Farmer fidelity in the Canary Islands revealed by ancient DNA from prehistoric seeds

Jenny Hagenblad, Jacob Morales, Matti W. Leino and Amelia C. Rodríguez-Rodríguez

Journal Article

N.B.: When citing this work, cite the original article.

Original Publication:
http://dx.doi.org/10.1016/j.jas.2016.12.001
Copyright: Elsevier
http://www.elsevier.com/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-136007
Title: Farmer fidelity in the Canary Islands revealed by ancient DNA from prehistoric seeds

Authors: Jenny Hagenblad\textsuperscript{a}, Jacob Morales\textsuperscript{b}, Matti W Leino\textsuperscript{a,c,d}, Amelia C. Rodriguez-Rodriguezb

Author affiliations:
\textsuperscript{a} Linköping University, IFM Biology, SE-581 83 Linköping, Sweden.
\textsuperscript{b} University of Las Palmas de Gran Canaria, Research Group Tarha, Department of Historical Sciences, Pérez del Toro 1, Las Palmas 35003, Spain.
\textsuperscript{c} Nordiska museet, Swedish Museum of Cultural History, Box 27820, SE-115 93, Stockholm, Sweden.
\textsuperscript{d} The Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, SE-106 91 Stockholm, Sweden

Corresponding author: Jenny Hagenblad, email: Jenny.Hagenblad@liu.se
Abstract

The Canary Islands were settled in the first millennium AD by colonizers likely originating from North Africa. The settlers developed a farming economy with barley as the main crop. Archaeological evidence suggests the islands then remained isolated until European sea-travellers discovered and colonized them during the 14th and 15th centuries. Here we report a population study of ancient DNA from twenty-one archaeobotanical barley grains from Gran Canaria dating from 1050 to 1440 cal AD. The material showed exceptional DNA preservation and genotyping was carried out for 99 single nucleotide markers. In addition 101 extant landrace accessions from the Canary Islands and the western Mediterranean were genotyped. The archaeological material showed high genetic similarity to extant landraces from the Canary Islands. In contrast, accessions from the Canary Islands were highly differentiated from both Iberian and North African mainland barley. Within the Canary Islands, landraces from the easternmost islands were genetically differentiated from landraces from the western islands, corroborating the presence of pre-Hispanic barley cultivation on Lanzarote. The results demonstrate the potential of population genetic analyses of ancient DNA. They support the hypothesis of an original colonization, possibly from present day Morocco, and subsequent isolation of the islands and reveal a farmer fidelity to the local barley that has lasted for centuries.

Keywords: aDNA; archaeobotany; barley (Hordeum vulgare); crop evolution; landrace; Canary Islands
Abbreviations

aDNA: Ancient DNA
PC: Principal component
PCA: Principal component analysis
SNP: Single nucleotide polymorphism
1. Introduction

The Canary Islands is an Atlantic archipelago of volcanic origin consisting of seven major islands, of which Fuerteventura is only some 100 km from mainland Africa (Fig. 1A, B). In spite of being close to the African continent, the islands were not settled until the early first millennium AD (Atoche Peña, 2013, Rodriguez-Rodriguez, et al., 2011). Archaeological remains, epigraphic and linguistic evidence (Onrubia-Pintado, et al., 1995, Springer, 2001) and genetic analyses of present day Canarians and pre-Hispanic remains suggest that the colonizers came from North Africa, but the exact region of origin or the frequency of settlements are yet unknown (Fregel, et al., 2009a, Maca-Meyer, et al., 2003).

Two main hypotheses concerning how the islands were populated have been postulated. The first emphasizes an early Phoenician-Punic enterprise followed by Roman settlements with a step-by-step colonization from the eastern islands to the western ones (Atoche Peña, 2013). The second hypothesis points to repeated migrations of North-African people of Berber origin with contacts among the islands at the beginning of the settlement (Fregel, et al., 2009b). After the first colonization

Fig. 1. Map of the studied area. A) Location of the Canary archipelago. B) Geography of the Canary Islands.
the islands remained isolated from the mainland, and to a large extent also from each other, until European sailors re-discovered the archipelago in the 14th century AD (Morales, et al., 2009). Towards the end of the 15th century AD the Castilian Crown conquered the islands, decimating the local population. The present human population contains a large majority of European ancestry, and to a lesser extent North African and pre-Hispanic ancestry, with a minor genetic contribution from sub-Saharan African (Flores, et al., 2001, Flores, et al., 2003, Pino-Yanes, et al., 2011, Rando, et al., 1999).

When the Europeans rediscovered the islands the indigenous population had developed an economy based on the cultivation of cereals and husbandry of goat, sheep and pig (Morales, et al., 2009). Archaeological findings show that six-row hulled barley (*Hordeum vulgare* ssp. *vulgare*) was cultivated as a major crop in pre-Hispanic times (4th – 15th century AD) on the islands of Gran Canaria, Tenerife, La Gomera and El Hierro (Morales, et al., 2009, Morales, 2010, Morales, et al., 2011, Morales, et al., 2014). On the island of La Palma barley seeds have been found at early sites dated to 4th – 11th century AD, but not in later contexts, and it is thought that people abandoned the practice of agriculture in general before any contact with Europeans (Morales, et al., 2013). There are no archaeological findings suggesting that barley was cultivated on Fuerteventura or Lanzarote, but the first Europeans reported barley being cultivated by the indigenous population on Lanzarote (de la Salle, 1404-19/1980).

On Gran Canaria, complex silo systems, dug out in the volcanic rock, were used for long-term storage of the harvest. The practice of using silos for long-term grain storage was to a large extent abandoned shortly after the Hispanic conquest and the silos have remained unused since then (Morales, et al., 2014). The preservation conditions in the silos were, however, of such a high quality that still today preserved desiccated vegetal remains can be found (Velasco Vázquez, et al., 2001). During surveys of the silos it has been possible to recover plant remains such as chaff, spikes, rachises and in some cases also seeds (Morales, 2010, Morales, et al., 2014). In particular barley grains have been found at high proportions making studies on a population level possible. The seeds recovered have been shown to be very well preserved and genetic analysis has been carried out on wheat (*Triticum durum* /
aestivum) kernels (Oliveira, et al., 2012). Archaeological seeds from the major crop on Gran Canaria, barley, have, however, to date not been investigated genetically.

In this study we have used SNP genotyping to genetically analyse up to 1000 year old barley seeds recovered from silos from the pre-Hispanic era. We have compared the genotypes of populations of pre-Hispanic barley with those of extant Canarian and Western Mediterranean landrace barley to answer three questions: 1) In what ways was barley cultivation affected by the Hispanic colonization of the islands? 2) To what extent was barley cultivated on the different islands in pre-Hispanic times? 3) Can support be found for either a Phoenician – Punic or a North African colonization of the Canary Islands?

2. Materials and methods

2.1 Sampling and genotyping.
In 2011 barley seeds were recovered from pre-Hispanic grain silos at three archaeological sites on Gran Canaria: Guayadeque – Cuevas Muchas (henceforth Guayadeque), Cueva de las Estrellas – Acusa (henceforth Estrellas) and Temisas (Table 1; Fig. 2A-C). All the grains recovered from Guayadeque, Estrellas and Temisas had been preserved by desiccation, and were separated from the sedimentary matrix by dry-sieving with a 2 mm mesh. AMS C¹⁴ dating was carried out at Beta Analytics.

The site of Guayadeque is a complex granary composed by several platforms, chambers, and interconnected tunnels excavated from the rock (Velasco Vázquez, et al., 2001). The samples were dated with a range of ages from 1260 to 1430 cal AD (table 1, Supplementary file S1). The site of Estrellas is a granary located on a platform excavated in the Acusa plateau (Velasco Vázquez, et al., 2001). All the barley grains analyzed in this granary came from a single silo, and a single grain of barley from this silo was radiocarbon dated to 1050-1250 cal AD (table 1, supplementary data S1). Temisas is a large complex of caves carved into the ‘Risco Pintado’ cliff. Seven barley grains were recovered from five silos (table 1, supplementary data S1). Four silos were radiocarbon dated, using barley rachises recovered along the barley grains, to a range of ages from 1050 to 1440 cal AD.
Table 1. Archaeological specimens used for genotyping, their age, DNA concentration and genotyping success rate measured as the percentage of successfully genotyped markers after merging replicate samples. CuEs, Gua and Tem refers to samples from Estrellas, Guayadeque and Temisas respectively. Detailed information is given in S1.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Age (cal AD)</th>
<th>DNA concentration (ng/µl)</th>
<th>Genotyping success rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CuEs-1</td>
<td>1050 – 1250</td>
<td>3.58</td>
<td>61.62</td>
</tr>
<tr>
<td>CuEs-2</td>
<td>1050 – 1250</td>
<td>4.28</td>
<td>85.86</td>
</tr>
<tr>
<td>CuEs-3</td>
<td>1050 – 1250</td>
<td>2.92</td>
<td>70.71</td>
</tr>
<tr>
<td>CuEs-4</td>
<td>1050 – 1250</td>
<td>2.62</td>
<td>78.79</td>
</tr>
<tr>
<td>CuEs-5</td>
<td>1050 – 1250</td>
<td>1.55</td>
<td>40.40</td>
</tr>
<tr>
<td>CuEs-6</td>
<td>1050 – 1250</td>
<td>1.75</td>
<td>76.77</td>
</tr>
<tr>
<td>CuEs-7</td>
<td>1050 – 1250</td>
<td>2.10</td>
<td>37.37</td>
</tr>
<tr>
<td>CuEs-8</td>
<td>1050 – 1250</td>
<td>1.16</td>
<td>79.80</td>
</tr>
<tr>
<td>CuEs-9</td>
<td>1050 – 1250</td>
<td>1.64</td>
<td>39.39</td>
</tr>
<tr>
<td>Gua S3 P3</td>
<td></td>
<td>2.20</td>
<td>92.93</td>
</tr>
<tr>
<td>Gua S5 P3-1</td>
<td>1310 – 1430</td>
<td>3.00</td>
<td>95.96</td>
</tr>
<tr>
<td>Gua S5 P3-2</td>
<td>1310 – 1430</td>
<td>2.70</td>
<td>83.84</td>
</tr>
<tr>
<td>Gua S5 P4</td>
<td>1260 – 1380</td>
<td>7.14</td>
<td>94.95</td>
</tr>
<tr>
<td>Gua S9 P3</td>
<td>1280 – 1400</td>
<td>5.60</td>
<td>95.96</td>
</tr>
<tr>
<td>Tem S1 P2</td>
<td>1290 – 1410</td>
<td>5.44</td>
<td>90.91</td>
</tr>
<tr>
<td>Tem S3 P2</td>
<td>1050 – 1250</td>
<td>2.34</td>
<td>14.14</td>
</tr>
<tr>
<td>Tem S8 P1-1</td>
<td>1320 – 1430</td>
<td>4.20</td>
<td>74.75</td>
</tr>
<tr>
<td>Tem S8 P1-2</td>
<td>1320 – 1430</td>
<td>0.04</td>
<td>14.14</td>
</tr>
<tr>
<td>Tem S12 P1</td>
<td>1330 – 1440</td>
<td>2.79</td>
<td>60.61</td>
</tr>
<tr>
<td>Tem S18 P1-1</td>
<td></td>
<td>5.92</td>
<td>49.49</td>
</tr>
<tr>
<td>Tem S18 P1-2</td>
<td></td>
<td>2.58</td>
<td>17.17</td>
</tr>
</tbody>
</table>

Fig. 2. Origin of archaeological samples. A) Gran Canaria and archaeological excavation sites. B) Guayadeque silo. C) Archaeological barley grains.
DNA from archaeological barley seeds was extracted in a dedicated, chambered, ancient DNA extraction laboratory at Warwick University, UK. No previous work with modern barley or PCR had been performed in the laboratory and suitable precautions were taken to avoid all foreign contaminants. Extraction controls were included in each extraction round. DNA was extracted according to Palmer, et al. (2009) but with incubation for 6 days and with the Amicon® concentration step excluded. Extracted DNA was quantified using a Qubit® Fluorometer (Invitrogen) and a subset of samples tested for fragment length distribution using a BioAnalyzer (Agilent 2100).

Extant landrace barley from 101 accessions to use as reference material was obtained from the genebanks IPK (Gaterleben, Germany), INIA (Madrid, Spain), CCBAT (Teneriffe, Spain), CAP (La Palma, Spain), GRIN (US) and INRA (Clermond-Ferrand, France) and from the private collections of Jaime Gil CIG (La Gomera, Spain) and GIL (Lanzarote, Spain). In addition three historical accessions, collected prior to 1877, were obtained from the collections of the Swedish museum of cultural history (Supplementary file S2).

DNA from the extant reference material was extracted from six (in some cases five) individuals of each accession using the DNeasy 96 Plant Kit or the DNeasy Plant Mini Kit from Qiagen. DNA from historical samples was extracted according to (Leino, et al., 2009). All reference DNA extractions were carried out at Linköping University, Sweden.

2.2 Genotyping
In total 640 individual barley grains were genotyped in two separate batches. Genotyping of 99 single nucleotide polymorphism (SNP) markers was carried out by LGC Genomics, using the KASP assay method (He, et al., 2014, Semagn, et al., 2014). During genotyping oligos are extended from two competing allele-specific primers and amplification is detected from fluorescence resonance energy transfer (FRET). Genotyping is carried out on genomic DNA extracts with no further preparation (e.g. library construction) needed. Amplicons shorter than 50 bp can be readily genotyped and amplification and genotyping is carried out in single-step sealed reactions.
avoiding risks of contamination and making the method suitable for analysis of ancient DNA (aDNA) (Lister, et al., 2013).

92 of the SNPs were derived from the BOPA1 array (Kota, et al., 2008). The SNPs were chosen based on their ability to provide suitable assays for aDNA, i.e. short amplicons. The original array was optimized for European cultivars, which means our markers most likely underestimate the genetic diversity of the landraces studied here compared to the diversity in cultivars. Our SNP selection should, however, not have lead to additional ascertainment bias.

The remaining seven markers were located in causative or associated SNPs of the functional genes $Vrs1$ (positions A40>F.S.; F75>L; E152>F.S.) (Komatsuda, et al., 2007), $int-c$ (nucleotide 124) (Ramsay, et al., 2011), $Ppd-H1$ (SNP48) (Jones, et al., 2008), $HvNAM-2$ (nucleotide 798) (Cai, et al., 2013) and $Lhcb1$ (nucleotide 907) (Xia, et al., 2012). All archaeological samples were genotyped in duplicates or triplicates (referred to as repeats below) alongside their extraction controls.

2.3 Data cleaning
SNPs only working in one of the two genotyping batches and loci with more than 30\% missing data in either batch (in total 10 SNPs) were removed from analysis. The SNP $Vrs1\_E152>F.S.$, scoring the presence of an indel, was only used for predicting the phenotype of the archaeological specimens. To facilitate downstream analysis all individuals were considered as homozygous with heterozygous genotypes (157 out of 55847 successfully genotyped markers) treated as missing data. Due to the low frequency of heterozygous loci, less than 0.3\% their replacement with missing data should have negligible effect on the results of our analyses. Samples with a low success rate, i.e. extant and historical samples with more than 20\% missing data and archaeological samples with more than 50\% missing data, were removed from further analysis. Accessions with three or less individuals remaining ($N = 2$) and accessions that were invariant or nearly invariant ($h < 0.01$, $N = 24$, Supplementary file S2) were excluded from PC and $STRUCTURE$ analyses. After data cleaning, the final dataset, used for analyses of geographic structure, consisted of genotypes from 467 individual grains (15 archaeological, 12 historical and 440 extant individuals respectively) from
two historical and 76 extant accessions and three archaeological sites, genotyped for 88 SNPs.

2.4 Data analysis.
Genetic diversity, estimated as Nei’s h (Nei, 1973), and Wright’s FST (Wright, 1951) were calculated using purpose-written perl scripts. Geographic structuring of genetic diversity was assessed with principal component analysis (PCA) and the Baysian clustering algorithm implemented in the software STRUCTURE (v 2.3.4) (Falush, et al., 2003, Pritchard, et al., 2000). STRUCTURE was run using the haploid setting, as suggested for structure clustering for predominantly self-fertilizing species by Nordborg, et al. (2005), and a model with correlated allele frequencies and admixture. Although STRUCTURE works to minimize deviations from Hardy-Weinberg equilibrium, an assumption that is violated in a selfing organism such as barley, its detection of genetic structuring in selfing organisms with a haploid setting is highly similar to that of other methods not assuming Hardy Weinberg equilibrium (Forsberg, et al., 2015). Non-random mating has been shown to bias STRUCTURE towards spurious signals of admixture (Gao, et al., 2007) and the analysis should be robust towards the detection of general patterns with little admixture.

STRUCTURE was run with a burn-in length of 20 000 iterations followed by 50 000 iterations for estimating the parameters, with 10 repeated runs at each level of predetermined clusters (K) ranging from 1 to 15. The software CLUMPP (v 1.1.2) (Jakobsson and Rosenberg, 2007) was used to compare the outcome of individual runs with the Greedy algorithm for $4 < K < 6$ and with the LargeKGreedy algorithm for $K \geq 6$. The number of clusters best describing the data was evaluated from the CLUMPP H’ values and ΔK calculated according to (Evanno, et al., 2005). Results were visualized using DISTRUCT (v 1.1) (Rosenberg, 2004). In the PCA, data was analyzed on the accession level where the numbers of copies of each allele at each locus were treated as independent variables using the prcomp command. To investigate the effect of missing data individual PCAs were merged to a single one using Procrustes transformation according to Skoglund, et al. (2012).
3. Results

3.1 Genotyping of archaeological specimens

Radiocarbon dating (AMS) on cereal remains from the three archaeological sites indicated that the age of the seeds ranged from 1050 to 1440 cal AD (Table 1, Supplementary file S1), i.e. that they were most likely from the pre-Hispanic era. There are documents mentioning exchange of food, including seeds, between Europeans and the indigenous people as early as 1402. However, these texts report local barley being exchanged for European items such as metal tools (Onrubia-Pintado and González-Marrero, 2004). Hispanic agriculture on Gran Canaria was focused on the cultivation of sugar cane and to a lesser extent of grapevine, and it was fully developed only after the conquest of the island in 1483 (Morales, 2010). It is thus unlikely that European barley was introduced on Gran Canaria during the first contacts in the early 15 century AD.

DNA extracts contained between 1.5 and 7 ng/μl DNA (Supplementary file S1). Fragment length distribution was analyzed in a sub-set of samples and was found to peak at 50 bp, but the extracts also contained fragments up to 150 bp in length. Repeated successful genotyping of DNA extracted from 21 individual prehistoric seeds showed a high level of consistency with less than 1 % of successfully genotyped SNPs providing different genotypes in the repeats. No amplification was detected in the extraction controls included in the genotyping. We therefore concluded that our genotyping of the archaeological specimens was accurate and reliable.

Genotyping results for repeats were merged in order to obtain genotypes from the maximum possible number of markers. After merging, genotyping success rate ranged from 13 – 96 % with the highest success rate for specimens from the Guayadeque site (Table 1, Supplementary file S1). Genotyping success rate was positively correlated with DNA concentration (c = 0.498, p < 0.05). Genotyping known causative mutations in the Vrs1 and Int-c genes confirmed that the archaeological specimens had the vrs1.a1 and Int-c.a alleles consistent with a six-row phenotype. At the PpdHI locus all samples had the responsive allele resulting in early flowering at increased day-length while they at the HvNAM-2 and Lhcb1 loci carried the markers associated with high grain protein content and high number of grains per spike, respectively.
3.2 Past and present barley in the Canary Islands

We investigated the distribution of genetic diversity of present day barley in the Canary Islands by analysing 36 landraces accessions from the seven major islands (Supplementary file S2). PCA showed that four of the accessions (BGE007573, BGE007565, BGE007566 and HOR11187) were highly divergent from the remaining accessions (Supplementary file S3). Removing these revealed a genetic structuring that was partially correlated with geography. The first principal component (PC) primarily separated accessions from the two easternmost islands, Lanzarote and Fuerteventura, from accessions from western islands (Fig. 3A). Among the western island accession separation was primarily shown along the second PC with Tenerife and La Gomera accessions clustering apart from Gran Canaria accessions and accessions from El Hierro, while accessions from La Palma were located in between accessions from Tenerife and Gran Canaria (Fig. 3A).
Fig. 3. Results of PCA. A) PCA of extant Canary barley after removal of divergent accessions. B) PCA of extant and archaeological Canary barley using only the most successful markers. C) PCA of Canary and mainland barley.
The software *STRUCTURE* was used to identify genetic diversity based clustering among the data. In analysis of both extant and archaeological barley (with seeds from each archaeological site treated as a single population) both $\Delta K$ and $H'$ values suggested two clusters ($K = 2$) best described the data, but with almost equally strong support for $K = 3$ (Supplementary file S4). The first two clusters primarily separated Lanzarote and Fuerteventura accessions from accessions from the remaining islands but with no separation with respect to age (Supplementary file S5). At $K = 3$ the accessions from the western islands were split into one group clustering some of the accessions from Tenerife and La Gomera and one with accessions from Gran Canaria, El Hierro and La Palma, with archaeological specimen in both groups (Fig. 4). Higher levels of $K$ did not result in further population structuring. The results of PCA including both archaeological and extant samples showed strong concordance with the results of the *STRUCTURE* analysis, with the first two PCs separating accessions similarly as *STRUCTURE*. Archaeological remains from Estrellas and Temisas clustered separately from all other accessions along the second PC (Supplementary file S6). Procrustes transformation confirmed that this was a consequence of the lower success rate in the samples from Estrellas and Temisas. Procrustes transformation did not affect the general clustering of the extant accessions, but moved Guayadeque to the cluster containing extant accessions primarily from Lanzarote and Fuerteventura (Supplementary file S7). In a PCA based on solely the SNPs amplifying in all archaeological samples (with a success rate higher than 50 %) archaeological samples were located among the extant samples with the highest similarity to accessions from the island of La Palma (Fig. 3B).
Fig. 4. Results of STRUCTURE analysis of extant and archaeological Canary barley at K = 3. Each vertical line represent data from a single individual where different colours decode the proportion of identity of that individual to each of the three clusters explained by the investigated model.

FST-values between the archaeological sites (with each site treated as a single population) were low and non-significant (Supplementary file S8). FST values obtained between pairs of archaeological sites were of a similar magnitude as FST values between pairs of extant landraces from the same island (t-test, p > 0.05 for all comparisons). This remained true also when the divergent accessions (BGE007573, BGE007565, BGE007566 and HOR11187) were removed. Looking at all the Canarian pairwise FST comparisons, the archaeological specimens showed greatest similarity and had the lowest FST values when compared with some of the accessions from Lanzarote, La Palma and Tenerife (Supplementary file S8).

3.3 The relationship between Canarian and mainland barley

We compared the distribution of genetic diversity of the barley from the Canary Islands with that of landrace barley from the Mediterranean area. In particular we chose accessions from Algeria, Morocco and Tunisia, areas where human genetic
studies suggest similarities to the pre-Hispanic Canarian population (Fregel, et al., 2009a, González Antón and Tejera Gaspar, 1981, Maca-Meyer, et al., 2003, Pino-Yanes, et al., 2011, Springer, 2001). In the STRUCTURE analysis both ΔK and H’ strongly supported K = 2 as the level of structuring best describing the data (Supplementary file S4). At this level the accessions clearly grouped into a Canarian cluster and a mainland cluster (Fig. 5). The Canarian cluster contained all archaeological samples and all extant accessions from the different Canary Islands with the exception of BGE007573, BGE007565 and BGE007566. The accession HOR11187 showed mixed clustering. No accession with a non-Canarian origin clustered with the Canarian accessions.

Fig. 5. Results of STRUCTURE analysis of Canary and mainland barley at K = 2.

The PCA confirmed the results of the STRUCTURE analysis with separation between the Canarian and non-Canarian accessions along the first PC (Fig. 3C). The spread of accessions from the same country also showed a lower diversity among Canarian accessions than among the accessions from Algeria, Morocco and Tunisia, the countries from which data from more than two accessions had been obtained. The higher genetic diversity present among the latter countries was further shown by the total within country genetic diversity (Canary Islands: h = 0.175, Algeria: h = 0.234, Morocco: h = 0.236 and Tunisia: h = 0.248 respectively). In the PCA the accession located most closely to the Canarian accessions was HOR874 from France, which also showed some indications of clustering with the Canarian accessions in the STRUCTURE analysis. This accession had, however, fairly low genetic diversity, which may have affected our ability to correctly assign it to a cluster.
F_{ST} values between all pairs of accessions were calculated to deduce which mainland accessions showed the highest similarity to the Canarian accessions (Supplementary file S8). F_{ST} values in Canarian – mainland accession comparisons (excluding the divergent Canarian accessions) ranged from 0.175 (CBT02609 vs. HOR 13412) to 0.778 (CBT01409 vs. LPA-382). Among the archaeological specimens the F_{ST} values were lowest when were paired with the Moroccan accessions IG32022, IG32012 and HOR13443 (average F_{ST} values 0.200, 0.201 and 0.203 respectively). These Moroccan accessions also had low F_{ST} values when compared with many of the extant Canarian accessions (Supplementary file S8). STRUCTURE analysis of only archaeological specimens and Moroccan accessions also supported a closer connection between the Canarian accessions and IG32022 and HOR13443, but also IG32055 and IG32066 at the number of K most relevant for exploring Canarian and Moroccan population structure (Supplementary file S9).

4. Discussion

4.1 Genotyping of archaeobotanical remains

The analysis of aDNA, has received massive attention in recent years, (Hagelberg, et al., 2015). The majority of aDNA studies have, however, been performed on animal and human remains and much fewer on archaeobotanical specimens (reviewed in (Brown, et al., 2015, Palmer, et al., 2012). This is beginning to change and recently next generation sequencing methods have been applied to both charred and desiccated seeds (Allaby, et al., 2015, Bunning, et al., 2012, Mascher, et al., 2016, Nistelberger, et al., 2016) and SNP genotyping methods, such as the LGC Genomics’ KASP and the Illumina Golden Gate technology, have been successfully used on more than a hundred years old herbarium material (Forsberg, et al., 2015, Lister, et al., 2013).

Here we show that SNP genotyping methods, in this case KASP, can be used to successfully genotype well-preserved desiccated plant remains approaching 1000 years of age. Even samples with an estimated date of 1050 – 1250 cal AD yielded DNA concentrations of up to 4.28 ng/µl and could be successfully genotyped for more than 85 % of our markers. To our knowledge this is the first time SNP genotyping of this scale has been reported for archaeobotanical remains. We also
show that for traits with well-characterized causative SNPs, such as row type (Komatsuda, et al., 2007) or flowering time (Jones, et al., 2008), the properties of prehistoric crops can be revealed. Genotyping of selected SNPs in archaeobotanical remains allows for genotyping of a large number of samples at a moderate cost and makes population genetic analyses of archaeological samples, such as the one reported here, possible.

Poor DNA quality may increase the risk of allelic dropout, the failure to genotype one or both alleles at a locus. Though this can result in heterozygotes being mistakenly identified as homozygous, such instances should be rare in a selfing organism such as barley. Complete failure to genotype a locus will result in missing data and a lower success rate may affect the outcome of analyses such as STRUCTURE and PCA. However, since our missing data can be assumed to be caused by poor DNA quality, rather than the presence of null alleles, missing data should be independent of allelic status and have limited effect on the outcome of the STRUCTURE analyses. The effect of the more frequent genotype failure in the archaeological material on the result of the PCA is, however, evident from Supplementary file S6. Procrustes transformation had limited overall effect on the general outcome of the PCA (Supplementary file S7). The restricted number of markers may limit our power to draw conclusions and future studies based on larger number of markers may allow a more detailed analysis of the history of Canarian barley.

4.2 Farmer fidelity to indigenous barley
The genetic similarities between the barley cultivated on the Canary Islands today and in pre-Hispanic times were striking. When compared with barley from mainland Africa and Europe, the Canarian barley, regardless of age, clearly formed a distinct genetic cluster both in the STRUCTURE and PC analyses. The $F_{ST}$ values between the archaeological samples from Guayadeque, with the highest genotyping success rate, and several present day accessions from La Palma, Tenerife and Lanzarote were low and non-significant further underlining the continuous cultivation of indigenous barley in the archipelago. Following the Spanish conquest in the late 15th century the human population on the islands has received a strong influx of genetic diversity from mainland Europe and Africa (Fregel, et al., 2009a, Maca-Meyer, et al., 2003, Santana, et al., 2016) and the majority of the genetic diversity of today’s Canarians is of non-
indigenous origin (Pino-Yanes, et al., 2011). The post-conquest replacement of the indigenous human gene pool is thus sharply contrasted by an almost complete preservation of the indigenous barley gene pool. Our results also contrast previous research on archaeobotanical remains of Gran Canarian wheat which, although with limited number of markers and extant Canarian landrace accessions, suggested more influx of mainland cultivars (Oliveira, et al., 2012).

According to the modern day farmers on the Canary Islands, local barley is better adapted to the climate of the archipelago, with short and mild winters, than barley from mainland Europe (Gil González, 2011, Gil González, et al., 2014). Local barley is able to ripen even during extreme droughts, producing larger harvests than foreign barley. The fidelity of the Canarian farmers to the local barley is thus likely, at least in part, due to its better performance, being well adapted to the climate of the archipelago. It has been suggested that conquerors and immigrants in the Southern Levant similarly favoured locally adapted barley over homeland crops (Mascher, et al., 2016).

The fidelity of farmers to local barley is also evident from the distribution of genetic diversity within the Canary Islands. Barley from Gran Canaria almost exclusively forms a close cluster in the PCA, as do, to a large part, the accessions from several of the other islands. The Gran Canarian accession CBT02687 is an exception and clusters closely with accessions from Lanzarote both in the PCA and the structure analysis. It is likely that this is the result of a recent transfer of barley germplasm from Lanzarote to Gran Canaria. Interviews with modern farmers on Gran Canaria and the western islands confirms that barley seeds from Lanzarote were sometimes traded and cultivated in the lowlands of the western islands because of their faster ripening than local barley (Gil González, 2011).

In this study four of the 36 extant accessions from the Canary Islands clustered completely or partially with the mainland accessions. Three of these accessions are curated at the Spanish INIA genebank and were collected in the Canary Islands during the early 1940s. The integrity of genebank material has been questioned previously (Forsberg, et al., 2015, Hagenblad, et al., 2012) and it is possible that during their ex situ conservation, a Canarian identity has been mistakenly assigned to these
accessions and their true origin has been lost. The similarity of BGE007573 and BGE007566 to the Iberian accessions studied tentatively suggests the accessions instead have a Spanish or Portuguese origin (figure 3C).

The time depth displayed in phylogeographic analyses of extant landraces has been a matter of concern in several studies. Although it has been suggested that landraces can contain a genetic signature reflecting the original introduction of agriculture into an area (Jones, et al., 2013, Jones, et al., 2008) other studies caution against such interpretations made on ex situ preserved materials (Forsberg, et al., 2015, Hagenblad, et al., 2012, Roullier, et al., 2013). Comparative genetic data on ancient and extant landraces from the same area is very sparse, but the results presented here suggest that, in an isolated area such as the Canary Islands, the genetic composition of a crop could remain relatively intact in landraces for millennia.

4.3 Pre-Hispanic barley cultivation in the Canary Islands

Genotyping of putatively functional markers confirmed that pre-Hispanic barley cultivated on Gran Canaria was of the six-row type (Morales, et al., 2014). In addition the barley cultivated seems to have been early flowering and likely had a high grain protein content and a high number of grains per spike. The same genotypes were found in all extant Canarian accessions suggesting that the nutritional and production qualities of pre-Hispanic barley can be studied by cultivating present day landraces using ancient farming practices.

The archaeological samples were recovered from three different sites located on different parts of Gran Canaria (Fig. 2A). In spite of this, the barley from the three sites showed high genetic similarity and clustered together in the STRUCTURE and PC analyses. The pairwise FST values for the archaeological sites were of a similar magnitude as those between extant within-island comparisons. This suggests that pre-Hispanic cultivation in different parts of the island was not, necessarily, more isolated than present day cultivation.

From archaeological remains it is known that barley was cultivated in pre-Hispanic times on Gran Canaria, Tenerife, La Gomera and El Hierro and between 4th and 11th century AD also on La Palma (Morales, et al., 2009, Morales, 2010, Morales, et al.,
2011, Morales, et al., 2013, Morales, et al., 2014). To date there is no archaeological support for pre-Hispanic barley cultivation on Fuerteventura or Lanzarote although the first Europeans reported barley cultivation on Lanzarote, but not Fuerteventura (de la Salle, 1404-19/1980). From our results it is clear that the present day barley on Lanzarote and Fuerteventura is part of the general Canarian gene pool. However, the relatively conspicuous differences between accessions from Lanzarote and Fuerteventura, and accessions from the western islands strongly suggest that the separation of these two genetic clusters predates the Hispanic conquest. In contrast, we do not consider the barley of La Palma to have a distinct enough gene pool to suggest continuous cultivation on the island of the same genotypes since pre-Hispanic times. Instead barley seems to have been reintroduced to La Palma following the Hispanic colonization.

The location of the archaeological samples differ between the PCAs of all markers (Supplementary figure S6, S7) and the one using only markers with a high success rate in the archaeological specimens (Fig. 3B). In the former, the excavation sites with the most missing data fall away from the Guayadeque samples (average 92 % success rate) and the extant landraces. Estrellas with the lowest genotyping success rate (average 70 %) fell further away than the somewhat more successful Temisas samples (average 75 % success rate). It therefore seems that the location of the Estrellas and Temisas is at least partially caused by the higher presence of missing data. (The same is true of the location of the archaeological samples in the PCA of all data in figure 3C).

The archaeological samples are fairly centrally located in the PCA of the reduced marker set (Fig. 3B), just as Guayadeque is in the PCA of the full set of markers before Procrustes transformation (Supplementary figure S6). Interestingly, when comparing only highly successful markers, although being from excavation sites all located on Gran Canaria the archaeological material does not primarily cluster together with extant Gran Canarian accessions. Instead accessions cluster among accessions from La Palma, an island known for keeping the largest agricultural biodiversity in the archipelago (Gil González, et al., 2014). It should be noted, however, that the PCA is based on a limited number of markers (41) and that this may have biased the results. A Procrustes transformed PCA of the full dataset instead
showed similarities between Guayadeque and extant accessions from Lanzarote. In the STRUCTURE analysis the clustering of the archaeological specimens was less clear with different seeds from the same site and seeds from different sites being grouped into different clusters (Fig. 4).

4.4 On the origin of Canarian barley

The introduction of barley to the Canary Islands caused a genetic bottleneck and loss of genetic diversity visible as low within-country genetic diversity and a tight clustering in the PCA (Fig. 3C). The bottleneck and the founder effect resulted in close similarity between Canarian barleys, regardless of age, and relatively large divergence from accessions from mainland Europe and Africa. Recently, the phylogenetic structure of almost 1500 barley accessions from Europe, Asia and Africa was investigated (Pasam, et al., 2014). In that study all accessions from the Western Mediterranean, including the Canary Islands, fell into the same major cluster. It is likely that the low number of Canarian accessions (six accessions from Tenerife) and the single individual per accession studied limited the power to detect a Canarian cluster. Analysis of a subset of our data consisting of single individuals from the mainland accessions and five accessions from Tenerife also failed to detect a Canarian cluster (data not shown).

As a consequence of the strong introduction bottleneck, our ability to draw conclusions concerning from where barley was brought to the Canary Islands is limited. The PCA tentatively points towards a Moroccan origin with the Moroccan accessions being the most closely located of the mainland accessions with the exception of the low diversity French accession HOR874. The mainland accessions with the lowest F<sub>ST</sub> values when compared to the archaeological remains were also all from northern Morocco. Structure analysis of the Moroccan accessions and archaeological specimens suggested HOR13443, IG32022, IG32055 and IG32066 as the accessions clustering most closely with the archaeological material, but did not indicate a single region in Morocco.

Several different lines of evidence, including ancient human DNA, language, and material culture, have pointed to a North African origin of the indigenous people of the Canary Islands (Fregel, et al., 2009a, Maca-Meyer, et al., 2003, Onrubia-Pintado,
et al., 1995, Springer, 2001). Some scholars have instead suggested a Phoenician-Punic-Roman origin of the population (Atoche Peña, 2013). It seems a reasonable assumption that the barley cultivated by the indigenous people was brought along from the homeland of the original settlers and a close genetic similarity to mainland barley in the area from which the settlers originated is expected. Our results thus lend additional support to a Canarian colonization from the area of present day Morocco, possibly northern Morocco, rather than a Phoenician-Punic-Roman settlement, or at least that the barley brought to the islands was of a North West African origin. The relatively distinct gene pools of barley on different islands further support a limited contact between the different islands of the archipelago after the initial colonization.

5. Conclusions
Our comparison of archaeological and extant barley on the Canary Islands and the North African and European mainland reveals a farmer fidelity to the local barley that has lasted for many centuries. It also supports the presence of pre-Hispanic barley cultivation on Lanzarote and corroborates the hypotheses regarding an original colonization from North Africa and subsequent isolation of the islands.

Acknowledgements
The authors are grateful to Dr. Nils Forsberg, Ida Gustafson, Maria Lundström and Maja Krzewinska for help in the lab. Dr. Robin Allaby is acknowledged for kind permission to use his aDNA facilities at Warwick University as are the gene banks and Jaime Gil for kindly providing extant landrace barley.

Funding
This work was supported by the Olle Engkvist Byggmästare foundation, the Royal Swedish Academy of Letters, History and Antiquities and the Project HAR2013-41934 funded by the Spanish Ministerio de Economía y Competitividad. JM was funded by ERC 2013 CoG 614960.
Data accessibility

The data used in this study are available in Dryad: (to be added upon acceptance).

References


Atoche Peña, P., 2013. Consideraciones en relación con la colonización protohistórica de las Islas Canarias/Considerations in relation with the Protohistoric colonization of the Canary Islands, Anuario de Estudios Atlánticos.


Supplementary data

Supplementary file S1. Table of archaeological samples used in the study.

Supplementary file S2. Table of accessions used in the study.

Supplementary file S3. Results of principal component analysis of the genetic diversity of extant Canarian barley accessions.

Supplementary file S4. $\Delta K$ and $H$ values for $STRUCTURE$ analyses of different data sets.

Supplementary file S5. Result of $STRUCTURE$ analysis at $K = 2$ of archaeological and extant Canary barley after the exclusion of highly divergent extant accessions.

Supplementary file S6. Results of principal component analysis of the genetic diversity of extant accessions and archaeological barley remains after the exclusion of highly divergent extant accessions.

Supplementary file S7. Results of principal component analysis of extant accessions and archaeological barley remains after Procrustes transformation.

Supplementary file S8. Pairwise $F_{ST}$ values. Lines separate archaeological specimen from extant accessions and Canarian accessions from mainland accessions.

Supplementary file S9. Distruct and geographical visualisation of result of $STRUCTURE$ analysis of Moroccan accessions and archaeological specimens from Gran Canaria at $K = 5$, the level of clustering with the highest support.