

This is a repository copy of *Biomolecular insights into North African-related ancestry, mobility and diet in eleventh-century Al-Andalus*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/178343/>

Version: Published Version

---

**Article:**

Silva, Marina, Oteo-García, Gonzalo, Martiniano, Rui et al. (30 more authors) (2021) Biomolecular insights into North African-related ancestry, mobility and diet in eleventh-century Al-Andalus. Scientific Reports. 18121. ISSN: 2045-2322

<https://doi.org/10.1038/s41598-021-95996-3>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



OPEN

# Biomolecular insights into North African-related ancestry, mobility and diet in eleventh-century Al-Andalus

Marina Silva<sup>1,22,23</sup>✉, Gonzalo Oteo-García<sup>1,23</sup>✉, Rui Martiniano<sup>2,3</sup>, João Guimarães<sup>4</sup>, Matthew von Tersch<sup>5</sup>, Ali Madour<sup>1</sup>, Tarek Shoeib<sup>1,6</sup>, Alessandro Fichera<sup>1</sup>, Pierre Justeau<sup>1</sup>, M. George B. Foody<sup>1</sup>, Krista McGrath<sup>5,7</sup>, Amparo Barrachina<sup>8</sup>, Vicente Palomar<sup>9</sup>, Katharina Dulias<sup>1,5,10</sup>, Bobby Yau<sup>1</sup>, Francesca Gandini<sup>1</sup>, Douglas J. Clarke<sup>1</sup>, Alexandra Rosa<sup>11,12</sup>, António Brehm<sup>12</sup>, Antònia Flaquer<sup>13</sup>, Teresa Rito<sup>4,14,15</sup>, Anna Olivieri<sup>16</sup>, Alessandro Achilli<sup>16</sup>, Antonio Torroni<sup>16</sup>, Alberto Gómez-Carballa<sup>17,18</sup>, Antonio Salas<sup>17,18,19</sup>, Jaroslaw Bryk<sup>1</sup>, Peter W. Ditchfield<sup>20</sup>, Michelle Alexander<sup>5</sup>, Maria Pala<sup>1</sup>, Pedro A. Soares<sup>4,21</sup>, Ceiridwen J. Edwards<sup>1</sup> & Martin B. Richards<sup>1</sup>✉

Historical records document medieval immigration from North Africa to Iberia to create Islamic al-Andalus. Here, we present a low-coverage genome of an eleventh century CE man buried in an Islamic necropolis in Segorbe, near Valencia, Spain. Uniparental lineages indicate North African ancestry, but at the autosomal level he displays a mosaic of North African and European-like ancestries, distinct from any present-day population. Altogether, the genome-wide evidence, stable isotope results and the age of the burial indicate that his ancestry was ultimately a result of admixture between recently arrived Amazigh people (Berbers) and the population inhabiting the Peninsula prior to the Islamic conquest. We detect differences between our sample and a previously published group of contemporary individuals from Valencia, exemplifying how detailed, small-scale aDNA

<sup>1</sup>Department of Biological and Geographical Sciences, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK. <sup>2</sup>Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK. <sup>3</sup>School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool L3 3AF, UK. <sup>4</sup>Department of Biology, CBMA (Centre of Molecular and Environmental Biology), University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. <sup>5</sup>BioArCh, Department of Archaeology, University of York, York, UK. <sup>6</sup>Department of Forensic Science, Faculty of Biomedical Science, University of Benghazi, P.O. Box: 1308, Benghazi, Libya. <sup>7</sup>Department of Prehistory and Institute of Environmental Science and Technology (ICTA), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. <sup>8</sup>Servei d'Investigacions Arqueològiques i Prehistòriques - Museu Belles Arts de Castelló, Av. Germans Bou, 28, 12003 Castellón, Spain. <sup>9</sup>Museo Municipal de Arqueología y Etnología de Segorbe, Calle Colón, 98, 12400 Segorbe, Castellón, Spain. <sup>10</sup>Institut für Geosysteme und Bioindikation, Technische Universität Braunschweig, Langer Kamp 19c, 38106 Braunschweig, Germany. <sup>11</sup>Faculty of Life Sciences, University of Madeira, Campus of Penteada, 9000-390 Funchal, Portugal. <sup>12</sup>Human Genetics Laboratory, University of Madeira, Campus of Penteada, 9000-390 Funchal, Portugal. <sup>13</sup>Institute for Medical Information Processing, Biometry and Epidemiology - IBE, LMU University, Munich, Germany. <sup>14</sup>Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, 4710-057 Braga, Portugal. <sup>15</sup>ICVS/3B's, PT Government Associate Laboratory, 4710-057 Braga, Portugal. <sup>16</sup>Dipartimento di Biologia e Biotechnologie "L. Spallanzani, Università di Pavia, 27100 Pavia, Italy. <sup>17</sup>Grupo de Investigación en Genética, Vacunas, Infecciones y Pediatría (GENVIP), Hospital Clínico Universitario and Universidade de Santiago de Compostela, Galicia, Spain. <sup>18</sup>GenPoB Research Group, Instituto de Investigación Sanitaria (IDIS), Hospital Clínico Universitario de Santiago (SERGAS), 15706 Galicia, Spain. <sup>19</sup>Unidade de Xenética, Instituto de Ciencias Forenses (INCIFOR), Facultade de Medicina, Universidade de Santiago de Compostela, Galicia, Spain. <sup>20</sup>School of Archaeology, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK. <sup>21</sup>Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. <sup>22</sup>Present address: Ancient Genomics Laboratory, The Francis Crick Institute, London, UK. <sup>23</sup>These authors contributed equally to this work: Marina Silva and Gonzalo Oteo-García. ✉email: marina.silva@crick.ac.uk; gonzalo.oteo-garcia@hud.ac.uk; m.b.richards@hud.ac.uk

studies can illuminate fine-grained regional and temporal differences. His genome demonstrates how ancient DNA studies can capture portraits of past genetic variation that have been erased by later demographic shifts—in this case, most likely the seventeenth century CE expulsion of formerly Islamic communities as tolerance dissipated following the *Reconquista* by the Catholic kingdoms of the north.

The location of Iberia, bridging the Mediterranean and the Atlantic, and its proximity to Africa, has allowed contacts with populations of distinct ancestries over time, making the Peninsula a genetic and cultural crossroads. There is both archaeological and direct genetic evidence of contacts between Iberia and North African populations since at least the Late Neolithic<sup>1–6</sup>, and possibly as early as the postglacial period<sup>7,8</sup>. Prehistoric populations have been the focus of most of the ancient DNA (aDNA) work published on Iberia so far, including the study of Mesolithic individuals<sup>9,10</sup>, the impact of Neolithic dispersals<sup>11,12</sup>, and the incursions of individuals with Steppe-related ancestry at the time of the transition from the Chalcolithic to the Bronze Age<sup>5,6,13–15</sup>.

aDNA researchers have recently begun to explore in detail historical intervals of known population movements<sup>6</sup>. Although Iberia intensified contacts with North Africa through Phoenician traders, Carthaginians and Roman conquerors<sup>16</sup>, North-African genetic contribution seems to have been restricted to southern populations until the eighth century CE<sup>6</sup>. It is only with the Islamic conquest of Iberia in 711 CE that records start pointing towards a substantial influx of people from North Africa, involving the culturally and genetically differentiated Arab and Amazigh (Berber) peoples<sup>17,18</sup>. Attempts have been made to estimate their contribution to the genetic landscape of medieval Iberia using modern genomes, revealing a faint southwest-northeast pattern of decreasing North African-related ancestry<sup>19,20</sup>, which have recently been confirmed by means of aDNA analysis<sup>6</sup>.

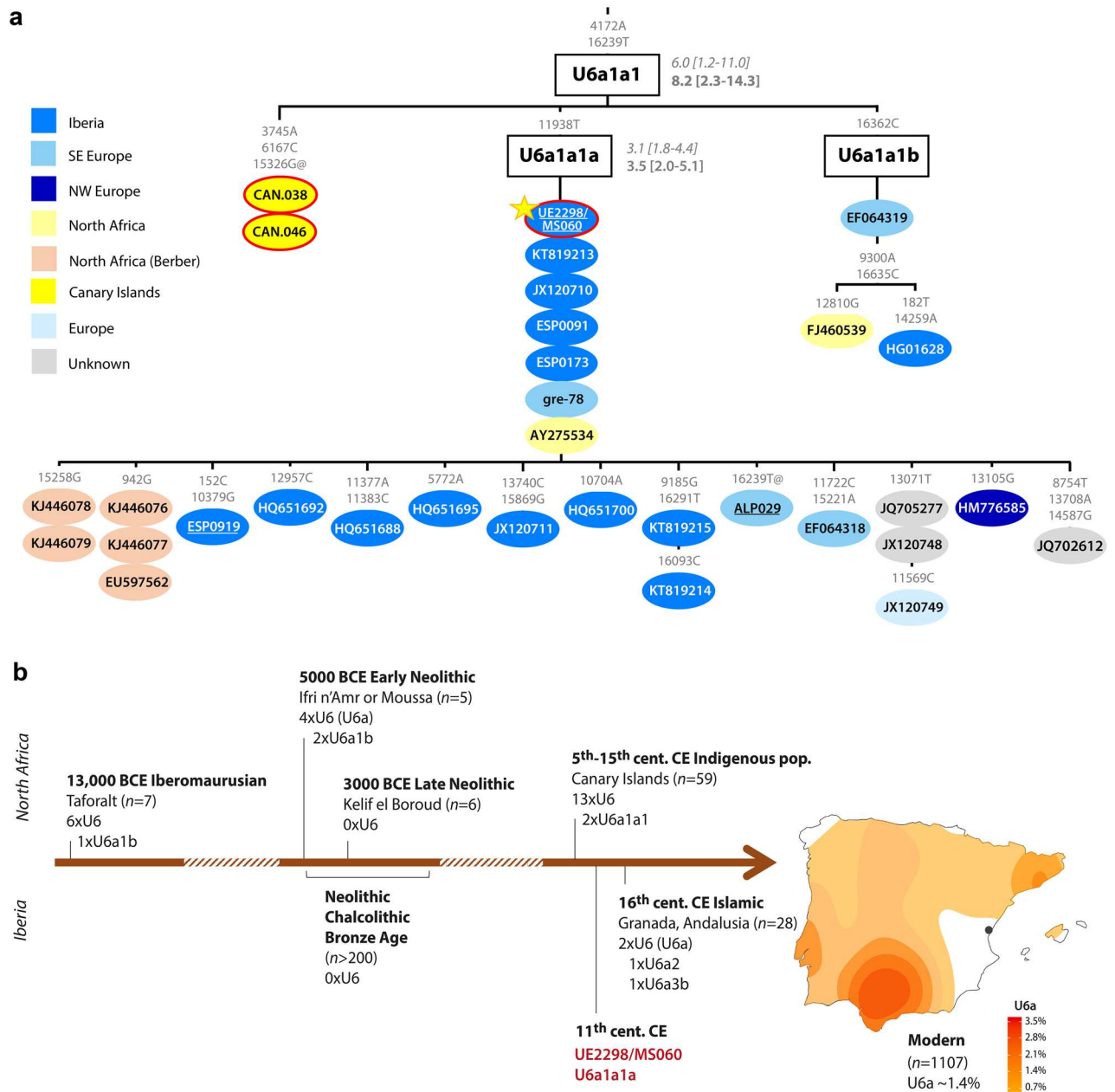
Although Arabs were the urban and political elite during the Umayyad Caliphate, ruling from 711 CE until the end of the Caliphate of Cordoba in 1031 CE, they are thought to have been a minority amongst the new settlers. Berbers formed the bulk of the army who first seized Visigothic Spain in the eighth century CE<sup>21</sup>. Berbers had converted to Islam as a result of the Arab conquest of North Africa in the preceding century and embarked in a slow and complex process of Arabisation that lasted centuries. However, they were far from culturally homogeneous; a deep division existed between nomadic and sedentary Berber groups, and it was the latter who first settled in the rural areas of Spain<sup>18</sup>. Although Berber numbers in Iberia were likely larger than those of the Arabs, they initially wielded no significant political power, but this changed during the eleventh–thirteenth centuries CE with the establishment of the Almoravid and Almohad Berber empires<sup>18</sup>.

After the southwards military expansion of the Catholic kingdoms ended in 1492, a large population of *Moriscos* (Muslims forcibly converted to Christianity) persisted in East Iberia (previously *Sharq al-Andalus*) until 1609 CE, when at least one third of the populace was forcibly expelled by the Spanish Crown and relocated to North Africa<sup>22</sup>. Historical documentation suggests that the population of the eastern Mediterranean provinces of Castellón, Valencia, Alicante, and—to a lesser degree—Murcia and parts of Andalusia (Almería and Granada), was greatly reduced, with subsequent resettlement from Aragon, Catalonia and Navarre to avoid economic and demographic collapse<sup>23</sup>. Many surnames currently widespread in the Valencian region are geographically structured and reflect their provenance from the colonizing regions (data from Spanish National Institute of Statistics, 2017; Supplementary Fig. S1). The hinterland was mostly repopulated by non-Catalan-speaking Aragonese, whereas the main coastal cities concentrated more Catalan-speakers from Catalonia<sup>23</sup>. This divide is thought to be still reflected in the genomic data today<sup>20</sup>. Thus, most of the existing genetic variation from the preceding eastern Iberian populations and the North African genetic variation potentially brought during Islamic rule had most likely disappeared by the late seventeenth century CE<sup>20</sup>, especially in the Valencian region<sup>24</sup>. Therefore, DNA from archaeological remains can provide an important tool to understand the demographic dynamics of the Islamic period in East Iberia<sup>25</sup>.

Here, we sequenced the genome of an individual (UE2298/MS060) who was buried in the Islamic *maqbara* (necropolis) of Plaza del Almudín in the city of Segorbe (province of Castellón, Comunidad Valenciana, Spain) (Supplementary Fig. S2). He was dubbed as “the Giant” by the archaeologists responsible for the excavation (here referred to as the “Segorbe Giant”), due to his unusual height (184–190 cm) compared with the other individuals found in the site (Barrachina 2004) (Supplementary Methods). Osteological assessment suggests that he had African ancestry, and he was postulated to be of possible Berber origin<sup>26,27</sup>. Although his uniparental lineages point to North African ancestry, at the autosomal level he displays both North African and European-related ancestries. The genetic analyses show differences in relation to contemporary individuals from Valencia<sup>6</sup> and highlight the contribution of admixture between people of North African origin and the populations inhabiting East Iberia prior to the Islamic period. We conducted a complementary analysis of stable isotopes on a total of thirteen individuals from the necropolis (Supplementary Table S1) to investigate mobility and diet patterns. We also generated more than 1000 new modern Iberian whole mitochondrial genomes to assess the potential impact of North African mitochondrial DNA (mtDNA) lineages in the modern Iberian maternal gene pool. As UE2298/MS060 belongs to mtDNA haplogroup U6, we also performed a detailed phylogeographic reanalysis of this haplogroup.

## Results

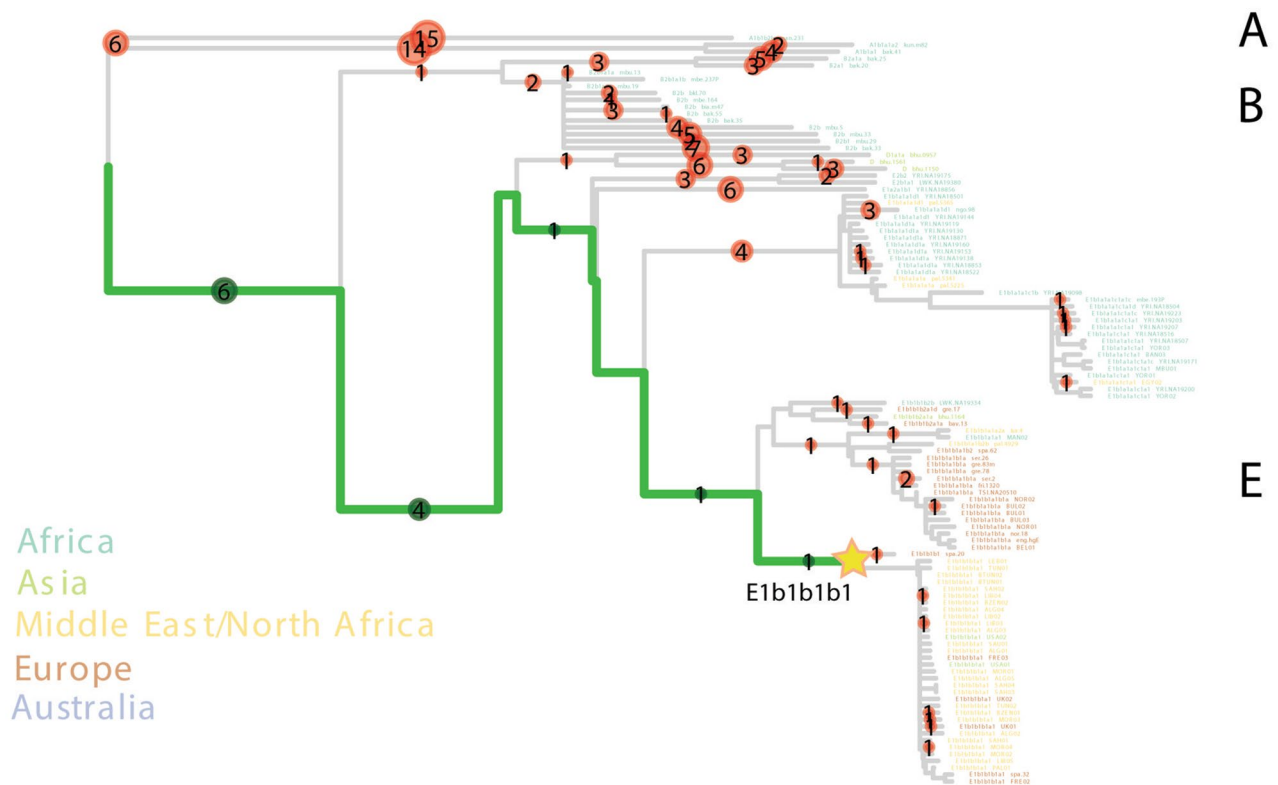
**Uniparental genetic background of the Segorbe Giant.** We confirmed that the individual was genetically male ( $R_Y > 0.077$ ; Supplementary Fig. S3), and both his uniparental markers point towards North African origins (Supplementary Table S2). He belongs to mtDNA haplogroup U6a1a1a (nomenclature according to Hernández et al.<sup>28</sup>). Although U6 in general, and U6a in particular, is present in higher frequencies in North and West Africa<sup>29,30</sup>, the complete mitochondrial genome dataset currently available is heavily biased towards Europe, and U6a1a1a, which dates to 3.5 thousand years ago (ka) (maximum-likelihood node estimation based



**Figure 1.** UE2298/MS060 maternal lineage. **(a)** Phylogenetic tree of mtDNA lineage U6a1a1.  $\rho$  and maximum-likelihood (ML) node age estimates (in ka) shown on the branches (in italics and in bold, respectively); sequences are coloured according to geography, with ancient sequences circled in red (position of UE2298/MS060 sequence is indicated by the star); underlined samples are newly reported; mutations relative to rCRS are indicated on the branches. The complete and more detailed tree for haplogroup U6 is shown in Supplementary Fig. S4. Details of the sequences used to build the tree are in Supplementary Table S4. **(b)** Timeline showing occurrence of haplogroup U6 in the archaeological record of North Africa and Iberia through time<sup>2,6,13–15,32–35</sup>, and a map of the frequency distribution of U6a in present-day Iberia, with a point indicating the location of Segorbe city. Density maps of additional mtDNA lineages are shown in Supplementary Fig. S5.

on modern variation), appears to have a more southern European distribution (Fig. 1a; Supplementary Fig. S4). However, in our Iberian mitogenome dataset, U6a1a1a occurs only at 0.3%, whereas the HVS-I (hypervariable segment I) subclade U6a1a1, defined by a transition variant at position 16239, which nests U6a1a1a, is found at ~14% in Algerian Mozabite Berbers<sup>31</sup>.

Haplogroup U6a1 has been found in Moroccan Iberomaurusian remains dating to 14–15 ka<sup>32</sup>, as well as in Early Neolithic Morocco (*i.e.* the pre-agricultural Holocene)<sup>2</sup> (Fig. 1b). Although U6 lineages have been retrieved from sixteenth century CE Islamic burials in Granada (Andalusia)<sup>6</sup>, to our knowledge, UE2298/MS060 (dating to the eleventh century CE) is the earliest documented finding of a U6 lineage in Iberia. Based on the results of our newly generated Iberian mitochondrial dataset ( $n = 1104$ : 1008 sequences from mainland Spain and the Balearic Islands, plus 96 from mainland Portugal), U6a can be found at a frequency of 1.6% in modern mainland Iberian



**Figure 2.** PathPhynder tree showing the position of UE2298/MS060 paternal lineage. Neighbour-joining phylogenetic tree estimated with 256 Y-chromosome sequences from worldwide populations<sup>45,46</sup>. Coloured circles indicate the number of derived (green) or ancestral (red) branch defining markers identified in the ancient individual. The branches coloured in green indicate the path with greatest support for the inclusion of UE2298/MS060 within a clade containing present-day Spanish, Near Eastern and North African individuals belonging to the E-M310 (E1b1b1b1) Y-chromosome lineage (indicated by the star). Label for haplogroups (A, B and E) provided on the right-hand side of the figure. The complete Y-chromosome tree is shown in Supplementary Fig. S7.

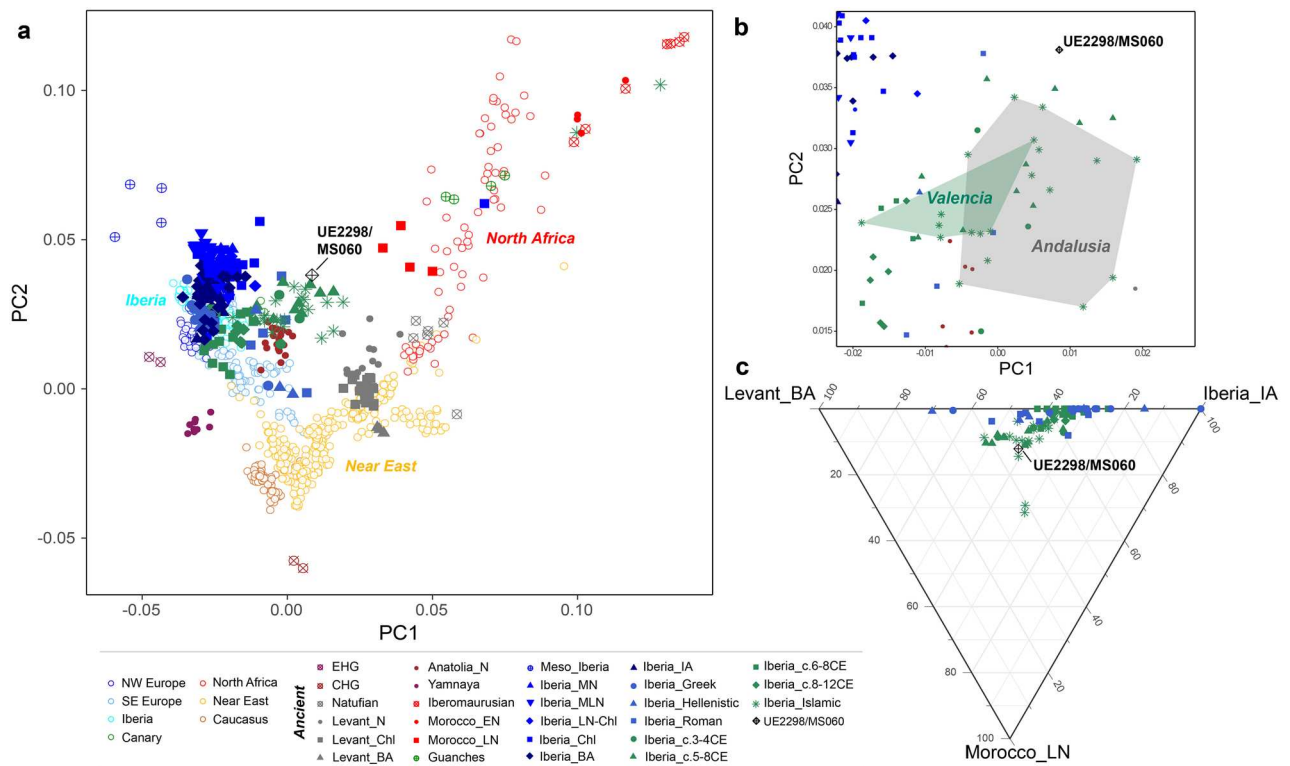
populations, with a peak of 3.6% in the south of Spain (Fig. 1b). This pattern contrasts with most mitochondrial lineages today in Iberia, although a peak of frequency in the south of the Peninsula is also observed for typically sub-Saharan African L lineages (but not for the predominantly northeast African haplogroup M1<sup>36</sup>) (Supplementary Fig. S5; Supplementary Table S5). UE2298/MS060 falls outside the modern geographic distribution of U6 lineages in Spain, suggesting that the present distribution might not reflect the medieval distribution of this haplogroup. A detailed phylogeographic analysis of U6 can be found in Supplementary Note 1.

We assigned UE2298/MS060 to the Y-chromosome haplogroup E1b1b1b1 (E-M310) (Supplementary Table S2), dating to ~13.9 [12.1–15.7] ka (Yfull, v.6.06.15) and immediately basal to the clade nesting E-M81 (E1b1b1b1a) (Fig. 2; Supplementary Figs. S6 and S7). E1b1b1 is very frequent in contemporary North Africa and has been found in North African and Levantine remains<sup>2,32,33,37</sup> (Supplementary Fig. S8). E-M81 (E1b1b1b1a), dating to ~2.8 ka (YFull, v.6.06.15), has been retrieved from early Islamic remains (seventh–eighth century CE) in southern France<sup>38</sup>, whereas the more derived E1b1b1b1a1 has been found in two individuals from an Islamic necropolis in the city of Valencia, dating to twelfth–thirteenth century CE<sup>6</sup>. E-M81 is today predominantly found in the Maghreb (where its average frequency is > 40%) and peaks in modern Berber populations, with frequencies reaching > 80%<sup>39–41</sup>, being almost fixed in some groups, such as the southern Moroccan Tachlhit-speakers<sup>42</sup> and the Chenini–Douiret and Jradou from Tunisia<sup>40</sup>. In Europe, it is found mostly in Iberia and Sicily at frequencies < 5%<sup>43</sup>.

Given that there are no reads covering any of its diagnostic positions, we cannot exclude the possibility that UE2298/MS060 could belong to the E-M81 lineage (Supplementary Fig. S6). Using pathPhynder<sup>44</sup> to investigate his Y-chromosomal affinity with present-day populations, UE2298/MS060 was positioned in a branch that harbours Iberian and North African E-M310-derived lineages, but with no support for membership to a more downstream lineage within this clade (Fig. 2; Supplementary Fig. S7).

**Genome-wide ancestry of the Segorbe Giant.** We investigated the autosomal ancestry of our ancient individual by calling ~74,200 autosomal SNPs (~72,300 when using a different approach to deal with post-mortem damage (Supplementary Table S2)). The PCA (Fig. 3a; Supplementary Fig. S9) shows that UE2298/MS060 occupies an intermediate position between present-day and ancient North African and Iberian populations in PC1, close to other Iberian Islamic individuals. Some differentiation between the Islamic individuals from



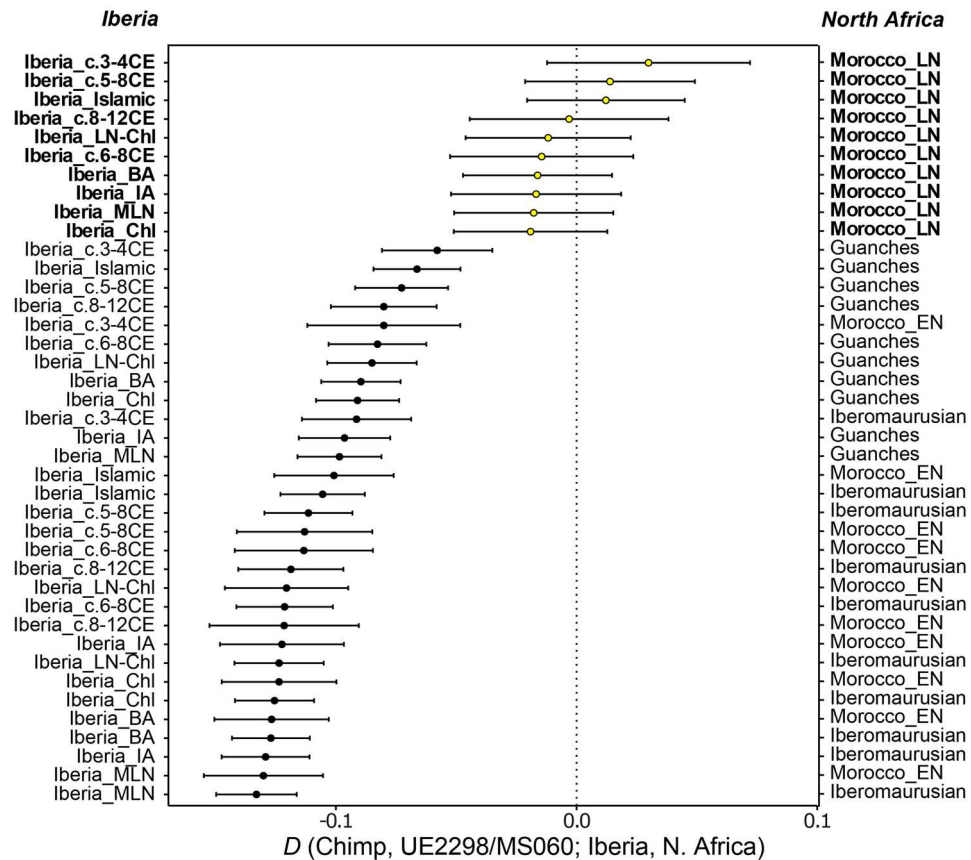


**Figure 3.** Overview of UE2298/MS060 autosomal ancestry. **(a)** PCA projecting 336 ancient samples on 702 modern individuals from North African, European, Near Eastern and Caucasian populations. **(b)** Zoom-in of PCA shown in **(a)** focussing on individuals from the Islamic period; individuals from Valencia and Andalusia (excluding two outliers that plot together with ancient North African individuals in **(a)**) within green and grey shapes, respectively. **(c)** Ternary plot showing supervised ADMIXTURE proportions ( $K=3$ ), using *Iberia\_IA*, *Morocco\_LN* and *Levant\_BA* as reference populations. Abbreviations as follows: E/CHG, Eastern/Caucasus Hunter-Gatherers; Meso, Mesolithic; (E/M/L) N, (Early/Middle/Late) Neolithic; ChI, Chalcolithic; BA, Bronze Age; IA, Iron Age; c., centuries.

Valencia and those from Andalusia is visible in the PCA, with the Andalusians mostly falling closer to North Africans and UE2298/MS060 falling outside both the Valencian and Andalusian clusters (Fig. 3b). However, this difference between UE2298/MS060 and the other Islamic individuals is not detected with ADMIXTURE in supervised mode ( $K=3$ ), using *Iberia\_IA*, *Levant\_BA* and *Morocco\_LN*/Guanches as reference populations (following the findings in Olalde et al.<sup>6</sup>) (Fig. 3c; Supplementary Fig. S10).

Outgroup- $f_3$  runs using different outgroups (*Mbuti*, *Ju\_hoan\_North* and *Ust\_Ishim*) consistently show a higher proportion of shared drift with Middle/Late Neolithic, Chalcolithic and Bronze Age Iberian populations, and with the Anatolian Neolithic (Supplementary Table S6), than with North African populations (although the proximity of North African groups, particularly Late Neolithic Morocco and the Guanches, to UE2298/MS060 changes when using *Ust\_Ishim*, a non-sub-Saharan African outgroup, suggesting that his genome may have some African-related ancestry).  $D$ -statistics consistently show UE2298/MS060 to be significantly closer to Iberian populations than to Iberomaurusians, Early Neolithic Morocco or the Guanches (Fig. 4; Supplementary Table S7). However, tests using Late Neolithic Morocco, in the form  $D(\text{outgroup}, \text{UE2298/MS060}; \text{Morocco_LN}, \text{Iberian population})$ , consistently generated results close to zero and non-significant ( $|Z|$ -score  $< 3$ ), which might be an indicator that a population genetically close to *Morocco\_LN* contributed to the ancestry of UE2298/MS060 in similar proportions to an Iberian source. We note that we did not observe any major differences in the patterns observed for outgroup- $f_3$  and  $D$ -statistics using different approaches to minimise the effects of post-mortem damage (“mapDamage —rescale” and “soft-clipping”) (Supplementary Tables S6 and S7), but additional  $qpAdm$  models are accepted using “mapDamage --rescale” (Supplementary Tables S8 and S9).

We tested different  $qpAdm$  1-way scenarios using different proximal Iberian sources as left populations. Models using populations from Andalusia (*Iberia\_c.5-8CE* and *Iberia\_c.3-4CE*, which already displayed North African-related ancestry<sup>6</sup>) are accepted ( $p$ -values: 0.092 and 0.343, respectively), whereas models using populations from Catalonia, in the northeast of the Peninsula, are rejected ( $p$ -value  $< 0.05$ ) (Supplementary Table S8). However, considering the genetic heterogeneity in different regions of Iberia through time, and given the complex history of population interactions in Iberia during the first millennium CE<sup>16,18</sup>, it is unlikely that UE2298/MS060 descends directly from Andalusian Visigothic populations and therefore we also explored 2-way admixture scenarios. Notably, 1-way  $qpAdm$  analysis was consistent with UE2298/MS060 descending from *Islamic\_Andalusia* ( $p$ -value = 0.327) but not from *Islamic\_Valencia* ( $p$ -value = 0.0005), in line with the position of UE2298/MS060 in the PCA (Fig. 3b) and highlighting regional genetic differences during this period.

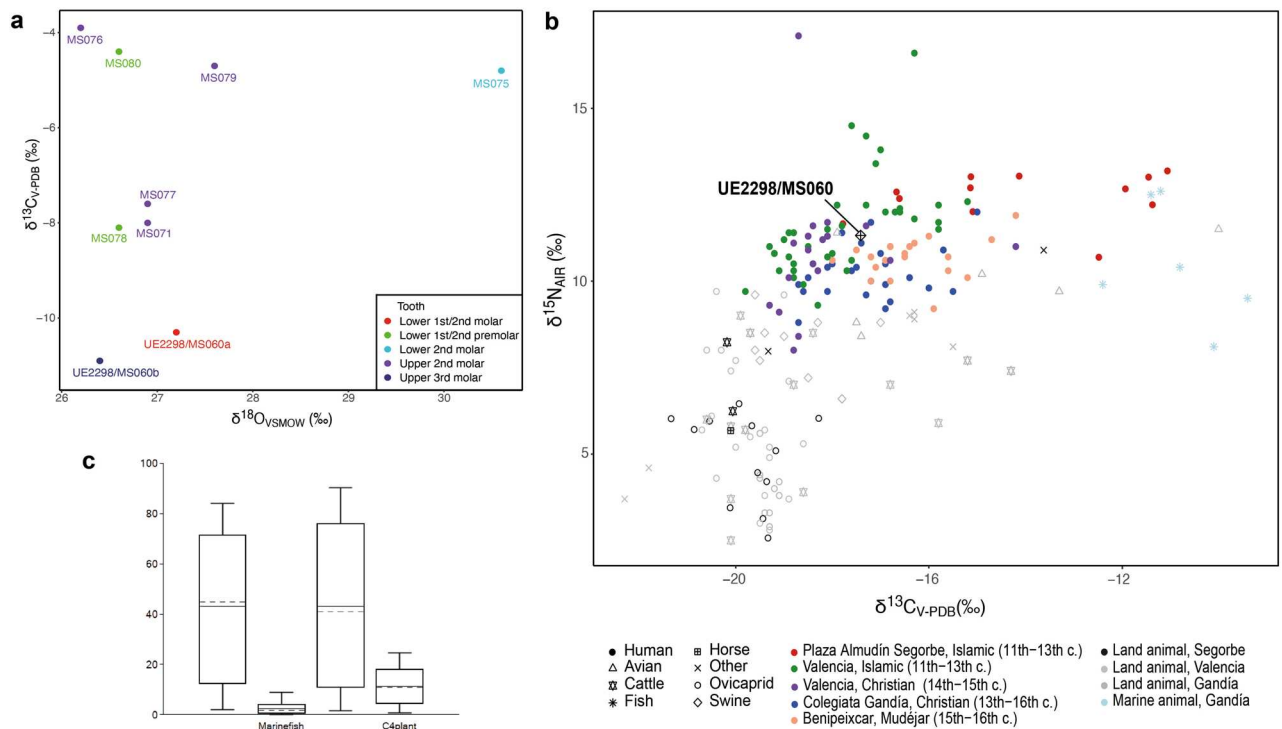


**Figure 4.** Detection of North African- and European-related ancestries in the genome of UE2298/MS060.  $D(\text{Chimp, UE2298/MS060; Iberian population, North African population})$ . A significant negative  $D$ -value indicates that UE2298/MS060 shares more genetic drift with the Iberian population; a significant positive  $D$  indicates more shared drift with the North African population. Non-significant  $D$  indicates that UE2298/MS060 is symmetrically close to both populations tested (shown in yellow, with labels in bold). Error bars correspond to 2 standard errors. Detailed output can be found in Supplementary Table S7. Abbreviations as follows: (E/M/L)N, (Early/Middle/Late) Neolithic; Chl, Chalcolithic; BA, Bronze Age; IA, Iron Age; c., centuries.

Target	Left populations		Admixture proportions		SE	$p$ -value
	Pop. 1	Pop. 2	Pop. 1	Pop. 2		
UE2298/MS060	Guanches	Iberia_Islamic	0.102	0.898	0.066	0.148617
<i>UE2298/MS060</i>	<i>Guanches</i>	<i>Iberia_c.8-12CE</i>	0.172	0.828	0.058	0.220096
UE2298/MS060	Guanches	Iberia_c.5-8CE	0.122	0.878	0.064	0.168943
UE2298/MS060	Guanches	Iberia_c.3-4CE	0.091	0.909	0.07	0.391308
UE2298/MS060	Guanches	Iberia_IA	0.349	0.651	0.048	0.078756
UE2298/MS060	Morocco_LN	Iberia_Islamic	0.235	0.765	0.149	0.146791
<i>UE2298/MS060</i>	<i>Morocco_LN</i>	<i>Iberia_c.8-12CE</i>	0.308	0.692	0.13	0.099908
UE2298/MS060	Morocco_LN	Iberia_c.6-8CE	0.593	0.407	0.059	0.053639
UE2298/MS060	Morocco_LN	Iberia_c.5-8CE	0.18	0.82	0.17	0.091952
UE2298/MS060	Morocco_LN	Iberia_c.3-4CE	0.094	0.906	0.23	0.287481

**Table 1.** Accepted 2-way  $qpAdm$  admixture models with standard errors (SE) and  $p$ -values. Models accepted using both datasets (“mapDamage --rescale” and “soft-clipping”) are shown in italics.

Alternatively, UE2298/MS060 could be modelled using 2-way combinations of distal and proximal Iberian populations (showing varied proportions of North-African related ancestry<sup>6</sup>) and either the *Guanches* or *Morocco\_LN* (Table 1; Supplementary Table S9).  $D$ -statistics comparing these two North African populations indicate that UE2298/MS060 is closer to *Morocco\_LN* ( $|Z| > 3$ ) (Supplementary Table S7) than to the *Guanches*.



**Figure 5.** Mobility and diet in Islamic Segorbe. **(a)** Mobility isotopes (oxygen and carbon) for UE2298/MS060 and other individuals from Plaza del Almudín. **(b)** Dietary isotopes (carbon and nitrogen) from Plaza del Almudín compared to other medieval Islamic and Christian sites from Gandía and Valencia<sup>51,52</sup>. **(c)** FRUITS model for UE2298/MS060; models for other individuals can be found in Supplementary Fig. S11.

**Mobility in Islamic Segorbe.** In order to assess whether or not UE2298/MS060 was likely to have spent their childhood in the local region, we performed stable oxygen analysis on eight individuals from Plaza del Almudín. Tooth enamel carbonate data is presented in Supplementary Table S10 and plotted in Fig. 5a. The  $\delta^{18}\text{O}_{\text{VSMOW}}$  values for the Segorbe population (excluding outlier MS075) range from 26.2 to 27.6‰ (range = 1.4‰,  $n = 7$ ), with a mean of  $26.8 \pm 0.5\text{‰}$  ( $1\sigma$ ). The converted  $\delta^{18}\text{O}_{\text{dw}}$  values (mean  $-6.0\text{‰}$ , excluding MS075) fit with the meteoric water values for eastern Iberian coast. The  $\delta^{18}\text{O}_{\text{VSMOW}}$  values from both teeth sampled from UE2298/MS060 are consistent with the rest of the population and the small difference in values between the different molars (M1/M2 and M3) provide no indication of movement between early childhood and adolescence. Overall, there is no evidence that UE2298/MS060 was an immigrant in East Spain, on the basis of his oxygen values.

By contrast, one other individual reported here (MS075) seems to be an outlier ( $\delta^{18}\text{O}_{\text{VSMOW}} = 30.6$ ;  $> 1.5$  times the interquartile range above quartile 3)<sup>47</sup>, and possibly a migrant from a warmer climate, with a  $\delta^{18}\text{O}_{\text{dw}}$  value similar to Africa or the Near East<sup>48</sup>. Detailed results and discussion of oxygen analysis can be found in Supplementary Note 2.

**Diet patterns in Islamic Segorbe.** The values for  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  dietary isotopes in the Islamic necropolis of Plaza del Almudín range between 10.7 to 13.2‰ and from  $-17.8$  to  $-11\text{‰}$ , respectively, for the 13 individuals studied (Fig. 5b; Supplementary Table S11). UE2298/MS060 has a  $\delta^{15}\text{N}$  value of 11.3‰ and a  $\delta^{13}\text{C}$  value of  $-17.4\text{‰}$ , showing lower  $\delta^{15}\text{N}$  and a more negative  $\delta^{13}\text{C}$  than the majority of the humans sampled from this assemblage. Application of a Bayesian mixing model (BMM), FRUITS (Food Reconstruction Using Isotopic Transferred Signals)<sup>49</sup>, supports the observation that  $\text{C}_4$  plants likely played a substantial part in the diet of some individuals and that marine fish consumption was variable (Supplementary Fig. S11). UE2298/MS060 (Fig. 5c) seems to have consumed limited amounts of  $\text{C}_4$ -plants (mean:  $11.4 \pm 6.5\%$  or  $4.8$ – $17.9\%$  of the diet) and marine protein (mean:  $2.4 \pm 2.4\%$  or  $0$ – $4.8\%$  of the diet) compared to the rest of the population analysed. On the other hand, he seems to have the highest levels of mammal and  $\text{C}_3$ -plant consumption amongst the analysed individuals (Supplementary Fig. S11).

Individual MS075, identified as a possible migrant due to their oxygen value, displays the lowest probability (close to zero) of marine fish consumption amongst the individuals studied here (Supplementary Fig. S11), and shows signals of a mixed  $\text{C}_3/\text{C}_4$  diet, which is also a possibility for Africa<sup>50</sup>. Detailed results and discussion of diet patterns inferred from individuals from the site of Plaza del Almudín can be found in Supplementary Note 2.

**Discussion.** We analysed individual UE2298/MS060 excavated from the Islamic necropolis of Plaza del Almudín, in Segorbe, dating to the eleventh century CE. The archaeologists responsible for the excavation in 1999 considered this individual unusual due to his considerable height compared with other individuals found



at the same site (despite periods of disease and/or malnutrition in childhood)<sup>27</sup>, and dubbed him the “Segorbe Giant”. The subsequent anthropological analysis suggested some African morphological features and a link was postulated to the Berber-speaking populations that settled in the region in medieval times<sup>26,27</sup>.

Analysis of the uniparental markers from UE2298/MS060 fits well with this assumption, pointing to an origin in the Maghreb, most likely from a Berber group. MtDNA lineage U6a is not only connected to modern Amazigh populations<sup>30</sup>, but has also been found in Moroccan remains associated with Iberomaurusian culture, and in the Moroccan Early Neolithic site of Ifri n’Amr or Moussa<sup>2,32</sup> (Fig. 1b). He also carries the Y-chromosome E1b1b1b1 (E-M310) lineage. E1b1b is extremely common amongst extant North Africans and has been found in ancient North African and Levantine remains<sup>2,32,33,37</sup> (Supplementary Fig. S7). Due to low coverage, we could only assign him to a basal position within E1b1b1b1, but it is possible that he may belong to a more derived subclade. One possibility would be E1b1b1b1a (E-M81), which is the most common haplogroup amongst modern Berber males today<sup>42,53</sup>, and has been linked to Islamic remains in southern France<sup>38</sup>. Another would be its descendant E1b1b1b1a1-M183 lineage, identified in three Guanche males, in two Islamic individuals from Granada, and in an earlier sixth century CE male from the Visigoth phase of Pla de l’Horta, in Catalonia<sup>6,33</sup>.

Although he carries both uniparental markers of North African origin, autosomal evidence paints a more complex picture. The individual is positioned in the PCA mid-way between modern/ancient Iberian populations, and Late Neolithic Moroccan, Guanches and modern North African individuals (Fig. 3a), and formal tests of admixture point to high proportions of Iberian-like ancestry (Fig. 4; Supplementary Table S7).

Considering the archaeological and historical records for this period in the region of Valencia, we envisage three possible scenarios to explain the observed ancestry in UE2298/MS060. One would be to assume that this individual is a direct migrant from North Africa (whose unique genetic composition has not yet been examined using aDNA), or derives from a population that moved into Iberia but retained its genetic identity. A second scenario is that he descends from pre-Islamic Iberian genetic diversity. Finally, the third scenario is that he is the result of admixture between Iberian and North African sources.

The first scenario would imply that pre-Islamic populations in North Africa would be genetically similar to UE2298/MS060 (or possibly to other contemporary individuals found in Spain<sup>6</sup>). The nearest temporal proxy available are the Guanches (from the seventh–eleventh centuries CE), who originated in the Maghreb but have been isolated in the Canary Islands since at least the early Iron Age. *D*-statistics, however, suggest that UE2298/MS060 is genetically closer to *Morocco\_LN* than to the *Guanches* (Supplementary Table S7). In any case, *qpAdm* rejects the hypothesis that UE2298/MS060 directly descends from a population resembling either the *Guanches* or *Morocco\_LN* (Supplementary Table S8). Additionally, the oxygen data for UE2298/MS060 (Supplementary Note 2) is consistent with someone who grew up in the region, and points towards low mobility between early childhood and adolescence. (In contrast, another individual from the same necropolis (MS075) does look non-local (Supplementary Note 2), possibly a migrant from a warmer climate outside the Mediterranean, with oxygen values similar to those of Africa or the Near East<sup>48</sup>). Nevertheless, one should note that aDNA sampling in North Africa is sparse and limited to a few individuals from very specific sites and periods, and we cannot rule out that a population with a similar genetic composition to that of UE2298/MS060 existed in the region around this period.

Although North African-related ancestry in present-day Spain is present at low values (typically ~3–8%), with a slight southwest-to-northeast decline<sup>19,20</sup>, increased African-related ancestry has been present in south Spain since the third century CE<sup>6</sup>. This North African influence is captured in our *qpAdm* analysis, with 1-way models using pre-Islamic Andalusian populations being accepted (Supplementary Table S8). However, it is unlikely that UE2298/MS060 descends directly from Andalusian Visigothic populations and ultimately these models, despite being statistically plausible, do not fully explain the ancestry of our individual. We note that there are no data available from or around the region of Valencia between the end of the Iron Age and the Islamic period, and post-Iron-Age genetic variation in Spain was most likely very heterogeneous across locations and centuries<sup>6</sup>. This heterogeneity is confirmed by our results showing that UE2298/MS060 forms a clade with *Islamic\_Andalusia*, but not with *Islamic\_Valencia* (Supplementary Table S8).

The third scenario would be that the genetic variation seen in UE2298/MS060 was a result of admixture between Amazigh people who migrated from North Africa to Iberia, and the local population inhabiting the Peninsula, at some point during either the Islamic conquest, the Caliphate period, or the Berber empires. This would explain UE2298/MS060’s intermediate position in the PCA and ternary plot (supervised ADMIXTURE) (Fig. 3). *D*-statistics support this scenario, with tests comparing Morocco Late Neolithic and Iberian populations from different periods not showing him to be significantly closer to one or the other (Fig. 4; Supplementary Table S7). We show that UE2298/MS060 can be modelled as admixture between Iberian and North African sources (either the Guanches from the Canary Islands or Late Neolithic Moroccans) (Table 1). The fact that he still carried both uniparental markers of North African origin suggests that the admixture may have happened only a few generations before his time, coinciding with the zenith of Berber power, rather than earlier during the conquest, in agreement with admixture dates inferred from modern Iberian genomes from Aragon and Catalonia<sup>20</sup>. However, we cannot rule out assortative mating, allowing these uniparental markers to be retained for longer, or the possibility that these lineages were common in some Iberian populations before the Islamic period. The date of the burial (eleventh century CE)<sup>27</sup> fits the historical narrative of Berber settlement in the region of Sharq al-Andalus<sup>18</sup>. Considering the genetic evidence, together with the stable isotope results and the historical accounts of intermarriage between local individuals and the North African newcomers, and in agreement with recent aDNA evidence from Iberia<sup>6</sup>, this third scenario seems the most plausible to explain the ancestry patterns seen in his genome.

Nevertheless, the original source populations are difficult to pinpoint. Due to lack of sampling in North Africa for this specific period and preceding centuries, the nearest proxies available for the North African source are the Guanches<sup>33</sup> and the Late Neolithic Moroccan population from Kelif el Boroud site<sup>2</sup>. There is high differentiation between present-day North African populations and ancient North African individuals available to date (seen

in PC3; Supplementary Fig. S9), which indicates that important population dynamics occurring after the Late Neolithic and/or Iron Age shaped extant genetic structure in the region. Modern North African populations show a signal of increased Levantine-related ancestry around the seventh century CE, as a result of movements from the Near East during the Islamic expansion into North Africa<sup>17</sup>; the impact of these movements was also seen in the Levant, as shown by the study of seventh–eighth century Islamic individuals in Syria<sup>54</sup>. Therefore, the North-African source of UE2298/MS060 might have already displayed this increased Near Eastern-related ancestry. Similarly, the population of Valencia in the immediately preceding centuries has yet to be studied.

A study in modern South Americans detected North African ancestry introduced at the early stages of European colonization<sup>55</sup>. The presence of individuals in medieval Spain with a genetic background similar to that of UE2298/MS060 would explain the source of this ancestry in America, suggesting that admixture with North Africans had a wider impact on medieval Spanish genetic variation, before virtually disappearing in the following centuries.

We found no U6 in our present-day whole-mtDNA dataset from the region of Valencia ( $n = 54$ ), or in a larger previously published HVS-I database ( $n = 123$ )<sup>56</sup>. This absence might be an echo of the brutality of the decree of expulsion of *Moriscos* (Muslims forcibly converted to Christianity), which may have effectively erased the population carrying North African-related ancestry that lived in the region in the preceding centuries. They were replaced by settlers from regions further north with little North African-related ancestry<sup>20</sup>. This is in sharp contrast with regions of the Crown of Castilla, where historical sources claim there was better integration of the *Morisco* identity into the general population, and where no mass deportations were recorded: the frequency of U6, M1 and L lineages are higher in these regions today (present-day central and south Spain) (Fig. 1b; Supplementary Fig. S5). This pattern is also visible at the genome-wide level<sup>20</sup>.

This study emphasises the importance of immigration during the Islamic period. In contrast to Andalusia, the region of Valencia is not geographically close to the Maghreb, and was under Islamic rule for a shorter time, but nonetheless developed strong links with the Arab–Berber world during the Islamic period<sup>57</sup>. A contemporary individual, MS075, is evidence of continued movement during Berber rule (Supplementary Note 2).

UE2298/MS060 is a single, low-coverage sample and although the results cannot be extrapolated to the population as a whole, recently published results<sup>6</sup> show a similar trend of admixture in Islamic Spain. The heterogeneity of genomic patterns that is now being uncovered by aDNA studies emphasises the need for much more detailed, high-resolution fine-scale studies. More individuals and a wider diversity of sites across the Peninsula should be studied to explore the population dynamics during the Islamic period in more detail and assess potential fine differences between geographical regions and periods, and between urban and rural societies.

## Methods

**Islamic Segorbe: aDNA and stable isotope analysis.** We collected teeth from thirteen individuals from the medieval Islamic necropolis of Plaza del Almudín in Segorbe<sup>27</sup> (province of Castellón, Spain) (Supplementary Fig. S2; Supplementary Table S1). Although the necropolis is dated to the eleventh–thirteenth centuries CE, the samples studied here come from a context dated to the eleventh century. We screened three individuals for aDNA, but only one, UE2298/MS060 (dubbed the “Segorbe Giant” due to his unusual height), excavated in 1999, yielded sufficient DNA for genomic analysis (Supplementary Fig. S2; Supplementary Table S2). We undertook stable isotope analyses on a total of thirteen individuals (including UE2298/MS060), and sixteen bone fragments from animals found in the site (although these might post-date the timeframe of the Islamic necropolis of Plaza del Almudín and belong instead to the later Christian context). All samples were collected from the Museo Municipal de Arqueología y Etnología de Segorbe, and permissions were agreed by the museum and granted by the Direcció General de Cultura i Patrimoni (Conselleria d'Educació, Investigació, Cultura i Esport de la Generalitat Valenciana).

We processed all the archaeological samples in clean rooms in the specialized Ancient DNA Facility at the University of Huddersfield. We sequenced one USER™-treated library on a tenth of an Illumina HiSeq4000 lane (100 cycles) to screen for endogenous aDNA content, and later sequenced three additional libraries (one of which was non-USER treated) in half an Illumina HiSeq4000 lane (100 cycles) (Macrogen, inc., Seoul, South Korea). We performed oxygen analysis and ZooMS (for taxonomic identification of the faunal assemblage) at the University of York, and dietary isotope analysis of carbon and nitrogen at the Research Laboratory for Archaeology, University of Oxford. Further details of ancient DNA, stable isotope and ZooMS analyses can be found in Supplementary Methods.

**Sequence data processing.** We assessed raw read quality with FastQC v.0.11.5<sup>58</sup>, and merged paired-end reads and removed sequencing adapters using leeHomp<sup>59</sup>. We mapped reads both to the human genome reference (hg19, modified to include rCRS (revised Cambridge Reference Sequence) instead of chrM) and to only the rCRS with BWA v.0.7.5a-r405<sup>60</sup> *aln* (using the optimized settings for aDNA mapping<sup>61</sup>) and *samse*. We performed quality control of the alignment with QualiMap v.2.2<sup>62</sup> and confirmed aDNA authenticity by checking contamination estimates (schmutzi<sup>63</sup> and ANGSD<sup>64</sup>) and post-mortem damage patterns (Supplementary Fig. S12), as well as consistency in mtDNA haplogroup and sex assignment across all libraries. To avoid SNP miscalls due to post-mortem damage, we followed two approaches: (1) downscaling base quality of positions likely affected by post-mortem misincorporations using mapDamage<sup>65</sup> *--rescale*; and (2) soft-clipping the terminal 3 base pairs of sequencing reads using the *trimBam* option in bamUtil package v. 1.0.14<sup>66</sup>, to control for potential reference bias resulting from downscaling base quality scores that could influence formal tests of admixture. We merged all libraries using picard MERGESAM (<https://github.com/broadinstitute/picard>). Detailed methods and parameters can be found in Supplementary Methods.

**Analysis of mtDNA and Y-chromosome variation.** We used HaploGrep 2.0<sup>67</sup> to classify mtDNA haplogroups. We performed Y-chromosome haplogroup classification using Yleaf<sup>68</sup>, and checked mutations against the ISOGG (International Society of Genetic Genealogy) SNP index (as of June 2018). We used pathPhynder<sup>44</sup> (<https://github.com/ruidlpm/pathPhynder>) to investigate the affinity of individual UE2298/MS060 with present-day Y chromosomes<sup>45,46</sup>. More details on the analysis of uniparental markers of UE2298/MS060 can be found in Supplementary Methods.

**Autosomal DNA analysis.** We called pseudo-haploid autosomal SNPs (Supplementary Table S2) against the 1240k SNP list for UE2298/MS060 (available at <https://reich.hms.harvard.edu/>) using samtools *mpileup* and *pileupCaller* (<https://github.com/stschiff/sequenceTools>). We used *convertf* and *mergeit* (both included in EIGENSOFT v.7.2.1 package<sup>69</sup>) to merge and convert files when necessary.

We compiled a dataset with ~ 1.2 M SNPs for analysis using only ancient samples. Published ancient samples were remapped to our reference and reanalysed alongside UE2298/MS060 to prevent possible batch effects due to differences in pipelines. Principal component analysis (PCA) of ~ 600 k autosomal SNPs was performed using smartpca (EIGENSOFT v.7.2.1) to project 336 ancient samples (Supplementary Table S3) on a selection of 702 modern individuals from North Africa, Europe, the Caucasus and the Near East<sup>37</sup>.

We filtered the ancient dataset for positions in linkage disequilibrium (LD) using the command *--indep-pairwise* (200, 25, 0.4) in PLINK v.1.07<sup>70</sup>. The LD pruned dataset (~ 450 k SNPs) was used to run ADMIXTURE v.1.3.0<sup>71</sup> for post-Iron Age Iberian individuals (shown to display different levels of North African and Levantine-associated ancestries<sup>6</sup>) in supervised mode for  $K = 3$  (with parameters: *--cv* and *--seed time*), using *Iberia\_IA*, *Morocco\_LN/Guanches* and *Levant\_BA* as reference populations (Supplementary Table S3).

We added outgroups to the ~ 1.2 M SNP dataset for the formal tests of admixture (ADMIXTOOLS v.4.1<sup>72</sup>), and ran the tests in the two datasets (generated using “mapDamage *--rescale*” and “soft-clipping”). We examined patterns of allele sharing between UE2298/MS060 and present-day and ancient populations using outgroup-*f*<sub>3</sub> statistics, as implemented in *qp3Pop*, testing three outgroups (*Mbuti*, *Ju\_hoan\_North*, *Ust\_Ishim*) to account for deeply divergent human ancestry. We computed *D*-statistics (using chimpanzee and Mbuti as outgroups) with *qpDstat* to untangle Iberian and North African-related contributions. For a more refined analysis, we ran a test with the formula *D(outgroup, UE2298/MS060; Islamic\_Valencia, Islamic\_Andalusia)*. In order to investigate admixture proportions in the genome of UE2298/MS060, we ran *qpAdm* (ADMIXTOOLS v.4.1), using *allsnps: YES* and testing 1- and 2-way models. Following 2-way *qpAdm* results, we ran a *D*-statistics test in the form *D(outgroup, UE2298/MS060; Morocco\_LN, Guanches)*. All plots were created with RStudio<sup>73</sup>. Detailed methods and parameters can be found in Supplementary Methods.

**Modern Iberian mtDNA dataset.** We newly sequenced a total of 1126 mitogenomes from present-day Spain and Portugal (including samples assigned to insular territories, Melilla and Ceuta) with Illumina MiSeq paired-end sequencing (size of fragment: 150 bp) (Earlham Institute, Norwich Science Park, UK). A detailed description of the long-range PCR protocol, sequencing and data analysis can be found in Supplementary Methods.

**Phylogeographic analysis of mtDNA haplogroup U6.** We performed a reassessment of phylogeographic patterns of mtDNA haplogroup U6 based on a total of 330 modern (35 of which are newly published here) and 32 ancient sequences (including UE2298/MS060) (Supplementary Table S4). Detailed description of the methods can be found in Supplementary Methods.

**Ethics statement.** All archaeological samples were collected from the Museo Municipal de Arqueología y Etnología de Segorbe, and permissions were agreed by the museum and granted by the Direcció General de Cultura i Patrimoni (Conselleria d'Educació, Investigació, Cultura i Esport de la Generalitat Valenciana). For the present-day dataset written informed consent was obtained from all sample donors. The research was performed in accordance with the relevant guidelines and regulations and was approved by the University of Huddersfield's School of Applied Sciences Ethics Committee, the Ethical Committee of the University of Santiago de Compostela, the Ethics Committee for Clinical Experimentation of the University of Pavia and the Ethics Committee of the University of Minho. Portuguese modern samples (PT-codes) were collected among army volunteers, upon approval of the Portuguese Army Chief of Staff, and were fully anonymized with the single purpose of use for population studies.

### Data availability

Sequence data for UE2298/MS060 can be downloaded from the European Nucleotide Archive (accession number: PRJEB47085). Newly reported present-day mtDNA sequences are deposited into GenBank (MZ920249 - MZ921390). Additional requests should be addressed to: [marina.silva@crick.ac.uk](mailto:marina.silva@crick.ac.uk); [gonzalo.oteo-garcia@hud.ac.uk](mailto:gonzalo.oteo-garcia@hud.ac.uk); [m.b.richards@hud.ac.uk](mailto:m.b.richards@hud.ac.uk).

Received: 3 May 2021; Accepted: 27 July 2021

Published online: 13 September 2021

### References

1. Anderung, C. *et al.* Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle. *Proc. Natl. Acad. Sci.* **102**, 8431–8435 (2005).

2. Fregel, R. *et al.* Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6774–6779 (2018).
3. Valera, A. C. The 'exogenous' at Perdigões. Approaching interaction in the late 4<sup>th</sup> and 3<sup>rd</sup> millennium BC in Southwest Iberia. in *Key Resources and Sociocultural Developments in the Iberian Chalcolithic* (eds. Bartelheim, M., Ramírez, P. B. & Kunst, M.) 201–224 (Tübingen Library Publishing, 2017).
4. Sanjuán, L. G., Triviño, M. L., Schuhmacher, T. X., Wheatley, D. & Banerjee, A. Ivory craftsmanship, trade and social significance in the Southern Iberian Copper Age: The evidence from the PP4-Montelirio sector of Valencina de la Concepción (Seville, Spain). *Eur. J. Archaeol.* **16**, 610–635 (2013).
5. González-Fortes, G. *et al.* A western route of prehistoric human migration from Africa into the Iberian Peninsula. *Proc. R. Soc. B Biol. Sci.* **286**, 20182288. <https://doi.org/10.1098/rspb.2018.2288> (2019).
6. Olalde, I. *et al.* The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
7. Ottoni, C. *et al.* Mitochondrial haplogroup H1 in North Africa: An Early Holocene arrival from Iberia. *PLoS ONE* **5**, e13378. <https://doi.org/10.1371/journal.pone.0013378> (2010).
8. Achilli, A. *et al.* Saami and Berbers—An unexpected mitochondrial DNA link. *Am. J. Hum. Genet.* **76**, 883–886 (2005).
9. Sánchez-Quinto, F. *et al.* Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr. Biol.* **22**, 1494–1499 (2012).
10. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
11. Olalde, I. *et al.* A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. *Mol. Biol. Evol.* **32**, 3132–3142 (2015).
12. Günther, T. *et al.* Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11917–11922 (2015).
13. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
14. Martiniano, R. *et al.* The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLOS Genet.* **13**, e1006852. <https://doi.org/10.1371/journal.pgen.1006852> (2017).
15. Valdiosera, C. *et al.* Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3428–3433 (2018).
16. Moorjani, P. *et al.* The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* **7**, e1001373. <https://doi.org/10.1371/journal.pgen.1001373> (2011).
17. Arauna, L. R. *et al.* Recent historical migrations have shaped the gene pool of Arabs and Berbers in North Africa. *Mol. Biol. Evol.* **34**, 318–329 (2016).
18. Watt, W. M. & Cachia, P. *A history of Islamic Spain. Islamic surveys* (University of Edinburgh Press, 1996).
19. Botigué, L. R. *et al.* Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11791–11796 (2013).
20. Bycroft, C. *et al.* Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula. *Nat. Commun.* **10**, 551. <https://doi.org/10.1038/s41467-018-08272-w> (2019).
21. Hitti, P. K. *The Arabs: A Short History*. (Regnery Publishing, 1990).
22. Colás Latorre, G. Los moriscos aragoneses: una definición más allá de la religión y la política. *Sharq al-Andalus* **12**, 147–161 (1995).
23. Cabanes Pecourt, M. de los D. La repoblación de los aragoneses en Valencia. In *Bajar al reino: relaciones sociales, económicas y comerciales entre Aragón y Valencia: siglos XIII-XIV* (ed. Sarasa Sánchez, E.) 13–30 (Institución Fernando el Católico, 2017).
24. de Tapia Sánchez, S. Los moriscos de Castilla la Vieja, ¿una identidad en proceso de disolución? *Sharq al-Andalus* **12**, 179–195 (1995).
25. Casas, M. J., Hagelberg, E., Fregel, R., Larruga, J. M. & González, A. M. Human mitochondrial DNA diversity in an archaeological site in al-Andalus: Genetic impact of migrations from North Africa in medieval Spain. *Am. J. Phys. Anthropol.* **131**, 539–551 (2006).
26. Forner, A. *Estudio antropológico y paleopatológico de un individuo de la necrópolis del Almudín*. (Universitat de València, 2002).
27. Barrachina, A. La necrópolis islámica de la plaça de l'Almudín, Sogorb (Alt Palància). Estudi antropològic i cronològic. *Quad. prehistòria i Arqueol. Castelló* **24**, 281–294 (2004).
28. Hernández, C. L. *et al.* Early Holocene and historic mtDNA African signatures in the Iberian Peninsula: The Andalusian region as a paradigm. *PLoS ONE* **10**, e0139784. <https://doi.org/10.1371/journal.pone.0139784> (2015).
29. Maca-Meyer, N. *et al.* Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. *BMC Genet.* **4**, 15; 0.1186/1471-2156-4-15 (2003).
30. Secher, B. *et al.* The history of the North African mitochondrial DNA haplogroup U6 gene flow into the African, Eurasian and American continents. *BMC Evol. Biol.* **14**, 109. <https://doi.org/10.1186/1471-2148-14-109> (2014).
31. Macaulay, V. *et al.* The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* **64**, 232–249 (1999).
32. van de Loosdrecht, M. *et al.* Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).
33. Rodríguez-Varela, R. *et al.* Genomic analyses of pre-European conquest human remains from the Canary Islands reveal close affinity to modern North Africans. *Curr. Biol.* **27**, 3396–3402.e5. <https://doi.org/10.1016/j.cub.2017.09.059> (2017).
34. Szécsényi-Nagy, A. *et al.* The maternal genetic make-up of the Iberian Peninsula between the Neolithic and the Early Bronze Age. *Sci. Rep.* **7**, 15644. <https://doi.org/10.1038/s41598-017-15480-9> (2017).
35. Fregel, R. *et al.* Mitogenomes illuminate the origin and migration patterns of the indigenous people of the Canary Islands. *PLoS ONE* **14**, e0209125. <https://doi.org/10.1371/journal.pone.0209125> (2019).
36. Olivieri, A. *et al.* The mtDNA Legacy of the Levantine Early Upper Palaeolithic in Africa. *Science* **314**, 1767–1770 (2006).
37. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
38. Gleize, Y. *et al.* Early Medieval Muslim graves in France: First archaeological, anthropological and palaeogenomic evidence. *PLoS ONE* **11**, e0148583. <https://doi.org/10.1371/journal.pone.0148583> (2016).
39. Cruciani, F. *et al.* 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 (2004).
40. Fadhloui-Zid, K. *et al.* Mitochondrial DNA heterogeneity in Tunisian Berbers. *Ann. Hum. Genet.* **68**, 222–233 (2004).
41. Pereira, L. *et al.* Population expansion in the North African late Pleistocene signalled by mitochondrial DNA haplogroup U6. *BMC Evol. Biol.* **10**, 390. <https://doi.org/10.1186/1471-2148-10-390> (2010).
42. Reguig, A., Harich, N., Eddoukalli Abdelhamid Barakat, C. & Rouba, H. Phylogeography of E1b1b1b-M81 haplogroup and analysis of its subclades in Morocco. *Hum. Biol.* **86**, 105–112 (2014).
43. Semino, O. *et al.* Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: Inferences on the Neolithization of Europe and later migratory events in the Mediterranean Area. *Am. J. Hum. Genet* **74**, 1023–1034 (2004).
44. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing ancient DNA sequences into reference phylogenies. *bioRxiv* 12.19.423614; <https://doi.org/10.1101/2020.12.19.423614> (2020).
45. Solé-Morata, N. *et al.* Whole Y-chromosome sequences reveal an extremely recent origin of the most common North African paternal lineage E-M183 (M81). *Sci. Rep.* **7**, 15941. <https://doi.org/10.1038/s41598-017-16271-y> (2017).



46. Hallast, P. *et al.* The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. *Mol. Biol. Evol.* **32**, 661–673 (2015).
47. Lightfoot, E. & O'Connell, T. C. On the use of biomineral oxygen isotope data to identify human migrants in the archaeological record: intra-sample variation, statistical methods and geographical considerations. *PLoS ONE* **11**, e0153850. <https://doi.org/10.1371/journal.pone.0153850> (2016).
48. Bowen, G. J. & Revenaugh, J. Interpolating the isotopic composition of modern meteoric precipitation. *Water Resour. Res.* **39**, 1299. <https://doi.org/10.1029/2003WR002086> (2003).
49. Fernandes, R., Millard, A. R., Brabec, M., Nadeau, M.-J. & Groote, P. Food reconstruction using isotopic transferred signals (FRUITS): A Bayesian model for diet reconstruction. *PLoS ONE* **9**, e87436. <https://doi.org/10.1371/journal.pone.0087436> (2014).
50. Turner, B. L., Edwards, J. L., Quinn, E. A., Kingston, J. D. & Van Gerven, D. P. Age-related variation in isotopic indicators of diet at medieval Kulubnarti, Sudanese Nubia. *Int. J. Osteoarchaeol.* **17**, 1–25 (2007).
51. Alexander, M. M., Gerrard, C. M., Gutiérrez, A. & Millard, A. R. Diet, society, and economy in late medieval Spain: Stable isotope evidence from Muslims and Christians from Gandía, Valencia. *Am. J. Phys. Anthropol.* **156**, 263–273 (2015).
52. Alexander, M. M., Gutiérrez, A., Millard, A. R., Richards, M. P. & Gerrard, C. M. Economic and socio-cultural consequences of changing political rule on human and faunal diets in medieval Valencia (c. fifth–fifteenth century AD) as evidenced by stable isotopes. *Archaeol. Anthropol. Sci.* **11**, 3875–3893 (2019).
53. Fadhlouzi-Zid, K. *et al.* Genetic structure of Tunisian ethnic groups revealed by paternal lineages. *Am. J. Phys. Anthropol.* **146**, 271–280 (2011).
54. Srigyan, M. *et al.* Bioarchaeological analysis of one of the earliest Islamic burials in the Levant. *bioRxiv* 2020.09.03.281261; <https://doi.org/10.1101/2020.09.03.281261> (2020).
55. Chacón-Duque, J.-C. *et al.* Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. *Nat. Commun.* **9**, 5388. <https://doi.org/10.1038/s41467-018-07748-z> (2018).
56. Barral-Arca, R. *et al.* Meta-analysis of mitochondrial DNA variation in the Iberian Peninsula. *PLoS ONE* **11**, e0159735. <https://doi.org/10.1371/journal.pone.0159735> (2016).
57. Coscollá Sanz, V. *La Valencia musulmana*. (Carena Editors, 2003).
58. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (2010). Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
59. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* **42**, e141. <https://doi.org/10.1093/nar/gku699> (2014).
60. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. (2013).
61. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13**, 178. <https://doi.org/10.1186/1471-2164-13-178> (2012).
62. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2015).
63. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224. <https://doi.org/10.1186/s13059-015-0776-0> (2015).
64. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356. <https://doi.org/10.1186/s12859-014-0356-4> (2014).
65. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
66. Jun, G., Wing, M. K., Abecasis, G. R. & Kang, H. M. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* **25**, 918–925 (2015).
67. Kloss-Brandstätter, A. *et al.* HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum. Mutat.* **32**, 25–32 (2011).
68. Ralf, A., Montiel González, D., Zhong, K. & Kayser, M. Yleaf: Software for human Y-chromosomal haplogroup inference from next-generation sequencing data. *Mol. Biol. Evol.* **35**, 1291–1294 (2018).
69. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190. <https://doi.org/10.1371/journal.pgen.0020190> (2006).
70. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
71. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
72. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
73. RStudio Team. RStudio: Integrated Development for R (2020). Accessible at: <http://www.rstudio.com>.

## Acknowledgements

We thank the Museo Municipal de Arqueología y Etnología de Segorbe for granting access to their collections and the Conselleria d'Educació, Investigació, Cultura i Esport de la Generalitat Valenciana for granting the permissions for the study. We thank Lara Cassidy, Valeria Mattiangeli and Dan Bradley for valuable advice and technical support. Part of this work was delivered via the BBSRC National Capability in Genomics and Single Cell Analysis (BBS/E/T/000PR9816) at Earlham Institute by members of the Genomics Pipelines and Core Bioinformatics Groups. We wish to acknowledge the use of the Orion High Performance Computing cluster at the School of Applied Sciences, University of Huddersfield. M.S., G.O.G., A.Fi., P.J., M.G.B.F., K.D., B.Y. were supported by a Leverhulme Doctoral Scholarship awarded to M.B.R. and M.P. P.S., M.B.R., and M.P. acknowledge FCT (Fundação para a Ciência e a Tecnologia) support through project PTDC/EPH-ARQ/4164/2014, partially funded by FEDER funds (COMPETE 2020 project 016899). P.S., A.Br., A.R. and T.R. acknowledge FCT support through project PTDC/SOC-ANT/30316/2017. P.S. acknowledges the “Contrato-Programa” UIDB/04050/2020 and contract CEECINST/0007772018 funded by FCT I.P. A.A., A.O., and A.T. acknowledge the support of the Italian Ministry of Education, University and Research for the projects “Dipartimenti di Eccellenza” Program (2018–2022) – Department of Biology and Biotechnology “L. Spallanzani,” University of Pavia and PRIN2017 20174BTC4R. The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.



## Author contributions

M.S. and G.O.-G. contributed equally to this work. M.S., G.O.-G. and M.B.R. designed the study. M.S. and G.O.-G. processed archaeological samples. M.S., G.O.-G., A.M., A.Fi., K.D., B.Y., F.G. and C.J.E. performed genetics lab work. M.S., G.O.-G., R.M., J.G. and P.J. performed data analysis. A.R., A.Br., A.Fl., T.R., A.O., A.A., A.T., A.G.C. and A.S. collected and/or extracted modern DNA samples; regarding North African Berber samples, A.M. and T.S. collected samples, A.M. performed genetics lab work/data analysis and D.J.C. supervised the work. V.P. curates and manages the archaeological samples; A.Ba. and V.P. excavated the samples and provided archaeological interpretation. M.v.T., M.G.B.F., K.M., P.W.D. and M.A. performed isotope work. J.B. supervised the computational work. M.P., P.A.S., C.J.E. and M.B.R. supervised the overall work. M.S., G.O.-G. and M.B.R. wrote the manuscript. All authors reviewed and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-95996-3>.

**Correspondence** and requests for materials should be addressed to M.S., G.O.-G. or M.B.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021