# LIVERPOOL JOHN MOORES UNIVERSITY

## LJMU Research Online

Rodríguez-Varela, R, Günther, T, Krzewińska, M, Storå, J, Gillingwater, TH, MacCallum, M, Arsuaga, JL, Dobney, K, Valdiosera, C, Jakobsson, M, Götherström, A and Girdland-Flink, L

 Genomic Analyses of Pre-European Conquest Human Remains from the Canary Islands Reveal Close Affinity to Modern North Africans

http://researchonline.ljmu.ac.uk/id/eprint/7546/

**Article**

# Current Biology

# Genomic Analyses of Pre-European Conquest Human Remains from the Canary Islands Reveal Close Affinity to Modern North Africans

## Highlights

- The first genome-wide data from the Guanches confirm a North African origin

- The Guanches were genetically most similar to modern North African Berbers

- Modern inhabitants of Gran Canaria carry an estimated 16%–31% Guanche autosomal ancestry

## Authors

Ricardo Rodríguez-Varela,
Torsten Günther, Maja Krzewińska, ...,
Mattias Jakobsson,
Anders Götherström,
Linus Girdland-Flink

## Correspondence

ricardo.rodriguez.varela@arklab.su.se
(R.R.-V.),
linusflink@hotmail.com (L.G.-F.)

## In Brief

Rodríguez-Varela et al. report the first genome-wide data from the aboriginals of the Canary Islands, the Guanches, confirming the long-held hypothesis that the Guanches originated from a North African Berber-like population and showing that modern inhabitants of Gran Canaria carry an estimated 16%–31% Guanche autosomal ancestry.

CrossMark

CellPress

# Genomic Analyses of Pre-European Conquest Human Remains from the Canary Islands Reveal Close Affinity to Modern North Africans

Ricardo Rodríguez-Varela,[1,2,11,*] Torsten Günther,[3] Maja Krzewińska,[1] Jan Storå,[1] Thomas H. Gillingwater,[4] Malcolm MacCallum,[4] Juan Luis Arsuaga,[2,5] Keith Dobney,[6,7,8] Cristina Valdiosera,[2,9] Mattias Jakobsson,[3] Anders Götherström,[1] and Linus Girdland-Flink[1,10,*]

[1]Department of Archaeology and Classical Studies, Stockholm University, 10691 Stockholm, Sweden
[2]Centro Mixto, Universidad Complutense de Madrid–Instituto de Salud Carlos III de Evolución y Comportamiento Humano, 28029 Madrid, Spain
[3]Department of Organismal Biology, Uppsala University, 75236 Uppsala, Sweden
[4]Anatomy, Edinburgh Medical School: Biomedical Sciences, University of Edinburgh, Edinburgh EH8 9AG, UK
[5]Departamento de Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, 28040 Madrid, Spain
[6]Department of Archaeology, School of Geosciences, University of Aberdeen, St. Mary's, Aberdeen AB24 3UF, UK
[7]Department of Archaeology, Classics and Egyptology, University of Liverpool, Liverpool L69 7WZ, UK
[8]Department of Archaeology, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[9]Department of Archaeology and History, La Trobe University, Melbourne, VIC 3086, Australia
[10]Research Centre in Evolutionary Anthropology and Palaeoecology, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[11]Lead Contact
*Correspondence: ricardo.rodriguez.varela@arklab.su.se (R.R.-V.), linusflink@hotmail.com (L.G.-F.)
https://doi.org/10.1016/j.cub.2017.09.059

## SUMMARY

The origins and genetic affinity of the aboriginal inhabitants of the Canary Islands, commonly known as Guanches, are poorly understood. Though radiocarbon dates on archaeological remains such as charcoal, seeds, and domestic animal bones suggest that people have inhabited the islands since the 5th century BCE [1–3], it remains unclear how many times, and by whom, the islands were first settled [4, 5]. Previously published ancient DNA analyses of uniparental genetic markers have shown that the Guanches carried common North African Y chromosome markers (E-M81, E-M78, and J-M267) and mitochondrial lineages such as U6b, in addition to common Eurasian haplogroups [6–8]. These results are in agreement with some linguistic, archaeological, and anthropological data indicating an origin from a North African Berber-like population [1, 4, 9]. However, to date there are no published Guanche autosomal genomes to help elucidate and directly test this hypothesis. To resolve this, we generated the first genome-wide sequence data and mitochondrial genomes from eleven archaeological Guanche individuals originating from Gran Canaria and Tenerife. Five of the individuals (directly radiocarbon dated to a time transect spanning the 7th–11th centuries CE) yielded sufficient autosomal genome coverage (0.21× to 3.93×) for population genomic analysis. Our results show that the Guanches were genetically similar over time and that they display the greatest genetic affinity to extant Northwest Africans, strongly supporting the hypothesis of a Berber-like origin. We also estimate that the Guanches have contributed 16%–31% autosomal ancestry to modern Canary Islanders, here represented by two individuals from Gran Canaria.

## RESULTS AND DISCUSSION

We processed twelve samples from twelve different individuals, of which eleven yielded genome-wide sequence data (Table 1). We sampled human teeth only from intact skulls. The ancient individuals, collected from caves on Tenerife and Gran Canaria (the two largest islands of the Canary archipelago) (Figure 1), were donated to the skull collection housed in the AMEU (Anatomical Museum Edinburgh University) in the 19th century (see Experimental Model and Subject Details).

The average read length (DNA fragmentation) and patterns of cytosine deamination, clustering at a significantly higher frequency at fragment termini, are consistent with expectations for ancient DNA (Figure S1A) [10]. Mitochondrial contamination estimates [11] for the five individuals included in further autosomal genomic analysis range from point estimates of 0.53% to 5.97%, showing that contamination is largely negligible (Table 1). The five individuals with the highest autosomal genome coverage (0.21× to 3.93×) were directly radiocarbon dated at the Svedberg Laboratory (Uppsala University) (Figure S1B) and span approximately 400 calendar years from the 7th to the 11th centuries CE (Table 1). All individuals predate the European colonization of the Canary Islands (15th century) and one predates the

**Table 1. Guanche Samples – Summary Statistics**

| Sample | Origin | Molecular Sex | Genome Coverage | Mitochondrial Genome Coverage | SNPs in HO Dataset | Mitochondrial Haplotypes | Y Chromosome Haplotypes | C$^{14}$ Radiocarbon Date Before Present | C$^{14}$ Radiocarbon Date, Calibrated Common Era | Mitochondrial Contamination Estimate/Confidence interval |
|---|---|---|---|---|---|---|---|---|---|---|
| gun002 | Tenerife | XY | 0.21 | 294.7 | 74,618 | H1cf | E1b1b1b1a1 E-M183 | 951 ± 26 | 1089.4 ± 65.5 | 3.63%/2.42%–4.85% |
| gun005 | Gran Canaria | XX | 0.47 | 341.1 | 140,873 | H2a | NA | 1082 ± 26 | 956 ± 61 | 2.41%/1.56%–3.27% |
| gun008 | Gran Canaria | XX | 0.30 | 690.9 | 101,216 | L3b1a | NA | 1116 ± 26 | 935.5 ± 56.5 | 1.65%/1.35%–1.95% |
| gun011 | Tenerife | XY | 3.93 | 931.6 | 370,465 | T2c1d2 | E1b1b1b1a1 E-M183 | 1216 ± 27 | 791.5 ± 96.5 | 0.53%/0.38%–0.69% |
| gun012 | Tenerife | XY | 0.54 | 214.3 | 157,104 | U6b1a | E1b1b1b1a1 E-M183 | 1421 ± 28 | 621 ± 39 | 5.97%/4.22%–7.73% |
| gun001 | Tenerife | XY | 0.016 | 9.6 | NA | U6b1a | NA | NA | NA | NA |
| gun004 | Tenerife | NA | 0.004 | 49.98 | NA | J1c3 | NA | NA | NA | 5.76%/3.57%–7.95% |
| gun006 | Gran Canaria | XY | 0.027 | 12.21 | NA | L3b1a | NA | NA | NA | 6.60%/1.87%–11.33% |
| gun007 | Gran Canaria | XY | 0.003 | 3.47 | NA | L3b1a | NA | NA | NA | NA |
| gun013 | Tenerife | NA | 0.005 | 14.2 | NA | U6b1a | NA | NA | NA | 13.79%/1.24%–26.34% |
| gun014 | Tenerife | XY | 0.008 | 4.2 | NA | U6b | NA | NA | NA | NA |

NA, not applicable. C$^{14}$ lab numbers in order of appearance in the table: Ua-55344, Ua-55345, Ua-55346, Ua-55347, Ua-55348. See also Table S1 and Figure S1B.
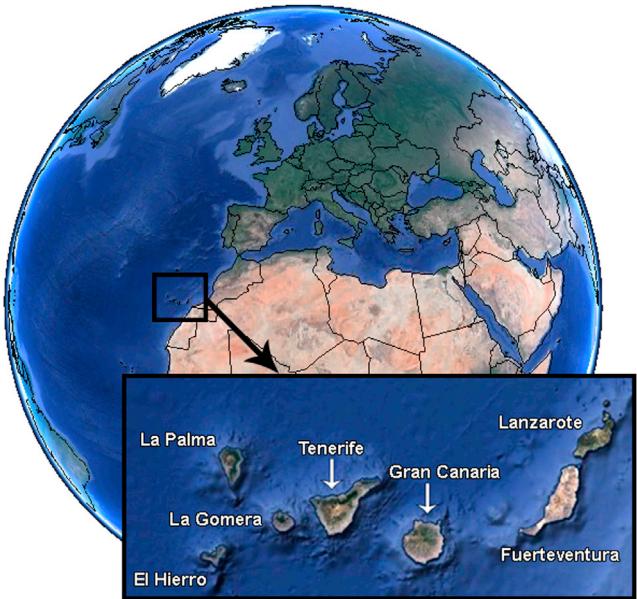


**Figure 1. Map of the Canary Islands**
Map template is modified from Google Earth (https://www.google.com/earth/).

Muslim conquest of the Maghreb (7th–8th centuries), events that have had significant impact on the gene pool of the Canary Islands and North Africa, respectively (Figure S1B) [6, 12]. Hence, our sample is a good representative of the pre-conquest aboriginals.

## Analysis of Uniparental Genetic Markers

The mitochondrial genome coverage for the eleven Guanche individuals ranges from 3.4× to 931×. The number of SNPs that support each haplotype varies between 38 and 58 and are reported as deviations from the Reconstructed Sapiens Reference Sequence (RSRS) [13] (Table S1). All positions supporting the predicted haplogroup calls and mutations missing for the absolute haplogroup assignment are listed in Table S1. We found six different mitochondrial haplogroups in eleven individuals (Tables 1 and S1). The Guanches analyzed here carried mitochondrial lineages such as J1c3, H2a, U6b, L3b1a, and T2c1d2 that are common across West Eurasia and/or North Africa [14] (Tables 1 and S1) and are consistent with previous studies on ancient Guanche mitochondrial DNA [6, 7]. Two individuals from Tenerife (gun001 and gun012) and one from Gran Canaria (gun013) carried the U6b1a haplotype, which is hypothesized to be endemic to and a founder lineage of the Canary Islands (Tables 1 and S1) [15]. We also found the H1cf haplotype in one individual from Tenerife (gun002), which is defined by a mutation at position 16260T (Table S1). H1-16260 is also considered a founder lineage of the Canary Islands, for two reasons: (1) it is present in all modern Canary populations [15] and rare outside the Canary Islands (found previously in only a single Algerian individual) [7, 16]; and (2) it was found in pre-Hispanic populations from Tenerife [6], La Palma [7], El Hierro [16], and La Gomera [17]. The three males from whom haplogroup-defining Y chromosome SNPs were retrieved carried the E1b1b1b1a1 (E-M183) haplotype (a major sub-clade of the
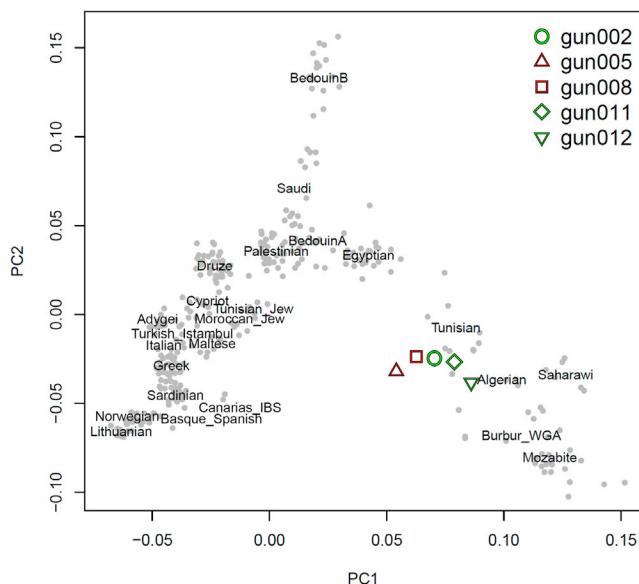
**Figure 2. PCA of the Five Guanche Individuals Used in the Autosomal Analysis**

Principal component analyses performed on the Guanches and a number of populations from Europe, the Middle East, and North Africa selected from the Human Origins dataset [19, 20]. Green symbols represent individuals from Tenerife; red symbols represent individuals from Gran Canaria. See also Figure S2.

haplogroup E1b1b1b1, defined by the derived M81 marker), which is again consistent with previous analyses of ancient Guanches [8]. This haplogroup is ubiquitous across modern North African populations and particularly common in Berber-speaking populations of North Africa [8, 18]. The derived mutations supporting the Y chromosome haplogroups are listed in Data S1, sheet 1.

**Population Genomic Analysis of Autosomal DNA**

A principal component analysis (PCA) of the five samples with the highest autosomal genome coverage, performed using genome-wide autosomal SNPs overlapping with Human Origins (HO) data [19, 20], reveals close affinity to modern Northwest African populations such as Tunisians and Algerians, but with a tendency (especially for individuals from Gran Canaria) to occupy a space outside modern Northwest African variation, closer to Europeans (Figures 2 and S2). However, outgroup $f_3$ statistics [19] suggest that the Guanches share more genetic drift with non-African test populations than with African test populations, including Northwest African populations of Berber origin (Data S1, sheet 2). This observation is inconsistent with the PCA and the uniparental genetic marker data, indicating that the outgroup $f_3$ statistic may be misleading, possibly due to the complex history of recent sub-Saharan admixture events in North African populations [12, 21] and the sensitivity of the $f_3$ estimator to such patterns. This issue seems to extend to other statistics based on allele frequency correlations such as the $D$ statistic [19] since $D$(Outgroup, Guanches; North African, Sardinian/Anatolian farmer) consistently produces highly significant positive values of $D$ (Z > 4), which would imply a closer rela-

tionship between Guanches and Sardinians and Anatolian farmers than between Guanches and North African populations (Data S1, sheet 3).

In order to resolve this, we used different statistics to measure genetic similarity between Guanches and modern populations, as well as between different modern North African populations and other populations (Table S2 and Quantification and Statistical Analysis). Both outgroup $f_3$ statistics and average pairwise differences identify Sardinians as the population sharing the most drift with both modern North Africans and Guanches (Table S2). This is in stark contrast to our expectation that North Africans would cluster with other North Africans. However, it replicates recent findings [21] showing that both ancient and modern Egyptians share more genetic drift with European populations than they do with other North African groups, while still showing evidence of strong genetic affinity to one another in other types of analyses. This behavior might be due to varying degrees of admixture from highly divergent sources (e.g., sub-Saharan populations) into different populations, which can affect the interpretation of $f_3$ statistics [22]. Therefore, we conclude that these statistics are not suitable to identify the modern population most similar to Guanches in the sense that we intend it for this study (Quantification and Statistical Analysis). In contrast, our extended analyses using $F_{ST}$ (fixation index) [23, 24] and $f_2$ [25] identify different North African and Near Eastern populations as most similar to modern North African populations and Guanches, results that are consistent with the PCA and the uniparental genetic analysis (Table S2; Figures 2 and S2). This finding is in agreement with anthropological, archaeological, linguistic [1, 4, 9, 26], and uniparental genetic data [6–8] and adds to a growing body of new evidence suggesting that some ancient domesticated plants [27] and animals [28] from the Canary Islands originated from North Africa.

The Guanches' Berber-like affinity is further supported by ADMIXTURE [29] analysis (Figures 3 and S3), where Guanches largely behave like modern Berbers across all values of K. At K = 10, a Northwest African-specific ancestry component makes up the greatest amount of autosomal ancestry in the Guanche and Berber populations in the HO dataset, such as the Mozabite and Saharawi. It is also ubiquitous across other Northwest African populations with Berber ancestry, such as Algerians and Tunisians, consistent with the PCA results. This ancestry component is also represented in present-day Canary Islanders and at a low proportion in some South European populations (Figures 3 and S3). Interestingly, it is also shared by Middle Eastern populations, including some Natufians (Figure 3). Y chromosome E1b1b haplotypes (though not M183 variants) were also common in Natufians (circa 11,000 BCE) and pre-pottery Neolithic male individuals from the Levant (circa 7,000 BCE), suggesting some affinity to North Africans [30].

The results of the ADMIXTURE analysis furthermore show that the Guanches carried early European farmer (EEF)-like ancestry; this ancestry component is widespread (though at varying proportions) in present-day North Africans and Middle Easterners but rare or largely absent in some Berber populations (Figure 3). The EEF component is strongly associated with early Neolithic farmers from Anatolia and Europe (as well as present-day Sardinians), hinting at a possible link between present-day North Africans and the expansion of Neolithic culture through the
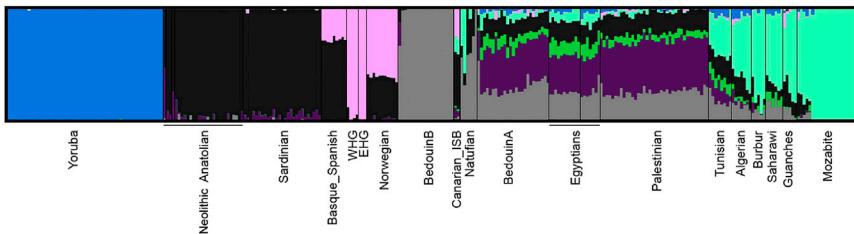
**Figure 3. ADMIXTURE Plot**

ADMIXTURE analysis at K = 10 of the five Guanche individuals that yielded the highest genome coverage (0.21× to 3.93×) and a selection of modern samples from the Human Origins dataset together with published ancient genomes. (See Figure S3 for complete admixture plot and references.) Within the Guanches cluster, the five individual bars represent (from left to right) gun005, gun008, gun011, gun012, and gun002.

Mediterranean [31–33], though it could also reflect post-Neolithic gene flow among Mediterranean groups [34]. The Guanches also appear to have carried varying proportions of Middle Eastern ancestry, best represented by the HO Bedouin_B population.

We also note that one Guanche individual (gun005) carried a greater proportion of hunter-gatherer (HG)-like ancestry than the other individuals, possibly suggesting low-level gene flow from a European source that predates the European conquest. Although our results are overall consistent with an origin from a single ancestral population in North Africa, the possible small-scale introgression from other sources postdating the earliest settling is consistent with archaeological finds of Phoenician-Punic amphora in Buena Vista (Lanzarote) and Roman amphora fragments retrieved from El Bebedero (Lanzarote), indicating that the islands (and local islanders) were in at least sporadic contact with other peoples and cultures prior to the European colonization in the 15th century CE [5].

The Europeans who colonized the islands in the 15th century CE found that the Guanches shared a similar language and material culture centered on goat herding and small-scale cultivation of domesticated crops, despite lacking boats and the ability to navigate the surrounding sea [1, 4, 5]. In the years following the European colonization, a great number of Guanches were killed in several wars throughout the islands [35]. Survivors integrated into the colonizing populations and left behind a genetic as well as cultural imprint on today's population, although the degree to which the former has shaped the gene pool of modern Canary Islanders remains unclear [6, 8, 26]. Since the ADMIXTURE analysis reveals that present-day Canary Islanders possess a small degree of North African ancestry, we performed admixture $f_3$ tests [19, 25] to determine whether modern Canary Islanders (n = 2 in the HO dataset) can be modeled as a mix between Guanches and a Western European population from the Human Origins panel. $f_3$ was significantly negative (Z < −2) for 12 out of 26 populations tested as second source (Table S3). The lowest $f_3$ value was obtained when modeling Canary Islanders as a mix of Guanches and modern-day Scots. Next, we used $f_4$ ratios [19, 25] to estimate the contribution of Guanches to modern Canary Islanders. The estimates range from 16.7% (SE 4.96%) when using Spanish_Murcia as the European source to 31.3% (SE 3.89%) when using modern Norwegians. This difference is presumably due to modern Spaniards showing higher affinities to North Africans than Northern Europeans [34]. However, other European source populations provided intermediate values, such as a Basque population from Spain (27.2%, SE 4.92%) (Table S3). These results are smaller than the estimates of genetic contribution obtained using mitochondrial DNA (42%–73%) [6] but higher than estimates based on Y chromosome data (17%) [8], suggesting that the male Guanche contribution to modern-day Canary Islanders is lower than the female Guanche contribution. This may have been due to the violent process of colonization by the Europeans, which led to the death of a proportionally greater number of Guanche males than females [8, 35]. However, we caution that Canary Islanders in the HO dataset may not be representative of all Canary Islands, not only because of the small sample size (n = 2) but also because previous studies have shown significant differences in the relative proportion of mitochondrial ancestry between the different islands [17].

**Inferring Phenotypes**

Lastly, we obtained phenotypically informative SNPs from the five individuals with the highest genome coverage; however, only individual gun011 yielded high enough coverage to infer genotypes (Table S4). We relied on the HirisPlex and 8-plex prediction systems, which are based on 24 and 8 SNPs respectively, for skin, hair, and eye color prediction [36, 37], as well as 3 SNPs involved in lactose tolerance [38–40] (Table S4). The results reveal that this individual likely was lactose intolerant and had brown eyes, dark hair, and light or medium skin color. These results are similar for the other individuals where SNP information is available, albeit with lower coverage, suggesting that—at least for this sample of Guanches—the dominating phenotype was lactose intolerant, dark hair, light or medium skin color, and brown eyes (Table S4).

In summary, by generating the first genome-wide sequence data from several individuals of the aboriginal population of the Canary Islands, the Guanches, we confirm the long-held hypothesis that they were genetically most similar to modern Berber populations from Northwest Africa [1, 4, 9]. Importantly, our data represent a genomic time transect spanning from the early 7th century to the early 11th century CE, allowing us to explore temporal structure and to test the extent to which the European conquest of the Canary Islands (15th century CE) replaced Guanche ancestry. Although we find no clear indication that the Muslim conquest of the Maghreb (mid to late 7th century) significantly impacted the genetic variation and ancestry of the Guanches, we caution that this does not reject the possibility of limited contribution to the Guanche gene pool from either non-African or even African populations throughout the studied time period. However, overall our data suggest that the Guanches on Tenerife and Gran Canaria were genetically similar during the 7th–11th centuries CE. On the other hand, we show that the European conquest led to a decline in the overall degree

of Guanche autosomal ancestry and provide an estimate that modern Canary Islanders (as represented by two individuals on the HO panel) carry between 16% and 31% autosomal ancestry derived from the Guanches (Table S3).

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - DNA extraction
  - Library construction
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Sequence processing and alignment
  - SNP genotype data
  - Principal Component Analysis
  - ADMIXTURE analysis
  - Affinity to modern-day North African populations
  - Mitochondrial DNA analysis
  - Y chromosome analysis
  - Damage patterns
  - Mitochondrial DNA authentication
  - Outgroup $f_3$ statistics
  - $D$ statistics
- DATA AND SOFTWARE AVAILABILITY

### SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, four tables, and the extended dataset and can be found with this article online at https://doi.org/10.1016/j.cub.2017.09.059.

## REFERENCES

1. Onrubia Pintado, J. (1987). Les cultures préhistoriques des Îles Canaries, état de la question. Anthropologie 91, 653–678.

2. del Arco, M.M., Peña, P.A., Armas, E.A., del Arco Aguilar, M.d.C., and Oval, M.M. (1997). Dataciones absolutas de la prehistoria de Tenerife. In Homenaje a Celso Martín de Guzmán (1946–1994) (Universidad de Las Palmas de Gran Canaria, Servicio de Publicaciones), pp. 65–77.

3. Galván, B., Hernández, C., Velasco, J., Alberto, V., Borges, E., Barro, A., and Larraz, A. (1999). Orígenes de Buenavista del Norte: de los primeros pobladores a los inicios de la colonización europea (Ayuntamiento de Buenavista del Norte).

4. Mederos, J.F.N. (1997). Arqueología de las islas Canarias. Espac. Tiempo Forma Ser. Prehist. Arqueol. 10, 447–478.

5. Atoche, P. (2013). Consideraciones en relación con la colonización proto-histórica de las Islas Canarias. Anu. Estud. Atl. 59, 519–562.

6. Maca-Meyer, N., Arnay, M., Rando, J.C., Flores, C., González, A.M., Cabrera, V.M., and Larruga, J.M. (2004). Ancient mtDNA analysis and the origin of the Guanches. Eur. J. Hum. Genet. 12, 155–162.

7. Fregel, R., Pestano, J., Arnay, M., Cabrera, V.M., Larruga, J.M., and González, A.M. (2009). The maternal aborigine colonization of La Palma (Canary Islands). Eur. J. Hum. Genet. 17, 1314–1324.

8. Fregel, R., Gomes, V., Gusmão, L., González, A.M., Cabrera, V.M., Amorim, A., and Larruga, J.M. (2009). Demographic history of Canary Islands male gene-pool: replacement of native lineages by European. BMC Evol. Biol. 9, 181.

9. Guatelli-Steinberg, D., Irish, J.D., and Lukacs, J.R. (2001). Canary islands-north African population affinities: measures of divergence based on dental morphology. Homo 52, 173–188.

10. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., and Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS ONE 7, e34131.

11. Green, R.E., Malaspinas, A.-S., Krause, J., Briggs, A.W., Johnson, P.L.F., Uhler, C., Meyer, M., Good, J.M., Maricic, T., Stenzel, U., et al. (2008). A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134, 416–426.

12. Arauna, L.R., Mendoza-Revilla, J., Mas-Sandoval, A., Izaabel, H., Bekada, A., Benhamamouch, S., Fadhlaoui-Zid, K., Zalloua, P., Hellenthal, G., and Comas, D. (2017). Recent historical migrations have shaped the gene pool of Arabs and Berbers in North Africa. Mol. Biol. Evol. 34, 318–329.

13. Behar, D.M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N.M., Kivisild, T., Torroni, A., and Villems, R. (2012). A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Am. J. Hum. Genet. 90, 675–684.

14. Jobling, M., Hollox, E., Hurles, M., Kivisild, T., and Tyler-Smith, C. (2014). Human Evolutionary Genetics, Second Edition (Garland Science).

15. Rando, J.C., Cabrera, V.M., Larruga, J.M., Hernández, M., González, A.M., Pinto, F., and Bandelt, H.-J. (1999). Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. Ann. Hum. Genet. 63, 413–428.

16. Ordóñez, A.C., Fregel, R., Trujillo-Mederos, A., Hervella, M., de-la-Rúa, C., and Arnay-de-la-Rosa, M. (2017). Genetic studies on the prehispanic population buried in Punta Azul cave (El Hierro, Canary Islands). J. Archaeol. Sci. 78, 20–28.

17. Fregel, R., Cabrera, V.M., Larruga, J.M., Hernández, J.C., Gámez, A., Pestano, J.J., Arnay, M., and González, A.M. (2015). Isolation and prominent aboriginal maternal legacy in the present-day population of La Gomera (Canary Islands). Eur. J. Hum. Genet. 23, 1236–1243.

18. Cruciani, F., La Fratta, R., Santolamazza, P., Sellitto, D., Pascone, R., Moral, P., Watson, E., Guida, V., Colomb, E.B., Zaharova, B., et al. (2004). Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. Am. J. Hum. Genet. 74, 1014–1022.

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

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

21. Schuenemann, V.J., Peltzer, A., Welte, B., van Pelt, W.P., Molak, M., Wang, C.-C., Furtwängler, A., Urban, C., Reiter, E., Nieselt, K., et al. (2017). Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat. Commun. *8*, 15694.

22. Rogers, A.R., and Bohlender, R.J. (2015). Bias in estimators of archaic admixture. Theor. Popul. Biol. *100C*, 63–78.

23. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution *38*, 1358–1370.

24. Hudson, R.R., Slatkin, M., and Maddison, W.P. (1992). Estimation of levels of gene flow from DNA sequence data. Genetics *132*, 583–589.

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

26. De Guzmán, C.M. (1986). La arqueología canaria: una propuesta metodológica. Anu. Estud. Atl. *1*, 575–682.

27. Hagenblad, J., Morales, J., Leino, M.W., and Rodríguez-Rodríguez, A.C. (2017). Farmer fidelity in the Canary Islands revealed by ancient DNA from prehistoric seeds. J. Archaeol. Sci. *78*, 78–87.

28. Olalde, I., Capote, J., Del-Arco, M.C., Atoche, P., Delgado, T., González-Anton, R., Pais, J., Amills, M., Lalueza-Fox, C., and Ramírez, O. (2015). Ancient DNA sheds light on the ancestry of pre-hispanic Canarian pigs. Genet. Sel. Evol. *47*, 40.

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

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

31. Omrak, A., Günther, T., Valdiosera, C., Svensson, E.M., Malmström, H., Kiesewetter, H., Aylward, W., Storå, J., Jakobsson, M., and Götherström, A. (2016). Genomic Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool. Curr. Biol. *26*, 270–275.

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

33. Olalde, I., Schroeder, H., Sandoval-Velasco, M., Vinner, L., Lobón, I., Ramirez, O., Civit, S., García Borja, P., Salazar-García, D.C., Talamo, S., et al. (2015). A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. Mol. Biol. Evol. *32*, 3132–3142.

34. Botigué, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., et al. (2013). Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. Proc. Natl. Acad. Sci. USA *110*, 11791–11796.

35. De Viera Clavijo, J. (1773). Noticias de la Historia General de las Islas Canarias (B. Román).

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

37. Hart, K.L., Kimura, S.L., Mushailov, V., Budimlija, Z.M., Prinz, M., and Wurmbach, E. (2013). Improved eye- and skin-color prediction based on 8 SNPs. Croat. Med. J. *54*, 248–256.

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

39. Ingram, C.J.E., Elamin, M.F., Mulcare, C.A., Weale, M.E., Tarekegn, A., Raga, T.O., Bekele, E., Elamin, F.M., Thomas, M.G., Bradman, N., and Swallow, D.M. (2007). A novel polymorphism associated with lactose tolerance in Africa: multiple causes for lactase persistence? Hum. Genet. *120*, 779–788.

40. Ye, K., and Gu, Z. (2011). Recent advances in understanding the role of nutrition in human genome evolution. Adv. Nutr. *2*, 486–496.

41. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, t5448.

42. Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., and Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques *59*, 87–93.

43. Kircher, M. (2012). Analysis of high-throughput ancient DNA sequencing data. Methods Mol. Biol. *840*, 197–228.

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

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

46. Skoglund, P., Mallick, S., Bortolini, M.C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M.L., Salzano, F.M., Patterson, N., and Reich, D. (2015). Genetic evidence for two founding populations of the Americas. Nature *525*, 104–108.

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

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

49. van Oven, M., Van Geystelen, A., Kayser, M., Decorte, R., and Larmuseau, M.H. (2014). Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome. Hum. Mutat. *35*, 187–191.

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

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

52. Behr, A.A., Liu, K.Z., Liu-Fang, G., Nakka, P., and Ramachandran, S. (2016). pong: fast analysis and visualization of latent clusters in population genetic data. Bioinformatics *32*, 2817–2823.

53. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

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

55. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.

56. Scottish Parliament (2006). Human Tissue (Scotland) Act 2006. https://www.legislation.gov.uk/asp/2006/4/contents.

57. Museums Galleries Scotland (2011). Guidelines for the Care of Human Remains in Scottish Museum Collections (Museums Galleries Scotland). https://www.museumsgalleriesscotland.org.uk/media/1089/guidelines-for-the-care-of-human-remains-in-scottish-museum-collections.pdf.

58. Historic Scotland (2006). The Treatment of Human Remains in Archaeology (Historic Scotland). https://www.historicenvironment.scot/archives-and-research/publications/publication/?publicationId=02e7320f-4fb2-4c4a-8aba-a58e00e3f22c.

59. Giesen, M., and White, L. (2013). International perspectives towards human remains curation. In Curating Human Remains: Caring for the Dead in the United Kingdom, M. Glesen, ed. (Boydell & Brewer), pp. 13–24.

60. The University of Edinburgh (2017). The Collection. http://www.ed.ac.uk/biomedical-sciences/anatomy/anatomical-museum/collection.

61. Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., and Allentoft, M.E. (2017). Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. PLoS ONE 12, e0170940.

62. Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., and Allentoft, M.E. (2015). Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. 5, 11184.

63. Adler, C.J., Haak, W., Donlon, D., and Cooper, A. (2011). Survival and recovery of DNA from ancient teeth and bones. J. Archaeol. Sci. 38, 956–964.

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

65. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA 110, 15758–15763.

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

67. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. 40, 4477–4482.

68. Wang, C., Szpiech, Z.A., Degnan, J.H., Jakobsson, M., Pemberton, T.J., Hardy, J.A., Singleton, A.B., and Rosenberg, N.A. (2010). Comparing spatial maps of human population-genetic variation using Procrustes analysis. Stat. Appl. Genet. Mol. Biol. 9, 13.

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

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

71. Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford, T.W., Jr., Orlando, L., Metspalu, E.,

et al. (2014). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature 505, 87–91.

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

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

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

75. Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez-Del-Molino, D., van Dorp, L., López, S., Kousathanas, A., Link, V., et al. (2016). Early farmers from across Europe directly descended from Neolithic Aegeans. Proc. Natl. Acad. Sci. USA 113, 6886–6891.

76. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65.

77. Peter, B.M. (2016). Admixture, Population Structure, and F-Statistics. Genetics 202, 1485–1501.

78. Henn, B.M., Botigué, L.R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J.K., Fadhlaoui-Zid, K., Zalloua, P.A., Moreno-Estrada, A., Bertranpetit, J., et al. (2012). Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS Genet. 8, e1002397.

79. Hansen, A., Willerslev, E., Wiuf, C., Mourier, T., and Arctander, P. (2001). Statistical evidence for miscoding lesions in ancient DNA templates. Mol. Biol. Evol. 18, 262–265.

80. Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., and Pääbo, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 29, 4793–4799.

81. Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA 104, 14616–14621.

82. Orlando, L., Ginolhac, A., Raghavan, M., Vilstrup, J., Rasmussen, M., Magnussen, K., Steinmann, K.E., Kapranov, P., Thompson, J.F., Zazula, G., et al. (2011). True single-molecule DNA sequencing of a pleistocene horse bone. Genome Res. 21, 1705–1719.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Biological Samples | | |
| gun001 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 52 |
| gun011 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 57 |
| gun005 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 47 |
| gun012 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 62 |
| gun008 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 46 |
| gun014 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 61 |
| gun006 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 48 |
| gun002 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 49 |
| gun007 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 45 |
| gun013 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 60 |
| gun004 | Anatomical Museum Edinburgh University (AMEU) | Cat#XXVI A 53 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| Proteinase K | VWR Sweden | Cat#1.24568.0100 |
| EDTA buffer solution pH 8.0 (0.5 mol/l) for biotechnology, sterile | VWR Sweden | Car#E522-100ML |
| Sodium acetate | Sigma Aldrich Sweden | Cat#S2889 |
| Guanidine hydrochloride | Sigma Aldrich Sweden | Cat#50933 |
| Isopropanol | Sigma-Aldrich Sweden | Cat#67-63-0 |
| Tween-20 | Sigma-Aldrich Sweden | Cat#9005-64-5 |
| Sodium phosphate PH 6.0 | VWR Sweden | Cat#101447-426 |
| ATP | Fermentas/Thermo Scientific | Cat#R0441 |
| T4 Polynucleotide Kinase | Thermo Scientific | Cat#EK0032 |
| T4 DNA Polymerase | Fermentas/Thermo Scientific | Cat#EP0062 |
| Bst polymerase (supplied with 10X ThermoPol reaction buffer) | NEB/BioNordika | Cat#M0275S |
| AmplitTaq Gold | Invitrogen/life technologies | Cat#4311816 |
| ATP | Fermentas/Thermo Scientific | Cat#R0441 |
| 10X Tango Buffer | Fermentas/Thermo Scientific | Cat#BY5 |
| High Pure Viral Nucleic Acid Large Volume Kit | Roche | Cat#5114403001 |
| T4 DNA Ligase | Fermentas/Thermo Scientific | Cat#EL0011 |
| Min Elute PCR Purification Kit | QIAGEN | Cat#28006 |
| PEG-4000 | Sigma | Cat#1546569 |
| Agencourt AMPure XP beads (60ml) | Beckman Coulter | Cat#A63881 |

*(Continued on next page)*

**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Critical Commercial Assays | | |
| High Sensitivity DNA (chips + reagents) (Bioanalyzer 2100) | Agilent Technologies | Cat#5067-4626 |
| Deposited Data | | |
| European Nucleotide Archive | https://www.ebi.ac.uk/ena/data/view/PRJEB86458 | ENA: PRJEB86458 |
| Oligonucleotides | | |
| IS1_adapter.P5: 5′-A*C*A*C*TCTTTCC CTACACGACGCTCTTCCG*A*T*C*T-3′ (* indicates a PTO bond) | [41] | Biomers |
| IS2_adapter.P7: 5′-G*T*G*A*CTGGAGTT CAGACGTGTGCTCTTCCG*A*T*C*T-3′ (* indicates a PTO bond) | [41] | Biomers |
| IS3_adapter.P5+P7: 50-A*G*A*T*CG GAA*G*A*G*C-30(* indicates a PTO bond) | [41] | Biomers |
| IS4: (5′-AATGATACGGCGACCAC CGAGATCTACACTCTTT CCCTACACGACGCTCTT 3′) | [41] | Biomers |
| P7 indexing: (5′-CAAGCAGAAGACGGC ATACGAGATxxxxxxxGTGACTGGAGTT CAGACGTGT 3′) where x is one of 228 different 7 bp indexes provided in [42] | [41] | Biomers |
| Software and Algorithms | | |
| MergeReadsFastq_cc.py | [43] | https://bioinf.eva.mpg.de/fastqProcessing/ |
| FilterUniqSAMCons_cc.py | [43] | https://bioinf.eva.mpg.de/fastqProcessing/ |
| Burrows-Wheeler Aligner BWA aln version 0.7.8 | [44] | http://bio-bwa.sourceforge.net/ |
| Samtools-0.1.19 | [45] | https://sourceforge.net/projects/samtools/files/samtools/0.1.19/ |
| mpileup from the Samtools package | [45] | https://sourceforge.net/projects/samtools/files/samtools/0.1.19/ |
| vcfutils.pl from the Samtools package | [45] | https://sourceforge.net/projects/samtools/files/samtools/0.1.19/ |
| POPSTATS | [46] | https://github.com/pontussk/popstats |
| Haplofind | [47] | https://haplofind.unibo.it/new/ |
| PhyloTree Build 17 (18th February 2016) | [48, 49] | http://www.phylotree.org/ |
| ISOGG database v. 11.349 (accessed 04, 2016). | International Society of Genetic Genealogy | https://www.isogg.org |
| PMDtools | [50] | https://code.google.com/archive/p/pmdtools/ |
| EIGENSOFT v.6.0.1 | [51] | https://github.com/DReichLab/EIG |
| qp3Pop v. 204. ADMIXTOOLS package | [19] | https://github.com/DReichLab/AdmixTools |
| qpDstat v. 450. ADMIXTOOLS package | [19] | https://github.com/DReichLab/AdmixTools |
| ADMIXTURE | [29] | https://www.genetics.ucla.edu/software/admixture/ |
| Pong | [52] | https://github.com/ramachandran-lab/pong |
| Integrative Genomics Viewer (IGV) | [53] | http://software.broadinstitute.org/software/igv/download |
| Plink v1.90 | [54, 55] | https://www.cog-genomics.org/plink/2.0/ |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ricardo Rodríguez Varela (ricardo.rodriguez.varela@arklab.su.se).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

All activities relating to this project complied with the Human Tissue (Scotland) Act 2006 [56], the University of Edinburgh's Research Ethics Framework, the Guidelines for the Care of Human Remains in Scottish Museum Collections [57], and Historic Scotland's guide for The Treatment of Human Remains in Archaeology [58]. All sampling of teeth was carried out by L.G.-F. at the Anatomical Museum in Edinburgh under the supervision of T.H.C. and M.M. All the sampled individuals were adults; the gender is provided in Table 1.

The Guanches who survived the Spanish conquest in the 15th century assimilated into Spanish culture and there is currently no ethnic minority claiming aboriginal status or groups of living descendants claiming ancestral ownership of the Guanche human remains held at the Anatomical Museum of Edinburgh. However, the history of the Guanche human remains analyzed for this study is in itself interesting and important from an ethical perspective: the materials were donated to the Anatomical Museum of Edinburgh in the late 19th century by various collectors and anthropologists. Western physicians, medical doctors, and antiquaries (as well as private collectors) amassed numerous collections of human remains during the Victorian or Edwardian eras (1837-1914). Extensive trading and poor, partial or often missing, documentation, in addition to growing ethical concerns regarding both their acquisition, ownership, and public display, has rendered these, often disparate, collections increasingly marginalized in terms of new research. Thus, potentially unique human remains may be largely forgotten by modern scientific as well as culture-historical enquiry [59]. As such, the current project forms part of the Anatomical Museum of Edinburgh's ongoing detailed inventory of the collections that comprise some 12,000 objects reflecting 300 years of teaching Anatomy at the University of Edinburgh [60].

## METHOD DETAILS

### DNA extraction

Samples were prepared in dedicated ancient DNA (aDNA) facilities at the Archaeological Research Laboratory, University of Stockholm, Sweden. We targeted the cementum-rich root tip of the teeth as it has been shown to preserve DNA better than most other types of bone [61–63]. We first wiped the teeth with 1% sodium hypochlorite and ddH$_2$O, and obtained ca. 75 mg tooth powder/sample using a multitool drill at the lowest possible rpm (Dremel). DNA extractions were carried out in batches of six plus one extraction blank. The tooth powder was incubated (washed) three times for 15 min under constant rotation in 500 μL 1M sodium phosphate buffer, pH 6, in order to remove microbial contaminants and enrich for endogenous DNA [42]. Buffers were replaced after each incubation step and pellets were washed with 1 mL water after the incubation. DNA was extracted by incubating the bone powder for 24 to 48 hr at 37°C in 1 mL of digestion buffer (0.45 M EDTA pH 8.0 and 0.25 mg/ml of proteinase K). After incubation, samples were centrifuged and 1 mL of supernatant was transferred to an Eppendorf Lo-Bind 1.5 mL tube. DNA was extracted and purified using a silica-binding method optimized for short DNA fragments [64, 65]. The 1 mL of supernatant extract was added to a 13 mL of a binding buffer containing 5 M guanidine hydrochloride, 40% (vol/vol) isopropanol, 0.05% Tween-20 and 90 mM sodium acetate (pH 5.2) in a 50 mL Falcon tube. The 14 mL solution containing the binding buffer and the extraction supernatant was transferred into a 50 mL silica column (Roche, High Pure Viral Nucleic Acid Large Volume Kit). We centrifuged the 50 mL Roche tube for 6 min at 4,000 rpm and remove the Roche falcon tube placing the inside silica column into a new 2 mL collection tube. After centrifuging the silica column tube for 1 min at 6,000 rpm we purified the DNA by adding 750 μL PE buffer (QIAGEN) to the silica column followed by centrifugation at 6,000 rpm. The flow-through was discarded and we repeated this step one time. After these two wash steps, we centrifuged the dry column for 1 min at max speed (16,000 rpm) and placed the silica column in a new 1.5 mL collection tube. We pipetted 22 μL of TE buffer (EB QIAGEN buffer plus 0.05% Tween) onto the silica membrane for elution, and after 5-min incubation at 37°C, the DNA eluate was collected by centrifugation for 30 s at maximum rpm. This step was repeated once for a total of 44 μL of DNA extract.

### Library construction

One Illumina double stranded library was built from 20 μl of extract using the blunt end ligation protocol described in [41] with modifications as in [66]. Due to the fragmented state of ancient DNA the initial nebulization step was omitted. Indexing PCRs were set up in a total volume of 25 μl with a final concentration of 1X Gold Buffer (Invitrogen/life technologies), 2.5 mM Magnesium Chloride (Invitrogen/life technologies), 250 μM dNTP (each), 3 μL of DNA library, 0.2 μM IS4 PCR primer (5′- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTT 3′) and 0.2 μM indexing primer (5′-CAAGCAGAAGACGGCATACGAGATxxxxxxxGTGACTGGAGTTCAGACGTGT 3′) (x is one of 228 different 7 bp indexes provided in [41]) and 0.1 U/μl of AmpliTaq Gold (Invitrogen/life technologies). Cycling conditions were performed as follows: a 2 min activation step at 94°C, followed by 8-15 cycles of 30 s at 94°C, 30 s at 60°C, 45 s at 72°C, with a final extension of 10 min at 72°C. qPCR was used to assess the optimal number of PCR cycles for amplification that varied between 14 and 21 [41]. Each library was amplified in qintaduplicates together with two PCR blanks with an index in the P7 primer. PCR products from the same library were pooled and purified with Agencourt AMPure XP beads (Beckman Coulter). The concentration and size profiles of the purified libraries were identified on a Bioanalyzer 2100 using

the High Sensitivity DNA chip (Agilent) for DNA. Purified libraries were pooled in equimolar concentration and sequenced on an Illumina HiSeq X at the SNP and SEQ technology platform at the SciLife Sequencing Centre in Stockholm.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Sequence processing and alignment
Paired-end reads were merged requiring an overlap of at least 11 bp and remaining adaptor sequences were trimmed using MergeReadsFastq_cc.py [43]. The fragments were then mapped to the human reference genome as single-end reads using BWA aln version 0.7.8 [44] with the non-default parameters -l 16500 -n 0.01 -o 2. BAM files of different sequencing runs were merged per library using Samtools [45]. PCR duplicates were then removed using a modified version of FilterUniqSAMCons_cc.py [43]. Fragments of at least 35 bp length, with less than 10% mismatches to the reference genome and minimum mapping quality of 30 were kept for downstream analysis. Biological sex was estimated based on the number of reads mapping to the sex chromosomes and compared to a reference panel [67]. We restricted the analysis to sequence alignments with mapping quality of at least 30.

### SNP genotype data
We built pseudo-haploid genomes by randomly choosing one read with minimum mapping quality 30 and a base quality of 30 or higher covering the site of interest [66]. Sites showing indels and transition were excluded to avoid potential post-mortem damage and only transversion sites were used in the analysis.

### Principal Component Analysis
The principal component analysis of the Guanches was performed using EIGENSOFT v.6.0.1 [51] with a selection of European, Middle East and North African populations from the Human Origins dataset [19, 20]. The PCAs for the modern populations were made using all SNPs and each of the ancient individuals using the merged SNP data. The number of overlapping SNPs between ancient individuals and the Human Origins dataset ranged from 74,642 (gun002) to 370,595 (gun011). We use Procrustes analysis [68, 69] to transform coordinates of each individual and to plot them together with each individual. The individual PCAs for each ancient individual are shown in the Figure S2.

### ADMIXTURE analysis
ADMIXTURE [29] was run on all African and western Eurasian populations of the Human Origins dataset [19, 20] together with the Guanches and 239 relevant published ancient genomes [20, 30–33, 66, 70–75]. To avoid an excess of drift in ancient samples due to post-mortem damages, they were coded missing at all transition sites. The data were pruned for linkage disequilibrium between markers using Plink v1.90 [54, 55] and the parameters–indep-pairwise 200 25 0.4. ADMIXTURE was then run for 2 to 15 clusters (K) with 20 independent runs (different random seeds) per K. Pong [52] was then used in greedy mode to identify common modes among the different ADMIXTURE runs per K and to align clusters between different numbers of clusters. We display K = 10 in the main text as it represents the lowest K with a separate North African cluster and the mode shown is common among 15 out of the 20 different runs. All Ks are shown in Figure S3.

Admixture $f_3$ statistics and $f_4$ ratios [19] were calculated using popstats (https://github.com/pontussk/popstats [46]) and the Human Origins dataset. Transition polymorphisms were excluded and standard errors were calculated using a weighted block-jackknife with block sizes of 5Mbp. The proportion of Guanche ancestry in modern Canary Islanders $\alpha$ was calculated using the following $f_4$ ratio

$$\alpha = 1 - \frac{f_4(FIN, Chimp; CanaryIslanders, Guanches)}{f_4(FIN, Chimp; X, Guanches)},$$

where X is a modern western European population and FIN the Finish population from the 1000 genome project [76].

### Affinity to modern-day North African populations
We used different statistics to measure genetic similarity between Guanches and modern populations. We calculate $f_2$ statistics [25], Hudson's FST [24], Weir and Cockerham's FST [23] as well as the commonly used outgroup $f_3$ statistics [19, 71] and the related measurement of average pairwise differences (as proposed by [77]) (Table S2). These statistics were calculated with custom scripts or popstats [46]; https://github.com/pontussk/popstats). The estimators differ in their sensibility to drift, rare alleles, and recent admixture – factors important to take into account for the analysis of North African populations and ancient DNA due to the complex history of admixture events [12, 78] and data quality of low coverage ancient DNA. As a sanity check, we calculated these statistics for modern North Africans as a 'good' estimator for our purposes should identify another North African population as closest relative.

### Mitochondrial DNA analysis
Consensus sequences for the mitochondrial genomes were called using mpileup and vcfutils.pl (vcf2fq) from the Samtools package [45]. Only reads with a minimum mapping score of 30 and a minimum base quality of 30 were used to call confident bases for the final consensus sequences. However, haplogroup diagnostic positions were checked manually with the program (IGV) [53]. Mitochondrial haplogroups were identified using HAPLOFIND [47] and PhyloTree Build 17 (18th February 2016) [48] (Table S1).

### Y chromosome analysis

The Y chromosome sequences were filtered out using mpileup from the Samtools package [45]. The pileup file was then merged with the PyloTree Y haplogroup definitions [49]. Y chromosome haplotypes were called using the nomenclature of the International Society of Genetic Genealogy ISOGG database (https://www.isogg.org) v. 11.349 (accessed 04, 2016) (Data S1, sheet 1). We excluded all non-SNP sites, transition sites (to avoid deamination damage), and A/T and G/C SNPs (to avoid strand misidentification).

### Damage patterns

The presence of 3′ and 5′ cytosine deamination patterns characteristic of ancient DNA [10, 79–82] was estimated using PMDtools [50]. All individuals presented deamination patterns consistent with presence of ancient DNA templates (Figure S1A).

### Mitochondrial DNA authentication

Contamination levels were estimated based on the analyses of polymorphic site distribution in mitochondrial sequences [11]. The contamination point estimates ranged between 0.53% up to 13.79% (0.53% to 5.97% for the five individuals included in further autosomal genomic analysis) rendering the obtained genomic data reliable (Table 1 and S1).

### Outgroup $f_3$ statistics

In order to check the genetic affinities between the Guanches and the modern populations we performed outgroup $f_3$ statistics using qp3Pop v. 204. of the ADMIXTOOLS package [19]. The outgroup $f_3$ statistics using of the form (O; A, B) [19, 71] represent the amount of genetic affinity shared between test populations, A and B. The analyses were performed with the Human Origins dataset (using between 21,864 and 108,190 SNPs). The outgroup (O) was Ju/'hoansi population, the tested individual was A, and any of the populations from the reference panel was the test population B. The analyses were performed with up to 705 jackknife blocks. The results for each of the individuals are listed in Data S1, sheet 2.

### *D* statistics

Deviations from tree-like population history between the analyzed individuals were tested using *D* statistics included in ADMIXTOOLS package qpDstat v. 450 [19]. We exclude transition sites in all analysis and used pseudo-haploid genomes [66].

The *D* statistics of the form D(O, X; A, B) [19] represent the amount of genetic affinity shared between test individual X, A and B, where O is the outgroup the Ju/'hoansi population; X Guanche individual; A North African population; B Sardinians or AnatoliaEN (Early Neolithics)

The analyses were performed with the Human Origins dataset (using between 20,227 and 108,190 SNPs). The standard errors were estimated by performing block jackknife over blocks of 705 SNPs. The results are summarized in Data S1, sheet 3.

## DATA AND SOFTWARE AVAILABILITY

The newly generated genome data have been deposited at the European Nucleotide Archive with the accession number ENA: PRJEB86458 (https://www.ebi.ac.uk/ena/data/view/PRJEB86458).