1 **Title:**

2 Mitogenomes illuminate the origin and migration patterns of the indigenous people of the

3 Canary Islands

4

5 **Authors:**

6 Rosa Fregel[1,2], Alejandra C. Ordóñez[3], Jonathan Santana-Cabrera[4], Vicente M. Cabrera[2],

7 Javier Velasco-Vazquez[5], Verónica Alberto[6], Marco A. Moreno-Benítez[6], Teresa

8 Delgado-Darias[7], Amelia del Carmen Rodríguez-Rodríguez[5], Juan Carlos Hernández[8],

9 Jorge Pais[9], Rafaela González-Montelongo[10], José M. Lorenzo-Salazar[10], Carlos

10 Flores[10,11,12], M. Carmen Cruz de Mercadal[7], Nuria Álvarez-Rodríguez[8], Beth

11 Shapiro[13,14], Matilde Arnay[3*] and Carlos D. Bustamante[1*]

12 *These authors contributed equally

13

14 **Affiliations:**

15
16 1. Department of Genetics, Stanford University, Stanford, California, United States of
17 America.
18 2. Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad
19 de La Laguna, San Cristóbal de La Laguna, Spain.
20 3. Department of Prehistory, Anthropology and Ancient History, Universidad de La
21 Laguna, San Cristóbal de La Laguna, Spain.
22 4. Department of Archaeology, Durham University, Durham, United Kingdom.
23 5. Department of Historical Sciences, Universidad de Las Palmas de Gran Canaria, Las
24 Palmas de Gran Canaria, Spain.
25 6. Tibicena Arqueología y Patrimonio, Las Palmas de Gran Canaria, Spain.
26 7. Museo Canario, Las Palmas de Gran Canaria, Spain.
27 8. Museo Arqueológico de La Gomera, San Sebastián de La Gomera, Spain.
28 9. Museo Arqueológico Benahoarita, Los Llanos de Aridane, Spain.
29 10. Genomics Division, Instituto Tecnológico y de Energías Renovables, Granadilla,
30 Spain.
31 11. Research Unit, Hospital Universitario N. S. de Candelaria, Santa Cruz de Tenerife,
32 Spain.
33 12. CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain

34   13. Department of Ecology and Evolutionary Biology, University of California Santa
35   Cruz, Santa Cruz, California, United States of America.
36   14. Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz,
37   California, United States of America

38

Abstract

The Canary Islands' indigenous people have been the subject of substantial archaeological, anthropological, linguistic and genetic research pointing to a most probable North African Berber source. However, neither agreement about the exact point of origin nor a model for the indigenous colonization of the islands has been established. To shed light on these questions, we analyzed 48 ancient mitogenomes from 25 archaeological sites from the seven main islands. Most lineages observed in the ancient samples have a Mediterranean distribution, and belong to lineages associated with the Neolithic expansion in the Near East and Europe (T2c, J2a, X3a…). This phylogeographic analysis of Canarian indigenous mitogenomes, the first of its kind, shows that some lineages are restricted to Central North Africa (H1cf, J2a2d and T2c1d3), while others have a wider distribution, including both West and Central North Africa, and, in some cases, Europe and the Near East (U6a1a1, U6a7a1, U6b, X3a, U6c1). In addition, we identify four new Canarian-specific lineages (H1e1a9, H4a1e, J2a2d1a and L3b1a12) whose coalescence dates correlate with the estimated time for the colonization of the islands (1st millennia CE). Additionally, we observe an asymmetrical distribution of mtDNA haplogroups in the ancient population, with certain haplogroups appearing more frequently in the islands closer to the continent. This reinforces results based on modern mtDNA and Y-chromosome data, and archaeological evidence suggesting the existence of two distinct migrations. Comparisons between insular populations show that some populations had high genetic diversity, while others were probably affected by genetic drift and/or bottlenecks. In spite of observing interinsular differences in the survival of indigenous lineages, modern populations, with the sole exception of La Gomera, are homogenous across the islands, supporting the theory of extensive human mobility after the European conquest.

Introduction

66

67   The Canaries archipelago is located off the southern coast of Morocco (Figure 1).

68 Due to their oceanic volcanic origin, they have probably never been connected to the

69 continent. Mediterranean sailors discovered several groups of islands in the Atlantic

70 Ocean in the 13th century, but only the Canary Islands were found to be inhabited by an

71 indigenous population [1]. European chroniclers recorded that different islands were

72 inhabited by populations exhibiting different ways of life and speaking distinct dialects of

73 what they believed to be a Berber language. Ethno-historical sources provided ethnonyms

74 for the native population of each island (e.g. Guanches for Tenerife, Benehaoritas for La

75 Palma, and Bimbapes for El Hierro). However, for clarity, we will refer to them in

76 general terms, as the Canarian indigenous or native population.

77   Chroniclers were amazed to discover that the Canarian natives were unaware of

78 navigational methods and had remained isolated from the African continent [2,3]. During

79 the 15th century, the Spanish kingdom of Castile gradually conquered all of the Canary

80 Islands, after previous European attempts. In most of the islands, the indigenous people

81 resisted the European conquest [4]. The crushing of the resistance, and subsequent

82 European colonization, had a great impact on the indigenous people [5]. In spite of the

83 abolishment of slavery on the Islands in 1498, a large number of natives were deported

84 during and after the conquest [6]. Those that survived and stayed progressively mixed

85 with the European colonizers, leading to the loss of indigenous culture and language.

86   The geographic origin of the Canarian indigenous people was initially inferred

87 from both the interpretation of historical written sources and the analysis of

88 archaeological evidence. Most archaeological and anthropological data support a North

89 African origin for the Canarian indigenous people, relating to the Berber populations

90 [7,8]. Key evidence supporting a Berber origin includes inscriptions belonging to the

91 Libyco-Berber and Lybico-Canarian alphabets [9,10], pottery [11], communal granaries

92 [12], and domestic species [13-15]. Non-metric dental traits [16-18] and morphological

93 analyses of cranial and long bones [19,20] also show similarities between current

94 inhabitants of Northwest Africa and the Canarian indigenous people.

95   In regards to the time of the arrival of the first population groups, some authors

96 have proposed the first millennium BCE as the upper bound for human presence in the

4

97   archipelago [21], based on radiocarbon dating of charcoal and sediment samples. In

98   addition, there is evidence of a Roman short-stay settlement in Lobos islet dated during or

99   before the first centuries of the present era [22], which did not, according to the

100   archaeological data, involve attempting to colonize the Canaries. Recently, there has been

101   an effort to review and contextualize radiocarbon dates in the Canary Islands to better

102   assess the time of the archipelago's indigenous colonization. Accelerator mass

103   spectrometry (AMS) analyses support a later colonisation of the Canary Islands dating to

104   the outset of the first millennium AD. If only AMS analyses performed on short-lived

105   samples are considered [23], the earliest dates from the eastern islands of Lanzarote and

106   Fuerteventura range between 100-300 cal AD [21,22], whereas those from the central

107   island of Gran Canaria range between 400-500 AD [3]. The oldest AMS dates from

108   Tenerife are around 660-880 cal AD [24], while the western islands of La Palma, El

109   Hierro and La Gomera yield AMS dates ranging respectively between 260-450 cal AD

110   [24], 420-610 cal AD [24], and 120-330 cal AD [25]. On the other hand, older

111   radiocarbon dates that place the arrival of human populations before the 1st century BCE

112   were obtained from sediment, wood and charcoal samples that could be older than the

113   archaeological site where they were excavated.

114   Mitochondrial DNA (mtDNA) is a powerful tool for inferring the geographic

115   origin of populations [26]. MtDNA is maternally inherited, does not undergo

116   recombination and its different lineages are geographically structured in human

117   populations. For those reasons, mtDNA has been widely applied in phylogeographic

118   studies. The analysis of current Canary Islands samples using mtDNA has provided

119   support for a North African origin for the indigenous people, based on the presence of the

120   mtDNA U6 haplogroup [27], which has a clear Berber ascription [28,29]. Within the U6

121   lineages observed in the current Canary Islanders, it is worth mentioning U6b1a, a

122   haplogroup that is not present today in North Africa and which is considered a Canarian

123   autochthonous lineage [30]. Interestingly, U6b1a's coalescence age (3,600 years ago)

124   predates the proposed time of arrival of the first inhabitants of the islands, suggesting an

125   origin in North Africa [30]. Other haplogroups observed in the current Canarian people

126   have Eurasian (H, T, J…), sub-Saharan African (L1, L2 and L3) and Amerindian (A2 and

127   C1) affiliations [31]. These results highlighted the multiethnic nature of the modern

128   population of the Canary Islands, correlating with historical events, such as the

129   implementation of a slave workforce for the sugar cane plantations, or the commercial

130   connection with the Americas in the colonial period [32]. The detailed analysis of current

131   mtDNA of the modern Canary islanders has also suggested possible origins for the

132   indigenous population, including Morocco, Tunisia, Algeria or Sahara, but an overall

133   agreement has not yet been reached [31,33].

134   Regarding the colonization model, linguistic research has pointed to at least two

135   migration waves from North Africa [10,34]. Also, the observation of different cultural

136   backgrounds affecting the island of La Palma has been interpreted as evidence of

137   consecutive migrations. The specific timing for those migrations is still unclear, except

138   for La Palma, where the second wave of migration has been proposed to have taken place

139   around the 10th century [7]. This idea has also been supported by asymmetrical

140   distribution of both mtDNA [31] and Y-chromosome lineages [35] in the modern

141   Canarian population. The first colonization wave may have affected the entire

142   archipelago, creating the substrate population and bringing mtDNA and Y-chromosome

143   haplogroups observed today in most of the islands, including the mtDNA lineages U6b1a

144   or H1cf. The second colonization would have brought new migrants to certain islands and

145   created an asymmetrical distribution of haplogroups, such as T2c1 and U6c1.

146   The direct analysis of ancient remains from the Canary Islands, using mtDNA by

147   means of PCR techniques, confirmed the presence of North African markers in the

148   indigenous people, including the U6b1a haplogroup, as well as some of the Eurasian

149   lineages observed in the modern population [36]. Admixture analysis based on mtDNA

150   data, using the natives as parental population, determined that 42% of modern Canarian

151   mtDNA lineages have an indigenous origin [36]. Ancient mtDNA results from four of the

152   seven islands found high diversities for Tenerife and La Palma [33,36,37], and the partial

153   and complete fixation of certain haplogroups in La Gomera [38] and El Hierro [39],

154   suggesting that the colonization of the archipelago was a heterogeneous process and that

155   different islands could have had different evolutionary histories.

156   Although previous ancient DNA (aDNA) studies have been fundamental to

157   understanding the origin and evolution of the Canarian population, most of the ancient

158   mtDNA data produced so far has been obtained using PCR amplification. This classical

6

159 aDNA technique has provided valuable information, but results have always been
160 hindered by the risk of sample contamination. This is due to the fact that aDNA from
161 warm climates is often extremely degraded and the PCR technique is highly sensitive,
162 thus minute amounts of modern contaminant DNA can be preferentially amplified [40].
163 Additionally, because the molecules are short and degraded, aDNA analyses based on
164 PCR amplification have tended to isolate small, but informative, regions of the
165 mitochondrial genome, such as the hypervariable region (HVR). This partial information
166 does not allow for refined classification within haplogroups, which is needed to
167 discriminate between close geographical regions. This is especially true within
168 haplogroup H, which comprises ~40% of the ancient Canarian mtDNA lineages. The
169 advent of next-generation-sequencing (NGS) has greatly expanded the capacity of aDNA
170 research. NGS allows damage patterns that are unique to aDNA, such as short fragment
171 size and post-mortem damage, to be detected easily, thus authenticating mtDNA results.
172 NGS also has the advantage of providing complete mtDNA genomes to allow a better
173 geographic assignment, compared to those obtained from partial HVR sequences.

174 A recent NGS study of the Canarian indigenous people presented the first
175 complete mtDNA genomes and low-coverage full genomes from this population, and,
176 more specifically, from the central islands of Tenerife and Gran Canaria [41]. However,
177 previous aDNA data [36-39] suggested that the indigenous populations from different
178 islands might have experienced different demographic processes. The inclusion of data
179 from all seven islands is therefore of paramount importance to accurately characterizing
180 the archipelago's indigenous population. Additionally, to fully benefit from the potential
181 of ancient mtDNA data, a more detailed phylogeographic analysis is required.

182 In order to obtain a comprehensive mtDNA perspective on the origin of the
183 indigenous people of the Canary Islands, we have applied aDNA protocols and NGS to
184 assemble ancient mtDNA genomes from all seven sub-populations. Since human remains
185 from warm regions like the Canary Islands are expected to have low endogenous DNA
186 content, we applied an enrichment technique [42] to improve mtDNA coverage and
187 reduce sequencing costs.

188

189 # Methods

7

Sample collection

Samples were collected in collaboration with both Canarian universities, La Laguna (Tenerife) and Las Palmas de Gran Canaria (Gran Canaria), as well as the insular museums of Gran Canaria (El Museo Canario), La Palma (Museo Arqueológico Benahorita) and La Gomera (Museo Arqueológico de La Gomera). A total of 25 archaeological sites were selected for this project (Figure 1). Radiocarbon calibrated dates are available for several sites (Figure S1): El Agujero (1030 - 1440 cal AD), La Angostura (1318 - 1394 cal AD), Las Arenas (540 - 650 cal AD), El Capricho (400 - 480 cal AD), Cascajo (1640 - 1700 cal AD), Cuermeja (1270 - 1316 cal AD), La Fortaleza (599 – 633 cal AD), Guayadeque (540 - 737 cal AD), El Hormiguero (1020 - 1160 cal AD), Huriamen (1015 - 1050 cal AD; 1080 - 1150 cal AD), Lomo Galeón (1260 - 1290 cal AD), Montaña Mina (1313 - 1365 cal AD), El Pescante (150 - 350 cal AD), Portillo (1500 - 1580 cal AD), Puente de La Calzada (1265 - 1312 cal AD; 1358 - 1388 cal AD), Punta Azul (1015 - 1155 cal AD) and El Salitre (1060 - 1179 cal AD). For those sites with no available calibrated dates (Antoncojo, Barranco Majona, El Cedro, Cueva del Agua, El Espigón, Huerto de Los Morales, Los Pasitos and Salto del Casimiro), their assignation to the indigenous population was based on general context, the archaeological remains themselves and the presence of specific funerary practices. Sample CAN.005 is a tooth sample that was taken from a private collection of ancient human remains donated to El Museo Canario (Gran Canaria, Spain). Although this sample is not associated with any specific archaeological site, its calibrated radiocarbon date (1265 - 1312 cal AD) is in agreement with a pre-Hispanic origin. It is also worth mentioning that some archaeological sites from Tenerife (Cascajo and Portillo) are from the post-conquest period [43], but they are associated with the so-called "Alzados", indigenous people that rebelled against the European colonizers and retired to the mountains, leaving all contact with the Europeans behind [44].

DNA extraction and library preparation

218   Best-conserved samples were selected for DNA extraction. Although the petrous
219 bone is considered the best source for aDNA [45], we used teeth and small bones (e.g.
220 phalanx) to avoid destroying valuable archaeological material.

221   Required precautions were taken during the handling of samples, and all
222 experiments that included aDNA were carried out in dedicated, clean lab facilities at the
223 Paleogenomics Lab, University of California Santa Cruz, to avoid contamination. DNA
224 extraction was performed following Dabney et al. [46]. Bone samples were sanded to
225 remove the external surface, and then one bone piece was cut with a Dremel tool and
226 pulverized using a bone mill. The surface of tooth samples was decontaminated using a
227 bleach solution, and then the teeth were cut down the midline and the cementum drilled
228 using a Dremel tool and a metallic bit. Pulverized bone and tooth samples were incubated
229 overnight, using a proteinase K/EDTA solution, and DNA extracted using a silica-based
230 and guanidine method. Ancient DNA was then built into double-stranded libraries, with
231 7-bp single-index barcoding to allow for multiplexing sequencing, following Meyer and
232 Kircher [47]. Libraries were sequenced for an initial screening on an Illumina NextSeq
233 500 apparatus for obtaining paired-end shotgun data (~1 M reads per library) with a
234 sequencing read length of 2 x 75 bp.

235

Enrichment

237   After the screening of shotgun libraries, those samples with an endogenous DNA
238 content lower than 10% were enriched using whole-genome in solution capture [42].
239 Briefly, aDNA libraries were captured in singleplex reactions using human genomic
240 RNA baits, with the aim of increasing endogenous DNA rates and reducing sequencing
241 costs. Although this method is directed at capturing the whole genome, multicopy regions
242 of the mtDNA become particularly enriched. Post-capture libraries were sequenced as
243 indicated before, to obtain at least ~5 M reads per post-capture library.

244

HVR analysis

246   In order to perform population-based analyses, we included in our study
247 previously published [36-39] and unpublished HVR data from the seven islands. Newly

248 reported HVR data from the islands of Gran Canaria (n = 75), Lanzarote (n = 8) and

249 Fuerteventura (n = 10) was obtained following the methodology described by Maca-

250 Meyer et al. [36] and Ordóñez et al. [39]. Briefly, after external decontamination, tooth

251 samples were extracted by means of a GuSCN-silica based protocol. MtDNA

252 quantification was performed on a 7500 Real Time PCR system (Applied Biosystem,

253 Foster City, CA, USA), using a human-specific mtDNA fluorescent probe [48], and

254 ~3,000 copies were submitted to PCR with the aim of reducing the effects of DNA

255 damage. The mtDNA HVRI (from positions 16,000 to 16,400) was amplified using seven

256 overlapping fragments, with sizes ranging from 82 to 124 bp, to improve the

257 amplification of endogenous DNA. All the sequencing reactions were prepared with the

258 BigDye v3.1 Terminator Cycle Sequencing kit (Applied Biosystems) and run on an ABI

259 PRISM 3130xl Genetic Analyzer (Applied Biosystems). Standard contamination

260 prevention and monitoring were conducted as described earlier [39].

261

Modern mtDNA genomes

263 We included in this study several current Canary Islands mtDNA genomes,

264 analyzed using both whole-genome and Sanger sequencing. Complete genomes were

265 obtained in Instituto Tecnológico y de Energías Renovables (ITER) by whole-genome

266 sequencing from a set of 18 unrelated Canarians. Briefly, DNA samples were processed

267 with a Nextera DNA Prep kit, with dual indexes following the manufacturer's

268 recommendations (Illumina Inc., San Diego, CA). Library sizes were checked on a

269 TapeStation 4200 (Agilent Technologies, Santa Clara, CA) and their concentration

270 determined by the Qubit dsDNA HS Assay (Thermo Fisher, Waltham, MA). Samples

271 were sequenced to a depth of 30X on a HiSeq 4000 instrument (Illumina) with paired-end

272 150-base reads. Sanger sequencing mtDNA genomes were obtained at University of La

273 Laguna following previously published methodologies [49], for samples classified as

274 T2c1 (determined by HVRI analysis). These samples were selected because of their

275 potential to define new sub-lineages within T2c1. Institutional review board approval for

276 the analysis of human subjects was obtained from Stanford University.

277

Data analysis

Mapping and filtering of ancient mtDNA reads

Shotgun sequencing reads were trimmed and adapters removed using AdapterRemoval version 1.5.4 [50]. Specifically, the paired-end reads were merged, and low-quality bases (BASEQ < 20) and short reads (< 30 bp) removed. Merged trimmed reads were then mapped to the human reference genome (hg19) using BWA version 0.7.12 [51], while unmerged reads were discarded. Unmapped, low-quality (MAPQ<30) and duplicate reads were removed using SAMtools version 0.1.19 [52]. The percentage of endogenous DNA was calculated by dividing the number of reads remaining after filtering by the total number of trimmed reads.

Authentication

Damage patterns were assessed using MapDamage v2.0 [53]. Insert size of libraries was obtained with SAMtools mpileup, and plotted using R software v.3.2.0 [54]. Contamination rates of libraries were calculated using contamMix v.1.0-10 [55] and Schmutzi [56].

Analysis of complete mtDNA genomes

MtDNA reads were directly mapped to the revised Cambridge Reference Sequence (rCRS) [57] and filtered as described before. MapDamage was used to rescale the quality of bases likely affected by post-mortem damage. Indel Realigner from the GATK pipeline version 2.5.2 was also used for improving alignment quality around indels [58]. MtDNA consensus sequences were obtained using SAMtools and BCFtools version 0.1.19 [52]. A list of variants was then obtained using SAMtools mpileup, with a minimum depth of 5. Haplogroups were determined with HaploGrep version 2.0 [59], using PhyloTree build 17 version (http://www.phylotree.org) [60]. MtDNA haplotypes were manually curated by visual inspection, using Tablet v.1.17.08.17 [61]. Modern DNA sequencing data was analyzed following the same protocol used for ancient samples, except for the MapDamage rescaling step. After retrieving all available mtDNA genomes belonging to the haplogroups of interest from NCBI (http://www.ncbi.nlm.nih.gov), phylogenetic trees were built using median-joining networks [62]. Indels around nucleotides 309, 522, 573 and 16193, and hotspot

11

310 mutations (e.g. 16519) were excluded from phylogenetic analysis. For estimating

311 coalescence ages for specific clades, we used the ρ statistic [63]. We established a

312 mutation rate for the complete mtDNA sequence of one substitution in every 3,624 years,

313 correcting for purifying selection as in Soares et al. [64]. Accompanying standard errors

314 were calculated as per Saillard et al. [65]. For highly frequent haplogroups, such as H1cf

315 and T2cd3, we only kept one sample per site, to avoid relatedness interfering with

316 coalescence age estimations.

317

318 **Analysis of HVRI data**

319 Newly reported HVR sequences were analyzed using BioEdit software v.7.0.9.0

320 [66], and haplotypes were obtained by means of HaploSearch software [67] and further

321 confirmed by manually inspecting the electropherograms. Haplogroup nomenclature was

322 assessed following the most updated mtDNA phylotree (Build 17) [60].

323 Genome-wide data was combined with previous HVRI sequencing data to

324 perform population-based analysis. Published samples used for comparisons are detailed

325 in Table S1. As we do not know if samples in the same burial can be related, when

326 several samples with the same haplogroup were observed from the same archaeological

327 site, only one was included in the analysis. Two-tailed Fisher's exact test was used to

328 assess differences in mtDNA haplogroup frequencies between eastern and western

329 islands. Gene diversity was calculated according to Nei [68]. Distances between

330 populations were estimated using haplogroup frequency-based linearized F_{ST} [69] as in

331 Arlequin v.3.5 [70]. Multidimensional Scaling (MDS) was performed using R software

332 and the "smacof" package [71]. Admixture estimates were calculated with the

333 WLSAdmix program [72], which was kindly provided by Dr Jeffrey Long.

334

Results and Discussion

336 The average endogenous DNA content for the Canarian indigenous samples is

337 7.92%, a relatively high value considering the warm and humid environmental conditions

338 of the archipelago (Table S2). However, endogenous DNA values varied within and

339 between archaeological sites, ranging between 0.02% and 39.0% (IQR= 0.67% - 11.5%).

340   All samples meet the standard aDNA authentication criteria, including observation of

341   DNA fragmentation and damage patterns at both ends of molecules, and low modern

342   DNA contamination rates (Figure 2). Those contamination rates calculated with

343   contamMix are larger than those produced with Schmutzi. One possible reason is that

344   contamMix estimations are more sensitive to low coverage values (Table S2). For

345   example, sample CAN.033, with a 7.9X mtDNA coverage, has a contamination rate of

346   10.2% based on contamMix and 1.0% on schmutzi. Schmutzi has been reported to be

347   able to obtain accurate contamination rates for coverage down to ~5X [56]. However, in

348   other cases, variable contamination estimations do not seem to be related to low

349   coverage, and other factors may be interfering.

350   After capture, we obtained complete mtDNA genomes from 48 ancient human

351   remains sampled in 25 different archaeological sites (Table S2). Our sample set covers

352   the entire archipelago and a time span of 1,200 years (Figure S1). The average mtDNA

353   depth is ~140X, with a minimum value of 8X (Table S2). Observed haplogroups agree

354   with previous studies [33,36-39,41], indicating the presence of North African (U6),

355   Eurasian (H, J2, T2 and X) and sub-Saharan African lineages (L1 and L3) in the Canarian

356   indigenous population (Figure S2). As delineated before [36], the majority of

357   haplogroups observed are of Eurasian origin, most with a Mediterranean distribution.

358   This result is expected, as recent aDNA data from North Africa has indicated the

359   presence of Neolithic European lineages as early as the Late Neolithic period (~5,000

360   BP) [73].

361   We also obtained complete mtDNA genomes from a set of 18 modern Canarians

362   (Figure S3). More than 50% of the samples belong to haplogroup H, with a higher

363   diversity of sub-haplogroups than the one observed in the indigenous population. In

364   addition to H1cf and H1e1a, we observe other H1 sub-lineages and other branches, such

365   as H6a1, H3c2 or H43, which are most likely of European origin. Other haplogroups

366   present in the indigenous people are also observed in the modern population, including

367   J2a2d, U6b1a and X3a. In line with previous analyses [27,31], a sub-Saharan African

368   (L3d1b3a) [74] and an Amerindian lineage (A2) [75] are observed in the current

369   population of the Canary Islands. Assuming that our set of 48 ancient genomes is

370   representative of the native population, we performed a rough admixture estimate of

13

371   27.8% of maternal lineages in members of the present-day population possessing

372   indigenous origins, while 61.1% would be of European ascription (Figure S3).

373

374   ## Population-based analysis

375   In order to compare our samples to previously published data, we combined the

376   newly generated mtDNA genomes with HVRI data from the Canarian indigenous

377   population (Table S1) [33,36-39]. Given that sample sizes for Lanzarote and

378   Fuerteventura are small and their indigenous populations are considered to be similar

379   based on archaeological data [76], these data sets were pooled together. It is worth

380   mentioning that those samples for which mitochondrial data were generated, using both

381   classical techniques and NGS sequencing, produced identical HVRI haplotypes, proving

382   our PCR-based approach generates authentic results.

383   As previously observed, the indigenous populations of the Canary Islands in the

384   past were not homogenous (Table 1; Figure 3). The islands of La Palma and Tenerife

385   show a relatively diverse mtDNA composition (>70%) [33,36,37], while the others show

386   signs of genetic drift and/or diversity reduction events, such as a bottleneck or a founder

387   effect. In La Gomera, mtDNA diversity was 54.2%, due to the high frequency of

388   haplogroup U6b1a [38], while in El Hierro, this value was 2.9%, with the almost

389   complete fixation of H1cf haplogroup in the Punta Azul site [39]. With new data on the

390   indigenous population of Gran Canaria, Lanzarote and Fuerteventura (Table S3), we

391   show that Gran Canaria had high mtDNA diversity, similar to Tenerife and La Palma,

392   while Lanzarote and Fuerteventura had low diversity (51.1%) because of the high

393   frequency of H*(xH1cf, H4a1a) lineages. These findings emphasize that results obtained

394   from the larger islands of Tenerife and Gran Canaria should not be extrapolated to the

395   entire archipelago. Estimations of population sizes during pre-colonial times based on

396   archaeological evidence agree with mtDNA results. Populations in Gran Canaria,

397   Tenerife and, to a lesser degree, La Palma, were large and able to sustain relatively high

398   diversity, while Lanzarote, Fuerteventura and El Hierro were almost depopulated at the

399   time of the conquest [77]. In the case of La Gomera, the population size was also reported

400   to be small [78].

14

401   By directly comparing the mtDNA types found in the indigenous population of

402   each island, we observe that H4a1e, L3b1a, U5 and U6c haplogroups are present only in

403   the eastern islands (Gran Canaria, Lanzarote and Fuerteventura). Differences between

404   eastern and western islands were shown to be significant for the four haplogroups, when

405   all the ancient samples were considered: H4a1 (p=0.0127), L3b1a (p=0.0037), U5

406   (p=0.0114) and U6c (p=0.0012). Though also present in the western islands, haplogroups

407   T2c1 (p=0.0164) and U6a (p=0.0028) appeared more frequently in the eastern islands.

408   However, these results can be artifacts caused by the high frequency of H1cf in El Hierro

409   and U6b1a in La Gomera. After removing these two populations from the western group,

410   only differences in the distribution of U6c remained significant (p=0.0328).

411   In contrast with the heterogeneity we observe in pre-Hispanic times, mtDNA

412   haplogroup frequencies in modern populations of the Canary Islands are homogenous

413   (Figure 3), with the sole exception of La Gomera [27]. The high frequency of haplogroup

414   U6b1a observed in the indigenous population of La Gomera is also detected in its

415   present-day population [38]. However, the same pattern is not observed for El Hierro. In

416   pre-colonial times, H1cf was almost fixated in El Hierro [39], while the frequency of this

417   haplogroup today is 4.6%, not significantly different from the average 2.4% observed in

418   the entire archipelago (p=0.2364).

419   In order to determine the admixture pattern at an insular level, we compared

420   modern Canarian samples with their principal parental populations: indigenous people,

421   Iberians, and sub-Saharan Africans (Table S1). Global admixture estimations using the

422   new mtDNA dataset (Table 2) confirm previous results on the survival of native lineages

423   in the modern population (55.9%). However, we observed that results within islands are

424   variable. When the miscellaneous ancient sample is used as one of the parental

425   populations, indigenous contribution to the modern population ranges from 30.8% in

426   Gran Canaria to 71.4% in La Gomera. However, this approach is not correct, as we know

427   that the indigenous population of the archipelago was heterogeneous and mtDNA

428   frequencies were variable. With our new data, we were able to estimate admixture, using

429   aDNA sampled directly from each island. Indigenous mtDNA contribution estimates are

430   lower when a direct comparison is performed, with values ranging from 0% in El Hierro

431   to 55.5% in La Gomera (Table 2). The extreme result observed in El Hierro is evidently

15

432   due to the marked difference between the ancient and current people. It is interesting that,

433   when the miscellaneous sample is used, the indigenous contribution increases to 36.2%.

434   This result is reasonable, given that the present-day sample from El Hierro is not

435   significantly different from other islands. This can be explained if we consider that El

436   Hierro was almost depopulated at the time of the European conquest [79]. In fact, it was

437   recounted in the chronicles that the indigenous population of El Hierro was decimated

438   due to razzias (raids for the purpose of capturing slaves) at the time of the Spanish

439   conquest, and was later repopulated with indigenous populations from other islands and

440   European colonizers [80,81].

441   To determine if a more specific origin for the Canarian indigenous population

442   could be ascertained, the ancient mtDNA sample was combined with a reference modern

443   DNA database containing samples from the Canary Islands, Europe, North Africa, Sub-

444   Saharan Africa and the Near East (Table S1). In the MDS analysis (Figure 4), the

445   indigenous sample from El Hierro and the indigenous and modern samples from La

446   Gomera act as outliers, due to the high frequency of H1cf and U6b1a, respectively.

447   When the two outliers were removed and all the remaining ancient samples were pooled

448   together, the first dimension differentiates sub-Saharan populations from Eurasian

449   populations, including North Africa and the Canary Islands. The second dimension places

450   Canarian and European/Near Eastern populations on both ends, with North Africans in an

451   intermediate position. The closest North African sample to the Canarian indigenous

452   population in the second dimension is West Sahara. However, the ancient sample is

453   differentiated from all current North African populations and placed closer to modern

454   Canarians. This is due to the fact that haplogroups occurring frequently in the Canarian

455   indigenous and current samples (e.g. U6b1a) are not present or appear in low frequencies

456   within the reference populations. This result concurs with later demographic processes

457   reshaping the mtDNA landscape of North Africa, and/or founder effects and isolation in

458   the Canary Islands. It is interesting that, compared to the other islands, the modern

459   populations of Tenerife and Gran Canaria are closer to the European populations. This

460   result is expected, because they each have capital cities of the two Canarian provinces

461   and, thus, have received substantial historical migration from the mainland.

462

Phylogeographic analysis of mitogenomes

463

464 The HVRI has been proven to be of limited value in providing a clear picture of

465 the origin of the indigenous people of the Canary Islands. In order to conduct a better

466 assignment of the geographic origin of the maternal Canarian indigenous lineages, we

467 performed detailed phylogeographic analysis of all the lineages observed in the aDNA

468 dataset (Figure S2), including those from Rodriguez-Varela et al. [41]. For detailed

469 information on our phylogenetic analysis, see Supplementary Text.

470 We observe five different H sub-lineages in the indigenous people of the Canary

471 Islands: H1cf, H1e1a9, H2, H3 and H4a1e. H1cf (Figure S4) seems to be restricted to

472 both the Canary Islands and Central North Africa, and shows a coalescence age (~3,400

473 years ago) that is in agreement with a continental origin before the colonization of the

474 islands (Figure S5). Newly defined haplogroups H1e1a9 (Figure S6) and H4a1e (Figure

475 S7) are both restricted to the Canary Islands, with a distribution similar to that observed

476 for U6b1a. However, in this case, H1e1a9 and H4a1e coalescence ages overlap with the

477 human occupation period (Figure S5) and are compatible with an origin in the islands.

478 The presence of lineages derived of H1e1a and H4a1 in both European Neolithic and the

479 Canary Islands indigenous samples corresponds with Eurasian prehistoric intrusions in

480 North Africa (Fregel et al. 2018). Two samples were classified within basal H2 and H3

481 haplogroups, preventing further phylogenetic analysis.

482 Two sublineages of haplogroup J are observed in the indigenous population of the

483 Canary Islands: J1c3 and the newly defined J2a2d1a1. J1c3 is present in Europe, North

484 Africa and the Near East, and more interestingly, in ancient Neolithic samples from Spain

485 and Sardinia (Figure S8). Although J2a2d1a* has been spotted in Central North Africa,

486 subhaplogroup J2a2d1a1 is exclusive to the Canary Islands and Brazil, the latter

487 representing an area with known historical migrations from the islands (Figure S9).

488 Accordingly, this new autochthonous Canarian lineage has a coalescence age that

489 overlaps with the indigenous occupation of the islands (Figure S5).

490 Phylogenetic analysis of the Canarian T2 sequences places them within T2b and

491 T2c1d, two haplogroups thoroughly observed in Neolithic and Bronze Age sites from

492 Europe. The inclusion of ancient and modern Canarian samples allows us to define four

493 new T2c1d subhaplogroups (Figure S10). T2c1d3 haplogroup is present in both Tunisia

17

494 and the Canary Islands. T2c1d1c1 and its two subclades (T2c1d1c1a and T2c1d1c1b) are

495 present in both North Africa and the current population of the eastern Canary Islands.

496 This distribution could be explained by an asymmetrical migration pattern, or, given its

497 absence in the indigenous people, by a higher impact of Moorish slave trade in the

498 eastern islands (Supplementary Text).

499   We identify several indigenous samples within macrohaplogroup L, belonging to

500 L1b1a and the newly defined L3b1a12. Although Later Stone Age [82], and Early and

501 Late Neolithic [73] samples from North Africa did not show any mtDNA lineage of sub-

502 Saharan origin, our results imply the presence of L1b and L3b1a in North Africa at the

503 time of the colonization of the Canary Islands. Regarding L3b1a12 (Figure S11), this

504 lineage can also be considered autochthonous of the Canary Islands, with a coalescence

505 age posterior to the proposed colonization date (Figure S5). Interestingly, this lineage was

506 only present in the eastern islands in ancient times, but has a wider distribution at the

507 present time, suggesting extensive movement of native people after the conquest.

508   Canarian indigenous sequences belonging to X haplogroup are classified within

509 the X3a clade (Figure S12). This lineage is present both in Europe, the Near East and

510 northeast Africa, as well as in the ancient and current populations of the Canary Islands.

511   Finally, several U6 sublineages are observed in the indigenous population of the

512 Canary Islands, including U6a1a1 (Figure S13), U6a7a1 (Figure S14), U6b1a (Figure

513 S16) and U6c1 (Figure S16). U6a1a1, U6a7a1 and U6c1 are present in the Maghreb,

514 southern Europe and the Canary Islands, and are most probably related to prehistoric

515 Mediterranean expansions (Figure S13, Figure S14 and Figure S16). As reported before,

516 the Canarian autochthonous U6b1a is also present in regions with recent Canarian

517 migration, including mainland Spain and Cuba (Figure S15). Given its coalescence age

518 and the oldest calibrated radiocarbon dates from human remains from the Canary Islands

519 (Figure S5), U6b1a most probably originated in North Africa and later migrated to the

520 Canaries. However, to date, this lineage has not been observed in the continent,

521 indicating the migrations occurred after the colonization of the Canary Islands reshaped

522 the North African mtDNA landscape.

523

524

18

Discussion

525

526   Our mtDNA results on the indigenous people of the Canary Islands shed light on

527   the prehistory of North Africa. Our data are in agreement with recent aDNA data from

528   Morocco [73] and further evidence of a complex pattern of Mediterranean migrations in

529   North Africa. Archaeological records in the Maghreb support this result, and also suggest

530   further European intrusions during the Chalcolithic and Bronze Age eras [83,84].

531   Additionally, Phoenicians, Carthaginians and Romans arrived in the North African region

532   in historical times [85-88]. The presence of haplogroups of Mediterranean distribution in

533   the indigenous people of the Canaries demonstrates the impact of these prehistoric and

534   historical migrations in the Berbers and that they were already an admixed population at

535   the time of the indigenous colonization of the islands [89].

536   In our phylogeographic analysis of complete mtDNA sequences from the

537   Canarian indigenous population, we found lineages that are only observed in Central

538   North Africa and the Canary Islands (H1cf, J2a2d and T2c1d3), while others have a

539   wider distribution including both West and Central North Africa, and, in some cases,

540   Europe and the Near East (U6a1a1, U6a7a1, U6b, X3a, U6c1). These results point to a

541   complex scenario, where different migration waves from a dynamic and evolving North

542   African population reached the islands over time. Every island experienced their own

543   evolutionary path, determined by the environmental conditions and limitations of

544   insularity. Those islands with the capability of sustaining large populations retained

545   variability, while others with more restricted means (La Gomera and probably El Hierro)

546   had to develop cultural practices to avoid inbreeding, like mandatory exogamic practices

547   [78,90].

548   Although the North African Berber origin is the most widely accepted hypothesis,

549   other lines of research have proposed that certain funerary practices and religious beliefs

550   observed in the indigenous population of the Canary Islands could be linked to Punic-

551   Phoenician influence [91], thus proposing the colonization of the Canary Islands as the

552   result of Phoenicians expanding their control to the Atlantic Ocean. Based on the limits

553   of the territorial occupation of the Atlantic West Africa by Phoenicians, Carthaginians

554   and Romans, most researchers consider it unlikely that there was a political occupation or

555   economic exploitation of the archipelago [92-94]. However, the islands were not

19

556   unknown to Mediterranean cultures, and Romans possessed the seafaring skills needed to

557   travel to the islands [22]. Some authors think Phoenicians also had the navigational

558   technology required to reach the Canary Islands [95,96], although this idea has been

559   challenged [97]. The first Phoenician aDNA sample published was a complete mtDNA

560   sequence of a child from Carthage dated to the 6th century BC [98]. This Carthaginian

561   sample was classified within U5b2c1 haplogroup. This result is interesting, given that U5

562   was more frequent in the indigenous population of the eastern islands, including the

563   island of Lanzarote, where a Punic-Phoenician influence has been proposed. As U5

564   haplogroup was not uncommon in Neolithic European samples, and its presence in North

565   Africa might be due to prehistoric migrations, an alterative explanation would be that

566   haplogroup U5 was incorporated into the Berber mtDNA pool before the Carthaginians

567   were established in Tunisia. Recently, Matisoo-Smith et al. [99] published thirteen

568   complete mitogenomes from Punic-Phoenician samples from Lebanon and Sardinia. The

569   only haplogroups in common with the indigenous population of the Canary Islands are

570   H3 and H1e1a, although, in this case, the Phoenician H1e1a sample is classified within

571   the sub-lineage H1e1a10. The lack of overlap between the mtDNA composition of

572   Phoenicians and the Canarian indigenous people disagrees with either a Punic-Phoenician

573   origin for the ancient islanders or sustained contact between the two populations.

574   Previous genetic analyses of the modern Canarian population detected an

575   asymmetrical distribution of maternal and paternal lineages in the archipelago [31,35].

576   Our aDNA results confirm the existence of asymmetrical distribution of mtDNA

577   haplogroups in pre-colonial times, with the presence of haplogroups H1e1a9, H4a1e,

578   L3b1a12 and U6c1 only in the eastern islands. However, it is worth mentioning that La

579   Palma, the island with the most anthropological evidence of two migrations waves, does

580   not show any of these lineages. If we consider the presence of H1e1a9, H4a1e, L3b1a12

581   and U6c1 haplogroups to be the result of further population movements from North

582   Africa to the eastern islands, we could approximate the date based on radiocarbon dates

583   of the sites where the sample was taken. Most sites where these lineages have been

584   observed have radiocarbon dates placed around the 13th century, and all except one are

585   from after the 10th century. The only site with an older date is Guayadeque; however, we

586   have to take into account that this is a large site, with evidence of human occupation

587 extending until the 14th centuries AD [100], and the dating was not performed directly on

588 the analyzed sample.

589   Archaeological data has evidenced significant changes in the productive strategies

590 of some islands around the 11th - 12th centuries [12,76,101-103]. In fact, recent data

591 indicates probable population growth in Gran Canaria at that time, suggesting the

592 appearance of new settlements associated with an exploitation model that intensified the

593 use of marine resources, the increase in the size of settlements linked to agricultural

594 nuclei, and changes in the production of some craftsmanships [12,104,105]. These

595 changes have been interpreted as part of an endogenous process, as it has been

596 determined that this population growth involved neither significant changes in the

597 structure of human settlements or burials, nor introduced differences in land management

598 or the types of domestic species that were exploited. However, it is also possible to

599 explain those changes as the result of the arrival of new migrants to the island of Gran

600 Canaria. Although it is still under study, there is evidence for transformations in the

601 configuration of some settlements in Lanzarote, between the 8th and 13th centuries [106].

602 Again, these modifications could be reflecting changes in the conception of domestic

603 space due to an endogenous process, or associated with the arrival of new colonizers.

604 Archaeological information from Fuerteventura is not abundant enough to determine

605 population size changes that could be related to the arrival of new migrants. Nevertheless,

606 it is clear from the archaeological record that Fuerteventura and Lanzarote maintained

607 frequent contact and shared both cultural and economic elements [76,107]. Future

608 paleogenomic efforts to obtain high-coverage genomes from all seven islands, in

609 combination with proper archaeological contextualization of the genetic data and detailed

610 radiocarbon dating, will be essential for improving our knowledge of the origins and

611 evolution of the indigenous population of the Canary Islands.

612

613 **DATA AVAILABILITY:**

614

615   Mitochondrial DNA sequence data are available through the European Nucleotide

616 Archive (PRJEB29569). Consensus mtDNA sequences are available at the National

617 Center of Biotechnology Information (Accession Numbers MK139577 - MK139649).

618   Requests for additional materials should be addressed to R.F. (e-mail:
619   rfregel@ull.edu.es).

620

621   **ACKNOWLEDGMENTS:**

622

REFERENCES

1. Farrujia de la Rosa J (2014) An archaeology of the margins: colonialism, amazighity and heritage management in the Canary Islands. New York: Springer.

2. Navarro JF (2005) Todo sobre Canarias: Los aborigenes. Santa Cruz de Tenerife: Centro de la Cultura Popular Canaria.

3. Velasco-Vázquez J (2015) Más allá del horizonte: Una "perspectiva humana" del poblamiento de Canarias. In: Farrujía de la Rosa J, editor. Orígenes Enfoques interdisciplinares sobre el poblamiento indígena de Canarias. Santa Cruz de Tenerife: Thesaurus Arqueología, Ediciones Idea. pp. 23-89.

4. Rumeu de Armas A (2006) La conquista de Tenerife (1494-1496). La Laguna: Instituto de Estudios Canarios.

5. Crosby AW (1988) Imperialismo ecologico. La expansion biologica de Europa 900-1900. Barcelona: Ed. Critica.

6. Suarez JJ, Rodriguez F, Quintero CL (1988) Historia popular de Canarias. Vol. 2. Conquista y Colonizacion: Centro de la Cultura Popular Canaria.

7. Navarro JF (1997) Arqueologia de las Islas Canarias. Espacio, tiempo y forma Serie I, Prehistoria y Arqueologia 10: 447-478.

8. Tejera-Gaspar A (2006) Canarias y el África Antigua. San Cristóbal de La Laguna: Centro de la Cultura Popular de Canarias.

9. Springer R (2014) Die libysch-berberischen Inschriften der Kanarischen Inseln in ihrem Felsbildkontext; Stroomer H, editor. Koln: Rudiger Koppe Verlag.

10. Tejera-Gaspar A, Perera-Betancort MA (2011) Las supuestas inscripciones púnicas y neopúnicas de las Islas Canarias. Spal Revista de Prehistoria y Arqueología 20: 175-184.

11. Navarro-Mederos J (1999) «El viaje de las loceras»: la transmisión de tradiciones cerámicas prehistóricas e históricas de África a Canarias y su reproducción en las islas. Anuario de Estudios Atlánticos 45: 61-118.

12. Morales J, Rodriguez-Rodriguez A, González-Marrero MC, Martín-Rodríguez E, Henríquez-Valido P, del Pino M (2014) The archaeobotany of long-term crop storage in northwest African communal granaries: a case study from pre-Hispanic

664   Gran Canaria (cal. ad 1000–1500). Vegetation history and archaeobotany 23: 789-
665   804.

666 13. Hagenblad J, Morales J, Leino M, Rodriguez-Rodriguez A (2017) Farmer fidelity in
667   the Canary Islands revealed by ancient DNA from prehistoric seeds. Journal of
668   Archaeological Science 78: 78 - 87.

669 14. Ferrando A, Manunza A, Jordana J, Capote J, Pons A, Pais J, et al. (2015) A
670   mitochondrial analysis reveals distinct founder effect signatures in Canarian and
671   Balearic goats. Anim Genet 46: 452-456.

672 15. Olalde I, Capote J, Del-Arco MC, Atoche P, Delgado T, Gonzalez-Anton R, et al.
673   (2015) Ancient DNA sheds light on the ancestry of pre-hispanic Canarian pigs.
674   Genet Sel Evol 47: 40.

675 16. Guatelli-Steinberg D, Irish JD, Lukacs JR (2001) Canary islands-north African
676   population affinities: measures of divergence based on dental morphology. Homo
677   52: 173-188.

678 17. Bermudez De Castro JM (1989) The Carabelli trait in human prehistoric populations
679   of the Canary Islands. Hum Biol 61: 117-131.

680 18. Bermudez de Castro JM (1989) Third molar agenesis in human prehistoric
681   populations of the Canary Islands. Am J Phys Anthropol 79: 207-215.

682 19. Billy G (1982) Le peuplement prehistorique de l'Archipel Canarien. El Museo
683   Canario 41: 59-74.

684 20. Santana-Cabrera J, Velasco-Vázquez J, Rodríguez-Rodríguez A (2012) Patrón
685   cotidiano de actividad física y organización social del trabajo en la Gran Canaria
686   prehispánica (siglos XI-XV): la aportación de los marcadores óseos de actividad
687   física. Tabona: Revista de Prehistoria y de Arqueología 19: 125-163.

688 21. Atoche P, Ramirez-Rodriguez MA (2011) Nuevas dataciones radiocarbonicas para la
689   Protohistoria canaria: el yacimiento de Buenavista (Lanzarote). Anuario de
690   Estudios Atlanticos 57: 139-170.

691 22. Arco-Aguilar MC, Arco-Aguilar MM, Benito-Mateo C, Rosario-Adrian MC (2017)
692   Un taller romano de púrpura en los limites de la Ecumene: Lobos 1. Santa Cruz de
693   Tenerife: Cabildo de Tenerife.

694   23. Bronk-Ramsay C (2009) Bayesian analysis of radiocarbon dates. Radiocarbon 51:
695   337-360.

696   24. Morales-Mateos J, Rodríguez-Rodríguez A, Herbig A (2017) Agricultura y
697   recolección vegetal en la arqueología prehispánica de las Islas Canarias (siglos
698   XI-XV dC): la contribución de los estudios carpológicos. In: Fernández-Eraso J,
699   Mujika-Alustiza JA, Arrizabalaga-Valbuena A, editors. Miscelánea en homenaje
700   a Lydia Zapata Peña (1965-2015). Vitoria: Universidad del País Vasco/Euskal
701   Herriko Unibertsitatea. pp. 189-218.

702   25. Hernández-Marrero J-C, Navarro-Mederos J-F, Trujillo-Mora J-M, Cancel S,
703   Machado-Yanes C, Pais J, et al. (2016) An approach to prehistoric shepherding in
704   La Gomera (Canary Islands) through the study of domestic spaces. Quaternary
705   International 414: 337-349.

706   26. Kivisild T (2015) Maternal ancestry and population history from whole mitochondrial
707   genomes. Investig Genet 6: 3.

708   27. Rando JC, Cabrera VM, Larruga JM, Hernandez M, Gonzalez AM, Pinto F, et al.
709   (1999) Phylogeographic patterns of mtDNA reflecting the colonization of the
710   Canary Islands. Ann Hum Genet 63: 413-428.

711   28. Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J (2001) High-
712   resolution analysis of human Y-chromosome variation shows a sharp
713   discontinuity and limited gene flow between northwestern Africa and the Iberian
714   Peninsula. Am J Hum Genet 68: 1019-1029.

715   29. Rando JC, Pinto F, Gonzalez AM, Hernandez M, Larruga JM, Cabrera VM, et al.
716   (1998) Mitochondrial DNA analysis of northwest African populations reveals
717   genetic exchanges with European, near-eastern, and sub-Saharan populations.
718   Ann Hum Genet 62: 531-550.

719   30. Secher B, Fregel R, Larruga JM, Cabrera VM, Endicott P, Pestano JJ, et al. (2014)
720   The history of the North African mitochondrial DNA haplogroup U6 gene flow
721   into the African, Eurasian and American continents. BMC evolutionary biology
722   14: 109.

723   31. Santos C, Fregel R, Cabrera VM, Gonzalez AM, Larruga JM, Lima M (2010)
724   Mitochondrial DNA patterns in the Macaronesia islands: Variation within and
725   among archipelagos. Am J Phys Anthropol 141: 610-619.

726   32. Santana J, Fregel R, Lightfoot E, Morales J, Alamon M, Guillen J, et al. (2016) The
727   early colonial atlantic world: New insights on the African Diaspora from isotopic
728   and ancient DNA analyses of a multiethnic 15th-17th century burial population
729   from the Canary Islands, Spain. Am J Phys Anthropol 159: 300-312.

730   33. Fregel R (2010) La evolucion genetica de las poblaciones humanas canarias:
731   determinacion mediante marcadores polimorficos y uniparentales. La Laguna:
732   Servicio de Publicaciones de la Universidad de La Laguna.

733   34. Farrujía de la Rosa AJ, Pichler W, Rodrigue A, Garcia-Marin S (2010) The Libyco–
734   Berber and Latino–Canarian Scripts and the Colonization of the Canary Islands.
735   African Archaeology Review 27: 13-41.

736   35. Flores C, Maca-Meyer N, Perez JA, Gonzalez AM, Larruga JM, Cabrera VM (2003)
737   A predominant European ancestry of paternal lineages from Canary Islanders.
738   Ann Hum Genet 67: 138-152.

739   36. Maca-Meyer N, Arnay M, Rando JC, Flores C, Gonzalez AM, Cabrera VM, et al.
740   (2004) Ancient mtDNA analysis and the origin of the Guanches. Eur J Hum
741   Genet 12: 155-162.

742   37. Fregel R, Pestano J, Arnay M, Cabrera VM, Larruga JM, Gonzalez AM (2009) The
743   maternal aborigine colonization of La Palma (Canary Islands). Eur J Hum Genet
744   17: 1314-1324.

745   38. Fregel R, Cabrera VM, Larruga JM, Hernandez JC, Gamez A, Pestano JJ, et al.
746   (2015) Isolation and prominent aboriginal maternal legacy in the present-day
747   population of La Gomera (Canary Islands). Eur J Hum Genet 23: 1236-1243.

748   39. Ordóñez AC, Fregel R, Trujillo-Mederos A, Hervella M, de-la-Rúa C, Arnay-de-la-
749   Rosa M (2017) Genetic studies on the prehispanic population buried in Punta
750   Azul cave (El Hierro, Canary Islands). Journal of Archaeological Science 78: 20-
751   28.

752   40. O'Rourke DH, Hayes MG, Carlyle SW (2000) Ancient DNA studies in physical
753   anthropology. Annual Review of Anthropology 29: 217-242.

754   41. Rodriguez-Varela R, Gunther T, Krzewinska M, Stora J, Gillingwater TH,
755   MacCallum M, et al. (2017) Genomic Analyses of Pre-European Conquest
756   Human Remains from the Canary Islands Reveal Close Affinity to Modern North
757   Africans. Curr Biol 27: 3396-3402 e3395.

758   42. Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, Sikora M,
759   et al. (2013) Pulling out the 1%: Whole-Genome Capture for the Targeted
760   Enrichment of Ancient DNA Sequencing Libraries. American Journal of Human
761   Genetics 93: 852-864.

762   43. Arnay-De-La-Rosa M, González-Reimers E, Yanes Y, Romanek CS, Noakes JE,
763   Galindo-Martín L (2011) Paleonutritional and paleodietary survey on prehistoric
764   humans from Las Cañadas del Teide (Tenerife, Canary Islands) based on
765   chemical and histological analysis of bone. Journal of Archaeological Science 38:
766   884-895.

767   44. Espinosa A (1980) Del origen y milagros de la Santa Imagen de Nuestra Señora de
768   Candelaria que apareció en la Isla de Tenerife con la descripción de esta Isla.
769   Santa Cruz de Tenerife: Ediciones Goya.

770   45. Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, et al.
771   (2015) Optimal Ancient DNA Yields from the Inner Ear Part of the Human
772   Petrous Bone. PLoS One 10: e0129102.

773   46. Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, et al. (2013)
774   Complete mitochondrial genome sequence of a Middle Pleistocene cave bear
775   reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci U S A 110:
776   15758-15763.

777   47. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly
778   multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010: pdb
779   prot5448.

780   48. Fregel R, Almeida M, Betancor E, Suarez NM, Pestano J (2011) Reliable nuclear and
781   mitochondrial DNA quantification for low copy number and degraded forensic
782   samples. Forensic Science International: Genetics Supplement Series 3: e303-
783   e304.

784   49. Gonzalez AM, Cabrera VM, Larruga JM, Tounkara A, Noumsi G, Thomas BN, et al.
785   (2006) Mitochondrial DNA variation in Mauritania and Mali and their genetic
786   relationship to other Western Africa populations. Ann Hum Genet 70: 631-657.

787   50. Lindgreen S (2012) AdapterRemoval: easy cleaning of next-generation sequencing
788   reads. BMC Res Notes 5: 337.

789   51. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler
790   transform. Bioinformatics 25: 1754-1760.

791   52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. (2009) The
792   Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.

793   53. Ginolhac A, Rasmussen M, Gilbert MT, Willerslev E, Orlando L (2011)
794   mapDamage: testing for damage patterns in ancient DNA sequences.
795   Bioinformatics 27: 2153-2155.

796   54. R_Development_Core_Team (2008) A language and environment for statistical
797   computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-
798   900051-07-0, URL http://www.R-project.org.

799   55. Fu Q, Mittnik A, Johnson PL, Bos K, Lari M, Bollongino R, et al. (2013) A revised
800   timescale for human evolution based on ancient mitochondrial genomes. Current
801   biology : CB 23: 553-559.

802   56. Renaud G, Slon V, Duggan AT, Kelso J (2015) Schmutzi: estimation of
803   contamination and endogenous mitochondrial consensus calling for ancient DNA.
804   Genome Biol 16: 224.

805   57. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N
806   (1999) Reanalysis and revision of the Cambridge reference sequence for human
807   mitochondrial DNA. Nat Genet 23: 147.

808   58. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. (2011) A
809   framework for variation discovery and genotyping using next-generation DNA
810   sequencing data. Nat Genet 43: 491-498.

811   59. Kloss-Brandstatter A, Pacher D, Schonherr S, Weissensteiner H, Binna R, Specht G,
812   et al. (2011) HaploGrep: a fast and reliable algorithm for automatic classification
813   of mitochondrial DNA haplogroups. Hum Mutat 32: 25-32.

28

814   60. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global
815   human mitochondrial DNA variation. Hum Mutat 30: E386-394.
816   61. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, et al. (2013) Using
817   Tablet for visual exploration of second-generation sequencing data. Briefings in
818   Bioinformatics 14: 193-202.
819   62. Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring
820   intraspecific phylogenies. Mol Biol Evol 16: 37-48.
821   63. Forster P, Harding R, Torroni A, Bandelt HJ (1996) Origin and evolution of Native
822   American mtDNA variation: a reappraisal. Am J Hum Genet 59: 935-945.
823   64. Soares P, Ermini L, Thomson N, Mormina M, Rito T, Rohl A, et al. (2009)
824   Correcting for purifying selection: an improved human mitochondrial molecular
825   clock. Am J Hum Genet 84: 740-759.
826   65. Saillard J, Forster P, Lynnerup N, Bandelt HJ, Norby S (2000) mtDNA variation
827   among Greenland Eskimos: the edge of the Beringian expansion. Am J Hum
828   Genet 67: 718-726.
829   66. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and
830   analysis program for Windows 95/98/NT. Nucleic acids symposium series 41: 95-
831   98.
832   67. Fregel R, Delgado S (2011) HaploSearch: a tool for haplotype-sequence two-way
833   transformation. Mitochondrion 11: 366-367.
834   68. Nei N (1987) Molecular evolutionary genetics. New York: Columbia University
835   Press.
836   69. Slatkin M (1995) A measure of population subdivision based on microsatellite allele
837   frequencies. Genetics 139: 457-462.
838   70. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to
839   perform population genetics analyses under Linux and Windows. Mol Ecol
840   Resour 10: 564-567.
841   71. de Leeuw J, Mair P (2009) Multidimensional Scaling Using Majorization: SMACOF
842   in R. Journal of Statistical Software 31: 1-30.
843   72. Long JC (1991) The genetic structure of admixed populations. Genetics 127: 417-
844   428.

845   73. Fregel R, Mendez FL, Bokbot Y, Martin-Socas D, Camalich-Massieu MD, Santana J,
846   et al. (2018) Ancient genomes from North Africa evidence prehistoric migrations
847   to the Maghreb from both the Levant and Europe. Proc Natl Acad Sci U S A in
848   press.

849   74. Heinz T, Cardenas JM, Alvarez-Iglesias V, Pardo-Seco J, Gomez-Carballa A, Santos
850   C, et al. (2015) The Genomic Legacy of the Transatlantic Slave Trade in the
851   Yungas Valley of Bolivia. PLoS One 10: e0134129.

852   75. Bandelt HJ, Herrnstadt C, Yao YG, Kong QP, Kivisild T, Rengo C, et al. (2003)
853   Identification of Native American Founder mtDNAs Through the Analysis of
854   Complete mtDNA Sequences: Some Caveats. Annals of Human Genetics 67:
855   512--524.

856   76. Cabrera-Pérez J, Perera-Betancort MA, Tejera-Gaspar A (1999) Majos La primitiva
857   población de Lanzarote. Islas Canarias. Teguise: Fundación César Manrique.

858   77. Macias-Hernandez AM (1988) Fuentes y principales problemas metodológicos de la
859   demografía histórica de Canarias. Anuario de Estudios Atlánticos 34: 51-157.

860   78. Navarro JF (1992) La Prehistoria de Canarias. Vol. 5. La Gomera y Los Gomeros.
861   Santa Cruz de Tenerife: Centro de la Cultura Popular Canaria.

862   79. Ruiz-Gonzalez MT (2008) El territorio de los antiguos bimbaches / Algunas
863   cuestiones en torno al territorio de Esero. VI Congreso de Patrimonio Histórico:
864   'Arqueología en Canarias, Territorio y Sociedad'. Lanzarote: Cabildo Insular de
865   Lanzarote.

866   80. Jimenez JJ (1992) La Prehistoria de Canarias. Vol. 6. El Hierro y los Bimbaches.
867   Santa Cruz de Tenerife: Centro de la Cultura Popular Canaria.

868   81. Pico B, Aznar E, Corbella D (2003) Le Canarien. Manuscritos, trancripción y
869   traducción, La Laguna. Santa Cruz de Tenerife: Instituto de Estudios Canarios.

870   82. van de Loosdrecht M, Bouzouggar A, Humphrey L, Posth C, Barton N, Aximu-Petri
871   A, et al. (2018) Pleistocene North African genomes link Near Eastern and sub-
872   Saharan African human populations. Science.

873   83. Souville G (1998) Contacts et échanges entre la péninsule Ibérique et le Nord-Ouest
874   de l'Afrique durant les temps préhistoriques et protohistoriques. Comptes rendus
875   des séances de l'Académie des Inscriptions et Belles-Lettres. pp. 163-177.

876   84. Camps G (1974) La civilisations préhistoriques de l'Afrique du Nord et du Sahara.
877   Paris: Doin.

878   85. Clark JD (1997) The Cambridge History of Africa,1. From the earliest times to c.
879   500BC. Cambridge: Cambridge University Press.

880   86. Aubet ME (1994) Tiro y las colonias fenicias de Occidente. Barcelona: Crítica.

881   87. Lopez-Pardo F (1993) La colonización fenicio-púnica en el África noroccidental.
882   Hispania Antiqua 17: 435-449.

883   88. Le Bohec Y (2005) Histoire de l'Afrique romaine (146 avant J.C. - 439 après J. C.).
884   Paris: Picard.

885   89. Camps G (1980) Berberes: aux marges de l´histoire. Paris: Hesperides.

886   90. Tejera-Gaspar A (1989) Parentesco, familia y matrimonio en las culturas
887   prehistóricas de las Islas Canarias. In: Bermejo JC, editor. Parentesco, familia y
888   matrimonio en la historia de Galicia. Santiago de Compostela: Torculo Ediciones.

889   91. Atoche P, Paz JA, Ramirez MA, Ortiz ME (1995) Evidencias arqueologicas del
890   mundo romano en Lanzarote. Arrecife: Servicio de publicaciones del Excmo.
891   Cabildo Insular de Lanzarote.

892   92. Delgado-Delgado JA (2001) Las islas de Juno: ¿hitos de la navegación fenicia en el
893   Atlántico en época arcaica? The Ancient History Bulletin 15: 29-43.

894   93. Delgado-Delgado JA (2012) Canarias en la antigüedad como problema histórico.
895   Tabona: Revista de Prehistoria y de Arqueología 19: 9-13.

896   94. Raven S (2003) Rome in Africa. London: Routledge.

897   95. Gonzalez-Anton R, del Arco-Aguilar MC, Berhmann RB, Bueno-Ramírez P (1998)
898   El poblamiento de un archipiélago Atlántico: Canarias en el proceso colonizador
899   del primer milenio a. C. Eres 8: 43-100.

900   96. González-Antón R, Del Arco-Aguilar MC (2009) Navegaciones exploratorias en
901   Canarias a finales del II milenio AC e inicios del primero: El cordón litoral de La
902   Graciosa (Lanzarote). Canarias Arqueológica: Arqueología-Bioantropología 17:
903   9-80.

904   97. Guerrero-Ayuso VM (2008) Canarias en el contexto de las navegaciones atlánticas de
905   la Antigüedad. SAGVNTVM Papeles del Laboratorio de Arqueología 40: 37-62.

906 98. Matisoo-Smith EA, Gosling AL, Boocock J, Kardailsky O, Kurumilian Y, Roudesli-
907   Chebbi S, et al. (2016) A European Mitochondrial Haplotype Identified in
908   Ancient Phoenician Remains from Carthage, North Africa. PLoS One 11:
909   e0155046.

910 99. Matisoo-Smith E, Gosling AL, Platt D, Kardailsky O, Prost S, Cameron-Christie S, et
911   al. (2018) Ancient mitogenomes of Phoenicians from Sardinia and Lebanon: A
912   story of settlement, integration, and female mobility. PLoS One 13: e0190169.

913 100. Delgado-Darias T, Alberto-Barroso V, Velasco-Vazquez J (2018) Violence in
914   paradise: Cranial trauma in the prehispanic population of Gran Canaria (Canary
915   Islands). Am J Phys Anthropol.

916 101. Delgado T (2009) La Historia en los dientes. Una aproximación a la Prehistoria de
917   Gran Canaria desde la Antropología Dental. Las Palmas: Cabildo Insular de Gran
918   Canaria.

919 102. Rodriguez-Rodriguez A, Morales-Mateos J, Pino-Curbelo MD, Naranjo-Mayor Y,
920   Martín-Rodriguez E, González-Marrero MC (2012) Espacios de producción
921   especializada, excedentes y estratificación social en la Gran Canaria pre-europea.
922   Tabona: Revista de Prehistoria y de Arqueología 19: 101-123.

923 103. Morales J, Rodriguez-Rodriguez A, Marrero A (2013) Prehistoric Plant Use on La
924   Palma Island (Canary Islands, Spain): an example of disappearance of agriculture
925   in an isolated environment. In: Stevens GJ, Nixon S, Murray MA, Fuller D,
926   editors. Archaeology of African Plant Use. California: Institute of Archaeology
927   Publications, UCL. pp. 195-204.

928 104. Moreno-Benítez M (2014) Tiempo y espacio: La cronología de Gran Canaria y el
929   territorio. Boletín electrónico de Patrimonio Histórico 2.

930 105. del Pino-Curbelo M, Rodriguez-Rodriguez A, Buxeda-Garrigos J, Mangas-Viñuela
931   J, González-Quintero P, Moreno-Benítez MA (2016) Las cerámicas aborígenes de
932   Gran Canaria (Islas Canarias) a través del yacimiento de La Cerera: materias
933   primas, tecnología y función. Trabajos de Prehistoria 73: 90-114.

934 106. Santana-Cabrera J, Moreno-Benítez M, Suarez-Medina I, Mendoza-Medina F,
935   Alberto-Barroso V (2018) Zonzamas: un yacimiento singular en la isla de
936   Lanzarote. Nuevos datos arqueológicos. Actas del XXII Coloquio de Historia

937   Canario Americana. Las Palmas de Gran Canaria: Cabildo de Gran Canaria. pp.

938   1-14.

939 107. Cabrera-Pérez JC (1996) La Prehistoria de Fuerteventura. Un modelo insular de

940   adaptación. Las Palmas de Gran Canaria: Cabildo Insular de Gran Canaria y

941   Fuerteventura.

942

943

944 **TABLES:**

945 **Table 1 – MtDNA haplogroup absolute frequencies for the indigenous population of**

946 **the Canary Islands.** Haplogroup frequencies and diversity were calculated using HVRI

947 sequence data from this study and previously published data. 1: This study; 2: Ordoñez et

948 al. 2017; 3: Fregel et al. 2009; 4: Maca-Meyer et al. 2004; 5: Fregel et al. 2014; 6:

949 Rodríguez-Varela et al. 2017.

950

Haplogroup	El Hierro[1,2]	La Palma[1,3]	Tenerife[1,4,6]	La Gomera[1,5]	Gran Canaria[1,6]	Lanzarote & Fuerteventura[1]	Total
H*	12	10	15	2	33	13	85
H1cf	57	8	6	2	1	-	74
H4a1e	-	-	-	-	3	1	4
HV0	-	-	1	-	-	-	1
J	-	3	4	5	2	-	14
K	-	1	1	1	2	-	5
L1/L2	-	2	2	1	1	-	6
L3	-	-	2	4	1	-	7
L3b1a12	-	-	-	-	5	-	5
HV0	-	-	-	-	1	-	1
Other T	-	1	-	-	3	-	4
T2c1	-	3	12	-	15	2	32
U5	-	-	-	-	3	1	4
U6a	-	-	2	-	6	2	10
U6b	-	2	8	38	4	-	52
U6c	-	-	-	-	5	1	6
U7	1	-	-	-	-	-	1
W1e1	-	1	-	-	-	-	1
X3a	-	4	-	4	1	-	9
Sample size	70	35	53	57	86	20	321
Haplogroup diversity	2.86% ± 2.76%	72.10% ± 7.63%	77.43% ± 4.02%	54.20% ± 7.50%	77.10% ± 3.78%	51.05% ± 12.84%	69.69% ± 2.37%

951

34

952 **Table 2 – Admixture results based on mtDNA haplogroup frequencies.** Admixture
953 results for the modern population of the Canary Islands using the three main parental
954 populations: Iberian Peninsula (IBP), sub-Saharan Africa (SSA) and the Canarian
955 indigenous population (CIP). Admixture calculations were performed using two
956 approximations: A) we used the whole ancient dataset (combining the ancient samples
957 from all the seven islands) as CIP for calculating admixture estimates for all islands; B)
958 we used each indigenous sample to calculate the admixture of its respective island (e.g. to
959 calculate admixture in the modern population of Gran Canaria we exclusively used the
960 ancient samples from Gran Canaria as CIP). Results are shown for: the whole Canary
961 Islands population (CAN) and the seven individual islands (FUE=Fuerteventura;
962 GCA=Gran Canaria; GOM=La Gomera; HIE=El Hierro; LAN=Lanzarote; PAL=La
963 Palma; TFE=Tenerife).

964

A: Whole indigenous sample								
Component	**IBP**			**SSA**			**CIP**		
FUE	0.4108	±	0.0071	0.0199	±	0.0015	0.5692	±	0.0070
GCA	0.6486	±	0.0039	0.0438	±	0.0012	0.3076	±	0.0038
GOM	0.2170	±	0.0181	0.0691	±	0.0075	0.7139	±	0.0186
HIE	0.6379	±	0.0122	0.0000	±	0.0000	0.3621	±	0.0122
LAN	0.3303	±	0.0084	0.0448	±	0.0026	0.6248	±	0.0084
PAL	0.5599	±	0.0106	0.0000	±	0.0001	0.4401	±	0.0107
TFE	0.5989	±	0.0041	0.0452	±	0.0012	0.3559	±	0.0040
CAN	0.3978	±	0.0100	0.0432	±	0.0029	0.5589	±	0.0100
B: Indigenous sample from each island								
Component	**IBP**			**SSA**			**CIP**		
FUE	0.6458	±	0.0095	0.0203	±	0.0020	0.3339	±	0.0093
GCA	0.6944	±	0.0040	0.0593	±	0.0015	0.2463	±	0.0038
GOM	0.3768	±	0.0049	0.0682	±	0.0024	0.5550	±	0.0049
HIE	1.0000	±	0.0087	0.0000	±	0.0000	0.0000	±	0.0087
LAN	0.7202	±	0.0116	0.0239	±	0.0027	0.2559	±	0.0113
PAL	0.5896	±	0.0134	0.0000	±	0.0116	0.4104	±	0.0118
TFE	0.7306	±	0.0030	0.0495	±	0.0012	0.2199	±	0.0029

965

966

967

35

968 **TITLES AND LEGENDS TO FIGURES:**

969 **Figure 1 – Map of the Canary Islands showing the geographical location of the**

970 **archaeological sites included in this study.** Codes are as follows: 1 – Cueva del Agua; 2

971 – Huerto de los Morales; 3 – Salto del Casimiro; 4 – El Espigón; 5 – Los Pasitos; 6 –

972 Punta Azul; 7 – Barranco de Majona; 8 – El Pescante; 9 – Antoncojo; 10 – Las Arenas;

973 11 – El Cedro; 12 – El Salitre; 13 – El Portillo; 14 – La Angostura; 15 – El Cascajo; 16 –

974 El Capricho; 17 – El Agujero; 18 – El Hormiguero; 19 – Guayadeque; 20 – La Fortaleza;

975 21 – Cuermeja; 22 – Lomo Galeón; 23 – Puente de la Calzada; 24 – El Huriamen; 25 –

976 Montaña Mina.

977 **Figure 2 – DNA authentication results for all the samples included in this study.** A)

978 Insert size density plot. B) Contamination rates estimated using contamMix and

979 schmutzi. C) Damage patterns.

980 **Figure 3 – MtDNA haplogroup frequencies for ancient and current populations of**

981 **the Canary Islands.**

982 **Figure 4 – MDS plot based on haplogroup frequency distances.** A) MDS analysis

983 comparing the individual ancient populations (FUI=Fuerteventura; GCI=Gran Canaria;

984 GOI=La Gomera; HII=El Hierro; LAI=Lanzarote; PAI=La Palma; TFI=Tenerife), with

985 modern Canarian (codes as in Table 2), Caucasus (CAU), North African (codes as in

986 Table S1), Sub-Saharan African (SSA), European (codes as in Table S1) and Near

987 Eastern populations (codes as in Table S2). B) MDS analysis as in Figure 4A, but

988 removing outliers (HIE, HII and GOM) and pooling all the remaining indigenous samples

989 together (CIP).

990

991

992 **TITLES AND LEGENDS TO SUPPLEMENTARY MATERIAL:**

993 **Table S1 –Populations used for comparisons in this study.**

994 **Table S2 – Summary of mtDNA results for all aDNA samples.**

995 **Table S3 – HVRI data used for this study, including new results on the islands of**

996 **Gran Canaria, Lanzarote and Fuerteventura.**

997 **Figure S1 – Combined calibrated radiocarbon per archaeological site (A) and per**

998 **mtDNA lineage (B).**

999 **Figure S2 – Phylogenetic tree of complete ancient Canarian mtDNA sequences.**

1000 Number along links refers to nucleotide changes, whereas "@", "d" and "i" indicates

1001 back mutations, deletions and insertions, respectively. Recurrent mutations, such as

1002 309iC, 315iC and 16519, have not been taken into account.

1003 **Figure S3 – Phylogenetic tree of complete modern Canary Islands sequences.** The

1004 most probable geographic origin of the sequences is indicated.

1005 **Figure S4 – Phylogenetic tree of complete haplogroup H1cf sequences.** GenBank

1006 accessions and geographic origin are indicated for each complete sequence taken from

1007 the bibliography.

1008 **Figure S5 – Coalescence ages for mtDNA haplogroups observed in the indigenous**

1009 **population of the Canary Islands.** All the coalescence ages have been calculated in this

1010 study, except for H2a, H3 and T2b, whose ages have been obtained from previous results

1011 (Behar et al. 2008).

1012 **Figure S6 – Phylogenetic tree of complete haplogroup H1e1a sequences.** GenBank

1013 accessions and geographic origin are indicated for each complete sequence taken from

1014 the bibliography. Sub-haplogroups in dark grey and white fonts indicate newly defined

1015 branches.

1016 **Figure S7 – Phylogenetic tree of complete haplogroup H4a1 sequences.** GenBank

1017 accessions and geographic origin are indicated for each complete sequence taken from

1018 the bibliography. Sub-haplogroups in dark grey and white fonts indicate newly defined

1019 branches.

1020 **Figure S8 – Phylogenetic tree of complete haplogroup J1c3 sequences.** GenBank

1021 accessions and geographic origin are indicated for each complete sequence taken from

1022 the bibliography.

1023   **Figure S9 – Phylogenetic tree of complete haplogroup J2a2d sequences.** GenBank

1024   accessions and geographic origin are indicated for each complete sequence taken from

1025   the bibliography. Sub-haplogroups in dark grey and white fonts indicate newly defined

1026   branches.

1027   **Figure S10 – Phylogenetic tree of complete haplogroup T2c1d sequences.** GenBank

1028   accessions and geographic origin are indicated for each complete sequence taken from

1029   the bibliography. Sub-haplogroups in dark grey and white fonts indicate newly defined

1030   branches.

1031   **Figure S11 – Phylogenetic tree of complete haplogroup L3b1a sequences.** GenBank

1032   accessions and geographic origin are indicated for each complete sequence taken from

1033   the bibliography. Sub-haplogroups in dark grey and white fonts indicate newly defined

1034   branches.

1035   **Figure S12 – Phylogenetic tree of complete haplogroup X3a sequences.** GenBank

1036   accessions and geographic origin are indicated for each complete sequence taken from
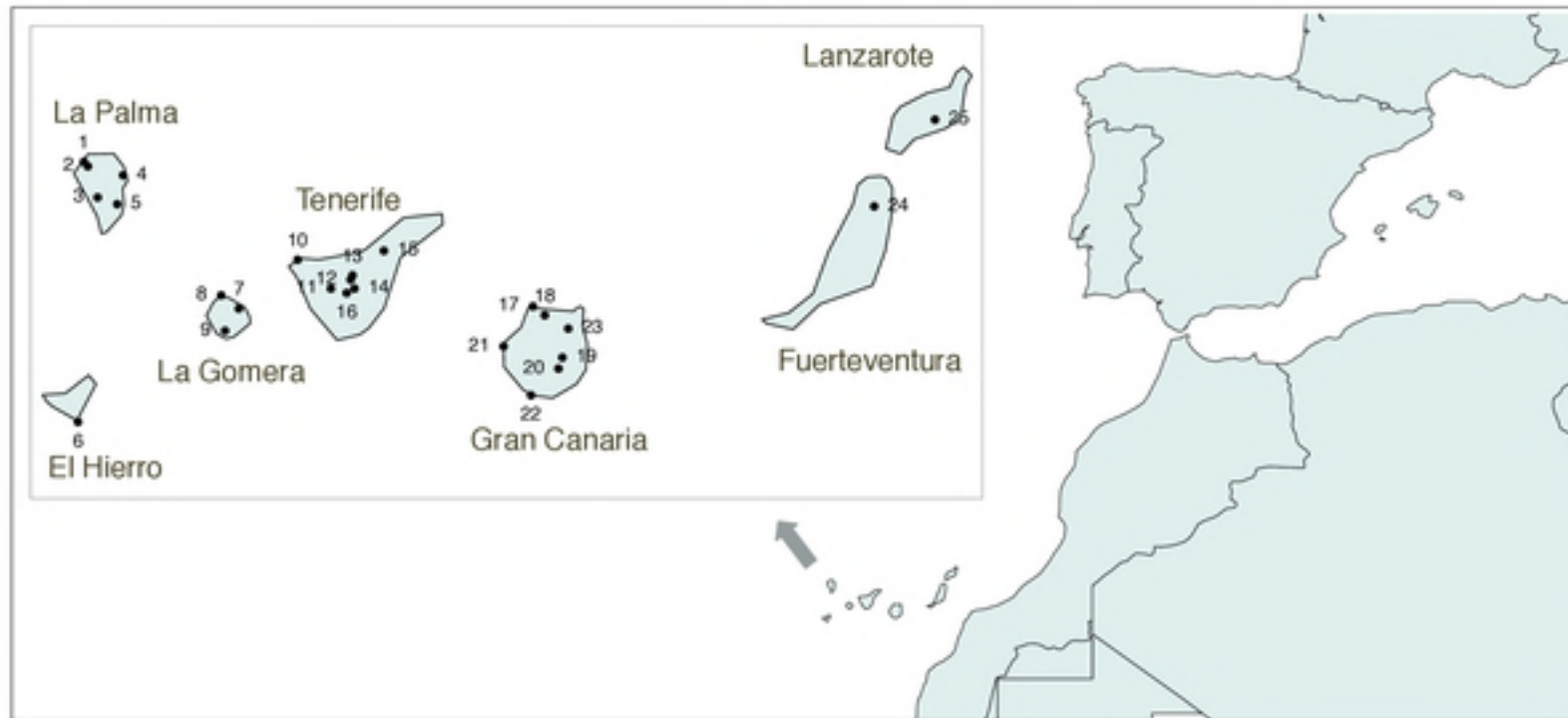
1037   the bibliography.

1038   **Figure S13 – Phylogenetic tree of complete haplogroup U6a1a1 sequences.** GenBank

1039   accessions and geographic origin are indicated for each complete sequence taken from

1040   the bibliography.

1041   **Figure S14 – Phylogenetic tree of complete haplogroup U6a7a1 sequences.** GenBank

1042   accessions and geographic origin are indicated for each complete sequence taken from

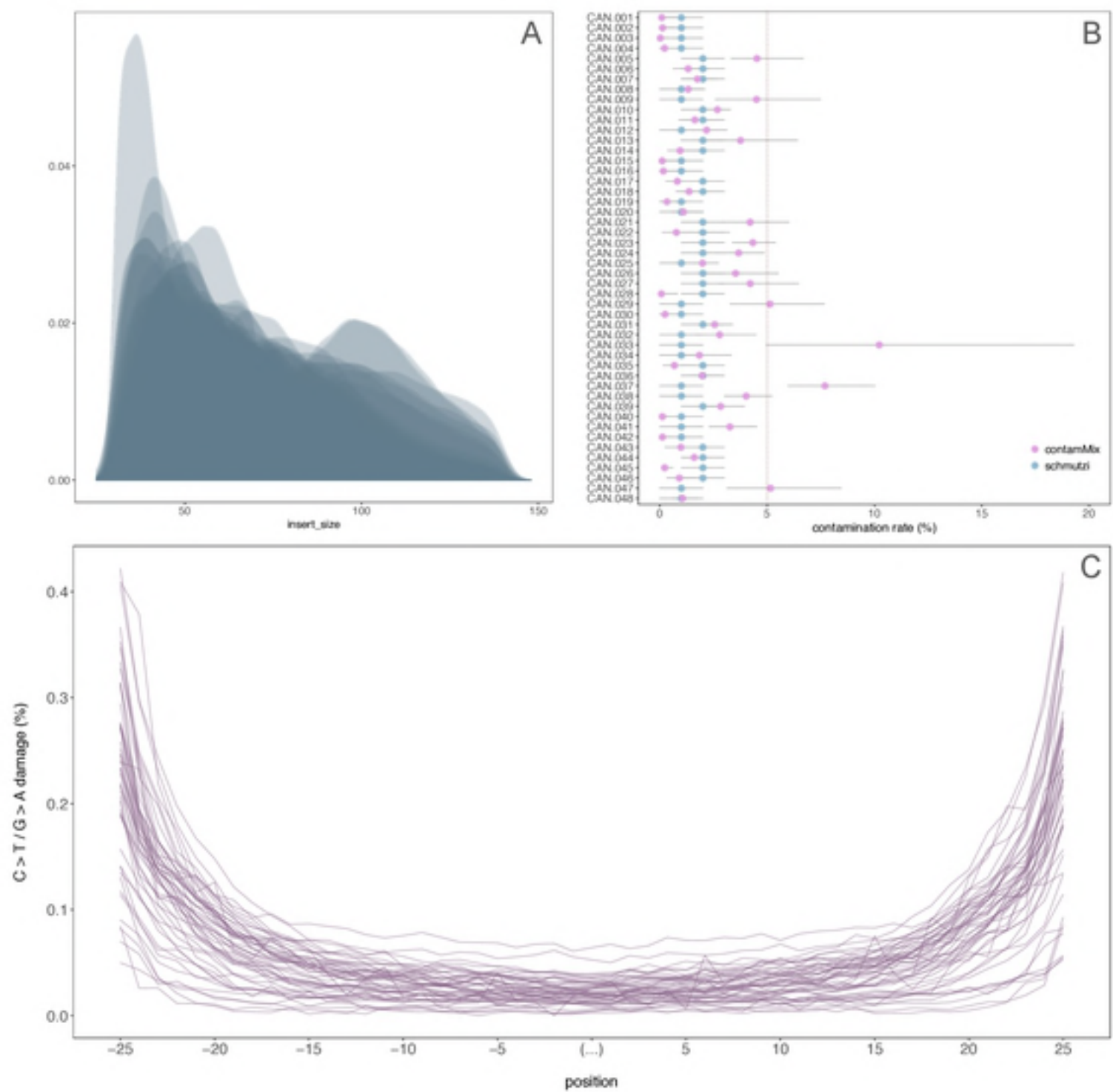1043   the bibliography.

1044   **Figure S15 – Phylogenetic tree of complete haplogroup U6b1a sequences.** GenBank

1045   accessions and geographic origin are indicated for each complete sequence taken from

1046   the bibliography.

1047   **Figure S16 – Phylogenetic tree of complete haplogroup U6c sequences.** GenBank

1048   accessions and geographic origin are indicated for each complete sequence taken from
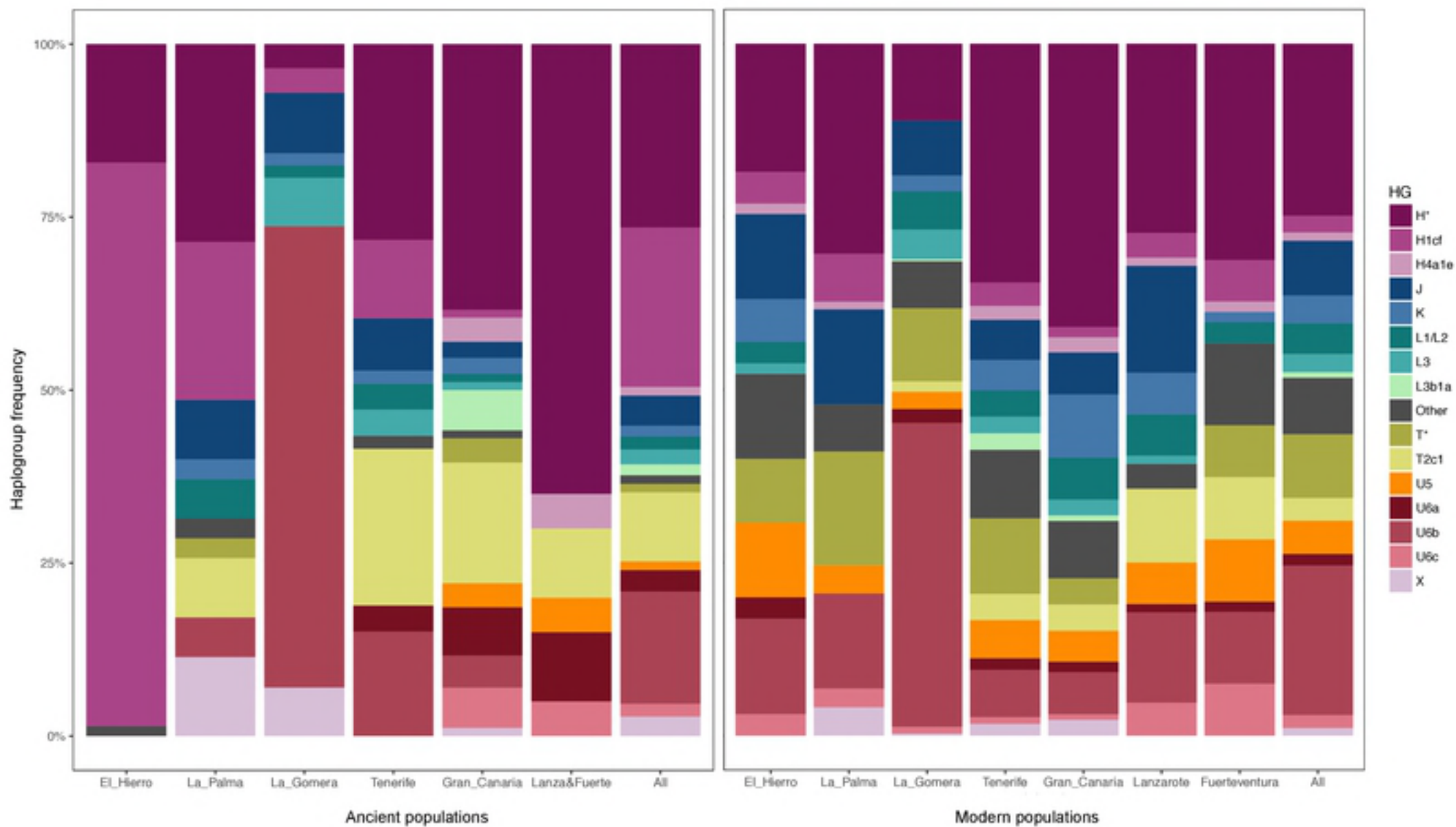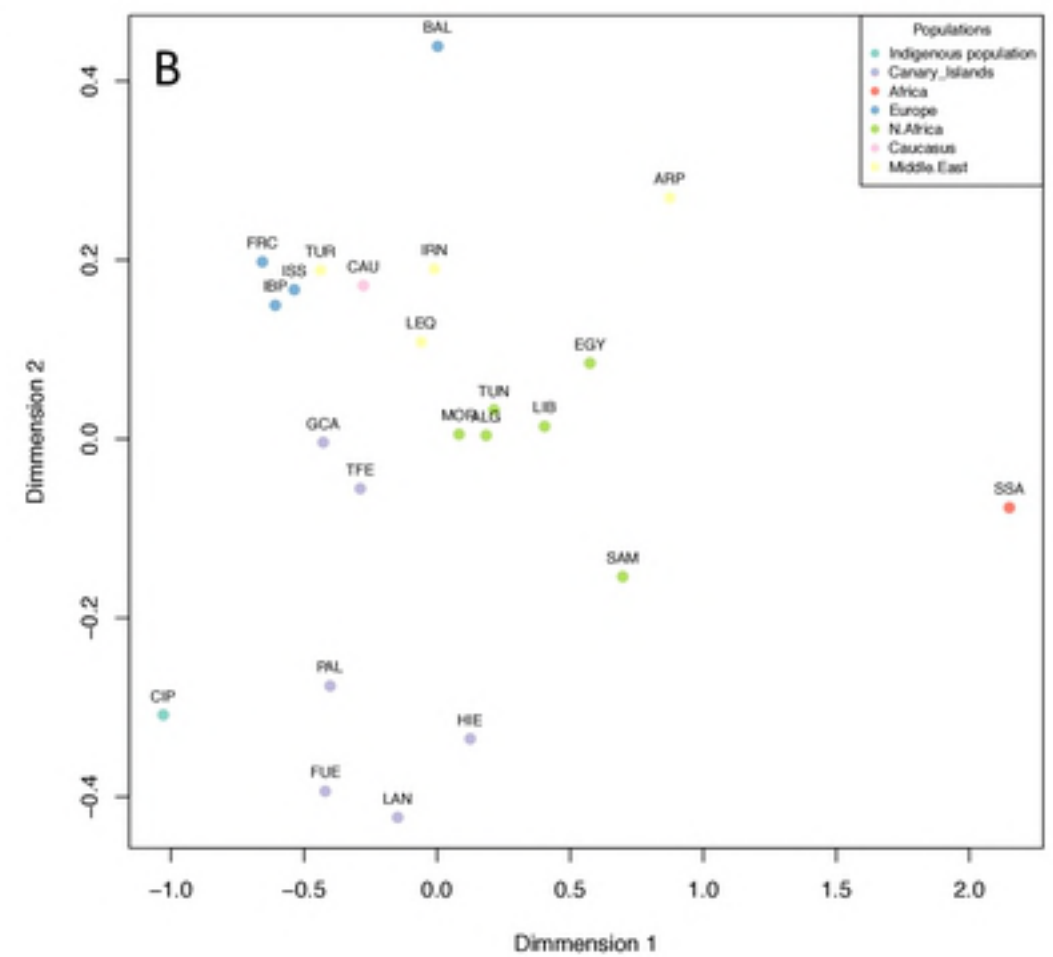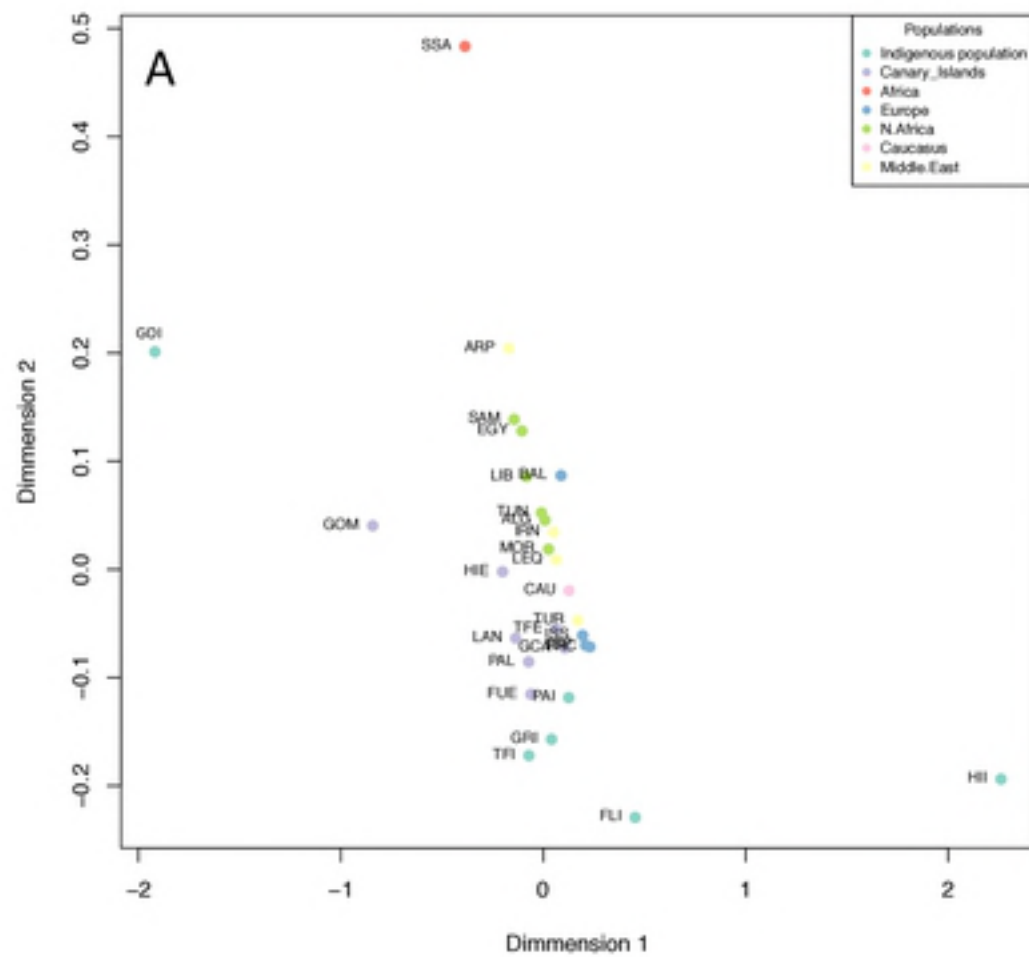
1049   the bibliography.

1050

Figure

Figure

Figure

Figure