Phylogeography of *Arenaria balearica* L. (Caryophyllaceae): Evolutionary history of a disjunct endemic from the Western Mediterranean continental islands

Javier BOBO-PINILLA[1, 5,*], Sara BARRIOS DE LEÓN[1], Jaume SEGUÍ COLOMAR[1], Giuseppe FENU[2], Gianluigi BACCHETTA[3], Julio PEÑAS DE GILES[4], María Montserrat MARTÍNEZ-ORTEGA[1, 5]

[1]Departamento de Botánica, Universidad de Salamanca, E-37007, Salamanca (Spain)

[2] "Sapienza" Università di Roma, Dipartimento di Biologia Ambientale, 00185 Roma (Italy)

[3]Università degli Studi di Cagliari, Dipartimento di Scienze della Vita e dell'Ambiente, Centro Conservazione Biodiversità (CCB), 09123 Cagliari (Italy)

[4]Departamento de Botánica, Universidad de Granada, E-18071 Granada (Spain)

[5] Banco de ADN vegetal, Banco Nacional de ADN, Salamanca (Spain)

[*]Correspondence: javicastronuevo@usal.es

Deleted: Disentangling the evolutionary

Deleted: the flora of

Deleted: since the Neogene: Phylogeography of *Arenaria balearica* L. (Caryophyllaceae)

Abstract: Although it has been traditionally accepted that *Arenaria balearica* (Caryophyllaceae) could be a relict Tertiary plant species, this has never been tested. Nor have the palaeohistorical reasons underlying the highly fragmented distribution of the species in the Western Mediterranean region been investigated. We have analysed AFLP data (213) and plastid DNA sequences (226) from a total of 250 plants from 29 populations sampled throughout the entire distribution range of the species in Majorca, Corsica, Sardinia, and the Tuscan Archipelago. The AFLP data analyses indicate very low geographic structure and population differentiation. Based on plastid DNA data, six alternative phylogeographic hypotheses were tested using Approximate Bayesian Computation (ABC). These analyses revealed ancient area fragmentation as the most probable scenario, which is in accordance with the star-like topology of the parsimony network that suggests a pattern of long term survival and subsequent *in situ* differentiation. Overall low levels of genetic diversity and plastid DNA variation were found, reflecting evolutionary stasis of a species preserved in locally long-term stable habitats.

Deleted: experimentally

Deleted: disjunct

Deleted: cpDNA

Deleted: 53

Deleted: 222

Deleted: Our results point to a Messinian origin of *A. balearica* and suggest that its present distribution is not directly related to the splitting of the Hercynian Massif.

Deleted: The

Deleted: based on cpDNA data

Deleted: radiative evolution and implies that all haplotypes were derived probably

Deleted: and many of them as recently as during the Pleistocene from a single ancient ancestor. The nested clade phylogeographic analysis performed identifies restricted gene flow with isolation by distance as the main historical process affecting the genetic structure.

Deleted: cpDNA

Deleted: Our data shed light on the complex phylogeographic patterns within the Western Mediterranean region.

Deleted: **Key words:** AFLPs - *Arenaria* – island evolution – Messinian - phylogeography - chloroplast DNA - Mediterranean. ¶

1. Introduction

Within the Mediterranean global biodiversity hotspot, the Tyrrhenian Islands represent ca. 22% of the total surface, and include a high percentage of endemic taxa (ca. 10-20%; Contandriopolous, 1990; Médail & Quézel, 1997; Bacchetta & Pontecorvo, 2005; Cañadas et al., 2014). Some of these endemic plant species show narrow distributions (Médail & Quézel, 1999; Thompson, 2005; Fenu et al., 2010; Bacchetta, Fenu & Mattana, 2012), but others are distributed in the major Western Mediterranean islands. Some endemic plant species shared by Corsica, Sardinia, and the Balearic Islands have been designated "Hercynian endemics" (Mansion et al., 2008) and are often considered palaeoendemic in the broad sense of the term (i.e., ancient or relict taxa often systematically isolated, Favarger & Contandriopolous, 1961; Greuter, 1995; Quézel, 1995). The present distribution of such Hercynian endemic species has been attributed to the Oligocenic connections among the Western Mediterranean islands (Greuter, 1995; Quézel, 1995; Thompson, 2005), but this has not been tested in all cases. Additionally, the term "palaeoendemic" has been restricted in concept (Thompson, 2005) to include only clearly ancient isolated species in large genera (or monotypic genera) that usually show little variability. There are some endemic species showing distribution patterns that seem to be concordant with the geological history of the Western Mediterranean continental fragments, which have been commonly considered palaeoendemics. But, as it has not been yet demonstrated that they are of ancient origin and do not seem to be highly isolated within large genera, they do not fit into the restrictive concept of palaeoendemism proposed by Thompson (2005). These species are referred to as disjunct endemics and *Arenaria balearica* L. (Caryophyllaceae) is a good example.

Deleted: lodge

Deleted: (

Deleted: ; Thompson, 2005; Molins et al., 2011).

Deleted:) but this has not been tested in all cases

Deleted: . from the family

The Mediterranean region has been affected by dramatic palaeogeographical events and by considerable climatic changes during the Neogene (Kadereit & Comes, 2005), which have influenced the structure and composition of the flora, determined plant species distributions, and influenced intraspecific genetic variability of species over the past few million years (Thompson, 2005; Médail & Diadema, 2009).

Like most Western Mediterranean islands, Corsica, Sardinia, and Majorca are of the continental type and have been separated from each other by tectonic and glacio-eustatic processes (Alvarez, 1972; Alvarez, Cocozza & Wezel, 1974; Rosenbaum, Lister & Duboz, 2002; Mansion et al., 2008; Mayol et al., 2012). The progressive, post-Oligocene fragmentation of land masses previously constituting part of the Hercynian belt has been described elsewhere (Alvarez, 1972; Alvarez et al., 1974; Rosenbaum et al., 2002; Speranza et al., 2002; Meulenkamp & Sissingh, 2003; Mansion et al., 2008; Salvo et al., 2010).

The Tuscan Archipelago consists of seven small islands and several islets of different geological origins, which are also tectonic fragments that were once integrated within the Hercynian massif (Salvo et al., 2010). The granitic basement of Montecristo appears also to be partly a result of the volcanic activity displayed in the area over the past 10 Ma, giving rise as well to other volcanic islands in the region, such as Capraia (Carmignani & Lazzarotto, 2004).

With the closure of the Strait of Gibraltar (ca. 5.59 Ma; Hsü, 1972; Garcia-Castellanos et al., 2009) the Messinian Salinity Crisis of the Late Miocene was initiated and some connections were established between North Africa, Corsica, Sardinia, and continental Europe, as well as between the Balearic Islands and Iberia; however, no evidence of direct terrestrial corridors between Corsica or Sardinia and Balearic Islands have been documented (Alvarez, 1972; Alvarez et al., 1974; Rosenbaum et al., 2002; Mansion et

Deleted: formidable

Deleted: bio

Deleted: Late Tertiary and Quaternary

Deleted: have contributed to shape

Deleted: have modelled

Deleted: post-Oligocene (which started ca. 30 Ma [million years ago]) progressive fragmentation of land masses previously constituting part of the Hercynian belt

Deleted: started

Deleted: but

al., 2008; Salvo et al., 2010). During the Messinian, the Tuscan Archipelago may have connected Corsica, Sardinia, and the Italian Peninsula. The cycles of desiccation and transgression of the Mediterranean Sea in this period enabled interchanges of lineages that predated the Messinian Salinity Crisis in all these territories (e.g. Salvo et al., 2010; Molins et al., 2011). The subsequent reopening of the Strait of Gibraltar (ca. 5.33 Ma; Krijgsman et al., 1999; Garcia-Castellanos et al., 2009) caused partial extinction and isolation of previously connected populations and seems to have promoted vicariant speciation and population divergence at least in some cases (e.g. *Quercus ilex* L. in Lumaret et al., 2002; *Anchusa crispa* Viv. in Quilichini, Debussche & Thompson, 2004; *Borago* L. in Selvi, Coppi & Bigazzi, 2006; *Abies* spp. in Terrab et al., 2007; *Anchusa* L. in Bacchetta et al., 2008; *Anchusa* L. in Coppi, Mengoni & Selvi, 2008; Rodríguez-Sánchez et al., 2008; Salvo et al., 2008; *Cephalaria* gr. *squamiflora* (Sieber) Greuter in Rosselló et al., 2009; Bacchetta et al., 2012; *Aquilegia* L. in Garrido et al., 2012).

The subsequent establishment of the Mediterranean climate (ca. 3-2 Ma) promoted the expansion of xerophytic elements and typically Mediterranean taxa (Suc, 1984; Thompson, 2005). Later, the cyclical climatic oscillations of the Pleistocene (ca. 1.8-0.01 Ma) also significantly shaped the genetic structure and spatial distribution of the biota, leading to population differentiation and eventually to speciation (Hewitt, 1999). Particularly, during the Pleistocene glacial maxima, sea level was approximately 120-150 m lower than at present (Yokohama et al., 2000; Church et al., 2001; Clark & Mix, 2002; Lambeck & Purcell, 2005) and the Corsican and Sardinian coastlines were directly connected by land bridges (Salvo et al., 2010). These connections facilitated exchanges of plant species and have alternatively limited or favoured gene flow between populations of species distributed in both islands and probably also among them and the Tuscan islets (Figure 1).

Deleted: of biota

Deleted: documented

Deleted: Quaternary

Deleted: of the

Several Mediterranean disjunct endemic species show high levels of morphological stability despite long-term isolation among populations distributed in different continental fragment islands (Molins et al., 2011, 3.2 Ma). The constancy of morphological characters over long periods has frequently been related to low molecular evolutionary rates, although this may not be completely clear in all cases (Casane & Laurenti, 2013). Recently, high levels of plastid DNA (cpDNA) diversity have been reported for the Tyrrhenian endemic *Thymus herba-barona* Loisel. (Molins et al., 2011). Also the apparent inconsistency between the fact that the Mediterranean region has undergone dramatic geological as well as climatic changes and the long persistence of Mediterranean endemic species has been explained as the result of reduced and isolated, but particularly stable, habitats (e.g. rocky habitats) suitable for species survival, within a sea of unsuitable landscapes (Hampe & Petit, 2005; Thompson, 2005; Youssef et al., 2010; Molins et al., 2011; Mayol et al., 2012). Although *A. balearica* has been cited as an example of evolutionary stasis (low levels of morphological variation paralleled with low sequence variation) (Molins et al., 2011), this has never been demonstrated.

Arenaria balearica is naturally distributed in Tyrrhenian islands of Majorca, Corsica, and Sardinia, including the surrounding minor islands of Tavolara, La Maddalena, Caprera, and Asinara, and in two of the main Tuscan Islands, Montecristo and Capraia (Diana Corrias, 1981). Most of the populations known from Majorca, Corsica and Sardinia are placed on the Hercynian basement of the corresponding island (Alvarez et al., 1974; Rosenbaum et al., 2002). The species is an alien plant in some European countries, where it is used as an ornamental. Due to its distribution pattern and to the fact that the plant usually inhabits plant communities having a notable relict character (Bolòs & Molinier, 1958), *A. balearica* has been traditionally considered to be a

Deleted: palaeoendemic

Deleted: time

Deleted: and, recently

Deleted: palaeoendemic

Comment [r1]: Rephrase this sentence: it is clumsy

Deleted: (Molins et al., 2011)

Deleted: L. (Caryophyllaceae) is an herbaceous perennial plant

Deleted: The species is an alien plant in some European countries, where it is used as an ornamental. It is a

Comment [r2]: What is meant by this: that most populations on each of the islands are associated with Hercynian basement (granite?)? Rephrase

Moved down [1]: delicate plant whose filiform, branched stems and small leaves form low, compact ever-green moss-like dense mats, preferentially on cool, moist soils in shaded rocky places (comophyte), although it can be secondarily found also on shady moist slopes, between 0 and 1800 m a.s.l. (Diana Corrias, 1981; López González, 1990). Although there are no available data on the reproductive biology of the species, its slender, short, upright stems that bear white, actinomorphic flowers suggest that it is probably partly wind, and partly insect pollinated.

Deleted: *Arenaria balearica* produces small seeds (0.5–0.6 mm) and lacks any evident adaptation to long-distance dispersal. Its chromosome number is 2n = 18 (Diana Corrias, 1981; López González, 1990). The available phylogenetic data based on the analysis of DNA sequences (Fior & Karis, 2007) indicate that this species is closely related to *A. bertolonii* Fiori, which is distributed primarily in mainland Italy (Iamonico, 2013) and Sardinia (Conti et al. 2005).¶
Due to its distribution pattern and to the fact that the species

Deleted: is

Deleted: palaeoendemic (Contandriopoulos, 1962). Long

Mediterranean paleoendemic in the broad sense of the term (Favarger & Contandriopolous, 1961), and a disjunct endemism by Thompson (2005). The plant produces small seeds (0.5–0.6 mm) and although it lacks any evident adaptation to long-distance dispersal (LDD), such events due to stochastic mechanisms, even human mediated (López González, 1990), cannot be *a priori* ruled out to explain its current distribution pattern.

Previous studies on phylogeopraphic patterns of Mediterranean disjunct endemic species have focused on examples from the Eastern Mediterranean region (e. g. Affre & Thompson, 1997; Widén, 2002; Bittkau & Comes, 2005; Edh, Widén & Ceplitis, 2007), as well as from the Western Mediterranean region, including species distributed in Majorca and Menorca (e.g., Sales et al., 2001; Molins, Mayol & Roselló, 2009) and Corsica and Sardinia (e.g., Falchi, 2009). Molins et al. (2011) have studied *T. herba-barona*, a disjunct endemic that shows a distribution similar to that of *A. balearica* except that the former is not as widespread neither in Majorca (only one population) nor in Sardinia as *A balearica*, and that it is absent from the islets of the Tuscan Archipelago.

Using both sequencing of plastid DNA (cpDNA) regions and amplified fragment length polymorphism (AFLP) fingerprinting, this study aims to reconstruct the phylogeographic patterns and differentiation of intraspecific lineages within the disjunct endemic plant *A. balearica*. More specifically, our objectives are: (1) test to which extent the observed distribution of *A. balearica* is concordant with the geological history of the continental fragment islands from the Western Mediterranean region; (2) assess how the colonization of the different islands and islets took place and 3) evaluate whether the low morphological variation observed among populations of *A. balearica*

Deleted:)

Deleted: the

Deleted: of this plant species.

Deleted: In this paper, we used

Deleted: for the facts

Deleted: and sequences of plastid DNA regions

Deleted: and evolutionary history of

Deleted: , we investigate the

Deleted: give a satisfactory answer to the question on

located in different islands is in correspondence with overall low levels of genetic diversity.

Deleted: structure and variability of the palaeoendemic species *A. balearica* with aim of assessing the impact that the main post-Oligocene palaeoclimatic events have had in shaping population structure and divergence.

2. Materials and methods

2.1. Reconstruction of the coastline during the Last Glacial Maximum in the study area

During the Last Glacial Maximum (LGM), ice sheets covered large areas in northern latitudes, and global temperatures were significantly lower than today (Yokohama et al., 2000). At the LGM, the Earth's ocean levels were at their lowest point and extensive reaches of dry land were exposed along the continents' coasts. Several analyses have substantially narrowed the uncertainties regarding total changes in ice sheets and sea level and their proxies, suggesting a net decrease in the eustatic sea level at the LGM ranging from 120 to 135 m a.s.l. (Church et al., 2001; Clark & Mix, 2002). The reconstruction of coastlines at 21 Ka (kiloyears before present) for the study area presented here (Figure 1) is derived from these references.

To map the past and current shorelines in detail, the present-day topographic and bathymetric data covering the area were taken from the ETOPO1, which is a 1 arc-minute global relief model of the Earth's surface that integrates land topography and ocean bathymetry. This model was built from numerous global and regional data sets, and is available in "Bedrock" (base of the ice sheets) versions (NOAA, 2009). Estimates of exposed land area at LGM with respect to the present-day are the result of the values of the Digital Elevation Model being raised by 120 m.

2.2. Study species

Deleted: **Sampling strategy, outgroup selection and monophyly test**

Arenaria balearica is an herbaceous perennial delicate plant whose filiform, branched stems and small leaves form low, compact ever-green moss-like dense mats, preferentially on cool, moist soils in shaded rocky places (comophyte), although it can also be found also on shady moist slopes, between 0 and 1800 m a.s.l. (Diana Corrias, 1981; López González, 1990). Although there are no available data on the reproductive biology of the species, its slender, short, upright stems that bear white, actinomorphic flowers suggest that it is probably partly wind, and partly insect pollinated. Its chromosome number is 2n = 18 (Diana Corrias, 1981; López González, 1990). Generation times are not known for the species. The available phylogenetic data based on the analysis of DNA sequences (Fior & Karis, 2007) indicate that this species is closely related to *Arenaria bertolonii* Fiori, which is distributed primarily in mainland Italy (Iamonico, 2013) and Sardinia (Conti et al. 2005). The most recent phylogeny published for the genus *Arenaria* L. (Sadeghian et al., 2015) concluded that. *A. balearica* should be excluded from *A.* sect. *Rotundifoliae* McNeill, where the species was traditionally included. Unfortunately these authors did not include *A. bertolonii* in the phylogeny and recovered *A. balearica* in a largely unresolved position (very low levels of statistical support).

Moved (insertion) [1]

Deleted: be secondari

Deleted: y

2.3. Sampling strategy, outgroup selection and monophyly test

Leaf material from a total of 250 plants from 29 sampling sites including the islands of Majorca (9), Corsica (8), Sardinia (9), Tavolara (1), and Montecristo (2), representing the entire distribution range of *A. balearica*, was collected and dried in silica gel (Table 1 and Figure 1). Each sampling site was geo-referenced with a GPS GARMIN GPSMAP 60, and voucher specimens were deposited at the herbaria of the University

Deleted: 222

of Salamanca (SALA), the University of Granada (GDA) in Spain and/or of the University of Cagliari (CAG) in Sardinia, Italy.

The intent was to include a minimum of 10-12 plants per population in the analysis, but sometimes the population sizes were small and it was not possible to collect such a quantity of well separated (> 5-10 m) individuals. Also further problems were encountered in some cases in the DNA extraction and amplification processes (the leaves are only 2-4 mm and it was often difficult to get an adequate quantity of DNA). In this situation a variable number of 1-16 individuals per sampling site were finally used (Table 1).

Three additional samples from *A. bertolonii* were selected to be used as outgroup in the plastid DNA haplotype analyses. Given the uncertain phylogenetic position of *A. balearica* within the genus according to the most recent data (Sadeghian et al., 2015), the selection of this outgroup was based on the results by Fior & Karis (2007). Furthermore, the monophyly of the study group was assessed in a parallel study (J. Bobo-Pinilla, J. Peñas de Giles & M. M. Martínez-Ortega, unpubl. data) through the phylogenetic analysis of nucleotide sequences of the nuclear ribosomal internal transcribed spacer (ITS) using 28 samples belonging to *A. balearica* and several other samples from the related species *A. funiculata*, *A. tejedensis*, and *A. suffruticosa*. These data further support the sister group relationship between *A. balearica* and *A. bertolonii* already proposed by Fior & Karis (2007).

2.4. DNA isolation, AFLP amplification, and data analysis

Total genomic DNA was isolated from crushed dried leaf material (ca. 25 mg) following the 2× CTAB (cetyl trimethyl ammonium bromide) protocol (Doyle & Doyle, 1987) with minor modifications. The quality of the extracted DNA was checked in 1%

Deleted: of

Deleted: AFLP

Deleted: many times

Comment [r3]: 4-16??

Deleted: 4

Deleted: plants

Deleted: cpDNA

Deleted: The

Deleted: ¶

Deleted: ¶

Deleted: 3

Deleted: ¶

TAE-agarose gel. A negative control sample was consistently included to test for contamination, and five randomly chosen samples were replicated to test for reproducibility.

Given the very small leaf size of *A. balearica*, it was not always possible to extract enough DNA to provide clear and reliable AFLP profiles. Therefore, five populations among the 29 initially sampled had to be excluded from the AFLP analysis (Table 1). AFLP profiles were finally drawn for 213 individuals following established protocols (Vos et al., 1995). An initial screening of selective primers was performed using 26 primer combinations. The four finally selected primer combinations (fluorescent dye in brackets), (6-FAM)*Eco*RI-ACT/*Mse*I-CAT, (6-FAM)*Eco*RI-AGA/*Mse*I-CTG, (VIC)*Eco*RI-AAG/*Mse*I-CAT, (VIC)*Eco*RI-AGG/*Mse*I-CC, were used for the selective polymerase chain reaction. These combinations were selected because they generated a relatively high number of clearly reproducible bands. A relatively high number of alleles per individual is desirable, given that AFLP are dominant markers (Lowe, Harris & Ashton, 2004). Samples (3µl) of the fluorescence-labelled selective amplification products were combined and separated on a capillary electrophoresis sequencer (ABI 3730 DNA Analyser; Applied Biosystems; Foster City, CA, USA), with GenScan ROX (Applied Biosystems) as an internal size standard.

Deleted: , for which homology was easy to ensure.

Raw AFLP data with amplified fragments from 150 to 500 base pairs (bp) were scored and exported as a presence/absence matrix using the software GeneMapper 4.0 (Applied Biosystems). As an initial approach to the global genetic relationships among the individuals analysed and possible structure of the data, a Neighbour-Joining (NJ) analysis including 1000 bootstrap pseudoreplicates based on a matrix of Nei-Li (Nei & Li, 1979) distances was conducted with the software Paup 4.0b10 (Swofford, 2003). An unrooted NeighbourNet was also produced using the program SplitsTree 4.12.3.

(Huson & Bryant, 2006) and based on Dice's coefficient, which is suitable for multilocus dominant genetic data (Dice, 1945; Lowe et al., 2004; results not shown as they are coincident with but less readable than NJ). Additionally, a Principal Coordinate Analysis (PCoA) based on a matrix of Dice's coefficient among individuals was performed with NTSYS-pc 2.02 (Rohlf, 2009).

Population genetic structure was additionally investigated using a Bayesian clustering method implemented in STRUCTURE v. 2.3.4 (Pritchard, Stephens & Donnelly, 2000) following the approach described by Falush, Stephens & Pritchard (2007) for dominant markers. This method uses a Markov chain Monte Carlo simulation approach to group samples into an optimal number of *K* genetic clusters and does not assume an *a priori* assignment of individuals to populations, nor to clusters. Analyses were based on an ancestral admixture model with correlated allele frequencies among populations. The proportion of membership of each individual and population to the *K* clusters was calculated by performing 20 runs for each *K* value between 2 and 9 with a run length of the Markov chain Monte Carlo of 1×10^6 iterations after a burn-in period of 1×10^6 iterations, with λ adjusted at 0.4523. The optimal number of *K* clusters was estimated using the *ad hoc* parameter (ΔK statistic) of Evanno, Regnatus & Goudet (2005), as implemented in the online application of Structure Harvester software (v0.63; Earl & VonHoldt, 2012).

Although aware that AFLP-based estimates of the level of genetic variation could be biased in this case by low sampling sizes and relative differences in sampling effort, Nei's (1987) gene diversity index was calculated for each population (or sampling site) using the R package AFLPDAT (Ehrich, 2006). This package was also used to calculate the frequency down-weighted marker values per population or sampling site (DW;

Schönswetter & Tribsch, 2005), which is an estimation of the genetic rarity of a population.

To test the comparative historical effects of the main biogeographical barriers, a hierarchical analysis of molecular variance (AMOVA) was performed with the software ARLEQUIN 3.5.1.2 (Excoffier & Lischer, 2010). For this, genetic variation was distributed into portions assignable to differences among predefined geographical groups (F_{CT}), among populations within these groups (F_{SC}), and among populations across the entire study area (F_{ST}) (Turner et al., 2000; Ortiz et al., 2009). Additionally, four alternative groupings were tested using AMOVA analysis: the first two tested the groups derived from PCoA and NJ analyses, respectively, while the third and fourth ones tested two additional geographical groupings [i.e. (Majorca) (Corsica) (Sardinia + Tavolara) and (Majorca) (Corsica + Sardinia + Tavolara), respectively].

2.4. Plastid DNA sequencing and data analysis

Three regions of the plastid DNA were sequenced and haplotype variation was explored to complement the information given by the mainly nuclear AFLPs. The plastid regions *trn*L^{UAA}-*trn*F^{GAA} (Taberlet et al., 1991), *psbA-3'trnK-matK* and *rpS16* (Shaw et al., 2005) showed the highest variability among seven surveyed regions (*trnQ(UUG)-rps16x1*, *trnL-rpl32F*, *atpI-atpH*, Shaw et al., 2007; *rpoB-trnC*, *trnH-psbA*, Shaw et al., 2005) and were used to analyse a total of 226 plants from 29 populations (Table 1) of *A. balearica*. PCR conditions and primers for DNA amplification are detailed in Table 2. PCR products were visualized on 1% agarose gel and purified using PCR Clean-Up with ExoSAP-IT Kit (AFFIMETRIX, Santa Clara, CA, USA) following the manufacturer's instructions. The cleaned amplification products were analysed with a

Deleted: **cpDNA**

Deleted: ¶

Deleted: 53

Deleted: i.e. two plants per sampling site except for populations 2, 20, 21, 25, and 28

437 3730 DNA Genetic Analyser capillary sequencer (Applied Biosystems). All sequences
438 were deposited in GeneBank (Table 1 PENDING).
439 Congruence in the phylogenetic signal of the different plastid DNA regions was tested
440 with the partition homogeneity test (ILD; Farris et al., 1995a, b). ILD significance
441 values were calculated in TNT v.1.1 (Goloboff, Farris & Nixon, 2003) with the
442 INCTST script—kindly provided by the authors of the program—with 1000 replicates.
443 The plastid DNA sequences were assembled and edited using GENEIOUS PRO™ 5.4
444 (Drummond et al., 2012) and aligned with CLUSTALW2 2.0.11 (Larkin et al., 2007);
445 further adjustments and optimisations were made by visual inspection. Sequences from
446 the three regions were concatenated based on the assumption that the plastid forms a
447 single linkage group into a single matrix to be analysed, considering also that the ILD
448 test did not report significant incongruities among DNA regions. Gaps
449 (insertions/deletions) were coded as single-step mutations and treated as a fifth
450 character state. Mononucleotide repeats of different sizes were excluded given that they
451 seem to be prone to homoplasy at large geographic scales (Ingvarsson, Ribstein &
452 Taylor, 2003).
453 The completeness of haplotype sampling across the range of *A. balearica* was estimated
454 using the Stirling probability distribution. It provides a way to evaluate the assumption
455 that all haplotypes have been sampled (Dixon, 2006).
456 As an approach to infer the genealogical relationships among haplotypes, an unrooted
457 haplotype network was constructed using the statistical parsimony algorithm
458 (Templeton, Crandall & Sing, 1992) as implemented in TCS 1.21 (Clement, Posada &
459 Crandall, 2000).
460 Six competing phylogeographic hypotheses were compared using a coalescent based
461 approximate Bayesian computation method (ABD approach), as implemented in

Deleted: cpDNA

Deleted: cpDNA

Deleted: The software BEAST v1.6.1 (Drummond & Rambaut, 2007) was used to date the divergence time between *A. balearica* and *A. bertolonii*. Input files for BEAST were created with BEAUti v1.6.1 (Drummond & Rambaut, 2007), using a relaxed lognormal clock model (Drummond et al., 2006).. The MCMC chain was run for 30 million generations with trees and parameter values saved every 1000th generation. Mutation rate was fixed at $1.1–2.9 \times 10^{-9}$ nucleotide substitutions per site per year (Wolfe, Li, & Sharp, 1987; Clegg et al., 1994). The tree files were summarized into one maximum credibility tree with median node heights using TreeAnnotator v1.4.8 (Drummond & Rambaut, 2007), the first 10% of the trees being eliminated as burn-in. The resulting trees were visualized using FigTree v.1.3.1 (Rambaut, 2010). Posterior probabilities ≥0.95 are considered as strong support, values of 0.80–0.94 as moderate support, and values below 0.79 as low support.¶

Deleted: The resulting haplotype network was nested into hierarchical clades using the automated implementation of Nested Clade Phylogeographic Analysis (NCPA, Templeton, Routman & Phillips, 1995) provided by the program ANECA v.1.1 (Panchal, 2007). ANECA software implements both TCS v.1.18 and GEODIS v.2.2 (Posada, Crandall & Templeton, 2000), providing a framework for replicating analyses in an objective way. After the identification of the major clades in the network, divergence times ($T=t_{div}/N_e$) and the migration rate ($M=N_em$) between them were estimated through a Markov Chain Monte Carlo (MCMC)

Deleted: ***using the program MDIV (Nielsen & Wakeley, 2001; available at http://people.binf.ku.dk/rasmus/webpage/mdiv.html), N_e is the effective population size and m is the migration rate. Initial runs were tested under a finite sites (HKY) model of evolution and default priors M (migration rate between populations since divergence) = 10, T (divergence time since two populations diverged from a common ancestral population) = 20 and θ (relative size of each population) = 10, to explore***

DIYABC v2.1 software (Cornuet et al., 2014). DIYABC allows testing the posterior probabilities of alternative scenarios involving complex population histories (i.e., any combination of population divergences and multifurcations, admixture events, population size changes, bottlenecks, etc., even with population samples potentially collected at different times and/or with unsampled populations, Cornuet et al., 2014). The logistic regression procedure (Fagundes et al., 2007) gives an estimate of the occurrence of each scenario among simulated data sets that are closest to the observed data. In our case, four different metapopulations (i.e. Majorca, Corsica, NE Sardinia and SW Sardinia, correspondingly MAJ, COR, NSA and SSA in Table 1) were considered. Due to low sample sizes and considering that only the most widely represented haplotype was present, populations 11, 28 and 29 were excluded from this analysis in order to avoid increasing exponentially computation times . The distinction between NE Sardinia and SW Sardinia (Table 1) was made considering relevant geological aspects, particularly the fact that the populations of *A. balearica* present in the island are located exclusively on two different geological units both located on the ancient Hercynian basement of the island and mainly separated by Oligocene and Miocene rift basins and Plio-Pleistocene basalts (Rosenbaum et al,. 2002). After some initial analysis and taking into account the haplotype network, the geographical distribution of the species and these geological aspects, six competing phylogeographic scenarios were designed. A list of all parameters and prior distributions used to model scenarios is summarized in Table 3. Prior distributions of the parameters were chosen as a first approach with a large interval due to the lack of ancestral information. Parameters were subsequently corrected according to values obtained after first tests. Population sizes were set equally in all cases; divergence times were taken unrestricted to allow the program to set the most likeable value. Uniform Mutation rate was set to $[10^{-9}- 10^{-7}]$. One million data sets

Deleted: ***M and T. The initial test showed that migration rate was closer to 0, so that the data were reanalysed with M = 0. MDIV analyses were run for 2 million generations following a burn-in period of 500,000 generations, and analyses were repeated twice to ensure convergence on the same posterior***

Deleted: ***for each parameter. Divergence time among significant clades was***

were simulated for each scenario (Cornuet et al., 2008, 2010). The posterior probabilities of each one were calculated by performing a polychotomous weighted logistic regression on the 1% of simulated data sets closest to the observed data set (Cornuet et al., 2008, 2010). The posterior distributions of parameters were evaluated under the best scenario using a local linear regression on the 1% closest simulated data sets with a logit transformation (Table 3). Bias and precision for the parameters estimations were also calculated. Divergence time between groups must be taken carefully, due to the lack of information about generation times for the species. Confidence in scenario choice has been tested by evaluating Type I and Type II error rates (Cornuet et al., 2010).

Deleted: ***using***

Deleted: ***previously indicated mutation rate. Despite***

Deleted: ***controversy on the use of this software (Panchal and Beaumont***

Deleted: ***), we used it to estimate the ages of demographic events.***

Deleted: ¶
¶

3. Results

3.1. Population structure based on AFLP

Deleted: ¶

The four primer combinations applied to 213 plants representative of the variation of the species *A. balearica* produced a total of 792 reproducible fragments.

Deleted: for which homology was easy to confirm.

Both the NJ and NeighbourNet diagrams conducted on all individuals revealed a relatively weak overall structure of the genetic variation into two main groups: one comprised the samples collected in Majorca ("group 1", represented in green in Figure 2A; populations 1–3, 5, 7–9; with not significant bootstrap support, BS < 75%) and a second poorly supported group (BS < 75%), which clustered together individuals from the remaining populations included in this study. Within the second group, three further subgroups were found: first, "group 2", which included samples collected mostly in C and S Sardinia (populations 14, 15, 18 and 19); second, "group 3", which grouped populations 10–13, plus 17 from W and NE Sardinia and Tavolara, together with populations 23–27 mostly from S Corsica; and third, "group 4", which included all the

579 individuals from population 16 in C Sardinia. None registered significant BS values (BS
580 < 75%).
581 Apparently a higher level of overall genetic structure was revealed by the PCoA (Figure
582 2B); in this case, the first two axes accounted for 55.31% and 5.33%, respectively, of
583 the total variance, although no evident geographic structure was found. Two groups
584 were roughly distinguished in the PCoA: the first one grouped populations 1–3, 5, 7–9
585 from Majorca with 10, 12, 15, 16, and 19 from Sardinia, while the second contained
586 populations 11, 13, 14, 17, and 18 from Sardinia and Tavolara, with 22–27 from
587 Corsica. This analysis indicated differentiation to a certain degree of the populations
588 from Majorca and Corsica, but not of those from Sardinia or Tavolara. The genetic
589 structure revealed by NJ and PCoA did not coincide except for the fact that the
590 populations from Majorca were slightly differentiated from the Corso-Sardinian ones.
591 Nei's gene diversity index (Table 1) ranged from 0.09 (populations 8, 1, and 2, all from
592 Majorca) to 0.20 (population 27 from Corsica, although this result may be biased due to
593 the small sampling size) and DW varied between 4.49 in population 2 and 14.83 in
594 population 7, both from Majorca. Overall, the genetically most distinctive and diverse
595 populations were found in Corsica, while the populations from Majorca displayed
596 generally low diversity and singularity values.
597 Bayesian clustering conducted using STRUCTURE estimated $K = 4$ as the most likely
598 number of genetic clusters in *A. balearica*, with a maximum modal value of $\Delta K =$
599 12.414075 (Figure 3). This clustering (Figure 2) showed that all four of these groups
600 were represented in the three main islands and also in Tavolara. In summary, Cluster A
601 (pink) was dominant in the populations from Majorca and S Sardinia (particularly in
602 population 16), was well represented in Tavolara, but its representation was poor in the
603 remaining populations, particularly in populations 23, 25, and 26 from Corsica; Cluster

B (purple) was also well represented −but consistently in a lower proportion than Cluster A− in Majorca (especially in population 5), southern Sardinia (particularly in population 16) and Tavolara, but it was present in a very low proportion in the remaining populations included in this study; Cluster C (yellow) was very well represented in all populations from Corsica, northern Sardinia, and Tavolara, but was almost absent from Majorca (completely absent from population 3); and Cluster D (orange) was best represented in Corsica, was present also in Tavolara and Sardinia (in an almost insignificant proportion in population 16), and had also a low representation in Majorca.

The hierarchical AMOVA (Table 4) showed that the genetic structure in four groups detected by NJ (and NeighbourNet, data not shown) [i.e. (populations 1, 2, 3, 5, 7, 8, 9) (populations 14, 15, 18, 19, 22) (populations 10–13, 17, 23–27) (population 16)] accounted for a comparatively higher amount of the total genetic variance (10.71%), among these groups. This amount was similar, although slightly lower, than that accounted for among populations within groups (11.41%). In the AMOVA analyses that evaluated other groupings the levels of genetic divergence were remarkably low among all groups considered and most of the variation was consistently found among populations within groups instead of among pre-established groups.

Deleted: 3

3.2. Plastid DNA variation in *Arenaria balearica* and geographical distribution of haplotypes

The length of the three plastid DNA regions for 226 individuals ranged between 846 and 704 bp, and resulted in an alignment of 2291 bp, 17 polymorphisms (12 substitutions / 5 indels) were detected across the whole dataset, 5 (4 substitutions / 1 indels), 8 (4 substitutions / 4 indels) and 4 substitutions were detected for the *trn*L^{UAA}-

Deleted: ¶
Deleted: **Chloroplast**
Deleted: **, inferred divergence time for**
Deleted: cpDNA
Deleted: 51
Deleted: 887
Deleted: 750
Deleted: 2422
Deleted: . In the *trn*L^{UAA}-*trn*F^{GAA} region, 14
Deleted: 3
Deleted: 11
Deleted: 6 (3 substitution / 3
Deleted: 13 (5 substitution / 8 indels) polymorphisms

*trn*F^{GAA}, *psbA-3'trnK-matK* and *rpS16*, respectively. All mutations together defined a total of 16 haplotypes (Table 1). The results of the ILD test did not reveal significant inconsistencies among the plastid-DNA regions studied. The completeness of haplotype sampling estimated using Dixon's (2006) method was 0.97 (the most likely value of haplotypes = 16), suggesting that all haplotypes present in the species had been sampled.

The statistical parsimony algorithm implemented in TCS inferred a 95% parsimony network with a maximum limit of four steps and star-like topology (Figure 1). As inferred from the networking analysis, *A. balearica* showed a single major haplotype (present in 24 from the 29 populations studied), probably ancestral (haplotype I), which occurred in all islands (including Tavolara and Montecristo). In addition, there were 15 haplotypes, nine haplotypes (II, III, V, VII, X, XI, XII, XIII and XVI) separated one step from the ancestral one, haplotypes VI and XIV derived one step from haplotypes V and XIII respectively and haplotype XV derived two steps from XIV, two haplotypes derived two steps from haplotype I (IV and VIII) and IX derived one step from VIII. The most derived haplotypes were endemic to one individual island and usually were restricted to single populations (except for haplotype XIV, which was found in two populations from Corsica). Apart from haplotype I, only haplotype V was shared by populations located in different islands (Corsica and Sardinia). *Arenaria bertolonii* is separated 50 steps from the *A. balearica* central haplotype. The levels of haplotypic variation found in Corsica and Sardinia seems to be in accordance with the high levels of overall genetic diversity revealed by AFLP markers.

3.4. DIYABC analysis

Deleted: 12
Deleted: 95
Deleted: 12
Deleted: According to the estimation performed using BEAST, *A. balearica* diverged from its sister species *A. bertolonii* during the Neogene, probably in the Messinian age (ca.6.75 Ma; 95% HPD: 3.77-11.99).¶
Deleted: six
Deleted: 21
Deleted: 28
Deleted: 11
Deleted: of which (
Deleted: VI, which
Deleted: from haplotype V) separated three
Deleted: the ancestral one and four haplotypes (IV, VIII, X, and XII, the latter two directly derived from haplotypes IX and XI, respectively) separated
Deleted: the central
Deleted: .
Deleted: VIII
Deleted: three
Deleted: Sardinia) were restricted to single populations.
Deleted: IX
Deleted:), and most haplotypes were exclusive to only one of the main islands: haplotypes II and III were found exclusively in Majorca, haplotypes IV, V, VI, VII, VIII, and X were present only in Sardinia, while haplotypes XI and XII were restricted to Corsica. *A.*
Deleted: low
Deleted: contrasts
Deleted: The hierarchical nested phylogeographic analysis identified (Table 4) five clades. Clade 1-1 was made up of haplotypes V and VI from Central Sardinia; clade 1-2 was composed of haplotypes XI and XII, which were represented in Corsica; clade 1-4 grouped haplotypes IX and X present in Corsica and Sardinia, and clade 1-5 included the central haplotype I plus other five plastid variants (II, III, IV, VII and VIII), which were represented in the three major islands. NCPA identified one clade (1-3) for which the null hypothesis (no geographic structuring of haplotypes) could not be rejected. No higher-level clades were identified. Additionally, the total cladogram indicated restricted gene flow and isolation by distance ($P<0.005$), while the same process was identified for clade 1-5 ($P<0.001$). For the remaining clades allopatric fragmentation was detected, although these results were not statistically significant. Lastly, the estimated scaled migration rate and divergence time between clades presented very low statistical support. The estimated divergence time between clades was significant in three cases: (1) the splitting between haplotypes IX and X, which, according to our data and assuming mutation rates of $1.1–2.9 \times 10^{-9}$ nucleotide substitutions per site per year, took place 0.03-0.01 Ma; (2) the divergence of the Corsican haplotypes IX and X from those in clade 1-5, which was inferred to have occurred 0.05-0.02 Ma; and (
… [1]

Scenario 1 (ancestral area fragmentation) was revealed as the most probable. The posterior probability of the logistic regression was 75%, while the alternative hypotheses (Figure 4) received less than 7%. Scenario 1 type I and type II errors resulted to be 21% and 17% respectively. DIYABC software places the fragmentation of the four areas 4730 generations ago.

4. Discussion

4.1. Phylogeography of the relict *Arenaria balearica*

Deleted: ¶
Deleted: **Evolutionary history**
Deleted: ¶
Deleted: Solid

Rigorous analysis in phylogeography should be based on the choice of appropriate study organisms and focal areas. Several requirements for reliable phylogeographic inference should be met, among a sound phylogenetic framework and the absence of obvious adaptations for LDD from the organism side, and the availability of good historical climatic and geographic data from the focal-area side (Salvo et al., 2010). *Arenaria balearica* and the Western Mediterranean region satisfy these prerequisites. One of the most basic questions related with Mediterranean plant populations that still remains open is what part of their present genetic diversity is, as generally assumed, due to isolation in refugia during the Pleistocene glaciations, and what part can be traced back to the Tertiary history of taxa (Magri et al., 2007; Médail & Diadema, 2009). Several authors (Thompson, 2005; Donoghue, 2008; Ackerly, 2009) have suggested that the filtering of elements from the ancient Tertiary geofloras that spread across the Northern Hemisphere during the Tertiary (Wolfe, 1975, 1978) played a crucial role in the assembly of the Mediterranean floristic diversity. Thus, traditionally, botanists have classified the floristic elements of the Mediterranean region into two main groups, depending on whether these were believed to have arisen before or after the development of Mediterranean-like climates (Thompson, 2005; Salvo et al., 2010).

Arenaria balearica was traditionally considered a Tertiary relict palaeoendemic species (Contandriopoulos, 1962) and has been particularly mentioned as a "Hercynian palaeoendemic" (Molins et al., 2011). Unfortunately, considering that the plant is perennial and that there is no information available on generation times, although we have obtained here an estimated divergence time for T1 (Table 3; Fig. 4), our results are not conclusive regarding the question on the age and hypothetic ancient origin of the species.

Several hypotheses may explain the presence of *A. balearica* in Majorca, Corsica, and Sardinia, plus minor Tyrrhenian continental fragment islands. This striking distribution may suggest that it could be a non-monophyletic lineage, but the phylogenetic analysis of ITS (nrDNA) and plastid DNA sequences, which included samples from all the Tyrrhenian islands where the species is represented, indicated that the study group is clearly monophyletic (J. Bobo-Pinilla, J. Peñas de Giles & M. M. Martínez-Ortega, unpubl. data). Additionally, both the careful review of herbarium materials prior to the sampling performed within this study, as well as the field observations, indicate very low morphological variation among populations (Lorite et al., unpubl. data). Both plastid and nuclear markers show the lack of a phylogeographic break among populations from different islands. Low levels of genetic structure are repeatedly found by the data analyses derived from the anonymous, mostly nuclear, DNA fingerprints (i.e. AFLP data; NJ, NNet and PCoA analyses; Figure 2) and by the plastid-DNA data. The AMOVA analyses also indicated moderate levels of divergence among populations of *A. balearica* considered as a unique group, which are even lower among the different groups tested with AMOVA. These results contrast with the expectation of high population or geographical group divergence in species that occur in spatially isolated territories, particularly when the species shows limited dispersal abilities (in these

Deleted: The experimental data presented here suggest, by contrast, a Messinian origin and, therefore, that the present distribution of *A. balearica* cannot be attributed to the Oligocenic connections among the Western Mediterranean islands.

Deleted: disjunct

Deleted: cpDNA

Deleted: M. M. Martínez-Ortega & J. Peñas de Giles,

Deleted:

Deleted: clustering performed by ANeCA using the

situations gene flow tends to be low and especially when population sizes are small, the effect of genetic drift is usually high). In the case of *A. balearica* the moderate levels of divergence found may represent remnants of Messinian contacts among the Tyrrhenian territories and long-term genetic stasis followed by recent differentiation in different stable habitats. Furthermore, the star-like arrangement of plastid DNA haplotypes (Figure 1) and DIYABC models suggest a pattern of long term survival and *in situ* differentiation. These results strongly agree with the idea of an ancient haplotype (I) widespread throughout the Tyrrhenian islands where the plant is present today, with different geographically scattered younger *in situ* derived haplotypes. In most cases they represent endemic local variants that originated in isolation from each other, probably due to insularity or geography, on the one hand, and to the scattered availability of rupicolous habitats, on the other.

The Messinian Salinity Crisis, which has been invoked to explain the distribution of many plant species in the Western Mediterranean (e.g. Molins et al., 2011), may also be invoked in this case, although the existence of Messinian terrestrial connections between the Corsica-Sardinia block and the Balearic Islands have never been documented (Alvarez, 1972; Alvarez et al., 1974; Rosenbaum et al., 2002). Also, although there is no evidence for further post-Messinian terrestrial connections between the major Tyrrhenian islands (Alvarez, 1972; Alvarez et al., 1974; Rosenbaum et al., 2002), direct land bridges existed during the Pleistocene glacial maxima between Corsica and Sardinia that allowed floristic exchanges (Salvo et al., 2010). This is also confirmed by the reconstruction of coastline during the LGM performed in this study (Figure 1). The slightly exerted small capsules, and very small seeds (López González, 1990), and the plant's preference for shaded rocky sites (comophyte) are features that probably favoured short-distance dispersal. LDD of *A. balearica*, appears to be

Deleted: cpDNA

Deleted: suggests

Deleted: radiative evolution

Comment [r4]: Rocky?

Deleted: Based on cpDNA mutation rates, it seems that at least two of the main clades detected by NCPA analysis (clade 1-1 and clade 1-4 from clade 1-5, respectively) split off as recently as in the Last Glacial Period. Restricted gene flow with isolation by distance appears to have been the main process that affected the genetic structure in *A. balearica*, and a migration rate of close to 0 further supports these ideas.

Deleted: cited

Deleted: causes

877 unfeasible during the Messinian when the Mediterranean Basin was a saline desert
878 (Hsü, 1972). The fact that the plant lacks adaptations for over-water dispersal suggests
879 also that LDD events between Majorca and the other Tyrrhenian islands (Corsica and/or
880 Sardinia) were unlikely even during the Quaternary glacial maxima. No random LDD
881 event was identified in the analyses performed in this study. Additionally, the star-like
882 parsimony network inferred from plastid DNA data compiled (Figure 1) is not
883 consistent with a range-expansion model after LDD events, and no evidence was found
884 for the existence of such events, either recent or ancient, between Majorca and the other
885 Tyrrhenian islands derived from the almost nuclear AFLPs.
886 Historical gene flow seems to have existed between Corsican and Sardinian
887 populations, as suggested by AFLPs. Both the NJ and PCoA analyses (Figure 2)
888 revealed no structuring of the overall genetic variability on a geographical basis. These
889 results are also confirmed by the AMOVA analyses, which show that the genetic
890 structure in four groups detected by NJ accounts for the comparatively highest amount
891 of the total genetic variance, thus supporting the idea that only those populations from
892 Majorca are to some extent genetically differentiated from the rest. The Bayesian
893 analysis of population structure reveals active historical gene flow and secondary
894 contacts between Corsican and Sardinian populations (Figure 2C). Particularly, clusters
895 B and D are well represented on both islands but almost absent from Majorca (Figure
896 2C) and the levels of admixture of these clusters tend to be higher among the
897 populations located in southern Corsica and northern Sardinia (Figure 2C). All these
898 facts agree with the hypothesis of recurrent connections between Corsica and Sardinia
899 in Miocene and Plio-Pleistocene times (Messinian Salinity Crisis: Gover, Meijer &
900 Krijgsman, 2009; Pleistocene glaciations: Lambeck et al., 2004; Lambeck & Purcell,
901 2005), which facilitated active exchanges of biota, as demonstrated for other organisms

Deleted: cpDNA

(Zachos et al., 2003; Salvi et al., 2010; Fritz, Corti & Päckert, 2012). By contrast, the plastid DNA data do not indicate significant post-Messinian floristic exchanges among Corsica, Sardinia, and the Tuscan Archipelago (only one haplotype is shared between Corsica and Sardinia), as proposed for other plant groups (e.g. Quilichini et al., 2004; Salvo et al., 2008; Zecca et al., 2011), a conclusion which may be biased by the fact that we were not able to establish good AFLP profiles for the plants collected in Montecristo and further highlights the importance of including anonymous hypervariable nuclear markers in phylogeographic studies.

4.2. Evolutionary stasis and habitat stability in Mediterranean disjunct endemic taxa

The low levels of genetic variation found in the maternally inherited plastid DNA (i.e. low number both of detected and of missing haplotypes, low variation common to all the plastid DNA regions tested, and a maximum limit of four steps from the inferred ancestral haplotype were detected in the haplotype network) are consistent with some of the criteria that usually characterized palaeoendemic species (at least in the traditional broad concept of Favarger & Contandriopolous, 1961). This low variation is usually interpreted as a consequence of long processes of adaptation in relative isolation to the intrinsic characteristics of the local refuge area (Mansion et al., 2008).

Molins et al. (2011) have emphasized that several relict endemic species show little or no morphological differentiation despite a long history of isolation on small continental fragments. Even though *A. balearica* was specifically cited in that work as an example of evolutionary stasis, this had never been demonstrated until now. The low mutation rates associated with the plastid genome in *A. balearica* probably correspond to low levels of genetic diversity detected also with AFLPs, thus revealing that stasis in this

Deleted: cpDNA

Deleted: ,

Deleted: **palaeoendemic**

Deleted: ¶

Deleted: cpDNA

Deleted: cpDNA

Deleted: three

Deleted: characterize

Comment [r5]: Break up into two sentencesd

Deleted: palaeoendemic

Deleted: chloroplast

Deleted: (although any conclusion on diversity based on AFLP data may be limited by the relatively low sampling sizes),

941 case agrees with generally low levels of genetic variation. A remarkable lack of
942 variation in all plastid DNA markers scored (including intron regions, intergenic
943 spacers, and plastid microsatellites) was detected for the Tertiary relict *Ramonda*
944 *myconi* (L.) Rchb. (Dubreuil, Riba & Mayol, 2008), which concords with previous
945 results for other relict species (e.g. *Zelkova abelicea* (Lam.) Boiss. and *Z. sicula* Di
946 Pasq., Garfì & Quézel by Fineschi et al., 2002; *Quercus suber* L. by Magri et al., 2007;
947 *Cephalaria squamiflora* (Sieber) Greuter by Rosselló et al., 2009). According to
948 Dubreuil et al. (2008), the absence or low variation in the plastid genome could be a
949 consequence of strong bottlenecks or genetic drift associated with small effective
950 population sizes for maternally inherited markers (Birky, Fuerst & Maruyama, 1989), of
951 slow population dynamics (Dubreuil et al., 2008) and/or of slowed sequence evolution
952 (Dubreuil et al., 2008; Molins et al., 2011). The latter has been repeatedly associated
953 with morphological stasis (Barraclough & Savolainen, 2001; Soltis et al., 2002; Molins
954 et al., 2011). Nevertheless, Casane & Laurenti (2013) have recently suggested that,
955 although a causal link between low molecular evolutionary rates and morphological
956 stasis has been generally assumed, it seems that low intra-specific molecular diversity
957 does not imply a low mutation rate, and also those intraspecific levels of molecular
958 diversity and morphological divergence rates are under different constraints and are not
959 necessarily correlated. As for *A. balearica*, independent markers suggest low levels of
960 intraspecific molecular diversity [i.e. low plastid DNA variation, that seems to parallel
961 the low overall genetic variability as revealed by a technique (AFLP) that covers the
962 whole genome and also with low ITS sequence variation (J. Bobo-Pinilla, J. Peñas de
963 Giles & M. M. Martínez-Ortega, unpubl. data) that covers a small proportion of the
964 nuclear DNA], but an explicit correlation between these data and either long-term

Deleted: cpDNA

Field Code Changed

Deleted: was found to be in accord

Deleted: cpDNA

morphological constancy or slowed mutation rates cannot be established with the available data.

Tertiary relict species have been forced to survive in refugia for long periods of time and their present genetic structure may therefore reflect the impact of a combination of ancient climatic and geographic changes. The ability to persist and resist overall adverse climatic conditions is probably coupled with the availability of relatively stable habitats, where intrinsic local properties have buffered the impact of historical climatic changes, thus allowing long-time persistence of particular species (Thompson, 2005; Médail & Diadema, 2009). The importance of local properties of refugia for survival of Tertiary relict taxa has previously been highlighted for other Mediterranean species, such as the rupicolous herb *R. myconi* (Dubreuil et al., 2008). Furthermore, several authors (e.g. Thompson, 2005; Peñas, Pérez-García & Mota, 2005; Rosselló et al., 2009; Youssef et al., 2010; Mayol et al., 2012) have commented on the long-term stability of rocky habitats in the Mediterranean region and their role at warranting species survival based on the relatively low incidence of disturbances and interspecific competition and the fact that it is probably not fortuitous that many Mediterranean endemic species occur in rocky habitats [e.g. *Cymbalaria aequitriloba* (Viv.) A. Chev., *Nananthea perpusilla* DC., *Naufraga balearica* Constance & Cannon, *Soleirolia soleirolii* (Req.) Dandy, etc.]. *Arenaria balearica* represents a further example of the importance of rocky sites as conservation habitats and as long-term reservoirs of plant diversity within the Mediterranean region.

Comment [r6]: comophytic

Deleted: rupicolous

Deleted: palaeoendemic

Acknowledgements

We are grateful to Dr. Andreas Tribsch for providing samples for this study and for his initial help with lab work, Teresa Malvar for the lab work and D. Nesbitt for the English-language revision.

Deleted: **and author contributions**

Deleted: ¶
This work has been financed by the Spanish Ministerio de Ciencia e Innovación through the projects CGL2010-16357, CGL2009-07555, and CGL2012-32574. SBL was supported by a predoctoral research grant financed by AECID (Agencia Española de Cooperación Internacional para el Desarrollo).

Deleted: .

References

Ackerly D. 2009. Some comments on the age, origin and evolution of the California and Mediterranean floras. *Journal of Biogeography* 36: 1221-1233.

Affre L, Thompson JD. 1997. Population genetic structure and levels of inbreeding depression in the Mediterranean island endemic *Cyclamen creticum*. *Biological Journal of the Linnean Society* 60: 527–549.

Alvarez W. 1972. Rotation of Corsica-Sardinia microplate. *Nature Physical Science* 235: 103-105.

Alvarez W, Cocozza T, Wezel FC. 1974. Fragmentation of Alpine orogenic belt by microplate dispersal. *Nature* 248: 309-314.

Bacchetta G, Pontecorvo C. 2005. Contribution to the knowledge of the endemic vascular flora of Iglesiente (SW Sardinia-Italy). *Candollea* 60: 481-501.

Bacchetta G, Coppi A, Pontecorvo C, Selvi F. 2008. Systematics, phylogenetic relationships and conservation of the taxa of *Anchusa* (Boraginaceae) endemic to Sardinia (Italy). *Systematics and Biodiversity* 6: 161-174.

Bacchetta G, Brullo S, Cusma Velari T, Feoli Chiapella L, Kosovel V. 2012. Analysis of the *Genista ephedroides* group (Fabaceae) based on karyological, molecular and morphological data. *Caryologia* 65: 47-61.

Bacchetta G, Fenu G, Mattana E. 2012. A checklist of the exclusive vascular flora of Sardinia with priority rankings for conservation. *Anales del Jardín Botánico de Madrid* 69: 81-89.

Barraclough TG, Savolainen V. 2001. Evolutionary rates and species diversity in flowering plants. *Evolution* 55: 677-683.

Birky CW, Fuerst P, Maruyama T. 1989. Organelle gene diversity under migration, mutation, and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121: 613-627.

Bittkau C, Comes HP. 2005. Evolutionary precesses in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. *Molecular Ecology* 14: 4065–4083.

1054 Bolòs O de, Molinier R. 1958. Recherches phytosociologiques dans l'Ile de Majorque.
1055 *Collectanea Botanica* 5: 699-865.
1056
1057 Cañadas E, Fenu G, Peñas J, Lorite J, Mattana E, Bacchetta G. 2014. Hotspots within
1058 hotspots: Endemic plant richness, environmental drivers, and implications for
1059 conservation. *Biological Conservation* 170: 282-291.
1060
1061 Carmignani L, Lazzarotto A. 2004. *Carta geologica della Toscana (scala 1:250.000)*.
1062 Università di Siena, Dipartimento Scienze della Terra, Centro di
1063 GeoTecnologie, Regione Toscana. Litografia Artistica Cartografica, Firenze.
1064
1065 Casane D, Laurenti P. 2013. Why coelacanths are not 'living fossils': a review of
1066 molecular and morphological data. *Bioessays* 35: 332–338.
1067
1068 Church JA, Gregory JM, Huybrechts P, Kuhn M, Lambeck K, Nhuan MT, Qin D,
1069 Woodworth PL. 2001. Changes in sea level. In: Houghton JT, Ding Y, Griggs
1070 DJ, Noguer M, Van der Linden PJ, Dai X, Maskell K, Johnson CA, eds.
1071 *Climate Change 2001. The Scientific Basis: Contribution of Working Group I*
1072 *to the Third Assessment Report of the Intergovernmental Panel*, 639-694.
1073
1074 Clark PU, Mix AC. 2002. Ice sheets and sea level of the Last Glacial Maximum.
1075 *Quaternary Science Reviews* 21: 1-7.
1076
1077 Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene
1078 genealogies. *Molecular Ecology* 9: 1657-1660.

Deleted: Clegg MT, Gautt BS, Learn GH *Jr*, Morton BR. 1994. Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America* 91: 6795-6801.¶
¶

Contandriopoulos J. 1962. Recherches sur la flore endémique de la Corse et sur ses origines. *Annales de la Faculté des sciences de Marseille* 32: 123-283.

Contandriopoulos J. 1990. Spécificité de l'endémisme corse. *Atti Convegni dei Lincei, Accademia Nazionale dei Lincei* 85: 393-416.

Conti F, Abbate G, Alessandrini A, Blasi C. (Eds.) 2005. An Annotated Checklist of the Italian Vascular Flora. Palombi Editori, Roma.

Coppi A, Mengoni A, Selvi F. 2008. AFLP fingerprinting of *Anchusa* (Boraginaceae) in the Corso-Sardinian system: Genetic diversity, population differentiation and conservation priorities in an insular endemic group threatened with extinction. *Biological Conservation* 141: 2000-2011.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A. 2014. DIYABC v2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30: 1187-1189

Cornuet JM, Ravigné V, Estoup A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 11: 401.

Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, Guillemaud T, Estoup A. 2008. Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24: 2713-2719

Diana Corrias S. 1981. Le piante endemiche della Sardegna: 94-95. *Bollettino della Società Sarda di Scienze Naturali* 20: 287-300.

Dice LR. 1945. Measures of the amount of ecologic association between species. *Ecology* 26: 297-302.

Dixon, CJ. 2006. A means of estimating the completeness of haplotype sampling using the Stirling probability distribution. *Molecular Ecology Notes* 6: 650–652.

Donoghue MJ. 2008. A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* 105: 11549-11555.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2012. Geneious v5.5.7. Biomatters Ltd. Available from http://www.geneious.com.

Dubreuil M, Riba M, Mayol M. 2008. Genetic structure and diversity in *Ramonda myconi* (Gesneriaceae): Effects of historical climatic change on a preglacial relict species. *American Journal of Botany* 95: 577-587.

Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361.

Edh K, Widén B, Ceplitis A. 2007. Nuclear and chloroplast microsatellites reveal extreme population differentiation and limited gene flow in the Aegean endemic *Brassica cretica* (Brassicaceae). *Molecular Ecology* 16: 4972–4983.

Ehrich D. 2006. AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes* 6: 603-604.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology* 10: 564-567.

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.

Fagundes NJR, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL, Excoffier L. 2007. Statistical evaluation of alternative models of human

Deleted: Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4(5): e88.¶ ¶ Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.¶ ¶

Deleted: Vonholdt

evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104: 17614-17619

Falchi A, Paolini J, Desjobert JM, Melis A, Costa J, Varesi L. 2009. Phylogeography of *Cistus creticus* L. on Corsica and Sardinia inferred by the TRNL-F and RPL32-TRNL sequences of cpDNA. *Molecular Phylogenetics and Evolution* 52: 538–543.

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.

Farris JS, Källersjö M, Kluge AG, Bult C. 1995a. Testing significance of incongruence. *Cladistics* 10: 315-319.

Farris JS, Källersjö M, Kluge AG, Bult C. 1995b. Constructing a significance test for incongruence. *Systematic Biology* 44: 570-572.

Favarger C, Contandriopoulos J. 1961. Essai sur l'endémisme. *Bulletin de la Societé Botanique Suisse* 77: 383-408.

Fenu G, Mattana E, Congiu A, Bacchetta G. 2010. The endemic vascular flora of Supramontes (Sardinia), a priority plant conservation area. *Candollea* 65: 347-358.

Fineschi S, Anzidei M, Cafasso D, Cozzolino S, Garfý G, Pastorelli R, Salvini D, Taurchini D, Vendramin GG. 2002. Molecular markers reveal a strong genetic differentiation between two European relic tree species: *Zelkova abelicea* (Lam.) Boissier and *Z. sicula* Di Pasquale, Garfý & Quézel (Ulmaceae). *Conservation Genetics* 3: 145-153.

Fior S, Karis PO. 2007. Phylogeny, evolution and systematics of *Moehringia* (Caryophyllaceae) as inferred from molecular and morphological data: a case of homology reassessment. *Cladistics* 23: 362-372.

Fritz U, Corti C, Päckert M. 2012. Mitochondrial DNA sequences suggest unexpected phylogenetic position of Corso-Sardinian grass snakes (*Natrix cetti*) and do not support their species status, with notes on phylogeography and subspecies delineation of grass snakes. *Organisms Diversity & Evolution* 12: 71-80.

Garcia-Castellanos D, Estrada F, Jiménez-Munt I, Gorini C, Fernàndez M, Vergés J, De Vicente R. 2009. Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature* 462: 778-781.

Garrido JL, Fenu G, Mattana E, Bacchetta G. 2012. Spatial genetic structure of *Aquilegia* taxa endemic to the island of Sardinia. *Annals of Botany* 109: 953-964.

Field Code Changed

1216 Goloboff PA, Farris JS, Nixon K. 2003. TNT: Tree Analysis Using New Technology,
1217 version 1.1. Program and documentation, available from the authors, and at
1218 http://www.zmuc.dk/public/phylogeny/.
1219
1220 Gover R, Meijer P, Krijgsman W. 2009. Regional isostatic response to Messinian
1221 salinity crisis events. *Tectonophysics* 463: 109-129.
1222
1223 Greuter W. 1995. Origin and peculiarities of Mediterranean island floras. *Ecologia*
1224 *Mediterranea* 21: 1-10.
1225
1226 Hampe A, Petit RJ. 2005. Conserving biodiversity under climate change: The rear edge
1227 matters. *Ecology Letters* 8: 461-467.
1228
1229 Hewitt GM. 1999. Post-glacial recolonization of European Biota. *Biological Journal of*
1230 *the Linnean Society* 68: 87-112.
1231
1232 Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary
1233 studies. *Molecular Biology and Evolution* 23: 254-267.
1234
1235 Hsü K. 1972. When the Mediterranean dried up. *American Scientist* 227: 27-37.
1236 Iamonico D. 2013. Taxonomical and chorological study on the central Mediterranean
1237 Basin endemic *Arenaria bertolonii* Fiori & Paol. (Caryophyllaceae). *Plant*
1238 *Biosystems* 147: 923-930.
1239

1240 Ingvarsson PK, Ribstein S, Taylor DR. 2003. Molecular evolution of insertions and
1241 deletion in the chloroplast genome of *Silene*. *Molecular Biology And Evolution*
1242 20: 1737-1740.
1243
1244 Kadereit JW, Comes HP. 2005. The temporal course of alpine plant diversification in
1245 the Quaternary. In: Bakker FT, Chatrou LW, Gravendeel B, Pelser PB, eds.
1246 *Plant species-level systematics new perspectives on pattern process. Regnum*
1247 *Vegetabile*. Germany: Koeltz Scientific Books, 143: 1-117.
1248
1249 Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and
1250 progressions of the Messinian salinity crisis. *Nature* 400: 652-655.
1251
1252 Lambeck K, Antonioli F, Purcell A, Silenzi S. 2004. Sea-level change along the Italian
1253 coast for the past 10,000 yr. *Quaternary Science Reviews* 23: 1567-1598.
1254
1255 Lambeck K, Purcell A. 2005. Sea-level change in the Mediterranean Sea since the
1256 LGM: model predictions for tectonically stable areas. *Quaternary Science*
1257 *Reviews* 24: 1969-1988.
1258
1259 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,
1260 Valentin F, Wallace IM, Wilm A, López R, Thompson JD, Gibson TJ, Higgins
1261 DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
1262
1263 López González G. 1990. *Arenaria* L. In: Castroviejo S. et al., eds. *Flora Iberica 2*.
1264 Madrid, Spain: Ediciones del Real Jardín Botánico de Madrid, CSIC, 172-224.

Lowe A, Harris S, Ashton P. 2004. Ecological Genetics: Design, Analysis and Application. Oxford, UK: Blackwell Science Ltd.

Lumaret R, Mir C, Michaud H, Raynal V. 2002. Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L). *Molecular Ecology* 11: 2327-2336.

Mansion G, Rosenbaum G, Schoenenberger N, Bacchetta G, Rosselló JA, Conti E. 2008. Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean Basin by the angiosperm family Araceae. *Systematic Biology* 57: 269-285.

Magri D, Fineschi S, Bellarosa R, Buonamici A, Sebastiani F, Schirone B, Simeone MC Vendramin GG. 2007. The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Molecular Ecology* 16: 5259-5266.

Mayol M, Palau C, Rosselló JA, González-Martínez SC, Molins A, Riba M. 2012. Patterns of genetic variability and habitat occupancy in *Crepis triasii* (Asteraceae) at different spatial scales: insights on evolutionary processes leading to diversification in continental islands. *Annals of Botany* 109: 429-441.

1289 Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the
1290 Mediterranean Basin. *Journal of Biogeography* 36: 1333-1345.
1291
1292 Médail F, Quézel P. 1997. Hot-spots analysis for conservation of plant biodiversity in
1293 the Mediterranean Basin. *Annals of the Missouri Botanical Garden* 84: 112-
1294 127.
1295
1296 Médail F, Quézel P. 1999. Biodiversity hotspots in the Mediterranean Basin: setting
1297 global conservation priorities. *Conservation Biology* 13: 1510-1513.
1298
1299 Meulenkamp JE, Sissingh W. 2003. Tertiary paleogeography and tectonostratigraphic
1300 evolution of the Northern and Southern Peri-Tethys platforms and the
1301 intermediate domains in the African–Eurasian convergent plate boundary zone.
1302 *Palaeogeography Palaeoclimatology Palaeoecology* 196: 209-228.
1303
1304 Molins A, Bacchetta G, Rosato M, Rosselló JA, Mayol M. 2011. Molecular
1305 phylogeography of *Thymus herba-barona* (Lamiaceae): Insight into the
1306 evolutionary history of the flora of the western Mediterranean islands. *Taxon*
1307 60: 1295-1305.
1308
1309 Molins A, Mayol M, Rosselló JA. 2009. Phylogeographical structure in the coastal
1310 species *Senecio rodriguezii* (Asteraceae), a narrowly distributed endemic
1311 Mediterranean plant. *Journal of Biogeography* 36: 1372–1383.
1312

Nei M. 1987. Molecular evolutionary genetics. New York, USA: Columbia University Press.

Nei M, Li W. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* 76: 5269-5273.

NOAA. 2009. ETOPO1 Bedrock Global Relief Model. Website http://www.ngdc.noaa.gov/mgg/image/2minrelief.html

Ortiz MA, Tremetsberger K, Stuessy T, Terrab A, García-Castaño JL, Talavera S. 2009. Phylogeographic patterns in *Hypochaeris* sect. *Hypochaeris* (Asteraceae, Lactuceae) of the western Mediterranean. *Journal of Biogeography* 36: 1384-1397.

Peñas J, Pérez-García FJ, Mota JF. 2005. Patterns of endemic plants and biogeography of the Baetic high mountains (south Spain). *Acta Botanica Gallica* 152: 347-360.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-59.

Quézel P. 1995. La flore du bassin méditerranéen: Origine, mise en place, endémisme. *Ecologia Mediterranea* 21: 19-39.

Deleted: Nielsen R, Wakeley JW. 2001. Distinguishing migration from isolation: an MCMC Approach. *Genetics* 158: 885-896.¶
¶

Deleted: Ospovat D. 1995. The Development of Darwin's Theory. Cambridge, UK: Cambridge University Press.¶
¶

Deleted: Panchal M. 2007. The automation of Nested Clade Phylogeographic Analysis. *Bioinformatics* 23: 509-510.¶
¶
Panchal M, Beaumont MA. 2010. Evaluating nested clade phylogeographic analysis under models of restricted gene flow. *Systematic Biology* 59: 415-432.¶
¶

Deleted: Posada D, Crandall KA, Templeton AR. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* 9: 487-488.¶
¶

Quilichini A, Debussche M, Thompson JD. 2004. Geographic differentiation of morphological traits and isozymes in the Mediterranean island endemic *Anchusa crispa*: implications for the conservation of a protected species. *Biological Conservation* 118: 651-660.

Rohlf FJ. 2009. NTSYSpc: Numerical Taxonomy System, v. 2.2 1n. Setauket, NY, USA: Exeter Software.

Rodríguez-Sánchez F, Pérez-Barreales R, Ojeda F, Vargas P, Arroyo J. 2008. The Strait of Gibraltar as a melting pot for plant biodiversity. *Quaternary Science Reviews* 27: 2100-2117.

Rosenbaum G, Lister GS, Duboz C. 2002. Reconstruction of the tectonic evolution of the Western Mediterranean since the Oligocene. *Journal of the Virtual Explorer* 8: 107-126.

Rosselló JA, Cosín R, Bacchetta G, Brullo S, Mayol M. 2009. Nuclear and chloroplast DNA variation in *Cephalaria squamiflora* (Dipsacaceae), a disjunct Mediterranean species. *Taxon* 58: 1242-1253.

Sadeghian S, Zarre S, Rabeler RK, Heubl G. 2015. Molecular phylogenetic analysis of *Arenaria* (Caryophyllaceae: tribe Arenarieae) and its allies inferred from nuclear DNA internal transcribed spacer and plastid DNA rps16 sequences. *Botanical Journal of the Linnean Society* 178: 648–669.

Deleted: Rambaut A. 2010. FigTree v. 1.3.1. http://tree.bio.ed.ac.uk/software/figtree/¶
¶

Sales E, Nebauer SG, Mus M, Segura J. 2001. Population genetic study in the Balearic endemic plant species *Digitalis minor* (Scrophulariaceae) using RAPD markers. *American Journal of Botany* 88: 1750–1759.

Salvi D, Harris DJ, Bombi P, Carretero MA, Bologna MA. 2010. Mitochondrial phylogeography of the Bedriaga's rock lizard, *Archaeolacerta bedriagae* (Reptilia: Lacertidae) endemic to Corsica and Sardinia. *Molecular Phylogenetics and Evolution* 56: 690-697.

Salvo G, Bacchetta G, Ghahremaninejad F, Conti E. 2008. Phylogenetic relationships of Ruteae (Rutaceae): new evidence from the chloroplast genome and comparisons with non-molecular data. *Molecular Phylogenetics and Evolution* 49: 736-748.

Salvo G, Ho SYW, Rosenbaum G, Ree R, Conti E. 2010. Tracing the temporal and spatial origins of island endemics in the Mediterranean region: A case study from the *Citrus* family (Ruta L., Rutaceae). *Systematic Biology* 59: 705-222.

Schönswetter P, Tribsch A. 2005. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* 54: 725-732.

Selvi F, Coppi A, Bigazzi M. 2006. Karyotype variation, evolution and phylogeny in *Borago* (Boraginaceae), with emphasis on subgenus *Buglossites* in the Corso-Sardinian system. *Annals of Botany* 98: 857-68.

1414 Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT,
1415 Schilling EE, Small RL. 2005. The tortoise and the hare II: relative utility of 21
1416 noncoding chloroplast DNA sequences for phylogenetic analysis. *American*
1417 *Journal of Botany* 92: 142-166.
1418
1419 Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast
1420 genome sequences to choose noncoding regions for phylogenetic studies in
1421 angiosperms: the tortoise and the hare III. *American Journal of Botany* 94:
1422 275-288.
1423
1424 Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough, TG. 2002. Rate
1425 heterogeneity among lineages of tracheophytes: Integration of molecular and
1426 fossil data and evidence for molecular living fossils. *Proceedings of the*
1427 *National Academy of Sciences of the United States of America* 99: 4430-4435.
1428
1429 Speranza F, Villa IM, Sagnotti L, Florindo F, Cosentino D, Cipollari P, Mattei M. 2002.
1430 Age of the Corsica-Sardinia rotation and Liguro-Provencal Basin spreading:
1431 New paleomagnetic and Ar/Ar evidence. *Tectonophysics* 347: 231-251.
1432
1433 Suc JP. 1984. Origin and evolution of the Mediterranean vegetation and climate in
1434 Europe. *Nature* 307: 429-432.
1435
1436 Swofford DL. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other
1437 Methods), version 4.0b10. Sunderland, MA, USA: Sinauer Associates.
1438

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*17: 1105-1109.

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.

Terrab A, Talavera S, Arista M, Paun O, Stuessy TF, Tremetsberger K. 2007. Genetic diversity and geographic structure at chloroplast microsatellites (cpSSRs) in endangered West Mediterranean firs (*Abies* spp., Pinaceae). *Taxon* 56: 409-516.

Thompson JD**.** 2005. Plant evolution in the Mediterranean. Oxford, UK: Oxford University Press.

Turner TF, Trexler JC, Harris JL, Haynes JL. 2000. Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics* 154: 777-785.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407-4414.

Deleted: Templeton AR, Routman E, Phillips CA. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140: 767-782.¶
¶

Widén B, Andersson S, Rao GY, Widén M. 2002. Population divergence of genetic (co)variance matrices in a subdivided plant species, *Brassica cretica*. *Journal of Evolutionary Biology* 15: 961–970.

Wolfe JA. 1975. Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* 62: 264-279.

Wolfe JA. 1978. A paleobotanical interpretation of Tertiary climates in the Northern Hemisphere. *American Scientist* 66: 694-703.

Yokohama Y, Lambeck K, Deckker PD, Johnston P, Fifiel LK. 2000. Timing of the Last Glacial Maximum from observed sea-level minima. *Nature* 406: 713-716.

Youssef S, Vela E, Baumel A, Tatoni T. 2010. Distribution, habitat and population size variation of *Genista lobelii* (Fabaceae) from the calcareous mountains of Basse Provence (S-E France). *Ecologia mediterranea* 36: 63-76.

Zachos F, Hartl GB, Apollonio M, Reutershan T. 2003. On the phylogeographic origin of the Corsican red deer (*Cervus elaphus corsicanus*): evidence from microsatellites and mitochondrial DNA. *Mammalian Biology* 68: 284-298.

Zecca G, Casazza G, Minuto L, Labra M, Grassi E. 2011. Allopatric divergence and secondary contacts in *Euphorbia spinosa* L.: Influence of climatic changes on the split of the species. *Organisms Diversity and Evolution* 11: 357-372.

Deleted: Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNA. *Proceedings of the National Academy of Sciences of the United States of America* 84: 9054-9058.¶

1502

1503
1504
1505
1506

The hierarchical nested phylogeographic analysis identified (Table 4) five clades. Clade 1-1 was made up of haplotypes V and VI from Central Sardinia; clade 1-2 was composed of haplotypes XI and XII, which were represented in Corsica; clade 1-4 grouped haplotypes IX and X present in Corsica and Sardinia, and clade 1-5 included the central haplotype I plus other five plastid variants (II, III, IV, VII and VIII), which were represented in the three major islands. NCPA identified one clade (1-3) for which the null hypothesis (no geographic structuring of haplotypes) could not be rejected. No higher-level clades were identified. Additionally, the total cladogram indicated restricted gene flow and isolation by distance ($P<0.005$), while the same process was identified for clade 1-5 ($P<0.001$). For the remaining clades allopatric fragmentation was detected, although these results were not statistically significant. Lastly, the estimated scaled migration rate and divergence time between clades presented very low statistical support. The estimated divergence time between clades was significant in three cases: (1) the splitting between haplotypes IX and X, which, according to our data and assuming mutation rates of $1.1–2.9 \times 10^{-9}$ nucleotide substitutions per site per year, took place 0.03-0.01 Ma; (2) the divergence of the Corsican haplotypes IX and X from those in clade 1-5, which was inferred to have occurred 0.05-0.02 Ma; and (3) divergence between haplotypes V-VI and the rest of the group, dated 0.11-0.04 Ma. All these divergences took place therefore within the Last Glacial Period (ca. 0.11-0.012 Ma.) and particularly the splitting between haplotypes IX and X represented in Corsica and Sardinia probably occurred in the LGM.