, Patterson N, Price AL, Singh L (2009) Reconstructing Indian population history. *Nature* 461(7263):489–494.
50. Meyer M, et al. (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338(6104):222–226.
51. Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14):1801–1806.
52. Rosenberg NA (2004) dstruct: A program for the graphical display of population structure. *Mol Ecol Notes* 4(1):137–138.