

Received Date : 05-Mar-2015

Revised Date : 16-Jun-2015

Accepted Date : 17-Jun-2015

Article type : Original Article

**Paleo-islands as refugia and sources of genetic diversity within volcanic archipelagos:  
The case of the widespread endemic *Canarina canariensis* (Campanulaceae).**

**M. MAIRAL<sup>\*1</sup>, I. SANMARTÍN<sup>1†</sup>, J. J. ALDASORO<sup>2</sup>, V. CULSHAW<sup>1</sup>, I. MANOLOPOULOU<sup>3</sup> AND  
M. ALARCÓN<sup>2†</sup>**

<sup>1</sup> *Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain*

<sup>2</sup> *Institut Botànic de Barcelona (IBB-CSIC-ICUB), 08038 Barcelona, Spain*

<sup>3</sup> *University College London, WC1E 6BT London, United Kingdom*

□ *Correspondence to be sent to: Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain;*

*Tel: +34914203017; Fax: +34914200157; e-mail: mariomairal@gmail.com*

† These senior authors contributed equally

Keywords: ancestral areas, extinction, oceanic islands, paleo-islands, volcanic refugia.

Running head: WITHIN-SPECIES DIVERSIFICATION IN VOLCANIC ISLANDS

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.13282

This article is protected by copyright. All rights reserved.

## ABSTRACT

Geographical isolation by oceanic barriers and climatic stability has been postulated as some of the main factors driving diversification within volcanic archipelagos. However, few studies have focused on the effect that catastrophic volcanic events have had on patterns of within-island diversification in geological time. This study employed data from the chloroplast (cpDNA haplotypes) and the nuclear (AFLPs) genomes to examine patterns of genetic variation in *Canarina canariensis*, an iconic plant species associated with the endemic laurel forest of the Canary Islands. We found a strong geographic population structure, with a first divergence around 0.8 Ma that has Tenerife as its central axis and divides Canarian populations into eastern and western clades. Genetic diversity was greatest in the geologically stable "paleo-islands" of Anaga, Teno and Roque del Conde; these areas were also inferred as the ancestral location of migrant alleles towards other disturbed areas within Tenerife or the nearby islands using a Bayesian approach to phylogeographic clustering. Oceanic barriers, in contrast, appear to have played a lesser role in structuring genetic variation, with intra-island levels of genetic diversity larger than those between-islands. We argue that volcanic eruptions and landslides after the merging of the paleo-islands 3.5 million years ago played key roles in generating genetic boundaries within Tenerife, with the paleo-islands acting as refugia against extinction, and as cradles and sources of genetic diversity to other areas within the archipelago.

## INTRODUCTION

Due to their small size, discrete geographic boundaries, substantial environmental heterogeneity and buffered climates, oceanic islands represent excellent natural laboratories to examine the role of ecological adaptations versus geographic isolation as drivers of diversification (Carson & Templeton 1984; Gillespie 2004). Patterns of genetic variation are

Accepted Article

expected to be hierarchical in islands, with between-island genetic differentiation stronger than within-islands (Bottin *et al.* 2005; García-Verdugo *et al.* 2010), because oceanic barriers are generally more effective than topographic barriers at promoting isolation in insular systems (Gillespie & Clague 2009). However, the birth and development of volcanic islands is usually followed by a large number of destructive events in the form of secondary eruptions, landslides, merging of paleo-islands, etc (Carracedo 1994; Fernández-Palacios *et al.* 2011). These events promote habitat fragmentation; the subsequent genetic isolation of populations; and they ultimately drive differentiation and speciation within-islands (Carson *et al.* 1990; Gillespie & Roderick 2002, 2014; Macías-Hernández *et al.* 2013). Therefore, volcanic archipelagos represent an ideal framework for studying patterns of diversification at different spatial scales (between- and within-islands) and also over different temporal levels (allopatric speciation, extirpations and recolonisations).

The Canary Islands are a volcanic archipelago formed by a chain of seven islands, located 110 km from the north-western coast of Africa. Their geological history and highly endemic biota have interested scientists since the early 19<sup>th</sup> century (von Humboldt 1814; Lyell 1855). The islands were formed in the last 21 million years (Ma) with an east to west pattern of chronological emergence due to a mantle plume (Carracedo *et al.* 1998; Zaczek *et al.* 2015) (Figure 1a). They are separated by deep oceanic trenches and have never been connected to the mainland. The Canary Islands have long been considered refugia for continental lineages that have survived the climatic changes of the Late Cenozoic, but also as cradles of biodiversity where multiple in-situ diversification events have taken place (Juan *et al.* 2000; Francisco-Ortega *et al.* 2000). In addition, these islands are regarded as a hotspot for plant diversity (Médail & Quézel 1997): approximately 40% of Canarian vascular plants are endemics (Santos-Guerra 2001) and the flora is generally characterised by high levels of inter-population differentiation in comparison to other archipelagos (Francisco-Ortega *et al.*

2000; de Paz & Caujapé-Castells 2013). Recently, they have been proposed as reservoirs and sources of genetic diversity (García-Verdugo *et al.* 2015; Patiño *et al.* 2015).

Most studies on the Canarian flora have focused on the pattern of inter-island diversification, particularly at the species level (Francisco-Ortega *et al.* 2002; Kim *et al.* 2008; Vitales *et al.* 2014a, b). Nevertheless, the complex topographies and long histories of avalanches and secondary eruptions of many of these islands are likely to have favoured within-island diversification (Juan *et al.* 2000; Brown *et al.* 2006). Tenerife has the most complex history of all the islands. It existed at first as three separate islands, dating back to the Late Miocene: Roque del Conde (11.9-8.9 Ma) in the southwest, Teno (6.2-5.6 Ma) in the northwest, and Anaga (4.9-3.9 Ma) in the northeast. Eruptive central volcanic episodes fused these islands 3.5 Ma and gave rise to the present island of Tenerife (Ancochea *et al.* 1990; Fig. 1b). The paleo-islands remained thereafter relatively stable, whereas the central part of Tenerife continued to be active until 0.13 Ma (Ancochea *et al.* 1990, 1999; Cantagrel *et al.* 1999; Guillou *et al.* 2004; Carracedo 2014; see Fig. 1b). Interestingly, the three paleo-islands of Tenerife, together with La Gomera – which has also remained geologically stable since the Pliocene – are presently home to the highest phylogenetic diversity and endemic richness of the Canarian Archipelago (Reyes-Betancourt *et al.* 2008). These areas also harbour the best-preserved laurel forests, considered as an ancient, unique flora restricted to Macaronesia, and they share several endemic and restricted species (see Table S1). Besides their geological (volcanic) stability, the three paleo-islands of Tenerife also exhibit a topographic complexity and variety of micro-climates that might have favoured their role as micro-refugia against climate- or human-induced extinction (Harter *et al.* 2015).

In reviewing the role of Tenerife paleo-islands as refugia across several plant lineages, Trusty *et al.* (2005) found that species endemic to the paleo-islands often occupied a derived position in the lineage's phylogeny. This position argued against the idea of these massifs as

Accepted Article

ancient refugia. However other studies, especially in animals, have reported ages for divergence events between taxa endemic to these paleo-islands, either at the species or at the infra-species level (Juan *et al.* 1996, 2000; Dimitrov *et al.* 2008; Macías-Hernández *et al.* 2013; Puppo *et al.* 2014;) that are contemporaneous or predate the age of merging of the precursor paleo-islands 3.5 million years ago (Ancochea *et al.* 1990). Few plant studies (Gómez *et al.* 2003; García-Verdugo *et al.* 2010) have focused on patterns of within-island genetic variation for widespread Canarian endemics, and none of them have provided estimates of lineage divergence times, which is necessary to relate within-island patterns to the island geological history. Species that are present in multiple islands ("multiple island endemics" or MIEs) are especially relevant to understand the role of paleo-islands as undisturbed areas that have acted as reservoirs and sources of genetic diversity not only within- but between-islands.

Here, we study patterns of genetic diversity and the demographic and spatial history of a multiple island endemic, *Canarina canariensis* (L.) Vatke. This "flagship" species of the Canary Islands, elected as its "national flower" (Kunkel 1991), is a diploid ( $2n = 34$ ) herbaceous plant that grows mostly in cleared areas surrounding the endemic laurisilva forest. It presently occurs in the central and western Canary Islands: Gran Canaria, Tenerife, La Gomera, La Palma and El Hierro. *C. canariensis* is a herbaceous plant that occasionally climbs on nearby plants; it is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011) and its fleshy fruits are dispersed by vertebrates (Valido *et al.* 2003, Rodríguez *et al.* 2008). Genus *Canarina* belongs to tribe Platycodoneae, a basal group within family Campanulaceae (Olesen *et al.* 2012; Mansion *et al.* 2012; Wang *et al.* 2013; Mairal *et al.* 2015). In addition to *C. canariensis*, the genus comprises two other species inhabiting the Afromontane forests of Eastern Africa, *C. eminii* and *C. abyssinica*, being this an example of a wide, continental-scale disjunction of 7000 km spanning across the Sahara. Mairal *et al.*

(2015) recently reconstructed the phylogeny and spatiotemporal evolution of *Canarina*. They inferred that *C. canariensis* diverged from its African relatives at the end of the Miocene (c. 7 Ma). This extraordinary temporal and spatial disjunction was explained as the result of vicariance and climate-driven extinction resulting in the fragmentation of an ancient widespread distribution. The colonization of the Canary Islands by the ancestors of *C. canariensis* apparently occurred much later, in the Pleistocene, probably from a now extinct and geographically closer North African population: the earliest event of population divergence is dated at only around 1 Ma (Late Pleistocene) and involved an east-west vicariance within Tenerife (Mairal *et al.* 2015).

This age and the presence of *Canarina canariensis* in several islands makes it an ideal candidate to evaluate patterns of within-island diversification in relation to the recent geological history of the archipelago. Our main aims were to: i) determine the geographic distribution of genetic variation within *Canarina canariensis*, ii) find evidence of extinction and diversification processes that may be related to geological events, iii) find ancestral areas and reconstruct inter-island migration events, iv) examine the putative role of the paleo-islands of Tenerife as refugia of genetic diversity, both relictual and recent.

Haplotype networks are commonly used in population-level studies because they provide a clearer picture of the reticulate relationships between genetic pools than a branching tree, especially when gene flow is present. These networks are often inferred using Statistical Parsimony (SP, Templeton *et al.* 1992) implemented in the software TCS (Clement *et al.* 2000), which allows estimation of the haplotype network while minimising the number of mutation events differing among haplotypes. However, this method fails to incorporate the uncertainty associated to the network inference, and therefore does not allow for statistical evaluation of alternative phylogeographic scenarios (Bloomquist *et al.* 2010). Moreover, unobserved events such as local population growth or past extinction of haplotypes may

mislead inference in parsimony-based methods. Here, we compare results from TCS with those obtained from a model-based, Bayesian statistical method, Bayesian Phylogeographic and Ecological Clustering (BPEC, Manolopoulou *et al.* 2011), which allows estimating the posterior probabilities for haplotype tree networks under a coalescent-based migration-mutation model (Manolopoulou & Emerson 2012). To our knowledge, this is the first study to use this method for island phylogeography.

## MATERIALS AND METHODS

### *Population sampling and DNA extraction*

*Canarina canariensis* has a significantly greater presence on Gran Canaria (GC) and Tenerife (TF), in comparison with La Gomera (LG), La Palma (LP) and El Hierro (EH). Seventeen populations of *C. canariensis* were sampled in several fieldtrips between 2009 and 2012: 4 in GC, 8 in TF, 2 in LG, 2 in LP and 1 in EH. Where possible, we collected a minimum of 10 samples per population. To reduce inflation in gene descriptors due to biased sampling (Caujapé-Castells 2010), samples were collected from individuals scattered across the whole occupancy area of each population. DNA from 160 individuals and preserved in silica gel was extracted using the DNeasy Plant Mini Kit (QIAGEN Inc., California, USA), from 20-25 mg of silica-gel dried leaves obtained from the fresh plant tissue collected from the field expeditions.

### *Chloroplast DNA sequencing*

We selected three cpDNA intergenic spacers regions for sequencing; these markers have proven to be useful for intraspecific analyses of population structure (Mairal *et al.* 2015). We generated 432 new sequences: *rpl32-trnL*<sup>UAG</sup> (144 sequences), *trnS*<sup>GCU</sup>-*trnG*<sup>UCC</sup> (144 sequences) and *petB*<sup>1365</sup>-*petD*<sup>738</sup> (144 sequences). PCR and sequencing protocols

followed those of Mairal *et al.* 2015. The sources of the material examined, the GenBank accession numbers and full references are detailed in Table S2.

### *Haplotype analyses*

Sequences for each region were aligned using MAFFT 6.814b (Kato *et al.* 2002), implemented in the software Geneious Pro 5.4.4. (Biomatters Ltd., Auckland, New Zealand). Sequences were checked and manually adjusted where necessary by following alignment rules described in Kelchner (2000). We analysed the three sequenced regions as three data partitions to perform phylogenetic analyses. MrModeltest v.2.2 software (Nylander, 2004) was used to determine the best fitting model of sequence evolution of each data partition.

Summary statistics for within-population genetic diversity were calculated as: the number of haplotypes  $H(n)$ , haplotype diversity ( $H_d$ ), nucleotide diversity  $\pi$ , nucleotide heterozygosity  $\theta$ ,  $G_{ST}$  and the number of migrants per generation ( $N_m$ ) were estimated for each population using DnaSP (version 5.10; Librado & Rozas 2009).

The relationships amongst lineages were investigated through haplotype network analysis, by using 6-12 individuals from different populations and examining the three sequenced regions. Genealogical relationships among haplotypes were inferred via the statistical parsimony algorithm (Templeton *et al.* 1992) implemented in TCS 1.21 (Clement *et al.* 2000). The number of mutational steps resulting from single substitutions among haplotypes was calculated with 95% confidence limits, gaps were represented as missing data.

The BPEC method (Manolopoulou *et al.* 2011; Manolopoulou & Emerson 2012) was implemented in the R package (Manolopoulou & Hille, in prep.) to identify genetically distinct geographical population clusters and ancestral locations. Like TCS, BPEC relies on parsimony in order to reduce the number of candidate trees to a manageable set. The BPEC

method, unlike Standard Parsimony, fits a prior over all possible trees in order to identify trees with high posterior probability in a fully model-based framework, thus accommodating for uncertainty in haplotype relationships, which is one of the main criticisms of TCS (Knowles 2008). Each possible tree defines a set of possible migration events that may have led to the observed population substructure. Different scenarios of trees and migration events are explored through Markov chain Monte Carlo (MCMC), similar to the method proposed by Sanmartín *et al.* (2008) for estimating rates of inter-island dispersal. Migration events were assumed to occur when a haplotype (with or without a mutation from its parent haplotype) migrates to a new geographical cluster. MCMC simultaneously estimates high probability trees, number of migration events and corresponding clusters. The method assumes that the migration rate and the population growth are constant. BPEC requires two main user-defined inputs: the maximum number of migrations (denoted as “MaxMig” in the software) and the parsimony relaxation parameter used to reconstruct the set of possible trees (denoted as “ds” in the software). MaxMig allows the user to set the upper bound for the number of migration events and hence the maximum number of clusters (MaxMig + 1). Larger values include more models but require much greater computation time. As the ds value is increased, the parsimony assumption is relaxed: if two observed sets of sequences have an unobserved intermediate missing sequence (an unobserved mutation), then any pair of sequences of distance  $\{1, \dots, ds\}$  nucleotides will be considered as the “missing path”. Two MCMC chains were run for three million iterations. The results were stable, with ds > 3 having no effect on the inferred tree, and similarly, any number of migrations above 4 converging to a 5–cluster model. The phylogeographic clustering obtained was superimposed upon a haplotype tree, and used to estimate ancestral locations for migration events. As a further exploration, we divided the dataset into two groups (eastern and western clades – for details see *coalescent dating* section in results). Haplotypes sampled from Roque del Conde

were quite divergent; thus, separate analyses were run in which we included and excluded this location from the two groups. We ran BPEC analysis on each of these four datasets (eastern and western groups with and without Roque del Conde).

Haplotype divergence times were estimated in BEAST v.1.7 (Drummond and Rambaut 2007). We carried out a first analysis under a strict clock model and a coalescent constant population tree prior, using a secondary age estimate (Mairal *et al.* 2015; normal prior: mean = 0.76 Ma, standard deviation (SD) = 0.327 Ma) to calibrate the root node of the “*C. canariensis* dataset”; this included all haplotypes detected in our sampling (N = 10). This analysis gave us very large 95%HPD (High Posterior Density) credibility intervals and poor ESS for posterior age estimates, probably due to the low information content at the population level (see Results) and presence of a single calibration point. We carried out a second analysis applying the “nested dating approach” described in Mairal *et al.* (2015), in which a higher-level dataset including representatives of all three species of *Canarina* and 9 outgroup taxa was used to inform the clock rate of a linked population-level dataset (*C. canariensis*) under a *mixed Yule-coalescent model* (Ho *et al.* 2005; Pokorny *et al.* 2011). The higher-level dataset was calibrated with fossil-derived secondary age estimates (see Mairal *et al.* 2015), while the tree prior was unlinked to apply a coalescent constant size model to the population-level dataset and a stochastic birth-death (Yule) prior to the species-level one (Mairal *et al.* 2015). The clock model was set to an uncorrelated log-normal prior to accommodate the change in mutation rate from species to populations, with a uniform distribution for the *uclid.mean* ( $10^{-4}$ - $10^{-1}$ ) and a default exponential distribution for the *uclid.stdev*; the substitution model was set to GTR+G; choice of priors was based on Bayes Factor comparisons using the path sampling method in BEAST (Baele *et al.* 2012); see Table S3 to see results from exploratory analyses to assess the reliability of our date estimates with reference to these settings. Two MCMC chains were run for 50 million generations, sampling

parameters every 1000 generation. We used Tracer v1.6 (Rambaut *et al.* 2013) to verify: whether a stationary distribution was attained; if there was convergence amongst chains; and that effective sample sizes (ESS values) were > 200 for all parameters. A 10% burn-in of the sampled populations was discarded (5 million). Post-burn-in trees were summarised into a maximum clade credibility tree using TreeAnnotator v.1.6.1, with mean values and 95% credible intervals for nodal ages, and were visualised in FigTree 1.3.1 (Rambaut and Drummond 2009). The resulting age estimates from this second analysis exhibited considerably larger ESS values and narrower 95%HPD intervals than in the first and are the ones reported here.

#### *Demographic history.*

Statistics used to describe demographic patterns may be biased by a strong genetic structure or lack of panmixia among populations. Since we detected strong genetic structure in our dataset (see results), we performed demographic analyses in subsetted datasets that were less genetically structured (including approximate panmitic populations). Three groups previously recognised by the haplotype network and BEAST analyses were used: ET-GC (including populations in east Tenerife and Gran Canaria), TENO-GO (including populations in Teno and La Gomera) and LP-EH (including populations in La Palma and El Hierro). For the same reasons as described above, the Roque del Conde population was included and discluded from the ET-GC and TENO-GO groups. Overall we ran the analyses on five groups.

We used three different approaches to infer the demographic processes shaping the genetic structure of *C. canariensis*. First of all, to test for evidence of population expansion, we carried out a neutrality test – Fu and Li’s tests (Fu & Li, 1993; Fu 1996) and Tajima’s *D* test (Tajima, 1989) – for each population group. We used the DNAsp program,

version 5.0 (Librado & Rozas 2009) and assessed the significance of each test with 10,000 coalescent simulations. Secondly, we plotted the mismatch distribution for each group using the observed number of differences between all pairs of sequences with the ARLEQUIN v.3.0 software (Excoffier *et al.* 2005). The goodness-of-fit of the observed mismatch distribution to the theoretical distribution under a constant population size model was tested with the raggedness index (*HRag*) (Harpending *et al.* (1994)). Thirdly, i) we created the Extended Bayesian Skyline Plot (EBSP), implemented in BEAST, for each population group, ii) we performed EBSP analyses for each of the 16 population groups. For each group from i) and ii): two independent chains were run simultaneously for 150 million generations, sampling every 1000 generations; a strict clock model was used, whereas all other parameters were set identical to those described above for the nested dating analysis; the root node was calibrated using a normal prior with a mean age estimate and 95% High Posterior Density (HPD) credible intervals obtained from this analysis.

#### *AFLP fingerprinting*

For the AFLP analysis we used a total of 97 individuals from 10 populations, which covered all of the islands: one population from Gran Canaria, five from Tenerife, one from La Gomera, two from La Palma and one from El Hierro. Laboratory molecular protocols for the AFLP analysis (Vos *et al.* 1995) were implemented using the AFLP plant mapping kit (Applied Biosystems®). To select the appropriate primers, we first carried out a pilot study combining fluorophores and restriction enzymes for five geographically distant individuals (one per island), using in total 32 primer combinations. One sample from each individual was duplicated as a blind sample to test for reproducibility and contamination. Reproducibility and the number of alleles per sample were calculated by choosing three combinations of primers: 1-*EcoRI*<sub>6-FAM</sub>-ACT/*MseI*-CAA, 2-*EcoRI*<sub>VIC</sub>-AGG/*MseI*-CTA, and 3-*EcoRI*<sub>VIC</sub>-

AGG/*Mse*I-CTT and using the Genemapper v3.7 software (Applied Biosystems, Foster City, CA, USA). These three primer pairs showed high reproducibility and homogeneously scattered bands, and produced polymorphic AFLP profiles and clear fragments. In the digestion phase, samples of DNA were digested with the restriction enzymes *Eco*RI and *Mse*I and linked to the primers *Eco*RI 5'-CTCGTAGACTGCGTACC-3'/5'-AATTGGTACGCAGTCTAC-3') and *Mse*I (5'- GACGATGAGTCCTGAC-3'/5'-ATCTCAGGACTCAT-3'). The three different AFLP reactions, were: i) restriction and ligation in a single reaction; (ii) and (iii) consecutive PCR amplifications (preselective and selective). PCR products were checked on 1% agarose gels.

#### *AFLP data analyses*

The resulting AFLP fragments were analysed using the GeneMapper 3.7 software. Peaks were recorded in 100 to 500 base pairs ranges. Shorter fragments were discarded because the majority of this size class have a high chance of being non-homologous fragments (Vekemans *et al.* 2002). For each primer combination, an automated size detection and peak binning was employed followed by manual editing of bins to exclude shoulder peaks and unreliable loci (variation between replicates). Peak height data were then exported and loaded into the R package AFLPScore version 1.4a (Whitlock *et al.* 2008) and the AFLP profiles were scored and the error rates were estimated. These rates were below the critical bound of 5% indicated in previous reports (Bonin *et al.* 2004) for each primer combination. Data reliability was assessed through comparison of duplicates, from one or two individuals per population. Data were converted into binary presence/absence scores for each locus. The resulting AFLP presence/absence matrix was analysed using a selection of different analyses. The AFLPSURV v.1.0 software (Vekemans *et al.* 2002) was used to estimate demographic statistics such as Nei's gene diversity ( $H_j$ ), pairwise differentiation among subpopulations

( $F_{ST}$ ) and the percentage of polymorphic fragments per population (P) (Nei and Li, 1979; Lynch and Michigan, 1994). This was done under the assumptions of either the Hardy-Weinberg equilibrium or partial self-fertilisation, based on a previous study on the reproductive biology of *C. canariensis* (Rodríguez-Rodríguez and Valido 2011). A Bayesian method in AFLPSURV was also used to estimate allelic frequencies through employment of a non-uniform prior distribution (Zhivotovsky 1999). Ten thousand permutations were run to calculate the  $F_{ST}$  parameter from which genetic distances between individuals, populations and geographic groups were calculated. To locate genetic clustering of individuals within the AFLP dataset, a pairwise similarity matrix for all individuals using the Dice's coefficient as similarity distance was constructed, and the resulting matrix subjected to a Principal Coordinates Analysis (PCA) implemented in Ntsys v.2.1 (Rohlf 1998). Next, genetic relationships among samples were visualised in SplitsTree v.4.10 (Huson and Bryant 2006) using Neighbour-Net analysis through the use of the split decomposition method. Finally, to quantify the amount of genetic differentiation attributable to geographic and population subdivision, a hierarchical analysis of molecular variance was performed using ARLEQUIN v.3.0 software (Excoffier *et al.* 2005). Exploratory analyses were performed considering, alternately, islands and paleo-islands as geographical units in order to investigate the distribution of genetic variance attributable to oceanic barriers.

Bayesian clustering methods implemented in STRUCTURE v.2.3 (Pritchard *et al.* 2000; Falush *et al.* 2007) were used to assess the genetic structure of populations. This model-based approach assumes that loci are in Hardy-Weinberg equilibrium and linkage equilibrium within populations. Analyses were performed under admixture conditions and correlated allele frequencies between groups. 500,000 MCMC generations (plus a burn-in of 100,000) were run for  $K$  values of 1-10, with 10 repetitions for each. The most likely  $K$  value was determined by the method from Evanno *et al.* (2005), which is implemented in

STRUCTURE HARVESTER (Earl 2012). We explored other values of K to detect further genetic substructure of populations, especially in Tenerife. To test the effect of the spatial distance on the genetic structure of the populations, correlations between genetic (measured as  $F_{ST}$ ) and spatial distances between pairs of populations were determined using the Mantel permutation procedure implemented in NTSYS v. 2.1. The genetic distance matrix used was based on the presence/absence matrix; the geographic distance matrix was based on the absolute distances between the geographic coordinates for each collected population. In addition, to identify possible geographic locations acting as major genetic barriers among *C. canariensis* populations, we computed barriers on a Delaunay triangulation using Monmonier's algorithm in BARRIER v.2.2 (Manni *et al.* 2004). The significance was examined by the mean of 1000 bootstrapped distance matrices obtained using AFLPsurv. Only barriers with support greater than 96% were considered as significant.

## RESULTS

### *Haplotype network analysis and coalescent dating*

The *pet B*<sup>1365</sup>-*pet D*<sup>738</sup> region consisted of 937 sites, *rpl32-trnL*<sup>UAG</sup> of 654 sites, and *trnS*<sup>GCU</sup>-*trnG*<sup>UCC</sup> of 658 sites. The final concatenated data matrix consisted of 144 sequences and 2249 nucleotide sites, of which 9 were polymorphic (Table S4). Among the 144 individuals sampled from the 16 populations, we observed 10 different plastid DNA haplotypes (H1 to H10, Figure 2a) with haplotype diversity of 0.6632. There were three haplotypes that dominated and were clearly geographically delimited: H1, H9 and H3. The most frequently observed haplotype was H1 (52.08%), which was dominant in GC and east Tenerife (ET). Haplotype H9 was restricted to West Tenerife (WT) and LG. Haplotype H3 was present in the western islands of LP and EH. Populations with the commonest and rarest haplotypes (frequencies  $\leq 2\%$ ) were in the most rugged areas of Tenerife: the paleo-

islands of Anaga (An-W, An-C, An-E – exclusive haplotypes H2 and H5), Teno (Ten-A, Ten-Ep – exclusive allele H8 and allele H10 shared with LG) and Roque del Conde (Conde – exclusive H7). The small western islands presented one unique haplotype, each: LP (LP-Lt – H4) and EH (H6).

Summary statistics for within-population genetic diversity are shown in Table 1. Populations situated in the three paleo-islands (Anaga (An-W), Teno (Ten-A) and Roque del Conde (Conde)) showed the highest values for the number of haplotypes  $H(n)$ , haplotype diversity  $H(d)$ , nucleotide diversity  $\pi$  and nucleotide heterozygosity  $\theta$ . The least diverse population areas were found in Gran Canaria and outside of the paleo-islands regions of Tenerife.  $G_{ST}$  and  $N_m$  values indicated high genetic cohesion between some islands: east Tenerife populations were highly connected with Gran Canaria populations; west Tenerife with La Gomera; and La Palma with El Hierro (Table S5). Interestingly, the highest  $G_{ST}$  values and lowest  $N_m$  were found between populations situated east and west of Tenerife, with the exception of Anaga An-W. This is the only population in east Tenerife that showed some genetic cohesion with populations in west Tenerife, in particular with the Teno and Adeje massifs.

The haplotype network constructed with TCS (Fig. 2b) shows a loop involving three dominant haplotypes, with haplotype H1 as the ancestral root haplotype according to coalescent criteria on haplotype frequency (Templeton *et al.* 1992). Dominant haplotypes at each side of Tenerife (East H1 and West H9) were separated by five mutational steps, with haplotype H10 as intermediate. Haplotype H3, present in the western islands, was separated by four mutational steps from H1 and three from H9.

BPEC results are summarised in Table 2. With mutational step limit equal to 10 and the maximum number of migrations equal to 8, we allowed for high gene flow, genetic divergence and numerous hidden mutations, which separate the haplotype clusters. The

Accepted Article

clustering with high posterior probabilities ( $pp = 1$ ) showed the existence of five clusters (contour regions in Figure 3a). The total number of haplotypes was 10 plus one missing haplotype (Fig. 3b). Location of the ancestral populations for migration events and inference of the root node were consistent across exploratory analyses of the complete dataset. The root node was inferred as missing (H11 – Fig. 3b), and carried a high uncertainty. Haplotypes H1, H2, H8 also carried significant posterior mass probabilities (Table 2). The most likely source for ancestral migration events within- and between-islands were populations located on: Conde ( $pp = 0.13$ ), An-W ( $pp = 0.10$ ) and LP-Lt ( $pp = 0.086$ ). The subsetting analysis, including Conde, always recovered this population as the ancestral area with the highest probability. Without Conde, other paleo-island populations were recovered as ancestral areas: Ten-A for the Western clade; and the three populations of Anaga for the Eastern clade, with An-W the most probable (Table 2, Figure S1).

The Bayesian chronogram of haplotypes showed a geographical pattern of divergence congruent with the groups detected by TCS and BPEC (Figure 4). The crown age or first divergence event among *Canarina* haplotypes was estimated to be 0.878 Ma (95% HPD = 0.452-1.365), corresponding to the Mid-Pleistocene period; it divided haplotypes into an eastern and western clade. The first divergence event in the eastern clade was estimated at 0.255 Ma (95% HPD = 0.041-0.633), while that in the western clade was dated older: 0.522 Ma (95% HPD = 0.161-0.967), separating La Palma and El Hierro from Teno and La Gomera. Mean age values and 95% HPD intervals for the BEAST analysis are shown in Table S6.

### *Demographic analyses*

Negative values of Fu's and Tajima's indicated a recent population expansion, although these values were only significant for the Fu's test of the East group (Table S7). The frequencies of pairwise differences in the mismatch distribution analysis resulted in unimodal distributions that were consistent with an expansion model. When we include the population of Conde in the East group, we obtained a second small peak. The raggedness statistics derived from the mismatch distribution were not significant and thus failed to reject the null hypothesis of recent population expansion. EBSP indicated a constant population size (Figure S2). EBSP for each of the individual populations also obtained a constant population size (results not shown).

### *AFLP polymorphism, genetic diversity and structure.*

These results are summarised in Table 1. The final dataset, after scoring, comprised 572 loci from 97 individuals and 10 populations. For AFLP analyses we removed populations from Gran Canaria and Roque del Conde, which had been extracted with a different (less rigorous) method (CTAB) to standardise the quality of extracted DNA.  $H_j$  and  $P$  were higher in the La Gomera (GO) population, followed by the Teno (Ten-A) population in west Tenerife. The largest number of private fragments was detected in Tenerife (7 fragments). In La Gomera and El Hierro islands a single fragment was detected. In Tenerife, five private fragments were detected in the Anaga massif (specifically in the An-W population – four fragments) and two in the Teno massif (Table 1).  $F_{ST}$  values with the nuclear data (Table S8) were consistent with the results obtained from the chloroplast markers (see above  $G_{ST}$  and  $N_m$ ), and showed the same genetic cohesions between-islands and paleo-islands.

A PCA analysis differentiated three groups (Figure S4): i) east Tenerife populations; ii) west Tenerife populations; iii) the An-W population, which demonstrate an intermediate position between i) and ii). The split network analysis outlined the same divergences between the two well separated groups, including the West group An-W population (Figure 5b).

Hierarchical AMOVA analyses showed the largest proportion of genetic variation to be found among groups 4 and 5 (Table 3 and S9). By analysing each island separately, we observed that differentiation within-islands (24.27%) is greater than between-islands (3.62%). A further analysis considering islands and paleo-islands as separate units showed that differentiation within-islands and within-paleo-islands is lower (11.30%) than between-islands and between-paleo-islands (16.63%). According to the method of Evanno *et al.* (2005), STRUCTURE indicated that the most likely number of genetic clusters  $K = 2$  ( $\Delta K = 460$ ) represented the optimal number of Bayesian groups within *C. canariensis*, separating the east populations from the west, but also detected the presence of admixture in the intermediate populations. K levels  $K = 4$  and  $K = 5$  (the latter identified as  $K = 6$ , with 5 defined clusters plus one "ghost" cluster with no individuals assigned so it was ignored, see Guillot *et al.* 2005) revealed a more complex genetic structure in these intermediate populations, resulting in additional clusters (Fig. 5a, Figure S5). The STRUCTURE analysis using only Tenerife populations also delimited two groups (inset  $K = 2$  in Fig.5a, Fig. S5b). Overall, the results revealed a strong inter-island genetic structure (Gran Canaria – Anaga massif and La Palma – El Hierro), admixture phenomena (among An-W, Teno populations, and La Gomera), and an even more complex substructure within the island of Tenerife. The STRUCTURE sublevels ( $K = 4$  and  $K = 5$ ) were consistent with the AMOVA analyses. Only La Gomera differed between analyses; this is explained due to its high genetic variability and mixed composition. These groups are also congruent with the BARRIER

Accepted Article

results, which revealed two or four major boundaries (Fig. 5a for K=2 and K = 5). All barriers showed values of 100% except LP-EH (96.8%). No linear relationship was found between pairwise  $F_{ST}$  and geographical distance with the Mantel analysis ( $r = 0.42, p = 1$ ).

## DISCUSSION

### *Paleo-islands as refugia against volcanic (catastrophic) extinctions*

The strong genetic structure often detected in island endemics has been postulated to respond to several interconnected physical and biotic factors (Stuessy *et al.* 2014); foremost among them are geographical isolation and extinction-recolonisation processes (Carson *et al.* 1990; Gübitz *et al.* 2005; Macías-Hernández *et al.* 2013). The strongest genetic structure and levels of genetic variation in *C. canariensis* were detected between two of the ancient paleo-islands that currently form Tenerife: Teno in the west and Anaga in the east. Populations in other islands were genetically associated to these two lineages: La Palma, La Gomera and El Hierro to Teno; and Gran Canaria to Anaga. The fact that these patterns are shared between chloroplast (cpDNA, Fig. 2) and nuclear (AFLPs, Fig. 5) markers supports a long history of isolation among populations (Zink & Barrowclough 2008). Furthermore, the hierarchical AMOVA analysis based on the AFLP data showed that genetic variation was notably higher among populations within Tenerife than those located in different islands (Table 3). This pattern has been found in a few other Canarian endemics (*Olea europaea guanchica* and *Pinus canariensis*), but usually not as marked as here (< 10%: Gómez *et al.* 2003; García-Verdugo *et al.* 2010; see review in García-Verdugo *et al.* 2014). Interestingly, considering the paleo-islands as separate areas reversed the pattern (Table 3). Together with our BARRIER analyses (Fig. 5a in K = 5), these results support the hypothesis that the geographical distribution of genetic diversity in *C. canariensis* is structured around the paleo-

islands, and that these ancient massifs could be considered as separate units in phylogeographic analyses, a sort of "islands within-islands". It indicates that geological barriers within islands – for example, those associated to catastrophic volcanic events – have probably been more important in structuring patterns of genetic differentiation within species than the oceanic barriers separating the islands (García-Verdugo *et al.* 2014).

Unfortunately, we could not obtain AFLP data for the population in Roque del Conde. Given the mixed composition of the chloroplast compartment in this population (Fig. 2), it is possible that inclusion of this population in our analysis would have led to higher levels of genetic admixture and lower K values in the STRUCTURE analyses – though given the marked east/west split among the remaining populations (Fig. 5), this is unlikely. On the other hand, the two most frequent haplotypes, H1 and H9, in the Conde population are also the most frequent or dominant within the eastern and western clades, respectively, while the divergence of these two haplotypes in the BEAST tree (Fig. 4) can be traced back to the basalmost split in *C. canariensis* (0.8 Ma). This, together with the presence of a unique haplotype (H7) and the fact that the Conde population is identified as the ancestral source of westward and eastward migration events in the BPEC analyses, suggests that the admixture detected in the chloroplast compartment for Roque del Conde is of ancient rather than recent origin. It is thus possible that including this population in the AFLP analysis would have increased rather than decreased levels of genetic diversity within Tenerife for *C. canariensis*.

Which might have caused this level of within-island genetic structure? Many Canary plant and animal taxa include sister lineages endemic to the Tenerife paleo-islands (Juan *et al.* 2000; Báez *et al.* 2001), especially among laurel forest species from Teno and

Accepted Article

Anaga (Table 4). While in some cases, the divergence between these taxa is found at the species level (e.g., *Trechus*, Contreras-Díaz *et al.* 2007; *Pericallis*, Jones *et al.* 2014) and/or predates the origin of the paleo-islands (*Micromeria*, Puppo *et al.* 2014; *Pholcus*, Dimitrov *et al.* 2008); in others, it is observed within species (e.g., *Tarentola delalandii*, Gübitz *et al.* 2000), and/or postdates the merging of the paleo-islands (*Eutrichopus*, Moya *et al.* 2004; *Calathus abaxoides*, Emerson *et al.* 1999). The fact that the pattern of divergence is not contemporary across taxa suggests that the role of Tenerife paleo-islands in structuring genetic variation has been long lasting. In *Canarina canariensis* the basal divergence between the basal eastern and western lineages was dated as 0.878 Ma (Fig. 4) substantially, postdating the age of the merging of the paleo-islands (c. 3.5 Ma, Ancochea *et al.* 1990). This ‘young’ east-west divergence might be explained by historical dispersal events between forest patches followed by in-situ diversification. Alternatively, it could be explained by allopatric speciation (vicariance) driven by catastrophic volcanic events within a previously more widespread distribution with low or no reconnection. This might explain the ‘young’ (< 3.5 Ma) east-west Tenerife divergence found in *Canarina*, and seen in other taxa (Table 4). The reciprocal monophyly, similar haplotype diversity levels and deep temporal divergence found between the eastern and western clusters of *C. canariensis* in Tenerife (Fig. 4), favour the vicariant, allopatric scenario. In the last one million years, several major landslides and volcanic events have affected the north of Tenerife, extending from the summit to the coast (Boulesteix *et al.* 2012, 2013; Carracedo 2014), e.g. the Güimar and La Orotava mega-landslides (Fig. 1b). These events could have fragmented the ancient laurel forest corridor that extended across the northern flank of the island (Moya *et al.* 2004), interrupting gene flow within species associated to this laurel forest (e.g., *C. canariensis*) and promoting differentiation among populations. Several studies have supported debris avalanches as important factors driving diversification within terrestrial Canarian organisms (Juan *et al.*

2000; Brown *et al.* 2006; Macías-Hernández *et al.* 2013); others have reported a temporal divergence within northern Tenerife (Thorpe *et al.* 1996; Moya *et al.* 2004) that is similar to the one found in *C. canariensis* and corresponds well with the time of the La Orotava megalandslide (Ancochea *et al.* 1990; Boulesteix *et al.* 2013).

Whereas central Tenerife was almost completely covered by catastrophic events until as recently as 0.13 Ma (Ancochea *et al.* 1999), the three paleo-islands of Teno, Anaga, and Roque del Conde, remained geologically stable since the mid-Pliocene (see Fig. 1). This suggests that these areas could have acted as refugia, allowing the survival of populations that otherwise disappeared in other parts of the island that were affected by volcanic activity. Support for this suggestion comes from the population genetic analysis. According to the central-marginal hypothesis (Eckert *et al.* 2008), spatial structure and genetic diversity should be higher in areas that have acted as refugia for the preservation of genotypes that went extinct in other areas and generally for the long-term persistence of populations (Hewitt 2000; Tzedakis *et al.* 2013; Gavin *et al.* 2014; Feliner 2014). The higher number of ancestral and endemic alleles, private fragments, and larger heterozygosity levels exhibited by the populations of the paleo-islands of Tenerife (Table 1), are congruent with the idea that these massifs acted as reservoirs of ancient genetic diversity and as refugia against volcanically induced extinction. Interestingly, La Gomera, an island that has been geologically quiescent since the Pliocene (Carracedo & Day, 2002), shows the highest  $H_j$  and percentage of polymorphic nuclear DNA fragments (Table 1), suggesting that this island might have acted in a similar way to the paleo-islands of Tenerife.

Further support for the extinction hypothesis comes from the BPEC analysis. Theoretical predictions of coalescent theory states that high frequency haplotypes have been present for a long time, and more recent ones are rare and derived from the commonest haplotypes (Posada & Krandall 2001). Additionally, a root haplotype is expected to have a

higher number of haplotype connections in the network, rather than being close to the tips. However, past extinction of haplotypes can obscure the inference, with younger haplotypes becoming the most prevalent, so accurately identifying the root haplotype is a challenging task. The fact that BPEC does not provide a single estimate of the haplotype network like Statistical Parsimony, but a finite (probability) distribution of haplotype trees - as well as the existence of the underlying migration model - allows this method to incorporate the uncertainty in the haplotype rooting. BPEC estimates the missing, extinct haplotype H11 as the most probable root of the haplotype network (Fig. 3b), although haplotypes H1, H2 and H8 are also associated with high posterior probabilities.

*Have paleo-islands acted as sources of genetic diversity within and towards other islands?*

The theory of Pleistocene climate refugia (Hewitt 2000) states that historically environmentally stable areas can act as sources of genetic diversity exporting migrant alleles to other, disturbed regions (Gavin *et al.* 2014). Paleo-islands could have played the same role in volcanic archipelagos, although in this case catastrophic geological events rather than climatic changes might be responsible for the observed patterns. BPEC provides support to this hypothesis, identifying the populations in the paleo-islands of Tenerife as the source areas of ancestral migration events to other adjacent islands, such as from Teno to La Gomera or from Anaga to Gran Canaria (Fig. 3a,b, Table 2). A third dispersal event from west Tenerife/La Gomera to La Palma is supported by the BEAST tree (Fig. 4), and a fourth dispersal event from La Palma to El Hierro is inferred by the BPEC analysis (Fig. 3). Similar patterns with the central Canaries as centres of dispersal events within the archipelago have been described in other animal and plant studies (Francisco-Ortega *et al.* 2002; Gómez *et al.* 2003; Sanmartín *et al.* 2008; Puppo *et al.* 2015; see also Mairal *et al.*, 2015). Moreover, a spatio-temporal pattern of colonization comparable to *Canarina*, showing Tenerife as the

center of dispersal events to adjacent islands in the last 1 Ma, can be found in the Canarian lineage of *Cistus* (0.33 (0.88-0.07) Ma; Guzmán & Vargas 2010), *Cheirolophus* (1 Ma; Vitales *et al.* 2014a), *Cistus monspeliensis* (0.93-0.20 Ma; Fernández-Mazuecos & Vargas 2011), and *Gallotia galloti* (0.8-0.9 Ma; Cox *et al.* 2010). This might be a consequence of the central geographic position of Tenerife within the archipelago – acting as a crossroad for dispersal events – but also of the concentration of plant genetic diversity in the paleo-islands.

The case of Gran Canaria is especially interesting. It shares the same haplotype with Tenerife (H1, Figs 2, 3). Low haplotypic diversity ( $H_n$  and  $H_d$  in Table 1) could be explained by a recent colonisation after a catastrophic event. The island was subject to intense volcanic activity during the Holocene (24 eruption events; Rodríguez-González *et al.* 2009), so extinction might explain its present low genetic diversity. Another possibility is related to the topography of Gran Canaria, where a network of ravines (locally known as "barrancos") connecting at their summits, could have facilitated gene flow among populations. Dispersal of *Canarina canariensis* seeds by *Gallotia* lizards (Rodríguez *et al.* 2008), probably using forest gaps and edges of roads as dispersal corridors (Delgado *et al.* 2007), might have helped to connect populations in the highly altered laurel forest of Gran Canaria.

In addition to dispersal events between-islands, the paleo-islands of Tenerife might have acted as sources of genetic diversity within Tenerife, exporting migrant alleles to other geologically unstable, disturbed areas. Our demographic analyses indicated a recent population expansion in two populations of east Tenerife close to Anaga (TF-Br and TF-Bj). Although this result should be taken with caution (the EBSP analysis supported a constant effective population size; Fig. S2), these two populations exhibited also a single cpDNA haplotype, which agrees with the idea of a recent colonisation. The areas where these populations are located (La Orotava Valley and Güimar Valley, respectively) have been

subjected to catastrophic volcanic events (Fig. 1b). Some authors (Thorpe *et al.* 1996; Gübitz *et al.* 2000) have proposed the existence of a corridor of suitable habitat along the northern coastal fringe of Tenerife to explain migrations of the reptiles *Gallotia galloti* and *Tarentola delalandii* from Anaga to the west. These dispersal events could also explain the patterns found in our BPEC analyses, which suggest Anaga populations as sources of migration events to other populations in eastern Tenerife and Gran Canaria (Fig. 3, Table 2).

Ongoing genetic connectivity between the populations in the paleo-islands of Tenerife and those from other areas is supported by the nuclear genome, which shows genetic admixture between west-Anaga, Teno and La Gomera (Fig. 5a). Admixture could be explained by the carrying of pollen by nectar-feeding birds between forest patches (e.g. ringing techniques have confirmed migration of the main pollinator (*Phylloscopus canariensis*) between Teno and Anaga; Alejandro González, pers. comm.). The fact that this connectivity is to some extent lost in the cpDNA might be explained by the cpDNA not being transported via pollen or, alternatively, by the small size and haploid nature of the chloroplast genome, which imply shorter coalescent times and less time to fix novel mutations for chloroplast markers (Avice 2000; Jakob & Blattner 2006). On the other hand, the widespread distribution of some cpDNA haplotypes across the archipelago (H1, H3 and H9, Fig. 2a) supports some gene flow among populations driven by seed dispersal. In the Canary Islands, birds have been cited as important vectors for the dispersal of fleshy fruits (Arevalo *et al.* 2007; Padilla *et al.* 2012), and the latter has been associated with frequent gene flow preventing speciation within widespread island endemics (García-Verdugo *et al.* 2014), such as in *Canarina*. This fits with what is known on the reproductive biology of *C. canariensis*, which is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011), while its fleshy fruits are dispersed by vertebrates such as *Gallotia* lizards (Valido *et al.* 2003, Rodríguez *et al.* 2008). An additional factor to explain frequent dispersal between-islands are

the eustatic sea-level shifts during the Pleistocene, which might have decreased geographic distance between-islands (Rijsdijk *et al.* 2014).

Finally, in addition to exporting migrant alleles to other islands and disturbed areas, the paleo-islands of Tenerife might have acted as cradles or sources of new genetic diversity. The higher haplotype and nucleotide diversity and higher Nei's gene diversity  $H_j$  exhibited by populations in these areas (Table 1) agree with their role as ancient refugia but also as sources of novel genetic diversity. Maximum topographic complexity is one of the main factors explaining species richness and high speciation rates within-islands (Whittaker *et al.* 2007). The rugged nature of the Tenerife paleo-islands has likely promoted genetic differentiation within these massifs. For example, divergences found between populations in east and west forest ranges within Anaga have been explained by the existence of deep ravines and the volcanic arc of Taganana (Fig. 1b), acting as geographical barrier to gene flow (Macías-Hernández *et al.* 2013). These divergences were also detected in *C. canariensis*.

## CONCLUSIONS

Traditionally, the distribution of genetic diversity within archipelagos is assumed to be structured around oceanic barriers, with between-island divergences expected to be larger than within-island differentiation. Here, we showed that within-island genetic patterns might be as strong as or stronger than those observed between-islands when they are associated with historical volcanic events. In *Canarina canariensis*, geographic patterns of genetic variation are structured around the paleo-islands of Tenerife, with a minor secondary effect due to oceanic barriers. Carine & Schaefer (2010) argued that although relatively short oceanic distances separate the Canary Islands, they might be responsible for the high diversity levels

found in the archipelago, acting as effective barriers to dispersal and promoting allopatric speciation. However, our results suggest that this hypothesis might not be valid for endemic species with widespread distributions across several islands (e.g., *Canarina*), for whom stretches of ocean are apparently less of a barrier than topographic relief within volcanic islands. Phylogeographic studies on other MIEs (multiple island endemics) are needed to confirm this hypothesis. The paleo-islands of Tenerife have probably acted as both genetic refugia and sources of new diversity within- and between-islands. The preservation of genotypes that became extinct everywhere else and the topographic complexity of the paleo-islands makes them potential “phylogeographical hotspots” (Médail & Diadema 2009) and reservoirs of unique genetic diversity, whose conservation should be prioritised.

#### ACKNOWLEDGEMENTS

We are grateful to Mitchell Cruzan and four anonymous reviewers, whose comments helped to significantly improve our manuscript. We would like to thank Alberto Herrero, Emilio Cano and Fatima Durán for their laboratory assistance. Field work could not have been conducted without the cooperation of the staff at the Jardín Botánico Canario Viera y Clavijo (Gran Canaria); Jacinto Leralta and Ángel Fernández (La Gomera); Félix Manuel Medina from the Cabildo of La Palma, and the Cabildo of Tenerife and El Hierro are thanked for helping with sampling permits and accommodation during field expeditions. We are grateful to Gonzalo Nieto-Feliner for a critical review of the manuscript. We also thank Águedo Marrero, Alfredo Valido, Aurelio Martín, Carlos García-Verdugo, Juli Caujapé-Castells, Lisa Pokorny, María Candelaria Rodríguez-Rodríguez, Marta Martínez, Moisés Soto, Nadir Álvarez, Nils Arrigo, Rubén Barone, Xavier Picó and Yulán Úbeda for all the fruitful discussions. This work was funded by Spanish Ministry of Science through projects CGL2009-1332-C03-01 and CGL2012-40129-C02-01). M. Alarcón was funded by the JAE-

Doc program (CSIC/FSE). MM and VC were supported by MINECO FPI predoctoral fellowships (BES-2010-037261 and BES-2013-065389, respectively).

## REFERENCES

- Ancochea E, Fuster J, Ibarrola E, Cendrero A, Coello J, Hernan F, Cantagrel JM, Jamond C (1990) Volcanic evolution of the island of Tenerife (Canary Islands) in the light of new K-Ar data. *Journal of Volcanology and Geothermal Research*, **44**, 231–249.
- Ancochea E, Huertas MJ, Cantagrel JM, Coello J, Fúster JM, Arnaud N, Ibarrola E (1999) Evolution of the Cañadas edifice and its implications for the origin of the Cañadas Caldera (Tenerife, Canary Islands). *Journal of Volcanology and Geothermal Research*, **88**, 177–199.
- Arevalo JR, Delgado JD, Fernández-Palacios JM (2007) Variation in fleshy fruit fall composition in an island laurel forest of the Canary Islands. *Acta oecologica*, **32**, 152–160.
- Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press.
- Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular biology and evolution*, **29**, 2157–2167.
- Báez M, JL Martín, P Oromí (2001) Diversidad taxonómica terrestre. En (Ed.) Fernández Palacios JM & Martín Esquivel JL, Naturaleza de las Islas Canarias. Ecología y Conservación. Tenerife: Ed. Turquesa, 65–76.
- Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms revisited. *American Journal of Botany*, **97**, 1296–1303.
- Bloomquist EW, Lemey P, Suchard MA (2010) Three roads diverged? Routes to phylogeographic inference. *Trends in ecology & evolution*, **25**, 626–632.
- Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular ecology*, **13**, 3261–3273.
- Boulesteix T, Hildenbrand A, Gillot PY, Soler V (2012). Eruptive response of oceanic islands to giant landslides: new insights from the geomorphologic evolution of the Teide–Pico Viejo volcanic complex (Tenerife, Canary). *Geomorphology*, **138**, 61–73.
- Boulesteix T, Hildenbrand A, Soler V, Quidelleur X, Gillot PY (2013) Coeval giant landslides in the Canary Islands: Implications for global, regional and local triggers of giant flank collapses on oceanic volcanoes. *Journal of Volcanology and Geothermal Research*, **257**, 90–98.
- Brown RP, Campos-Delgado R, Pestano J (2000) Mitochondrial DNA evolution and population history of the Tenerife skink *Chalcides viridanus*. *Molecular Ecology*, **9**, 1061–1067.
- Brown RP, Hoskisson PA, Welton JH, Baez M (2006) Geological history and within-island diversity: a debris avalanche and the Tenerife lizard *Gallotia galloti*. *Molecular Ecology*, **15**, 3631–3640.
- Brown RP, Yang (2010) Bayesian dating of shallow phylogenies with a relaxed clock. *Systematic biology*, **59**, 119–131.

- Cantagrel JM, Arnaud NO, Ancochea E, Fúster JM, Huertas MJ (1999) Repeated debris avalanches on Tenerife and genesis of Las Cañadas caldera wall (Canary Islands). *Geology*, **27**, 739-742.
- Carine MA, Schaefer H (2010) The Azores diversity enigma: why are there so few Azorean endemic flowering plants and why are they so widespread?. *Journal of Biogeography*, **37**, 77-89.
- Carracedo J (1994) The Canary Islands: an example of structural control on the growth of large oceanic-island volcanoes. *Journal of Volcanology and Geothermal Research*, **60**, 225-241.
- Carracedo JC, Day S, Guillou H, Rodríguez Badiola E, Canas JA, Pérez Torrado FJ (1998) Hotspot volcanism close to a passive continental margin: the Canary Islands. *Geological Magazine*, **135**, 591-604.
- Carracedo JC, Day S (2002) Canary Islands. Dunedin Academic Press Ltd.
- Carracedo JC, Badiola ER, Guillou H, Paterne M, Scaillet S, Torrado FP, ... Hansen A (2007) Eruptive and structural history of Teide Volcano and rift zones of Tenerife, Canary Islands. *Geological Society of America Bulletin*, **119**, 1027-1051.
- Carracedo JC (2014) The Teide Volcano, Tenerife, Canary Islands. *Landscapes and Landforms of Spain*, 257-272.
- Carson HL, Lockwood JP, Craddock EM (1990) Extinction and recolonization of local populations on a growing shield volcano. *Proceedings of the National Academy of Sciences*, **87**, 7055-7057.
- Carson HL, AR Templeton (1984) Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annual Review of Ecology and Systematics*, 97-131.
- Caujapé-Castells J (2010) General  $G_{ST}$  and  $\theta$  inflation due to biased intra-population sampling, and its consequences for the conservation of the Canarian Flora. *Conservation genetics* 11: 709-720.
- Clement M, Posada DCKA, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular ecology*, **9**, 1657-1659.
- Contreras-Díaz HG, Moya O, Oromí P, Juan C (2007) Evolution and diversification of the forest and hypogean ground-beetle genus *Trechus* in the Canary Islands. *Molecular Phylogenetics and Evolution*, **42**, 687-699.
- Cox SC, Carranza S, Brown RP (2010) Divergence times and colonization of the Canary Islands by *Gallotia* lizards. *Molecular Phylogenetics and Evolution*, **56**, 747-757.
- de Paz JP, J Caujapé-Castells (2013) A review of the allozyme data set for the Canarian endemic flora: causes of the high genetic diversity levels and implications for conservation. *Annals of botany* **111**, 1059-1073.
- Delgado JD, Arévalo JR, Fernández-Palacios JM (2007) Road edge effect on the abundance of the lizard *Gallotia galloti* (Sauria: Lacertidae) in two Canary Islands forests. *Biodiversity and Conservation*, **16**, 2949-2963.
- Dimitrov D, Arnedo MA, Ribera C (2008) Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: Evidence for long-term occupancy yet rapid recent speciation. *Molecular phylogenetics and evolution*, **48**, 596-614.
- Drummond AJ, A Rambaut (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, **7**, 214.
- Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, **4**, 359-361.
- Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-1188.

- Emerson, BC, Oromí P, Hewitt GM (1999) MtDNA phylogeography and recent intra-island diversification among Canary Island *Calathus* Beetles. *Molecular phylogenetics and evolution*, **13**, 149-158.
- Emerson BC, P Oromí (2005) Diversification of the forest beetle genus *Tharpius* on the Canary Island, and the evolutionary origins of island endemics. *Evolution*, **59**, 586-598.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online*, **1**, 47.
- Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology notes*, **7**, 574-578.
- Feliner GN (2014) Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 265-278.
- Fernández-Mazuecos M, Vargas P (2011) Genetically depauperate in the continent but rich in oceanic islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS One*, **6**, e17172.
- Fernández-Palacios JM, de Nascimento L, Otto R, Delgado JD, García-del-Rey E, Arévalo JR, Whittaker RJ (2011). A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226-246.
- Francisco-Ortega J, Santos-Guerra A, Kim SC, Crawford DJ (2000) Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany*, **87**, 909-919.
- Francisco-Ortega J, Fuertes-Aguilar J, Kim SC, Santos-Guerra A, Crawford DJ, Jansen RK (2002) Phylogeny of the Macaronesian endemic *Crambe* section *Dendrocrambe* (Brassicaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *American Journal of Botany*, **89**, 1984-1990.
- Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693-709.
- Fu YX (1996) New statistical tests of neutrality for DNA samples from a population. *Genetics*, **143**, 557-570.
- García-Verdugo C, Forrest AD, Fay MF, Vargas P (2010) The relevance of gene flow in metapopulation dynamics of an oceanic island endemic, *Olea europaea* subsp. *guanchica*. *Evolution*, **64**, 3525-3536.
- García-Verdugo C, Baldwin BG, Fay MF, Caujapé-Castells J (2014) Life history traits and patterns of diversification in oceanic archipelagos: a meta-analysis. *Botanical Journal of the Linnean Society*, **174**, 334-348.
- García-Verdugo C, Sajeva M, La Mantia T, Harrouni C, Msanda F, Caujapé-Castells J (2015) Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular ecology*, doi: 10.1111/mec.13060.
- Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodríguez-Sánchez F, Dobrowski SZ, ...Williams JW (2014) Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist*, **204**, 37-54.
- Gillespie RG, GK Roderick (2002) Arthropods on islands: colonization, speciation, and conservation. *Annual Review of Entomology*, **47**, 595-632.
- Gillespie R (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science*, **303**, 356-359.
- Gillespie RG, Clague DA (2009) Encyclopedia of Islands. University of California Press, Berkeley, CA.

- Gillespie RG, Roderick GK (2014) Evolution: Geology and climate drive diversification. *Nature*, **509**, 297-298.
- Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics*, **170**, 1261-1280.
- Gómez A, González-Martínez SC, Collada C, Climent J, Gil L (2003) Complex population genetic structure in the endemic Canary Island pine revealed using chloroplast microsatellite markers. *Theoretical and Applied Genetics*, **107**, 1123-1131.
- Gübitz T, Thorpe RS, Malhotra A (2000) Phylogeography and natural selection in the Tenerife gecko *Tarentola delalandii*: testing historical and adaptive hypotheses. *Molecular Ecology*, **9**, 1213-1221.
- Gübitz T, Thorpe RS, Malhotra A (2005) The dynamics of genetic and morphological variation on volcanic islands. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 751-757.
- Guillou H, Carracedo JC, Paris R, Pérez Torrado FJ (2004) Implications for the early shield-stage evolution of Tenerife from K/Ar ages and magnetic stratigraphy. *Earth and Planetary Science Letters*, **222**, 599-614.
- Guzmán B, Vargas P (2010) Unexpected synchronous differentiation in Mediterranean and Canarian *Cistus* (Cistaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, **12**, 163-174.
- Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, 591-600.
- Harter DE, Irl SD, Seo B, Steinbauer MJ, Gillespie R, Triantis KA., ...Beierkuhnlein C (2015) Impacts of global climate change on the floras of oceanic islands—Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics*, **17**, 160-183.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.
- Hochkirch A, Y Goerzig (2009) Colonization and speciation on volcanic islands: phylogeography of the flightless grasshopper genus *Arminda* (Orthoptera, Acrididae) on the Canary Islands. *Systematic Entomology*, **34**, 188-197.
- von Humboldt A (1814) *Voyage aux régions équinoxiales du nouveaux continent, fait en 1799-1804 par Al. de Humboldt et A. Bonpland*. Part 1. Paris.
- Huson DH, D Bryant (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254-267.
- Jakob SS, Blattner FR (2006) A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Molecular Biology and Evolution*, **23**, 1602-1612.
- Jones KE, Reyes-Betancort JA, Hiscock SJ, Carine MA (2014) Allopatric diversification, multiple habitat shifts, and hybridization in the evolution of *Pericallis* (Asteraceae), a Macaronesian endemic genus. *American journal of botany*, **101**, 637-651.
- Juan C, Ibrahim KM, Oromi P, Hewitt GM (1996) Mitochondrial DNA sequence variation and phylogeography of *Pimelia* darkling beetles on the island of Tenerife (Canary Islands). *Heredity*, **77**, 589-598.
- Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution*, **15**, 104-109.
- Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, **30**, 3059-3066.

- Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*, **87**, 482-498.
- Kim SC, McGowen MR, Lubinsky P, Barber JC, Mort ME, Santos-Guerra A (2008) Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS One*, **3**, e2139.
- Knowles LL (2008) Why does a method that fails continue to be used?. *Evolution*, **62**, 2713-2717.
- Kunkel G (1991) Flora y vegetación del Archipiélago Canario. Tratado florístico. 2ª parte. Dicotiledóneas. 312 pp. Edirca. Madrid.
- Librado P, J Rozas (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.
- Lyell C (1855) A manual of elementary geology. John Murray, London.
- Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91-99.
- Macías-Hernández N, Bidegaray-Batista L, Emerson BC, Oromí P, Arnedo M (2013) The imprint of geologic history on within-island diversification of woodlouse-hunter spiders (Araneae, Dysderidae) in the Canary Islands. *Journal of Heredity*, **104**, 341-356.
- Mairal M, Pokorny L, Aldasoro JJ, Alarcón M, Sanmartín I (2015) Ancient vicariance and climate-driven extinction continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular ecology*, **24**, 1335-1354.
- Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human biology*, **76**, 173-190.
- Manolopoulou I, Legarreta L, Emerson BC, Brooks S, Tavaré S (2011) A Bayesian approach to phylogeographic clustering. *Interface focus*, rfs20110054.
- Manolopoulou I, BC Emerson (2012) Phylogeographic Ancestral Inference Using the Coalescent Model on Haplotype Trees. *Journal of Computational Biology*, **19**, 745-755.
- Manolopoulou I, Hille A (in prep.) BPEC: An R package for Bayesian Phylogeographic and Ecological Clustering.
- Mansion G, Parolly G, Crowl AA, Mavrodiev E, Cellinese N, Oganessian M, ... Borsch, T (2012) How to handle speciose clades? Mass taxon-sampling as a strategy towards illuminating the natural history of *Campanula* (Campanuloideae). *PloS one*, **7**, e50076.
- Médail F, P Quézel (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden*: 112-127.
- Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333-1345.
- Moya Ó, Contreras-Díaz HG, Oromí P, Juan C (2004) Genetic structure, phylogeography and demography of two ground-beetle species endemic to the Tenerife laurel forest (Canary Islands). *Molecular Ecology*, **13**, 3153-3167.
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, **76**, 5269-5273.
- Nylander J (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University 2.
- Olesen JM, Alarcon M, Ehlers BK, Aldasoro JJ, Roquet C (2012) Pollination, biogeography and phylogeny of oceanic island bellflowers (Campanulaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, **14**, 169-182.
- Padilla DP, González-Castro A, Nogales M (2012) Significance and extent of secondary seed dispersal by predatory birds on oceanic islands: the case of the Canary archipelago. *Journal of Ecology*, **100**, 416-427.

- Patiño J, Carine M, Mardulyn P, Devos N, Mateo RG, González-Mancebo JM, Shaw J, Vanderpoorten A (2015) Approximate Bayesian Computation Reveals the Crucial Role of Oceanic Islands for the Assembly of Continental Biodiversity. *Systematic Biology*, doi: 10.1093/sysbio/syv013.
- Pokorny L, Oliván G, Shaw AJ (2011) Phylogeographic patterns in two southern hemisphere species of Calyptrochaeta (Daltoniaceae, Bryophyta). *Systematic Botany*, **36**, 542-553.
- Posada D, KA Crandall (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in ecology & evolution*, **16**, 37-45.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Puppo P, Curto M, Velo-Antón G, Pérez de Paz PL, Meimberg H (2014) The influence of geological history on diversification in insular species: genetic and morphological patterns of *Micromeria* Benth.(Lamiaceae) in Tenerife (Canary archipelago). *Journal of Biogeography*, **41**, 1871-1882.
- Puppo P, Curto M, Gusmão-Guedes J, Cochofel J, de Paz PL P, Bräuchler C, Meimberg H (2015) Molecular phylogenetics of *Micromeria* (Lamiaceae) in the Canary Islands, diversification and inter-island colonization patterns inferred from nuclear genes. *Molecular phylogenetics and evolution*, **89**, 160-170.
- Rambaut A, A Drummond (2009) FigTree v1. 3.1. Computer program and documentation distributed by the author at <http://tree.bio.ed.ac.uk/software>.
- Rambaut A, Drummond AJ, Suchard M (2013) Tracer v1. 6.
- Reyes-Betancort JA, Santos-Guerra A, Guma IR, Humphries CJ, Carine MA (2008). Diversity, rarity and the evolution and conservation of the Canary Islands endemic flora. *Anales del Jardín Botánico de Madrid*, **65**, 25.
- Rijsdijk KF, Hengl T, Norder SJ, Otto R, Emerson BC, Ávila SP, ... Fernández-Palacios JM (2014) Quantifying surface-area changes of volcanic islands driven by Pleistocene sea-level cycles: biogeographical implications for the Macaronesian archipelagos. *Journal of Biogeography*, **41**, 1242-1254.
- Rodríguez A, Nogales M, Rumeu B, Rodríguez B (2008) Temporal and spatial variation in the diet of the endemic lizard *Gallotia galloti* in an insular Mediterranean scrubland. *Journal of Herpetology*, **42**, 213-222.
- Rodríguez-Gonzalez A, Fernandez-Turiel JL, Perez-Torrado FJ, Hansen A, Aulinas M, Carracedo JC, ... Paterne M (2009) The Holocene volcanic history of Gran Canaria island: implications for volcanic hazards. *Journal of Quaternary Science*, **24**, 697-709.
- Rodríguez-Rodríguez MC, Valido A (2011) Consequences of plant-pollinator and floral-herbivore interactions on the reproductive success of the Canary Islands endemic *Canarina canariensis* (Campanulaceae). *American Journal of Botany*, **98**, 1465-1474.
- Rohlf F (1998) NTSYS-pc version 2.0. Numerical taxonomy and multivariate analysis system. Exeter software, Setauket, New York.
- Salomone N, Emerson BC, Hewitt GM, Bernini F (2002) Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. *Molecular Ecology*, **11**, 79-89.
- Sanmartín I, Van Der Mark P, Ronquist F (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35**, 428-449.
- Sanmartín I, Anderson CL, Alarcón M, Ronquist F, Aldasoro JJ (2010) Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters*, **6**, 703-707.
- Santos-Guerra A (2001) Flora vascular nativa. Naturaleza de las Islas Canarias. Ecología y conservación (ed. by J.M. Fernández-Palacios and J.M. Martín-Esquível), pp. 185-198. Publicaciones Turquesa, Santa Cruz de Tenerife, Spain.

- Accepted Article
- Stuessy TF, Takayama K, López-Sepúlveda P, Crawford DJ (2014) Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society*, **174**, 276-288.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619-633.
- Thorpe RS, Black H, Malhotra A (1996) Matrix correspondence tests on the DNA phylogeny of the Tenerife lacertid elucidate both historical causes and morphological adaptation. *Systematic Biology*, **45**, 335-343.
- Trusty JL, Olmstead RG, Santos-Guerra A, Sá-Fontinha S, Francisco-Ortega J (2005) Molecular phylogenetics of the Macaronesian-endemic genus *Bystropogon* (Lamiaceae): palaeo-islands, ecological shifts and interisland colonizations. *Molecular Ecology*, **14**, 1177-1189.
- Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696-704.
- Valido A, Nogales M, Medina FM (2003) Fleshy fruits in the diet of Canarian lizards *Gallotia galloti* (Lacertidae) in a xeric habitat of the island of Tenerife. *Journal of Herpetology*, 741-747.
- Vekemans X, Beauwens T, Lemaire M, Roldán-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139-151.
- Vitales D, Garnatje T, Pellicer J, Vallès J, Santos-Guerra A, Sanmartín I (2014a) The explosive radiation of *Cheirolophus* (Asteraceae, Cardueae) in Macaronesia. *BMC Evolutionary Biology*, **141**, 118.
- Vitales D, García-Fernández A, Pellicer J, Vallès J, Santos-Guerra A, Cowan RS, ... Garnatje T (2014) Key Processes for *Cheirolophus* (Asteraceae) Diversification on Oceanic Islands Inferred from AFLP Data. *PLoS one*, **9**, e113207.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, ... Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, **23**, 4407-4414.
- Wang Q, Zhou SL, Hong DY (2013) Molecular phylogeny of the platycodonoid group (Campanulaceae s. str.) with special reference to the circumscription of *Codonopsis*. *Taxon*, **62**, 498-504.
- Whitlock R, Hipperson H, Mannarelli M, Butlin RK, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, **8**, 725-735.
- Whittaker RJ, Ladle R J, Araújo MB, Fernández-Palacios JM, Delgado J, Arévalo J (2007) The island immaturity–speciation pulse model of island evolution: an alternative to the “diversity begets diversity” model. *Ecography*, **30**, 321-327.
- Zaczek K, Troll V, Cachao M, Ferreira J, Deegan F, Carracedo JC, ...Burchardt S (2015) Nannofossils in 2011 El Hierro eruptive products reinstate plume model for Canary Islands. *Scientific Reports*.
- Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, **8**, 907-913.
- Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107-2121.

#### DATA ACCESSIBILITY.

DNA sequences: Genbank accessions KP797991 to KP798432.

GenBank accessions, sampling locations and/or online-only appendices uploaded as online supplemental material.

NEXUS files for the concatenated cpDNA dataset and original AFLPs Matrix of *Canarina canariensis*, Dryad doi:10.5061/dryad.j4103.

#### AUTHOR CONTRIBUTIONS

M.M., M.A., J.J.A. and I.S. designed the study; M.M. carried out the field work, M.A. and

J.J.A. contributed some samples; M.M. performed the research and analyzed the data under

the supervision of M.A. and I.S.; I.M., V.C., I.S., and M.M. performed the BPEC analyses;

M.M. wrote the paper with help from I.S, M.A and J.J.A. M.M. and I.S. revised the paper. All

authors contributed with comments and approved the final version.

#### TABLES

**Table 1.** Descriptors of within-population genetic diversity in the cpDNA haplotypes and AFLP markers for each population studied of *C. canariensis*. Abbreviations: H(n): number of haplotypes; H(d): haplotype diversity;  $\pi$ : nucleotide diversity;  $\theta$ : nucleotide heterozygosity; H<sub>j</sub> (se): Nei's gene diversity (standard error). Geographical locations for population codes are shown in Figure 2.

Island (Paleo-island)	Population	Haplotypes						AFLPs			
		Nº samples	Haplotype	H (n)	H (d)	$\pi$	$\theta$	Nº samples	Nº of polymorphic fragments (% in brackets)	H <sub>j</sub> (se)	Number of private fragments
Gran Canaria	GC-Cc	8	H1	1	0	0	0	8	409 (68.7)	0.2344 (0.0085)	0
Gran Canaria	GC-Tm	11	H1	1	0	0	0	NA	NA	NA	NA
Gran Canaria	GC-Es	10	H1	1	0	0	0	NA	NA	NA	NA
Gran Canaria	GC-Ag	11	H1	1	0	0	0	NA	NA	NA	NA
<b>Tenerife (Anaga)</b>	An-W	6	H1, H2, H5, H9	4	0.8	0.00112	0.00136	7	401 (67.4)	0.2677 (0.0088)	4
Tenerife (Anaga)	An-E	10	H1, H2	2	0.2	0.00018	0.00031	10	345 (58.0)	0.2303 (0.0080)	0
Tenerife (Anaga)	An-C	11	H1, H5	2	0.327	0.00015	0.00015	11	375 (63.0)	0.2453 (0.0078)	1
Tenerife	TF-Bj	4	H1	1	0	0	0	NA	NA	NA	NA
Tenerife	TF-Br	6	H1	1	0	0	0	NA	NA	NA	NA
<b>Tenerife (Roque del Conde)</b>	Conde	11	H1, H9, H7	3	0.618	0.00121	0.00091	NA	NA	NA	NA
Tenerife (Teno)	Ten-Ep	8	H9	1	0	0	0	9	351 (59.0)	0.2449 (0.0082)	0
<b>Tenerife (Teno)</b>	Ten-A	7	H8, H9, H10	3	0.524	0.00038	0.00054	7	415 (69.7)	0.2997 (0.0088)	2
Gomera	GO	12	H8, H9	2	0.303	0.00013	0.00015	12	473 (79.5)	0.3012 (0.0077)	1
La Palma	LP-Ba	10	H3	1	0	0	0	9	285 (47.9)	0.1923 (0.0080)	0
La Palma	LP-Lt	10	H3, H4	2	0.2	0.00009	0.00016	10	348 (58.5)	0.2309 (0.0079)	0
El Hierro	EH	9	H3, H6	2	0.5	0.00022	0.00016	9	303 (50.9)	0.2115 (0.0082)	1

**Table 2.** Summary of results from the Bayesian Phylogeographic and Ecological Clustering (BPEC) analyses. Selected values for parameters MaxMig (maximum number of migrations) and ds (relaxed parsimony assumption) are shown. Ancestral locations (Anc. Loc.) for migration events are shown with their associated posterior probabilities.

	Maxmig	ds	Anc. Loc. 1	Anc. Loc. 2	Anc. Loc. 3	Root node	Root node probability	n° of clusters
<b>TOTAL</b>	8	10	Conde (0.13)	An-W (0.10)	LP-Lt (0.086)	H11 (missing)	0.17	5
<b>Western Clade (WT + LP + GO + EH)</b>	4	3	Ten-A (0.40)	Lp-Lt (0.31)	EH (0.18)	H11 (missing)	0.18	3
<b>Eastern Clade (ET + GC)</b>	1	3	An-W (0.212)	An-C (0.122)	An-E (0.116)	H11 (missing)	0.16	2
<b>Western Clade + Conde</b>	4	3	Conde (0.416)	Ten-A (0.197)	GO (0.190)	H11 (missing)	0.18	3
<b>Eastern Clade + Conde</b>	1	3	Conde (0.23)	An-W (0.17)	An-C (0.086)	H11 (missing)	0.16	3

**Table 3.** Hierarchical analysis of molecular variance (AMOVA) for *C. canariensis* based on allelic variation at different levels: (A) among groups, (B) among populations within groups and (C) within populations. Geographical locations of populations are shown in Figure 2.

AMOVA groups	N° of groups (K)	Levels			F-statistics		
		A	B	C	Fsc	Fst	Fct
[GC + Anaga] [An-W + Ten-A] [Ten-Ep] [GO + LP + EH]	4	17.44	13.06	69.5	0.15820	0.305	0.17439
[GC + An-E] [An-C] [Ten-Ep] [An-W + Ten-A] [GO + LP + EH]	5	18.92	11.21	69.88	0.13819	0.30123	0.18918
[GC] [TF] [GO] [LP] [EH]	5	3.62	24.27	72.10	0.25186	0.27897	0.03624
[GC] [An-E + An-C] [An-W] [Ten-Ep] [Ten-A] [GO] [LP] [EH]	8	16.63	11.30	72.07	0.13558	0.27934	0.16631

**Table 4.** Phylogeographic breaks with divergence times reported in the literature between the paleo-islands of Tenerife.

Organism	Phylogeographical disjunction and groups	Divergence (Ma)	Markers	Reference
Species complex in darkening beetle <i>Pimelia</i>	East Tenerife / West Tenerife	5.5 Ma	Mitochondrial (COI) and nuclear (ITS-1)	Juan <i>et al.</i> 1996
Lizard <i>Gallotia galloti</i>	Western/ North-eastern lineages	0.7 Ma	Mitochondrial cytochrome b	Thorpe <i>et al.</i> 1996
Beetle <i>Calathus abaxoides</i>	Teno / Anaga	350.000 years	Two mitochondrial (COI and COII)	Emerson <i>et al.</i> 1999
Skink <i>Chalcides viridanus</i>	Teno/ Anaga	1.1 Ma	Two mitochondrial (12S and 16S)	Brown <i>et al.</i> 2000
Gecko <i>Tarentola delalandii</i>	Clade 1. Teno + Roque del Conde/Anaga . Clade 2. Teno/Roque del Conde	Clade 1. 9–10 Ma Clade 2. 7.6 Ma	Mitochondrial cytochrome b	Gübitz <i>et al.</i> 2000
Mite <i>Steganacarus carlosi</i>	Clade 1. Teno/Anaga. Clade 2. Roque del Conde / Anaga	Clade 1. 3.2 Ma Clade 2. 25 - 3.6 Ma	Mitochondrial cytochrome oxidase 1	Salomone <i>et al.</i> 2002
Beetles <i>Eutrichopus</i>	Teno ( <i>E. gonzalezi</i> ) / Anaga ( <i>E. canariensis</i> )	0.7 Ma	Mitochondrial (COII marker)	Moya <i>et al.</i> 2004
species complex in beetle <i>Tarphius</i>	Teno / Anaga	1.2 Ma (1-1.4)	two mitochondrial (COI and COII)	Emerson & Oromí 2005
Beetle <i>Trechus</i>	Clade 1. Anaga + Teno ( <i>T. antonii</i> , <i>T. tenoensis</i> , <i>T. felix</i> ) / Anaga ( <i>T. fortunatus</i> ) Clade 2. <i>T. flavocintus</i> ; Teno / East Tenerife	Clade 1. 1.73 Ma (HPD:1.48–2.01) Clade 2. aprox 0.75 Ma	Part of mitochondrial genes cytochrome oxidase I and II (Cox1 and Cox2), and nuclear (ITS 2)	Contreras-Díaz <i>et al.</i> 2007
Spider <i>Pholcus</i>	Anaga ( <i>P. malpaisensis</i> , <i>P. knoeseli</i> ) / Teno ( <i>P. intricatus</i> , <i>P. mascaensis</i> / <i>P. tenerifensis</i> and <i>P. roquensis</i> )	3.93 Ma (HPD: 2.2–5.88)	Four mitochondrial (COI, 16S, NADH and tRNA <sup>Leu</sup> ). Morphological data.	Dimitrov <i>et al.</i> 2008
Grasshoper <i>Arminda brunneri</i>	Anaga + Güimar / Teno + Roque del Conde	1 – 0.17 Ma	Two mitochondrial (12s rRNA, ND5) and two nuclear gene fragments (28s rRNA, ITS2)	Hochkirch & Goerzig 2009
Spider <i>Dysdera verneui</i>	Teno / Anaga	3.94 Ma (HPD: 5.1-2.7)	mitochondrial ( <i>cox1</i> )	Macías-Hernández <i>et al.</i> 2013
Plant <i>Pericallis</i>	Teno ( <i>P. echinata</i> ) / Anaga ( <i>P. tussilaginis</i> )	2.87 Ma (HPD: 1.55–4.76)	nuclear ITS	Jones <i>et al.</i> 2014
Plant <i>Micromeria</i>	Teno ( <i>M. densiflora</i> ) / Anaga ( <i>M. teneriffae</i> , <i>M. glomerata</i> and <i>M. rivas-martinezii</i> ) + Central group	5.2 Ma	8 nuclear loci	Puppo <i>et al.</i> 2014

## Figure Captions

**Figure 1.** a) Geographic map of the Canary Island archipelago, with the code for each island used in the text and their age of emergence (million years, Ma). b) Map of Tenerife showing main Pleistocene landslides and other geological events cited in the text. The old basaltic series corresponding to the paleo-islands are shown in the three contoured massifs at the edges of the island. Lines show valleys created after landslides. Dashed lines indicate the post-erosional Las Cañadas volcano. Arrows indicate the directions of the landslides. The graph on the right shows the main Tenerife catastrophic events that are commented in the text; the numbers in the legend refer to those in the map. Ages and maps adapted from: Carracedo *et al.* 1998; Ancochea *et al.* 1999; Cantagrel *et al.* 1999; Guillou *et al.* 2004; Boulesteix *et al.* 2012; Zaczek *et al.* 2015.

**Figure 2.** Haplotype distribution and network inferred for the chloroplast markers by TCS. a) Pie charts show the geographical location of populations and the frequency of occurrence of each haplotype, circle size is proportional to population size. Population codes are given in

Table S2. b) Statistical Parsimony network inferred by TCS. Each haplotype is shown in a different colour; circle size is proportional to its frequency among populations.

**Figure 3.** Results from the Bayesian Phylogeographic and Ecological Clustering (BPEC) analyses for the cpDNA markers. a) Phylogeographical clusters (coloured blobs) and ancestral location for migration events (denoted by arrows). The contour regions are centred at the 'centre' of each population cluster, and the shaded areas show the radius of 50% concentration contours around it. Locations situated beyond the clusters could also belong to these clusters, but with low probability; in the case of Conde, it suggests a mixed composition, with this population as source of migrant alleles to the east and western clusters. b) Haplotype network receiving the highest posterior probability. The small black circle H11 indicates an unobserved (missing) ancestral haplotype.

**Figure 4.** Maximum Clade Credibility (MCC) tree obtained from the BEAST analysis of cpDNA haplotypes, showing mean ages (above branches) and 95% HPD credible intervals. Numbers below branches indicate Bayesian posterior clade support values. Codes for tips (H1 to H10) correspond to the haplotypes shown in Fig. 2.

**Figure 5.** Results from the analysis of AFLP markers. a) Histograms showing the Bayesian clustering of individuals within populations (STRUCTURE); colours represent the proportion of individual membership to each inferred Bayesian group. Dotted lines indicate barriers to gene flow and their percentage, as inferred by BARRIER. b) Split network with points coloured according to location, as indicated in the legend. Codes for populations found in Table S2.

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Figure 1

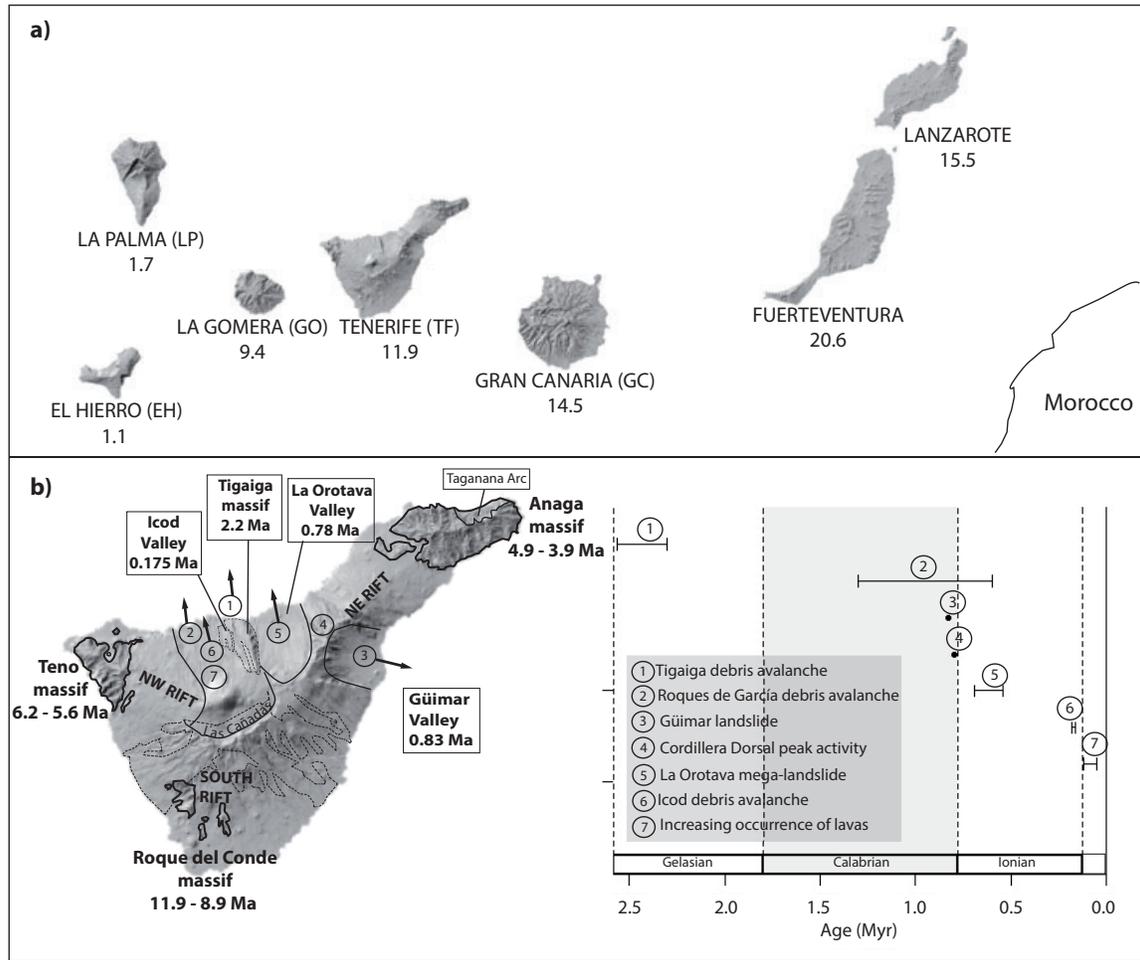


Figure 2

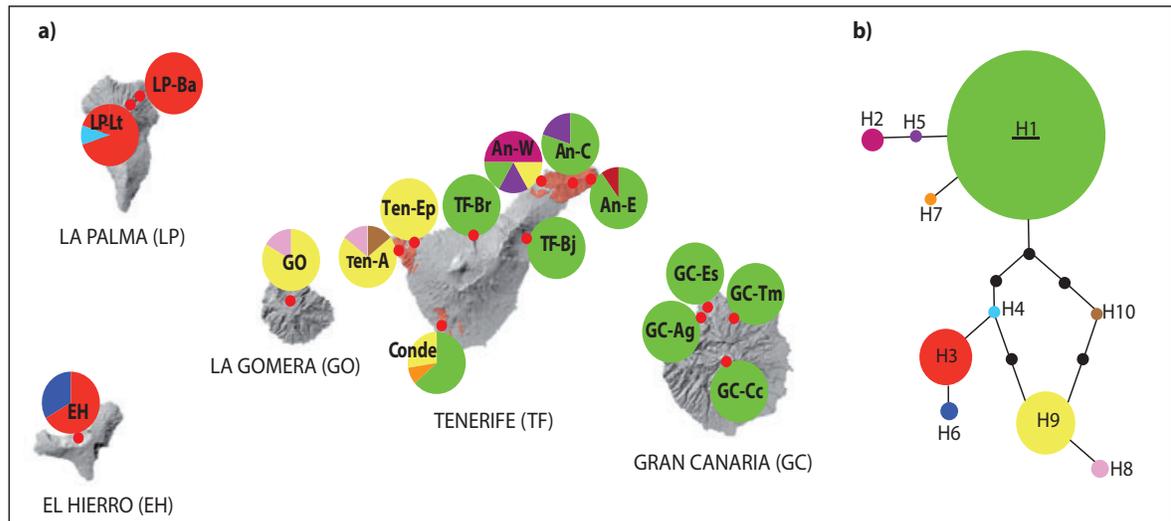


Figure 3

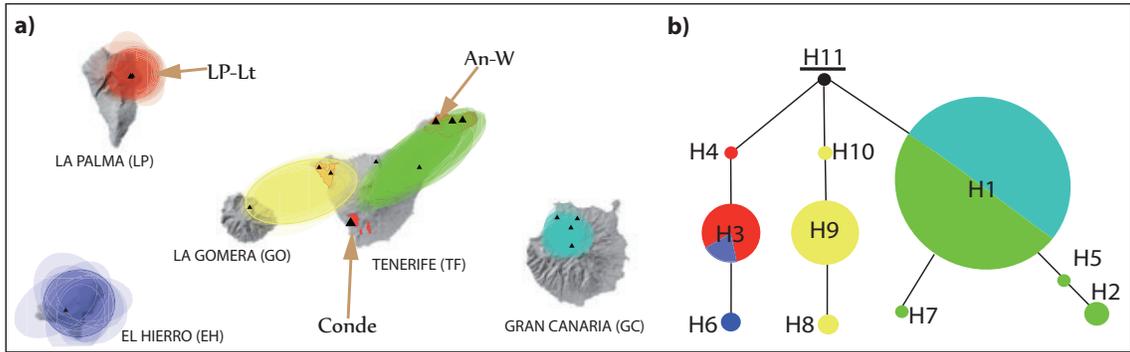


Figure 4

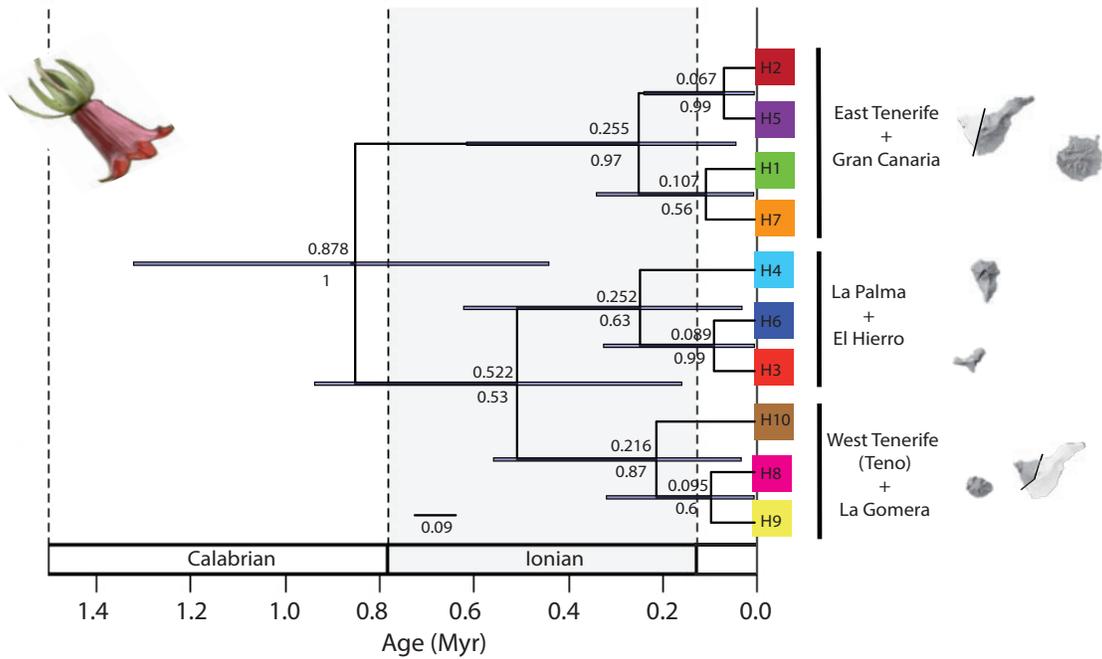


Figure 5

