Researchers Unravel The Mysteries Of The Basque People

September 9, 2015 | by Josh L Davis
http://www.iflscience.com/environment/researchers-unravel-mysteries-basque-people

With their distinct genetic make-up and ancient language, the origin of people from the Basque Country in northern Spain and southern France has long been an enigma. Previously thought to be a population of unmixed hunter-gatherers that survived the influx of farmers from the Middle East around 6,000 years ago, new genetic evidence suggests that things aren’t quite so clear-cut.

Instead, it appears that the Basques’ closest ancestors are early Iberian farmers who were descended from that first wave of Neolithic farmers, which then mixed with the local hunter-gatherers. The researchers think that the Basque people’s unique characteristics survived in isolation as, after this first initial mingling of hunter-gatherers and farmers, the Basque people became separated from subsequent waves of farmers, first from the East then from the South, and continued instead to mix further with the local hunter-gatherers.

“We show that the hunter-gatherer genetic component increases with time during several millennia, which means that later farmers were genetically more similar to hunter-gatherers than their
forefathers who brought farming to Europe,” explains Dr Torsten Günther, one of the lead authors of the study, in a statement. The research will be published in The Proceedings of the National Academy of Sciences.

Because the Basques are genetically distinct from the rest of Europe, and their language is what’s classed as a “pre-Indo-European” language (one that existed in prehistoric Europe before the dominant language type swept across the continent from the east), the region has always puzzled anthropologists. The researchers of this new study hoped to find answers in the 3,500 to 5,500-year-old remains of eight Iberian farmers discovered in the El Portalon cave in northern Spain.

The site shows remarkable preservation of artifacts, giving an incredible snap shot of Neolithic life. Archeologists have unearthed stone tools, ceramics, metal objects, and human and animal bones. But what’s so impressive is the preservation of organic remains, which allowed the scientists to extract and sequence the DNA of these ancient people.

“Our results show that the Basques trace their ancestry to early farming groups from Iberia, which contradicts previous views of them being a remnant population that trace their ancestry to Mesolithic hunter-gatherer groups,” says Professor Mattias Jakobsson of Uppsala University, who led the study. “The difference between Basques and other Iberian groups is these latter ones show distinct features of admixture from the east and from north Africa.”

The results of the study highlight the intricacies of the movements within Europe during the Neolithic, and how certain communities can become isolated and separated from others who aren’t far away geographically. It also helps to solve the longstanding puzzle of the Basques, and shows that they’ve remained relatively isolated for the last 5,000 years, but not much longer than that.

Main Image: Terre et Côte Basques/Flickr CC BY-NC-ND 2.0

Ancient DNA cracks puzzle of Basque origins

Image copyright AFP

caption. Two men saw a tree trunk with a "tronza" (traditional Basque saw) during a rural sports championship
DNA from ancient remains seems to have solved the puzzle of one of Europe's most enigmatic people: the Basques.

The distinct language and genetic make-up of the Basque people in northern Spain and southern France has puzzled anthropologists for decades.

One theory proposed that they were an unmixed pocket of indigenous hunters. Now, a study in *PNAS* journal suggests they descend from early farmers who mixed with local hunters before becoming isolated for millennia.

The Basques have unique customs and a language - Euskera - that is unrelated to any other spoken in Europe, or indeed the world.

Nestled in a mountainous corner of Atlantic Europe, they also show distinct genetic patterns to their neighbours in France and Spain.

It seemed logical that they were representatives of an older layer of population settlement, but just how far back their roots went has been a topic of debate.

Mattias Jakobsson from Uppsala University in Sweden analysed the genomes of eight Stone Age human skeletons from El Portalón in Atapuerca, northern Spain.

These individuals lived between 3,500 and 5,500 years ago, after the transition to farming in southwest Europe.

The results show that these early Iberian farmers are the closest ancestors to present-day Basques.

**Go west**

Comparisons with other ancient European farmers show that agriculture was brought to Iberia by the same migrant groups that introduced it to central and northern Europe. These pioneers expanded from a homeland in the Near East, sweeping across Europe about 7,000 years ago to usher in the period known as the Neolithic.

Once the farmers settled down, they mixed with local hunter-gatherers - the descendants of people who lived in Europe during the last Ice Age.

Indeed, the El Portalón individuals had more hunter-gatherer ancestry than pioneer farmers from Germany, Hungary and Spain who lived several thousand years earlier.

The new study also goes some way to explaining some of the differences between the Basques and their neighbours in France and Spain.

After the initial farmer-hunter mixture was set, the ancestors of the Basques became isolated from surrounding groups - perhaps due to a combination of geography and culture.

"It's hard to speculate, but we've been working with Basque historians and it's clear from the historical record that this area was very difficult to conquer," Prof Jakobsson told BBC News.

This means the Basque area was largely unaffected by subsequent migrations that shaped genetic patterns elsewhere in Europe.
Migration and isolation

One of these movements occurred in the Bronze Age, when pastoralists from the Steppe - on the eastern periphery of the continent - travelled west en masse. This migration probably spread Indo-European languages across Europe, affecting the central and northern parts of the continent to a greater extent than the south.

While the genomes of French and Spanish individuals showed evidence of this eastern genetic input, those of Basques did not.

Another migration served to further differentiate Basques from their Spanish neighbours. In AD 711, a Muslim army crossed from North Africa into Iberia, beginning an occupation that lasted more than 700 years.

Again, while a small amount of North African and Sub-Saharan ancestry can be detected in the Spanish, it is largely absent from the Basques.

Previous studies have shown that people native to the Italian island of Sardinia are most genetically similar to the pioneer farmers of central Europe.

The Sardinians also became isolated after the agricultural transition, but they lack the additional hunter-gatherer ancestry that characterises the Basques.

Paradoxically, while archaeology shows that Europe's earliest farmers hailed from the Near East, populations living in that region today do not particularly resemble them genetically.

This is because genetic patterns in Turkey and the Middle East were re-shaped by migrations from surrounding areas after the Early Neolithic.

The proportion of hunter-gatherer ancestry varies across Europe, peaking at about 30% in Estonians and Lithuanians, but no "pure" indigenous Europeans remain. They appear to have been assimilated by the Neolithic migrants, who probably had greater numbers.
Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques

Torsten Güntner,1,2 Cristina Valdiosera,3,4,5,6 Helena Malmström,7,8 Irene Ureña,9,10 Ricardo Rodríguez-Varela,11,12 Ödndy Ösk Sverrisdóttir,13 Evangelia A. Daskalaki,14,15 Pontus Skoglund,13,16 Thijessen Naido,17 Emma M. Svensson,18,19 José María Bermúdez de Castro,12 Eudald Carbonell,12 Michael Dunn,12 Jan Storå,12 Eneko Iriarte12, Juan Luis Arsuaga11, José-Miguel Carretero11, Anders Götherström10, and Mattias Jakobsson11,2

*Department of Evolutionary Biology, Uppsala University, 75236 Uppsala, Sweden; †Department of Archaeology and History, La Trobe University, Melbourne, VIC 3086, Australia; ‡Centro Mixto, Universidad Complutense de Madrid–Instituto de Salud Carlos III de Evolución y Comportamiento Humano, 28029 Madrid, Spain; §Department of Paleontology, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, 28040 Madrid, Spain; †Department of Archaeology and Classical Studies, Stockholm University, 10691 Stockholm, Sweden; ‡Department of Genetics, Harvard Medical School, Boston, MA 02115; §Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden; †Centro Nacional de Investigación sobre la Evolución Humana, 08002 Burgos, Spain; ‡Institut Català de Paleoeconomia Humana i Evolució Social, CEIP Marcel·lí Domingo, 43007 Tarragona, Spain; §Department of Linguistics and Philology, Uppsala University, 75238 Uppsala, Sweden; †Laboratorio de Evolución Humana, Departamento de Ciencias Históricas y Geografía, Universidad de Burgos, 09001 Burgos, Spain; and *Science 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 across 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 ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques.

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 co-led the project; A.G. and M.J. performed research; C.V. and C.E. contributed equally to this work; E.C. and M.J. contributed 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).

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.
human remains were associated with offerings, such as domestic animals and pottery vessels corresponding to the pre-Bell Beaker culture, and were directly radiocarbon-dated to between \( \sim 5,500 \) (Chalcolithic) and \( \sim 3,500 \) cal yBP (Bronze Age) (SI Appendix, Table S1 and Fig. S2). Seven of the burials contained fragmentary human remains whereas one burial was a near-complete skeleton of a male child showing signs of chronic malnutrition (15) (SI Appendix, section S1).

**Results and Discussion**

We generated genome-wide sequence data from all eight El Portalón individuals (Table 1 and SI Appendix, Table S2) ranging from 0.01\( \times \) to 4.08\( \times \), with four individuals with a depth of coverage at \( >0.4\times \). All samples displayed features characteristic of ancient DNA (16): (i) The DNA was fragmented and (ii) cytosine deamination was consistently higher at fragment termini (SI Appendix, section S3). Mitochondrial (mt) DNA-based contamination estimates were <1.5% for all four \( >0.4\times \) coverage individuals (Table 1). The eight individuals, genetically inferred to be four males and four females (Table 1), carried mtDNA haplogroups associated with early farmers of Europe (e.g., haplogroups K, J, and X) (10, 17), with hunter-gatherers (e.g., haplogroup U5) (18), or with both groups (e.g., haplogroup H) (17) (SI Appendix, section S4 and Dataset S2). Two Y-haplogroups were also determined (haplogroups H2 and I2a2a), but, given the small sample size, it is difficult to draw any solid conclusion (SI Appendix, section S5). The mtDNA- and Y-haplogroup composition of the El Portalón individuals is consistent with admixture between the incoming farmers and local HG groups, with contributions from both sexes (SI Appendix, sections S4 and S5). In addition, we inferred phenotype information from known genotype–phenotype associations, and we note that the
Table 1. Sequence information for the eight ancient individuals in this study

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

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. 1C and SI Appendix, section S7) revealed a distinct separation of hunter–gatherers and early farmers irrespective of their geographic location (Fig. 1C). 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. 1C). In contrast, early farmers from Iberia, Scandinavia, and Central Europe grouped with modern-day Southern Europeans, consistent with outgroup f3 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 (NEI and Stuttgart) (Fig. 2A), and this increase in HG admixture across Europe was significant as a function of time (Fig. 2B) (R2 = 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. 24) whereas contemporary farmers from Central Europe (Iceman) and Scandinavia received their (best-fit) HG admixture from Scandinavian hunter–gatherers (Fig. 24). 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 (f3-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 R2 = 0.69, P = 0.001).
Figure 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. 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.

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, similar to other modern-day Iberian groups, but little or no admixture from North Africa and the Caucasus/Central Asia (1, 23) (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 site 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 f3 and f4 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 Northern 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-gatherers, early farmer, North African, and Near Eastern components. The plot for all Ks 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).

5. Fano MA, Cubas M, Wood R (2014) The first Neolithic groups in a complete Chalcolithic child skeleton from the El Portalón site (Sierra de Atapuerca, Burgos, Spain).