

SCIENTIFIC REPORTS



OPEN

Phylogeography of western Mediterranean *Cymbalaria* (Plantaginaceae) reveals two independent long-distance dispersals and entails new taxonomic circumscriptions

Pau Carnicero^{1,2}, Peter Schönswetter², Pere Fraga Arguimbau³, Núria Garcia-Jacas⁴, Llorenç Sáez^{1,5} & Mercè Galbany-Casals¹

The Balearic Islands, Corsica and Sardinia (BCS) constitute biodiversity hotspots in the western Mediterranean Basin. Oligocene connections and long distance dispersal events have been suggested to cause presence of BCS shared endemic species. One of them is *Cymbalaria aequitriloba*, which, together with three additional species, constitute a polyploid clade endemic to BCS. Combining amplified fragment length polymorphism (AFLP) fingerprinting, plastid DNA sequences and morphometrics, we inferred the phylogeography of the group and evaluated the species' current taxonomic circumscriptions. Based on morphometric and AFLP data we propose a new circumscription for *C. fragilis* to additionally comprise a group of populations with intermediate morphological characters previously included in *C. aequitriloba*. Consequently, we suggest to change the IUCN category of *C. fragilis* from critically endangered (CR) to near threatened (NT). Both morphology and AFLP data support the current taxonomy of the single island endemics *C. hepaticifolia* and *C. muelleri*. The four species had a common origin in Corsica-Sardinia, and two long-distance dispersal events to the Balearic Islands were inferred. Finally, plastid DNA data suggest that interspecific gene flow took place where two species co-occur.

Islands in the Mediterranean Basin harbour both high species diversity and endemism. For instance, from the around 5000 islands scattered in the Mediterranean Sea¹ all the largest ones (i.e. Balearic Islands, Corsica, Crete, Cyprus, Sardinia and Sicily) constitute Mediterranean biodiversity hotspots². Most of the Mediterranean islands are of continental origin and have been separated from the mainland and from each other through progressive geomorphological processes, often with posterior re-connections due to sea level variation³. Such processes are thought to have played an important role in the evolution and diversification of biota⁴.

The Balearic Islands, Corsica and Sardinia (hereafter termed BCS) originated from progressive splitting and rotation of the Hercynian belt, an Oligocene mountain range linking the Alps with the Iberian Peninsula that started to disintegrate around 30 Million years ago (Ma)⁵. The split between Corsica-Sardinia and the Balearic Islands occurred around 20 Ma⁶. Endemic plant species shared by Corsica, Sardinia and the eastern Balearic Islands (Mallorca and Menorca) have traditionally been considered relict palaeoendemic remnants of

¹Systematics and Evolution of Vascular Plants (UAB) – Associated Unit to CSIC. Departament de Biologia Animal, Biologia Vegetal i Ecologia, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain.

²Department of Botany, University of Innsbruck, Sternwartestrasse 15, 6020, Innsbruck, Austria. ³Institut Menorquí d'Estudis, Camí des Castell 72, 07702, Maó, Spain. ⁴Institut Botànic de Barcelona (IBB-CSIC-ICUB), Pg. del Migdia s/n, ES-08038, Barcelona, Spain. ⁵Societat d'Història Natural de les Balears (SHNB), C/Margarida Xirgu 16, E-07011, Palma de Mallorca, Balearic Islands, Spain. Correspondence and requests for materials should be addressed to P.C. (email: pau.carnicero@gmail.com)

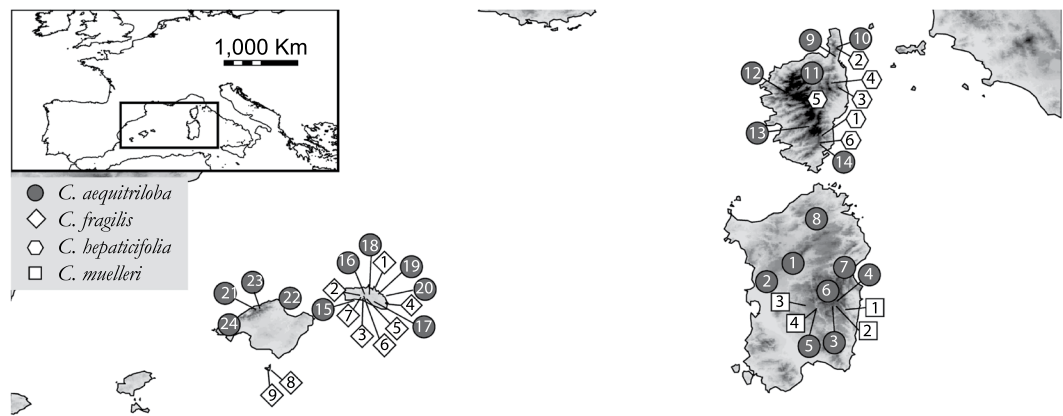


Figure 1. Sampling sites and population numbers. Taxonomy follows the taxonomic concept adopted in the present article, population numbers in the main text are preceded by a letter according to taxonomy (a: *C. aequitriloba*, f: *C. fragilis*, h: *C. hepaticifolia*, m: *C. muelleri*, see Supplementary Table S1). Map generated from the altitude layer obtained from <http://www.worldclim.org>¹³² and modified using QGIS v2.18.2¹³³.

the Oligocene connections^{3,7,8}. They were formerly named “Tyrrhenian endemisms” by Contandriopoulos and Cardona⁷, but we prefer to avoid this term given that the Tyrrhenian Sea is situated between Corsica, Sardinia, the Italian Peninsula and Sicily. To our knowledge, a date of origin consistent with the paleoendemism hypothesis has only been confirmed for *Helicodicerus muscivorus* (L.f.) Engler⁹. Bobo-Pinilla *et al.*¹⁰ have recently proposed a relict origin for the BCS endemic *Arenaria balearica* L., but the lack of absolute dating estimates did not allow for conclusive results. In most other dated phylogenies of plants showing this distribution pattern (e.g. *Arum pictum* L.f.⁹; *Thymus herba-barona* Loisel.¹¹; *Cymbalaria aequitriloba* (Viv.) A.Chev.¹², or other similar disjunctions in the western Mediterranean Basin^{13,14}, palaeoendemism hypotheses had to be rejected (but see Magri *et al.*¹⁵). Instead, long-distance dispersal (LDD) was invoked as a probable explanation for the present distribution pattern, since no land connections have been reported between the Balearic Islands on the one hand and Corsica and Sardinia on the other hand after their split.

Pleistocene climatic oscillations caused massive changes of the sea level that led to repeated appearance of land bridges between some islands during cold stages, and to transgressions reducing the area of islands during interglacials^{3,16}. During glaciation periods Corsica-Sardinia and the eastern Balearic Islands (Mallorca, Menorca and Cabrera), respectively, formed single landmasses. The resultant connections have often been used to explain phylogeographic patterns and shared endemisms within each archipelago (e.g. Mansion *et al.*¹⁷; Salvi *et al.*¹⁸; Mayol *et al.*¹⁹).

One of the genera with highest number of endemics in BCS is *Cymbalaria* Hill (Plantaginaceae), a Mediterranean genus comprising ca. 17 taxa, four of which are endemic to BCS²⁰ (Fig. 1). The BCS endemics form an independent lineage of presumably polyploid origin that diversified during the Pliocene and the Pleistocene, as revealed by a nuclear ribosomal DNA (nrDNA) phylogeny¹². However, interspecific phylogenetic relationships were poorly resolved, and therefore using more variable molecular markers such as AFLPs would be desirable. Chromosome counts suggest that the BCS endemics are hexa- to octoploids ($2n = 42, 56$), whereas all other *Cymbalaria* species are diploids or tetraploids ($2n = 14, 28$ ²⁰ and references therein). While *C. aequitriloba* is widely distributed throughout BCS and occurs in all main islands except for the western Balearic Islands Eibissa and Formentera, the other three species are single island endemics according to the current taxonomic treatments^{20,21}. This lineage is therefore an excellent example to understand the mechanisms underlying the endemism patterns in BCS.

The two single island endemics from Corsica and Sardinia have been traditionally considered as clearly distinct species. *Cymbalaria hepaticifolia* Wettst. ($2n = 56$ ²⁰) occurs in the mountains of Corsica, from subalpine habitats to fresh and humid forests^{22,23}. *Cymbalaria muelleri* (Moris) A.Chev., comprising ca. ten known populations, is a hexaploid ($2n = 42$ ²⁴) endemic of the mountains of central Sardinia²⁵.

Cymbalaria fragilis (J.J.Rodr.) A.Chev. is a critically endangered (CR) octoploid ($2n = 56$ ²⁶) species endemic to Menorca²⁷. It has been treated either as a subspecies of *C. aequitriloba*^{20,28} or as a separate species²¹. Furthermore, the circumscription of this taxon is uncertain. According to Güemes²¹, *C. fragilis* is characterized by several distinctive features such as white to pale violet flowers, fleshy leaves and fragile stems, dense hairiness and finely alveolate to nearly smooth seeds. However, some populations in Menorca, formerly identified as *C. fragilis*²⁹, and Cabrera exhibit all distinctive features but have deeply alveolate seeds similar to those of *C. aequitriloba*²¹. Seed ornamentation has been regarded as highly relevant in the taxonomy of the tribe Antirrhineae^{20,30–32} and this character was later considered sufficient by Güemes²¹ to attribute these populations to *C. aequitriloba*. The latter is the most widespread species among the four BCS endemics, occurring in the eastern Balearic Islands, Corsica, Sardinia and the Tuscan Archipelago. Octoploid counts ($2n = 56$) have been reported from Corsica and the Balearic Islands^{33–35}; older hexaploid counts ($2n = 42$ ²⁰) were considered erroneous by Verlaque *et al.*³⁵. Indeed, the reference cited by Sutton²⁰ for the hexaploid counts³⁶ does not report any chromosome number for *C. aequitriloba*, and we were unable to trace any other hexaploid count in the literature. *Cymbalaria aequitriloba*

exhibits high morphological variability, which is the reason why sparsely hairy forms from Mallorca have sometimes been confused with *C. hepaticifolia*³⁷. Some morphological extremes were described as separate taxa, but they were later synonymised with *C. aequitriloba*²⁰. Nuclear ribosomal and plastid DNA sequences, however, did not support the species' monophyly¹², and it is thus doubtful whether *C. aequitriloba* is indeed a single taxon. Some authors, based on chromosome number information, considered *C. aequitriloba* an example of a palaeoendemism^{3,7,8}, although some of its aforementioned features, such as its high morphological variability, do not fit with the classical definition of the term³⁸. Most importantly, its recent origin inferred from a dated phylogeny rejected that hypothesis¹².

Here, we focus on the phylogeography and systematics of the BCS endemics of the genus *Cymbalaria*. We use a combination of AFLPs, plastid DNA sequences and morphometrics to tackle the following questions: (1) Are genetic and morphological data congruent and do they support the current taxonomy? (2) Specifically, is *C. aequitriloba* monophyletic based on AFLP data? (3) What is the appropriate taxonomic rank and circumscription of *C. fragilis*? (4) What was the pattern of colonization of the BCS endemics among Sardinia, Corsica and the Balearic Islands?

Results

AFLPs. In total, 834 AFLP fragments were scored for 304 individuals (Supplementary Table S1), of which 779 polymorphic, high-quality, reproducible AFLP-fingerprints were obtained. Fifty-five fragments with singular presences or absences were excluded. The initial error rate³⁹ before the exclusion of non-reproducible fragments was 3.17%. The final dataset for *C. aequitriloba* comprised 175 individuals and 557 AFLP fragments.

The NJ analysis revealed that the four taxa analysed constituted a monophyletic group with high bootstrap support (BS 100%, Fig. 2). The main split in the NJ tree separated *C. aequitriloba* and *C. fragilis* from *C. hepaticifolia* and *C. muelleri*. Populations of *C. aequitriloba* formed a clade (BS 62%) excluding all populations with distinctive features of *C. fragilis*, but bearing alveolate seeds (f1, f3, f4, f5, f8 and f9), which instead grouped with *C. fragilis* (BS 100%). One well supported clade each grouped populations of *C. hepaticifolia* (BS 100%) and *C. muelleri* (BS 96%). Individuals from the same population – with the exception of populations a14, a23 and h4 – formed monophyletic groups. The *C. fragilis* clade showed a main split between the populations from Menorca (f1 to f7, BS 79%) and those from Cabrera traditionally included in *C. aequitriloba* (f8 and f9, BS 100%).

The most frequent species tree topology (61%) obtained by the SNAPP analysis coincided with the NJ analysis (Supplementary Fig. S1). However, a second most frequent topology (30%) indicates different relationships among the species: *C. hepaticifolia* diverged first, followed by the differentiation of *C. muelleri*, which is sister to the *C. aequitriloba*-*C. fragilis* clade. Both topologies showed how the diversification of the four western *Cymbalaria* species occurred roughly simultaneously. Three alternative topologies represented less than 10% of the total number of trees.

The results obtained by the Bayesian analyses with BEAST were congruent with the NJ analysis, minor differences concerned nodes with low support (Supplementary Fig. S2). The origin of diversification was estimated to be in northern Sardinia (Fig. 3), from where dispersal events to southern Sardinia, Corsica and the Balearic Islands occurred. Later, further dispersals from Sardinia to Corsica and the Balearic Islands (Mallorca) took place.

A hierarchical AMOVA (Supplementary Table S2) assigned 28.7% of the entire variation to the among-taxa component. In *C. aequitriloba* the highest variation among populations within an island was found in Sardinia (44.7%), and the lowest in Corsica (27.7%). Gene diversity expressed as π_n (Supplementary Table S3) varied from 0.028 in population a17 of *C. aequitriloba* from Menorca to 0.164 in population h5 of *C. hepaticifolia* from Corsica. No differences were found between taxa. For *C. aequitriloba*, π_n was significantly lower in the Balearic Islands than in Corsica and Sardinia.

In the separate NJ analysis for *C. aequitriloba* (BS 100%, Fig. 4) the populations from the Balearic Islands formed a clade (BS 100%), which contained a subclade comprising populations from Menorca (a15–a20, BS 91%). Populations from Corsica (a9–a14) formed a supported clade (BS 99%) whereas populations from Sardinia were distributed in three main clades (a1, BS 100%; a2, a7 and a8, BS 70%; a3–a6, BS 97%). Nonhierarchical K-means clustering resulted in an optimal separation of the dataset into four groups (Fig. 4) composed by 1) the Sardinian populations a3–a6; 2) the Sardinian populations a1, a2, a7 and a8; 3) all Corsican populations; and 4) all Balearic populations. At K = 6 populations a1–a8 from Sardinia were split into four clusters, whereas the other two groups included all Corsican and all Balearic populations, respectively (Fig. 4).

Plastid DNA Sequences. The concatenated *ndhF* and *rpl32-trnL(UAG)* sequences consisted of 2776 aligned positions. We found 26 plastid haplotypes that differed by one to 14 substitutions including codified indels (Fig. 5). There was no clear correlation with taxonomy; instead, some geographically close populations of different taxa shared haplotypes. For instance, haplotype 4 was found in almost all populations of *C. aequitriloba* and *C. fragilis* from the Balearic Islands, haplotype 9 was found in two populations of *C. aequitriloba* and one population of *C. hepaticifolia* from Corsica, and haplotype 22 was found in two populations of *C. aequitriloba* and one population of *C. muelleri* from southern Sardinia (Fig. 5). The Balearic Islands clearly showed low haplotype diversity in comparison to Corsica and Sardinia. Relationships inferred by MP and Bayesian analyses were congruent and consequently we only show the topology from the Bayesian analysis (Fig. 5). Two main clades were found in the tree: the first contains the strongly divergent haplotype 26 from Sardinia and a haplotype of *C. muralis* (PP 1; BS 100%), and it is sister to all other haplotypes found in BCS (PP 0.89). Phylogenetic relationships among haplotype groups (colour-coded in Fig. 5) were in general statistically weakly supported, whereas relationships among closely related haplotypes reflected the pattern of the parsimony network.

Morphometric analyses. Correlation coefficients did not exceed 0.95 for any pair of characters; therefore, all characters were retained for further analyses (Supplementary Table S4).

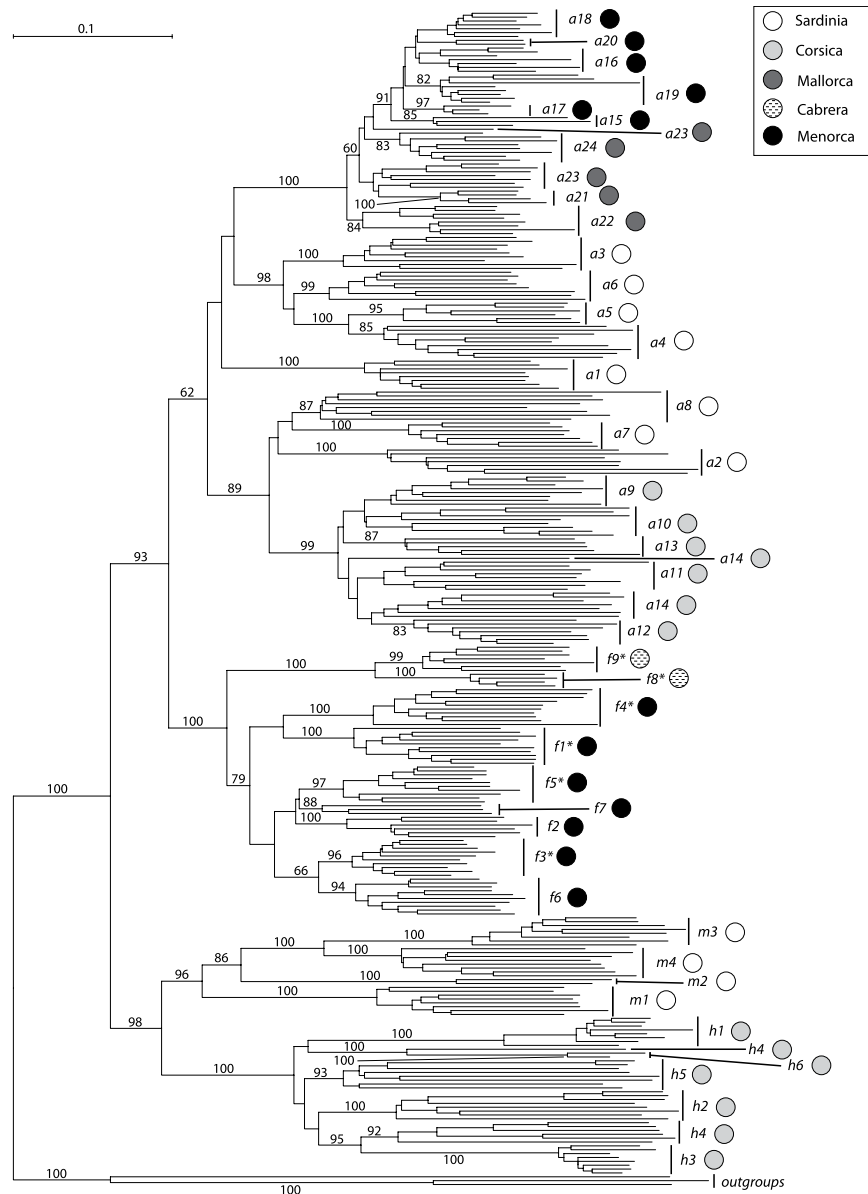


Figure 2. AFLP analyses of *C. aequitriloba*, *C. fragilis*, *C. hepaticifolia* and *C. muelleri*, Neighbour-Joining tree. Numbers above branches are bootstrap support values (>60%) derived from a Neighbour-Joining analysis. Asterisks indicate *C. fragilis* populations previously considered *C. aequitriloba*. Populations are numbered as in Fig. 1 and Table S1. Circles next to population names indicate the island of origin.

Dataset 1. In the PCA, the first axis accounted for 34% of the variation and the second axis for 28%. The ordination diagram (Fig. 6a) suggested three clusters, two corresponded to *C. aequitriloba* and *C. hepaticifolia*, respectively, and a third one grouped *C. fragilis* and *C. muelleri*, including the controversial *C. aequitriloba* populations f1, f3, f4, f5, f8 and f9 from Cabrera and Menorca. The characters with most weight in the separation of the three groups were those related with indumentum density. The CDA (Fig. 6b–e) supported morphological differentiation of the four main AFLPs clades. The first two axes (41% and 35% of the explained variation) supported the three groups observed in the PCA, with some overlap between *C. aequitriloba* and *C. fragilis* plus *C. muelleri*. The characters with the highest contribution to these axes were derived from the leaf indumentum. The third axis (24%) supported the differentiation between *C. fragilis* and *C. muelleri*. The characters with the highest contribution to the third axis were spur length and width.

Dataset 2. The first PCA axis accounted for 44% of the variation and the second axis for 14%. The ordination diagram showed a structure similar to that retrieved by the analysis of dataset 1, albeit with stronger overlap of *C. aequitriloba* and the *C. fragilis* – *C. muelleri* group (Supplementary Fig. S3).

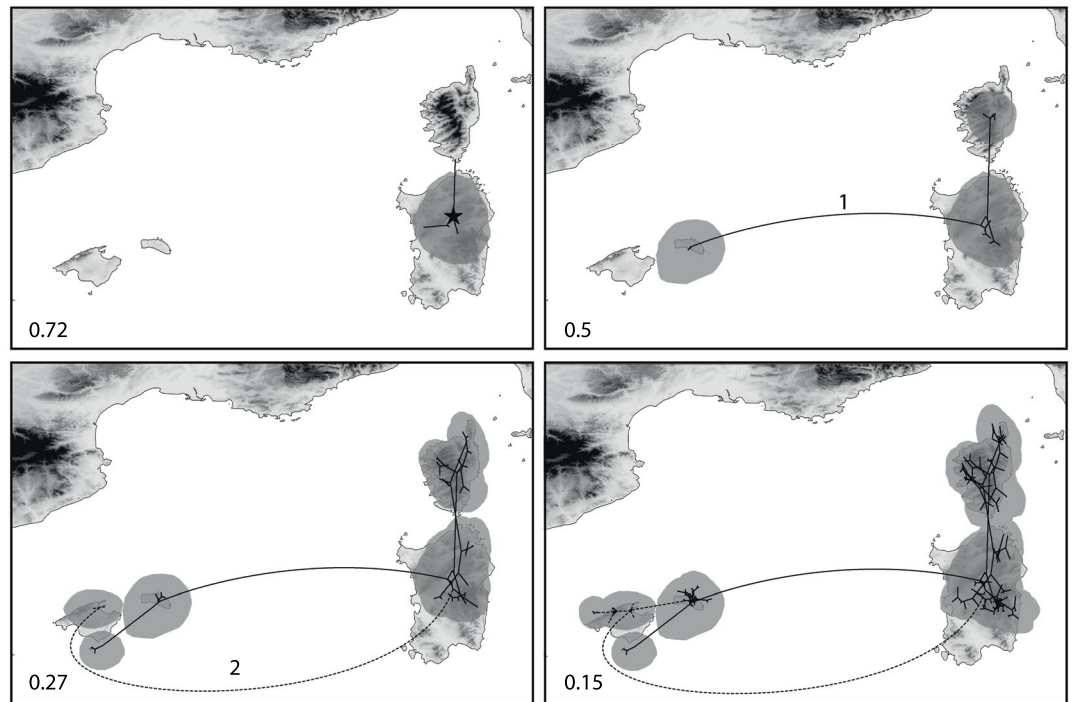


Figure 3. Snapshots of estimated ancestral node areas in the Maximum Clade Credibility tree obtained from the BEAST continuous phylogeographical analysis of AFLP data at different time horizons as visualized using the software SPREAD. The 80% highest posterior density areas for nodes are indicated as grey polygons. The star in the upper left panel indicates the starting point of diversification and the relative time scale of diversification is indicated by numbers (1 = origin, 0 = present). Numbers indicate LDD events to the Balearic Islands leading to (1) the origin of *C. fragilis* and (2) a range expansion of *C. aequitriloba*. Lines indicating the two LDD events were modified to prevent them from crossing. Maps generated from the altitude layer obtained from <http://www.worldclim.org>¹³², and modified using QGIS v2.18.2¹³³.

Discussion

The combination of molecular and morphological data has been widely used to reevaluate systematics at different taxonomic levels (e.g.^{32,40–42}). Specifically, many studies have used AFLP data and morphometrics to address taxonomic conflicts in groups of closely related plant taxa (e.g.^{43–47}). In the case of *Cymbalaria*, AFLP (Fig. 2) and morphometric analyses (Fig. 6) allow for the recognition of four distinct entities within a previously identified monophyletic group of *Cymbalaria* species endemic to the Balearic Islands, Corsica and Sardinia¹². These entities coincide with the current taxonomic circumscription of *C. hepaticifolia* and *C. muelleri*, but require novel circumscriptions of *C. aequitriloba* and *C. fragilis* (Figs 2 and 6).

A novel circumscription of *C. fragilis* and *C. aequitriloba*. Güemes²¹ suggested that *C. fragilis* is restricted to three populations in southwestern Menorca, and assigned reports from the North and East of the island (populations f1 and f4) as well as those from Cabrera (populations f8 and f9) to *C. aequitriloba*, based on their alveolate seeds. However, these populations exhibit all distinctive morphological features of *C. fragilis* such as (1) villous indumentum throughout the vegetative parts, (2) highly fragile stems, petioles and pedicels, (3) fleshy leaves, (4) white to pale violet corollas, (5) spur proportionally shorter and wider than in *C. aequitriloba* (ratio spur length/spur width 0.9–2.8), (6) seeds, which are often smaller than those of *C. aequitriloba* (0.6–0.9 mm long, vs. 0.8–1.2 mm long in *C. aequitriloba*, see identification key), as well as (7) a late flowering time compared to *C. aequitriloba* (June as compared to early spring in *C. aequitriloba* from the same altitude). Interestingly, in population f6 we observed both types of seed ornamentation. In the AFLP (Figs 2 and 4) as well as in the morphometric analyses (Fig. 6), these populations clustered with *C. fragilis*. Therefore, *C. fragilis* must be redefined to include populations with the before-mentioned distinctive morphological characters from Menorca and Cabrera in spite of their aberrant seeds. Seed ornamentation – in this particular case – must thus be considered a variable character without taxonomic value. The strongly supported monophyly in the AFLP analyses (Fig. 2) together with the clear morphological differentiation (Fig. 6) and the overlapping distributions of *C. fragilis* and *C. aequitriloba* in Menorca do not match the definition of subspecies^{48–50}. In contrast, the available evidence strongly suggests treating *C. fragilis* as a separate species in accordance with the original description⁵¹, rather than as a subspecies of *C. aequitriloba*^{20,52}. The proposed change of circumscription brings about a change in the species' threat category according to IUCN⁵³ criteria. We sampled all known populations of *C. fragilis* to date, updating the number of populations from three to nine, and therefore suggest that Near Threatened (NT) would be a more adequate category than the formerly applied Critically Endangered (CR)²⁷.

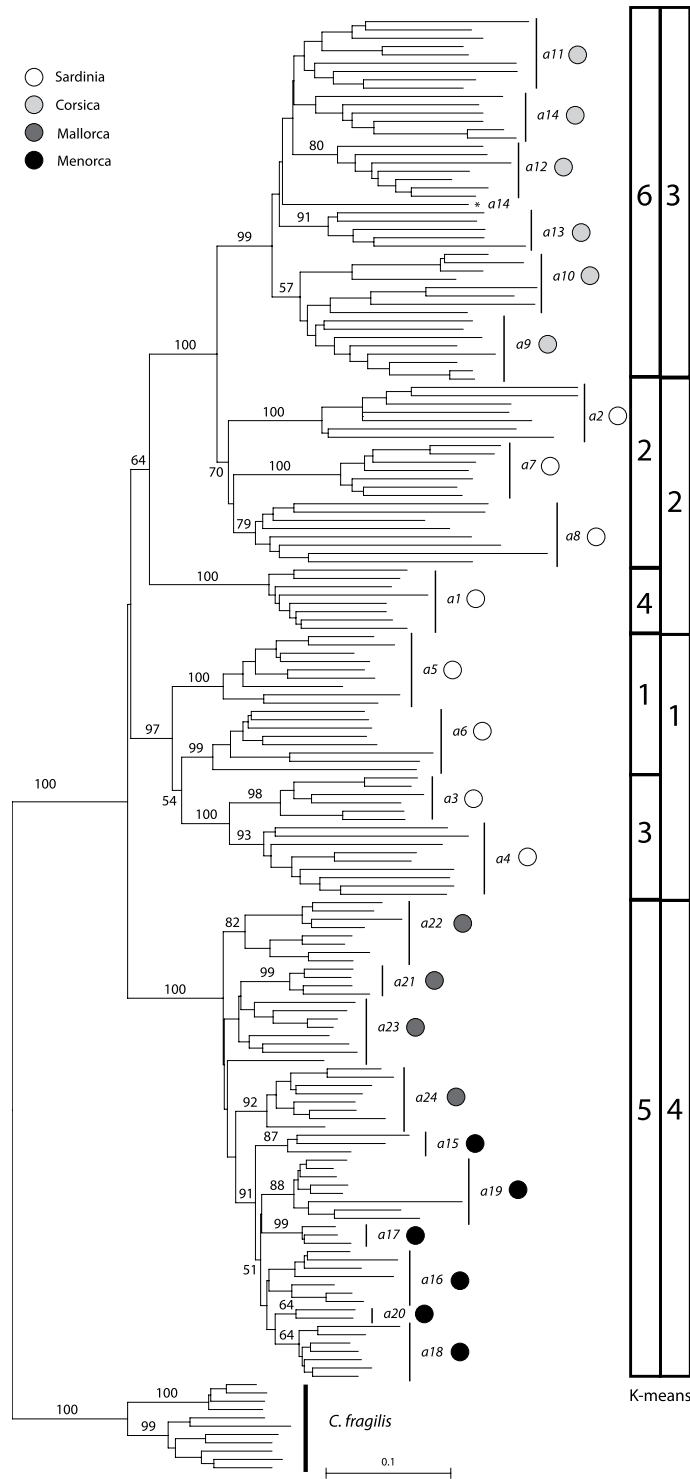


Figure 4. Neighbor-Joining tree based on AFLP data generated for *C. aequitriloba* following the taxonomic concept adopted in the present study. Distributions of clusters identified by nonhierarchical K-means clustering at K = 4 and K = 6 are shown to the right. Circles next to population names indicate the island of origin.

A Corso-Sardinian common ancestor for the BCS endemics. Spatial reconstruction of evolutionary dynamics using relaxed random walks (Fig. 3) suggests a Sardinian origin of the BCS endemic species of *Cymbalaria*. The low haplotype diversity found in the Balearic Islands (Fig. 5) is congruent with a Sardinian origin, but does not allow to reject the alternative hypothesis of a Corsican origin. An east (Corsica-Sardinia) to west (Balearic Islands) pattern of colonization has been also suggested for the BCS endemics *Arum pictum*⁹ and *Thymus herba-barona*¹¹.

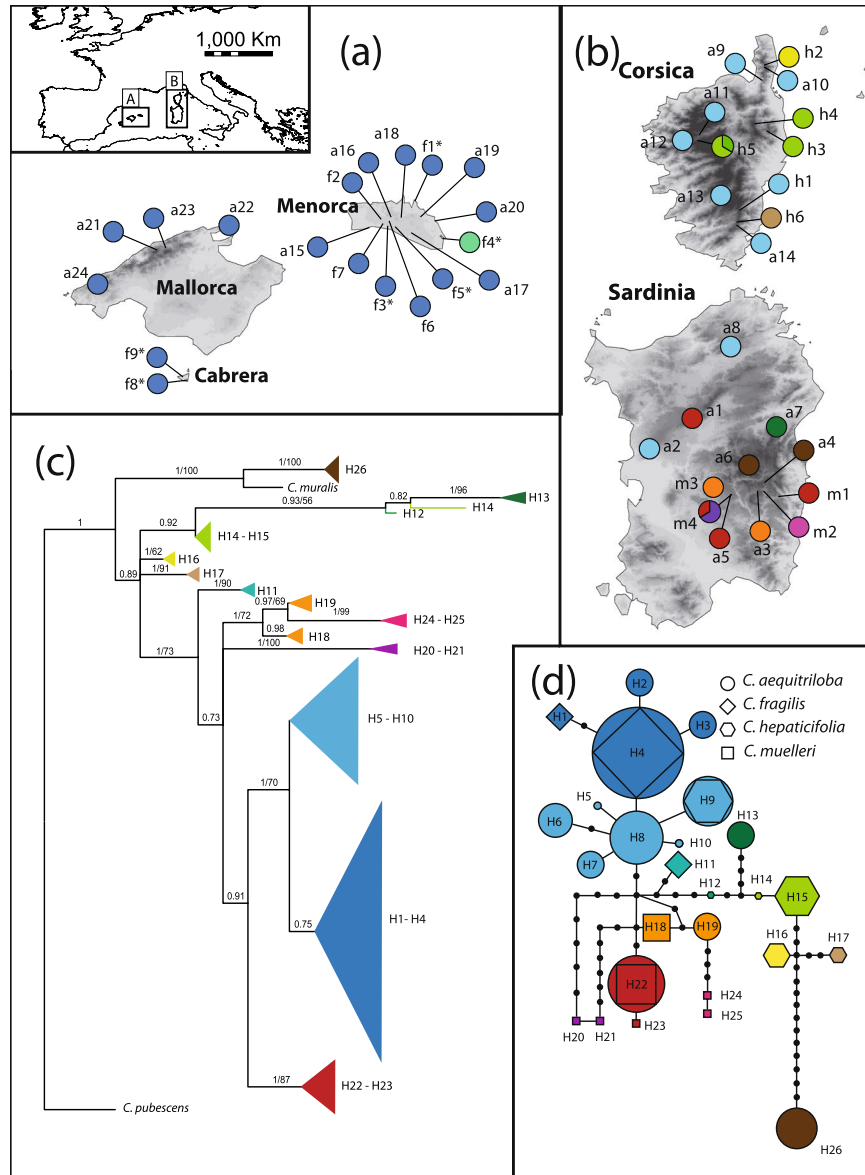


Figure 5. Sampled populations and patterns of plastid DNA (*ndhF* and *rpl32-trnL(UAG)*) variation. Colours correspond to groups of closely related haplotypes. Taxonomy follows the taxonomic concept adopted in the present article (a: *C. aequitriloba*, f: *C. fragilis*, h: *C. hepaticifolia*, m: *C. muelleri*). Maps generated from the altitude layer obtained from <http://www.worldclim.org>¹³², and modified using QGIS v2.18.2¹³³. (a,b) Pie charts indicate the proportions of haplotype groups sampled in each population. Asterisks indicate populations exhibiting mostly distinctive features of *C. fragilis* but previously assigned to *C. aequitriloba* owing to the presence of alveolate seeds. (c) Phylogram from the Bayesian Inference analysis. Numbers above branches indicate Bayesian posterior probabilities/Maximum parsimony bootstrap support values. Haplotypes are given for each lineage next to the tips, numbering corresponds to Table S1. (d) Statistical parsimony haplotype network. Black dots indicate hypothetical unsampled haplotypes, numbering corresponds to Table S1.

AFLP and morphological data confirmed *C. hepaticifolia* and *C. muelleri* as two well differentiated species (Figs 2 and 6). They diverged from a common ancestor probably via an allopatric speciation event, as suggested by their current distribution areas (Fig. 1, Supplementary Fig. S1). However, the SNAPP analysis revealed an alternative topology supported by 30% of trees, which resolved *C. hepaticifolia* as sister of the other BCS endemics, and *C. muelleri* as sister of *C. fragilis* and *C. aequitriloba*. This uncertainty could be caused by rapid simultaneous diversification of the BCS endemics, as suggested for the initial diversification of *Cymbalaria*¹². Indeed, rapid diversification rates after colonisation of new areas by polyploid lineages have been previously suggested^{54,55}. Alternatively, the closer relationship of *C. muelleri* with the *C. aequitriloba*-*C. fragilis* clade could have been caused by gene flow between *C. muelleri* and *C. aequitriloba* or its ancestor in Sardinia. Plastid sequences revealed, indeed, close relationships among the haplotypes found in *C. hepaticifolia* and *C. muelleri* and those of sympatric *C. aequitriloba* populations (Fig. 5).

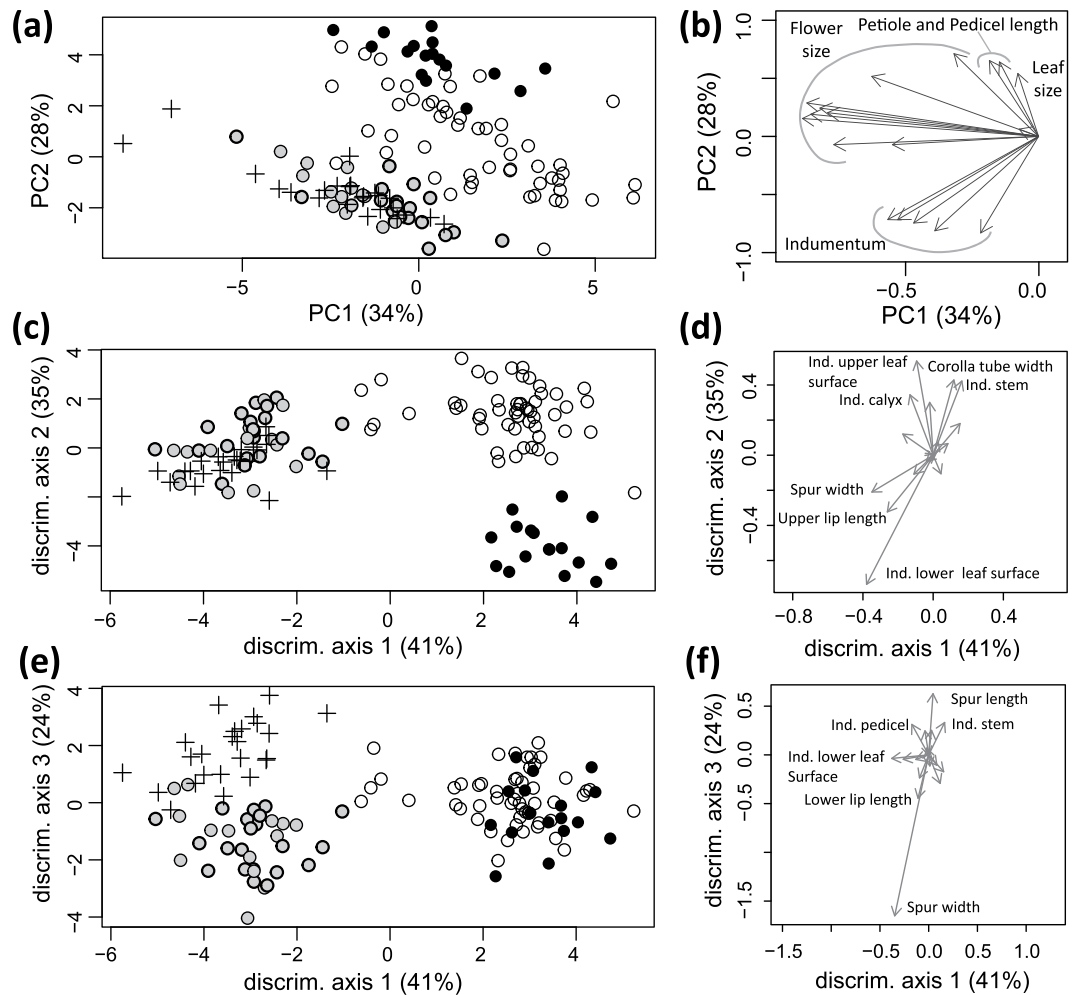


Figure 6. Ordination diagrams of morphometric analyses of 18 floral and vegetative characters obtained for 129 specimens (dataset 1) of *C. aequitriloba* (empty circles), *C. fragilis* (grey dots), *C. hepaticifolia* (black dots) and *C. muelleri* (crosses). Grey dots with thicker outline indicate *C. fragilis* specimens previously considered *C. aequitriloba*. **(a,b)** Principal component analysis. **(a)** Scatter plot of principal component scores for the first two components of morphological variation. **(b)** Character relationships and contribution to the first two axes. **(c–f)** Canonical discriminant analysis. **(c,e)** Scatter plots of discriminant scores for two combinations of the three discriminant axes. **(d,f)** Character relationships and contribution in the same ordination space as **(c,e)**. Characters with high contributions in each combination are identified. (ind. = indumentum).

Both the strongly reduced haplotype diversity of *C. fragilis* in comparison with Corsican and Sardinian populations of the three other studied species (Fig. 5) and the continuous phylogeographic analysis (Fig. 3) support the hypothesis that a LDD event to the Balearic Islands led to the origin of *C. fragilis*. Genetic diversity of *C. fragilis* was not significantly lower than that of *C. aequitriloba* (Supplementary Table S3), suggesting that the LDD event was ancient enough to allow for a recovery of genetic diversity⁵⁶. *Cymbalaria aequitriloba* has its centre of genetic diversity in Sardinia, as shown by the highest among-population variation in the AMOVAs (Supplementary Table S2), the significantly higher genetic diversity observed in Sardinian populations as compared to populations from the other islands (Supplementary Table S3), the high haplotype diversity (Fig. 5) and the high number of separable AFLP gene pools (Fig. 4).

Phylogeography of *C. aequitriloba* and *C. fragilis*. *Cymbalaria aequitriloba* exhibits a higher genetic diversity in Sardinia than in Corsica, as shown by the higher number of plastid haplotypes and AFLP clusters (Figs 4 and 5). This is most probably explained by Pleistocene glacial periods, which had a stronger effect in Corsica than in Sardinia as illustrated by the large glaciers documented in Corsica during the Last Glacial Maximum⁵⁷ (18,000 years ago), and could have caused local extinctions^{58,59}. Later range expansion might have occurred from either lowland Corsican or Sardinian refuge areas, via land bridges connecting both islands multiple times during the Pleistocene^{60,61} or across the narrow Bonifacio strait separating Corsica and Sardinia⁶². Populations from northern and western Sardinia showed a high genetic similarity with Corsican populations (Figs 4 and 5), supporting contacts between the populations of the two islands as shown for *Cistus creticus* L.⁶².

Cymbalaria reached the Balearic Islands via two independent LDD events, as supported by the NJ tree (Fig. 2) and the continuous phylogeographic analysis (Fig. 3), combined with evidence of a relatively young origin of *Cymbalaria* BCS endemics, ca. 15–20 myr later than the last land connection between Corsica-Sardinia and the Balearic Islands¹². The first LDD events ultimately led to the origin of *C. fragilis* and the second caused a range expansion of *C. aequitriloba*. These two events were also supported by the shorter terminal splits depicted in the NJ tree for all Balearic populations indicating strong similarity among individuals (Figs 2 and 4). Significantly lower average genetic diversity of Balearic populations of *C. aequitriloba* as compared to those from Corsica and Sardinia also support the hypothesis of a recent LDD event (Supplementary Table S4). In addition, the strongly reduced haplotype diversity in the Balearic Islands (Fig. 5) may evidence a strong bottleneck resulting from a founder event (e.g., Schönswetter *et al.*⁶³, Burnier *et al.*⁶⁴). With the exception of haplotype H11, the close phylogenetic relationships among haplotypes from the Balearic Islands (Fig. 5c) suggest a common ancestral haplotype for *C. aequitriloba* and *C. fragilis*. This common ancestor might have occurred in western and northern Sardinia as shown by the BEAST continuous phylogeographic analysis and the close phylogenetic relationship between haplotypes H1–H4 from the Balearic Islands and H5–H10 from Corsica and Sardinia (Figs 3 and 5). In the same line, the geographic setting supports a dispersal from western Sardinia, for instance from the surroundings of population a2. It is the closest to the Balearic Islands and thrives on the summit of the Monte Ferru massif, at 1050 m altitude only 12 km from the coast. Alternatively, or in addition, at a later stage hybridization with *C. fragilis* could have led to the sharing of haplotype H4 across species boundaries⁶⁵.

The history of the Balearic populations of both species was likely influenced by sea level variation during the Pleistocene. Land connections during glaciation periods¹⁶ might have aided *C. aequitriloba* and *C. fragilis* to disperse among the Eastern Balearic Islands until both species achieved their current distributions. However, during early Pleistocene interglacials (1.6–0.7 Ma) the sea rose for up to 100 m above current level, leading to considerable reduction of the surface of Menorca, division of Mallorca into two islands and transgression of most of Cabrera^{19,66}. This was likely a major obstacle for the expansion of *C. fragilis* to Cabrera, since only a small area of that island protruded from the sea, suggesting that the colonization most probably occurred after this period. In any case, presence of well-supported clades in Menorca and Cabrera (Fig. 2) supports the current isolation.

In contrast to other members of the tribe Antirrhineae exhibiting winged seeds or adaptation to marine dispersal (e.g. *Linaria* Mill., *Cymbalaria longipes* (Boiss. & Heldr.) A.Chev.)²⁰ neither *C. aequitriloba* nor *C. fragilis* show any apparent adaptation for LDD. Moreover, both species are at least partly achorous and actively prevent seed dispersal by positioning mature capsules in rock crevices²⁰. Notably, very similar features were used as an argument against LDD for the BCS endemic *Arenaria balearica*¹⁰. However, the importance of non-standard vectors has been recognised^{67,68} and rare stochastic events such as LDDs have been highlighted as important drivers of organisms' histories^{69,70}. In fact, LDD has been invoked in several cases in spite of the apparent lack of particular adaptations (e.g. Guzmán and Vargas⁷¹; Dixon *et al.*⁷²; Escudero *et al.*⁷³; Santos-Gally *et al.*⁷⁴; Piñeiro *et al.*⁷⁵). In the case of *Cymbalaria*, the small size of seeds could have favoured LDD as a consequence of rare stochastic events, as have been invoked for *Androsace*⁶⁷ or suggested as an *a priori* possibility for *A. balearica*¹⁰. The success in colonization after a LDD is thought to be governed by the availability of adequate niches rather than by dispersal abilities⁴. In addition, polyploids are better buffered against genetic bottlenecks and subsequent inbreeding resulting from founder effects⁷⁶, a feature that could have helped *C. aequitriloba* and *C. fragilis* to successfully establish in the Balearic Islands.

Finally, our previous results¹² strongly suggested that *C. aequitriloba* is not a palaeoendemic that originated and expanded its range when Corsica-Sardinia and the Balearic Islands were connected⁶ (~20 Ma), but rather colonized the Balearic Islands much later, 1–4 Ma. The results presented here support this statement and furthermore show that *C. aequitriloba* does not meet the criteria attributed to relict taxa, i.e., low levels of infraspecific morphological and genetic variation as a result of a long process of adaptation in refugial isolation, and taxonomic isolation due to extinction of close relatives^{8,37}. In contrast, *C. aequitriloba* is highly morphologically variable^{20,21} (Fig. 6), genetically equally diverse as its closest relatives (Supplementary Table S3) and not taxonomically isolated. Obviously, knowledge of the temporal framework is crucial for any study of palaeoendemism, since it allows for evaluating the congruence between a biogeographic hypothesis and, for instance, the geomorphological distribution of landmasses. In at least two cases focussing on BCS, dating analyses have rejected the hypothesis of a continuous palaeodistribution in the Oligocene landmass followed by fragmentation (*A. pictum*⁸, *T. herba-barona*¹¹). To date, molecular dating has resulted in a temporal framework congruent with a continuous palaeodistribution only in the case of *H. muscivorus*⁸. Although Bobo-Pinilla *et al.*¹⁰ suggested *A. balearica* to be a palaeoendemic, the lack of absolute time estimates prevented them from firmly rejecting a more recent origin. Often, dating analyses are revealing more recent divergent times than formerly thought (e.g. Salvo *et al.*¹³, Fernández-Mazuecos *et al.*¹⁴) and are thus increasingly shifting the biogeographic paradigm for the Mediterranean area.

Incongruence between AFLP and sequence data. The monophyly of each of the four species in the AFLP analyses (Fig. 2) suggests that reproductive isolation is effective at present. However, interspecific gene flow appears possible due to the close proximity between populations of *C. aequitriloba* and *C. hepaticifolia* in Corsica, and between *C. aequitriloba* and *C. muelleri* in Sardinia (P.C. & M.G., personal field observations). A few incongruities between mostly nuclear-derived AFLPs⁷⁷ and maternally inherited plastid DNA sequences⁷⁸, could indicate past reticulation events. Although incomplete lineage sorting (ILS) of ancient polymorphisms cannot be ruled out as cause for these incongruities^{79–83} hybridization appears more likely due to the grouping of haplotypes retrieved from geographically close populations of different taxa^{84,85}. The most obvious case is the occurrence of haplotype 26 otherwise characteristic of the phylogenetically distant (cf. Carnicero *et al.*¹²) *C. muralis* in populations a4 and a6 of *C. aequitriloba* from Sardinia (Fig. 4). *Cymbalaria muralis*, a diploid species native to the Apennine and northern Balkan Peninsulas, is naturalized almost worldwide in temperate areas²⁰. The population

of *C. muralis* used here is from the same area in Sardinia where haplotype 26 was found; hybridization resulting in chloroplast capture of a *C. muralis* haplotype is, therefore, likely.

Two statistically supported incongruities were found between the AFLP dataset and the previous phylogeny¹². The latter retrieved a clade formed by *C. aequitribiloba* accessions from Corsica and *C. hepaticifolia*, and inferred a common ancestor for *C. fragilis* and *C. muelleri*; these two clades did not appear in the AFLP analysis (Fig. 2). Although nrDNA and AFLPs have been shown to yield mostly congruent signals⁸⁶, incongruities have also been reported^{87,88}. Here, we suggest that hybridization followed by concerted evolution of nrDNA towards one parental copy^{79,89}, ILS and/or too little variation in nrDNA sequences are the most probable causes for the incongruities. As above argued for the plastid DNA data, we consider hybridization the most probable cause for the grouping of *C. aequitribiloba* and *C. hepaticifolia* in the nrDNA tree, since they occur in the same areas and exhibit no obvious reproductive barriers (P.C. & M.G., personal field observations). In contrast, ILS or insufficient variation among nrDNA sequences seem the most probable causes for incongruence affecting *C. muelleri* and *C. fragilis*, since they have non-overlapping distribution areas. In the same line, low variability of DNA sequences has previously been reported as cause for incongruence between DNA sequences and AFLP fingerprints⁹⁰. Therefore, we consider the aforementioned relationships retrieved from the nrDNA phylogeny to be spurious. Similarly, it has been stated that AFLPs explain better phylogenetic relationships among closely related species, especially when dealing with complex evolutionary processes such as polyploidy and hybridization⁸⁶.

Identification Key. Here we provide a dichotomous key to all *Cymbalaria* species occurring in BCS.

- 1 Leaves and stems glabrous or only with occasional hairs on young organs..... 2
Leaves and stems hairy..... 3
- 2 Leaves entire to 5-lobed. Corolla 9.7–12.6(13.9) mm long..... *C. hepaticifolia*
Leaves mostly 5- to 9-lobed. Corolla 8–11 mm long..... *C. muralis* subsp. *muralis*
- 3 Lower lip of the corolla 8.2–16.6(17.7) mm wide. 9–12 seeds per capsule. Seeds 1.1–1.3 mm long, alveolate to tuberculate..... *C. muelleri*
Lower lip of the corolla 4.2–13 mm wide. 4–56 seeds per capsule. Seeds 0.6–1.1(1.2) mm long, smooth to alveolate..... 4
- 4 Stems, petioles and pedicels highly fragile when fresh, villous. Corolla white to pale violet, ratio spur length/spur width 0.9–2.8. Seeds 0.6–0.9 mm long, smooth to finely alveolate..... *C. fragilis*
Stems, petioles and pedicels not fragile, sub-glabrous to villous. Corolla pale violet to violet or pink, ratio spur length/spur width 1.6–5.6. Seeds 0.8–1.1 (1.2) mm long, alveolate..... *C. aequitribiloba*

Methods

Plant material and DNA extraction. Leaf material for molecular analyses was collected in the field in 2012 and 2013, dried and stored in silica gel. The sampling localities are shown in Fig. 1 and voucher details are provided in Supplementary Table S1. Total genomic DNA was extracted from ca. 10–30 mg silica gel dried leaf material following the CTAB protocol⁹¹ with some modifications⁹². When less than 10 mg of dried material was available, no sorbitol washing was applied.

AFLP fingerprinting. AFLP fingerprinting was performed with usually eight to ten individuals per population (Supplementary Table S1), using *C. microcalyx* (Boiss.) Wettst. and *C. pubescens* (C.Presl) Cufod. as out-group taxa. AFLP profiles were generated following established protocols⁹³ with modifications^{94,95}. Three blanks (DNA replaced by water) were included to test for contamination, and 32 samples (10.5%) were used as replicates between PCR batches to test the reproducibility of AFLP fingerprinting. Based on an initial primer trial the following three selective primer combinations were chosen for selective PCR (fluorescent dye in brackets): EcoRI (6-FAM)ACA/MseI-CAT, EcoRI (VIC)AGG/MseI-CAC, and EcoRI (NED)AAC/MseI-CAG (6-FAM labelled primers: Sigma-Aldrich; NED and VIC labelled primers: Applied Biosystems). Selective PCR products were purified using Sephadex G-50 Fine (GE Healthcare Bio-Sciences, Uppsala, Sweden) applied to a MultiScreen-HV plate (Millipore, Molsheim, France) in three steps of 200 µl each and packed at 600 g for 1, 1 and 5 min, respectively. Then 0.8 µl of the elution product was mixed with 10 µl formamide (Applied Biosystems) and 0.125 µl GeneScan 500 ROX (Applied Biosystems) and run on an ABI 3130 automated capillary sequencer at the Department of Botany & Biodiversity Research of the University of Vienna, Austria.

Analyses of AFLP data. Electropherograms were analysed with Peak Scanner version 1.0 (Applied Biosystems) using default peak detection parameters except employing light peak smoothing. The minimum fluorescent threshold was set to 50 relative fluorescence units. Automated binning and scoring of the AFLP fragments were performed using RawGeno 2.0-1⁹⁶ for R 2.15.0⁹⁷ with the following settings: scoring range 140–500 bp, minimum intensity 50 relative fluorescence units, minimum bin width 1 bp, and maximum bin width 1.5 bp. We chose not to score AFLP fragments shorter than 140 bp as short fragments are more prone to homoplasy than long ones⁹⁸. Fragments with a reproducibility lower than 80% based on sample-replicate comparisons were eliminated. The error rate³⁹ was calculated as the ratio of mismatches (scoring 1 versus 0) over phenotypic comparisons in AFLP profiles of replicated individuals. Fragments present/absent in only one individual were excluded.

A Neighbour-Joining (NJ) analysis based on Nei-Li genetic distances⁹⁹ was conducted and bootstrapped (2000 pseudo-replicates) with TREECON v.1.3b¹⁰⁰.

As an alternative approach to infer the relationships among the studied species, we conducted a species tree analysis using SNAPP¹⁰¹. SNAPP is a Bayesian MCMC sampler, which interfaces with the BEAST package¹⁰² to

estimate species trees using a coalescent multispecies approach from biallelic input data. Due to the long run times, we reduced the sampling to two individuals per population. Three MCMCs were initially intended to run for 10 million generations, sampling trees every 1000 generations. However, runs were stopped before reaching the final generation number after checking the convergence of most parameters in Tracer 1.6.0¹⁰³, due to the long run times. Therefore, the first MCMC was run for 153000 generations, the second for 126000 generations, and the third for 392000 generations. After discarding the first 10% of trees of each run, the resulting 980 trees were combined with LogCombiner v.1.8.2, and the topologies were visualized in DensiTree v2.2.5¹⁰⁴.

A Bayesian phylogeographical diffusion model in continuous space¹⁰⁵ was performed on the AFLP dataset using BEAST v1.8.2¹⁰⁶ to infer the main pattern of colonization between the Balearic Islands and Corsica-Sardinia. Geographic input data were coordinates of each population (Supplementary Table S1). We used a simple substitution model with estimated state frequencies for phylogeny inference and the Bayesian skyline coalescent prior¹⁰⁷ and a piecewise-linear skyline model was set to model population growth. A strict molecular clock was used due to the simple presence-absence structure of AFLP data. The diffusion process was modelled by a lognormal relaxed random walk process, which is an extension of a phylogenetic Brownian motion process that rescales the precision matrix by a branch-specific scalar that is drawn independently from an identical distribution¹⁰⁸, in our case a lognormal distribution centred on 1. We specified a prior exponential distribution on the standard deviation of the lognormal distribution with a mean of 5. We added random jitter with a window size 1.0 to the tips, as more individuals were sampled from the same location. The analysis of the diffusion inference was run for 200 million generations, logging parameters every 5,000 generations. The performance of the analysis as well as mixing and effective sample size values for all parameters were checked in Tracer 1.6.0¹⁰⁹. A maximum clade credibility tree (MCC) was produced and annotated with Tree Annotator (part of the BEAST package) after removing 25% of trees as burnin and visualized with FigTree 1.4.2. The diffused MCC tree with annotated diffusion estimates was visualized in SPREAD v.1.0.6¹¹⁰ and projected together with polygons representing ancestral areas using ArcGIS 10.3.

Analyses of molecular variance (AMOVAs) and genetic diversity were computed with ARLEQUIN 3.5. Groups were defined according to taxa and – for *C. aequitriboloba* – to geographic distribution (islands and archipelagos). Average gene diversity over loci (π_n) was obtained for populations and groups after removing populations with fewer than four sampled individuals. Comparison of gene diversity among groups was performed with Student's t-tests after testing for normality with Shapiro-Wilk normality tests.

A separate dataset for *C. aequitriboloba* was prepared and, after independent automated binning and scoring, analysed in the same way. It comprised all populations of the *C. aequitriboloba* clade (see Results) plus *C. fragilis* populations f8 and f9 as outgroup, based on the results of the NJ analysis for the entire dataset. In addition, we performed nonhierarchical K-means clustering¹¹¹, a model-free clustering approach. We used a script of Arrigo *et al.*¹¹² in R⁹⁷ to identify genetically homogeneous groups in a dataset pruned to *C. aequitriboloba*. We performed 100,000 independent runs (i.e., starting from random points) for each assumed value of K (i.e., the number of groups) ranging from 2 to 10.

Amplification and sequencing of plastid DNA markers. We amplified and sequenced the plastid *ndhF* region and the *rpl32-trnL*^(UAG) spacer for three individuals per population, with the exceptions of populations a20 and m1 – for which we could only obtain sequences from two individuals –, and the outgroup (*C. muralis* G.Gaertn., B.Mey. & Scherb. and *C. pubescens*) – with a single individual per population. We used the primers 3'F¹¹³, +607¹¹⁴ and the internal specific primers *ndhF* CymbF and *ndhF* CymbR¹² for the *ndhF* region and *rpl32F* and *trnL*^{(UAG)115} for the *rpl32-trnL* spacer. All reactions were carried out in a MasterCycler Gradient thermocycler (Eppendorf).

For both plastid regions, the reaction mix (total volume 25 μ L) contained 9 μ L of RedTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich), 1.05 μ L bovine serum albumin (BSA; 1 mg mL⁻¹; Promega), 1.05 μ L of each primer (10 μ M) and 1.5 μ L template DNA. We followed the PCR profiles described in Galbany-Casals *et al.*¹¹⁶ for the *ndhF* and Magauer *et al.*⁴⁵ for the *rpl32-trnL*^(UAG) spacer.

The quality of the PCR products was checked on 1% TBE agarose gels. Subsequently, the amplification products were purified enzymatically using Exonuclease I and Shrimp Alkaline Phosphatase (Fermentas) according to the manufacturer's instructions. Sanger sequencing was conducted by a commercial sequencing facility (Eurofins MWG Operon, Munich, Germany) using the primers 3'F, +607, *rpl32F* and *trnL*(UAG). Sequences were examined and aligned by hand using Chromas Lite 2.0 (Technelysium Pty Ltd, Tewantin, Australia) and Mega 6.06¹¹⁷. GenBank accession numbers are given in Supplementary Table S1.

Analyses of sequence data. The alignment of the concatenated plastid markers was analysed using statistical parsimony as implemented in TCS 1.21¹¹⁸ with the connection limit set to 95%. For this analysis, an indel longer than 1 bp was reduced to a single base pair column allowing this structural mutation to be counted as single base pair mutation only. Two overlapping, but clearly non-homologous indels were detected and considered different. Differences on the indel length (1 bp) caused by different lengths of poly-T regions were not considered due to the high likelihood of homoplastic evolution¹¹⁹. Specimens with missing data in polymorphic positions were excluded to avoid ambiguous haplotype assignments. After these modifications, gaps were treated as fifth character state in this analysis.

Maximum parsimony (MP) analyses as well as MP bootstrap analyses of both data sets were performed using PAUP 4.0b10¹²⁰. Indels were coded as binary characters using the simple indel coding method¹²¹ using SeqState 1.4.1¹²². The most parsimonious trees were searched heuristically with 1000 replicates of random sequence addition, TBR swapping, and MulTrees on. The swapping was performed on a maximum of 1000 trees (nchuck = 1000). All characters were equally weighted and unordered. The data set was bootstrapped using full

heuristics, 1000 replicates, TBR branch swapping, MulTrees option off, and random addition sequence with five replicates. *Cymbalaria pubescens* was used as outgroup based on a previous study¹².

Bayesian Inference (BI) analysis was performed employing MrBayes v.3.2¹²³ applying the GTR + G substitution model proposed by the Akaike information criterion implemented in JModelTest 0.1.1¹²⁴. Indels were coded as binary characters using the simple indel coding method¹¹⁹. We generated 10,000 trees running MrBayes for 5,000,000 generations and sampling one of every 500 generations. After ensuring that the Markov chain Monte Carlo (MCMC) reached stationarity, we discarded the first 2500 trees as burn-in.

Morphometric analyses. Twenty-seven characters were scored on the basis of previous studies on the tribe Antirrhineae^{20,125}, of which 18 were quantitative and seven were ordinal (Supplementary Table S4). Given that ordination analyses perform better with quantitative characters¹²⁶, qualitative characters such as stem fragility and leave turgescence, which are diagnostic characters in the group, have not been included in these analyses. In *Cymbalaria*, indumentum and seed ornamentation, treated here as ordinal characters, have been considered of main diagnostic importance^{21,23,127,128}. Terminology used for indumentum characters follows Payne¹²⁹ and that for seed characters follows Sutton²⁰. Seven characters corresponded to vegetative features, 11 to floral features, and seven to fruit and seed features. For the vegetative characters, three measurements per specimen were averaged when possible. As it was not possible to score both floral and fruit or seed characters for all populations, we analysed two datasets. Dataset 1 included 11 floral and seven vegetative characters and dataset 2 included seven fruit and seven vegetative characters. Dataset 1 comprised 129 specimens from 31 populations and dataset 2 comprised 85 specimens from 33 populations (Supplementary Table S1). All specimens used in these analyses were also included in the molecular analyses, with the exception of *C. muelleri* and some individuals from population a5 of *C. aequitriloba* included in dataset 1, which were added later in order to provide floral measurements. Floral measurements were performed in the field with a caliper. Features of indumentum, calyx, fruits and seeds were examined under a binocular stereoscopic microscope. The other characters were measured on scanned specimens using Image J¹³⁰. Analyses were conducted using a set of R functions contained in MorphoTools ver. 1.01¹³¹. Pearson and Spearman correlation coefficients were computed to reveal correlation structure among the characters and to ensure that no strong correlations ($>|0.95|$) were present. After standardization to zero mean and unit variance, principal component analysis (zero-centred PCA based on a covariance matrix) was applied to datasets 1 and 2 to display the overall variation pattern along the first two components. A canonical discriminant analysis (CDA) was performed on dataset 1 to assess the morphological differentiation among the four taxa. Morphological intermediate specimens between *C. fragilis* and *C. aequitriloba* from populations f1, f3, f4, f5, f8 and f9 were included in *C. fragilis* according to the results of the molecular analyses and the PCA.

Data Accessibility

DNA sequences have been deposited in GenBank (see Supplementary Table S1). AFLP data matrix, alignments of concatenated plastid sequences and the xml files with details of the BEAST analyses are available at Dryad: <https://doi.org/10.5061/dryad.0cr574s>.

References

- Vogiatzakis, I. N., Pungetti, G. & Mannon, A. M. *Mediterranean island landscapes: natural and cultural approaches* (Springer Publishing, New York, 2008).
- Médail, F. & Quézel, P. Biodiversity Hotspots in the Mediterranean Basin: Setting Global Conservation Priorities. *Conserv. Biol.* **13**, 1510–1513 (1999).
- Thompson, J. D. *Plant evolution in the Mediterranean*. (Oxford University Press, Oxford, 2005).
- Nieto Feliner, G. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspect. Plant Ecol. Evol. Syst.* **16**, 265–278 (2014).
- Rosenbaum, G., Lister, G. S. & Duboz, C. Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. *J. Virt. Ex.* **8**, 107–130 (2002).
- Speranza, F. *et al.* Age of the Corsica–Sardinia rotation and Liguro–Provençal Basin spreading: new paleomagnetic and Ar/Ar evidence. *Tectonophysics* **347**, 231–251 (2002).
- Contandriopoulos, J. & Cardona, M. A. Caractère original de la flore endémique des Baléares. *Bot. Helv.* **94**, 101–132 (1984).
- Contandriopoulos, J. Spécificité de l'endémisme Corse. *Atti Dei Convegni Lincei* **85**, 394–416 (1990).
- Mansion, G. *et al.* Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean Basin by the angiosperm family Araceae. *Syst. Biol.* **57**, 269–85 (2008).
- Bobo-Pinilla, J. *et al.* Phylogeography of *Arenaria balearica* L. (Caryophyllaceae): evolutionary history of a disjunct endemic from the Western Mediterranean continental islands. *PeerJ* **4**, e2618 (2016).
- Molins, A., Bacchetta, G., Rosato, M., Rosselló, J. A. & Mayol, M. Molecular phylogeography of *Thymus herba-barona* (Lamiaceae): insight into the evolutionary history of the flora of the western Mediterranean islands. *Taxon* **60**, 1295–1305 (2011).
- Carnicero, P., Sáez, L., García-Jacas, N. & Galbany-Casals, M. Different speciation types meet in a Mediterranean genus: the biogeographic history of *Cymbalaria* (Plantaginaceae). *Taxon* **66**, 393–407 (2017).
- Salvo, G., Ho, S. Y. W., Rosenbaum, G., Ree, R. & Conti, E. Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae). *Syst. Biol.* **59**, 705–22 (2010).
- Fernández-Mazuecos, M., Jiménez-Mejías, P., Rotllan-Puig, X. & Vargas, P. Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered *Naufraga* (Apiaceae). *Perspect. Plant Ecol. Evol. Syst.* **16**, 190–202 (2014).
- Magri, D. *et al.* The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Mol. Ecol.* **16**, 5259–5266 (2007).
- Vesica, P. L. *et al.* Late Pleistocene Paleoclimates and sea-level change in the Mediterranean as inferred from stable isotope and U-series studies of overgrowths on speleothems, Mallorca, Spain. *Quat. Sci. Rev.* **19**, 865–879 (2000).
- Mansion, G., Selvi, F., Guggisberg, A. & Conti, E. Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral area reconstruction. *J. Biogeogr.* **36**, 1282–1296 (2009).
- Salvi, D., Harris, D. J., Bombi, P., Carretero, M. A. & Bologna, M. A. Mitochondrial phylogeography of the Bedriaga's rock lizard, *Archaeolacerta bedriagae* (Reptilia: Lacertidae) endemic to Corsica and Sardinia. *Mol. Phylogenet. Evol.* **56**, 690–697 (2010).

19. Mayol, M. *et al.* 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. *Ann. Bot.* **109**, 429–441 (2012).
20. Sutton, D. A. *A revision of the tribe Antirrhineae* (British Museum - Natural History, London, 1988).
21. Güemes J. *Flora iberica*. Vol. VIII. *Cymbalaria* Hill (eds Benedi, C., Rico, E., Güemes, J. & Herrero A.) 210–215 (CSIC, Madrid, 2009).
22. Gamsans, J. & Marzocchi, J. F. 2003. *La Flore endémique de la Corse*. (Edisud, Aix-en-Provence, 2003).
23. Jeanmonod, D. & Gamsans, J. *Flora Corsica*. (Edisud, Aix-en-Provence, 2007).
24. Onnis, A. & Floris, C. Il numero cromosomico di *Linaria mulleri* Moris. *Atti Soc. Tosc. Sci. Nat. Pisa Mem. ser. B* **74**, 76–81 (1967).
25. Arrigoni, P. V., Corrias, B., Corrias, S. D., Raffaelli, M. & Valsecchi, F. Le piante endemiche della Sardegna. *Boll. Soc. Sarda Sci. Nat.* **18**, 223–295 (1979).
26. Castro, M. & Rosselló, J. A. New chromosome numbers for plant taxa endemic to the Balearic Islands. *Folia Geobotanica* **41**, 433–451 (2006).
27. Bañares, A., Blanca, G., Güemes, J., Moreno, J. C. & Ortiz, S. *Atlas y libro rojo de la flora vascular amenazada de España*. (Dirección General para la Biodiversidad, Publicaciones del O.A.P.N., Madrid, 2004).
28. Bolòs, O. de, Vigo, J., Masalles, R. M. & Ninot, J. M. *Flora manual dels Països Catalans*. (Pòrtic, Barcelona, 2005).
29. Fraga, P., Gradaille, J. L., Pallicer, X. & Sastre, B. Notes floristiques de les Illes Balears (X). *Boll. Soc. Hist. Nat. Balears* **40**, 151–155 (1997).
30. Elisens, W. J. & Tomb, A. S. Seed morphology in New World Antirrhineae (Scrophulariaceae): systematic and phylogenetic implications. *Plant Syst. Evol.* **142**, 23–47 (1983).
31. Elisens, W. J. The systematic significance of seed coat anatomy among New World species of tribe Antirrhineae (Scrophulariaceae). *Syst. Bot.* **10**, 282–299 (1985).
32. Vigalondo, B., Fernández-Mazuecos, M., Vargas, P. & Sáez, L. Unmasking cryptic species: morphometric and phylogenetic analyses of the Ibero-North African *Linaria incarnata* complex. *Bot. J. Linn. Soc.* **177**, 395–417 (2015).
33. Dahlgren, R., Karlsson, T. & Lassen, P. Studies on the flora of the Balearic Islands I. *Bot. Notiser* **124**, 249–269 (1971).
34. Cardona, M. A. & Contandriopoulos, J. 1983. IOPB chromosome number reports 79. *Taxon* **32**, 320–324 (1983).
35. Verlaque, R., Contandriopoulos, J. & Aboucaya, A. Recherches cyto-biogéographiques sur quelques taxons rares ou endémiques de Corse: I. *Candollea* **48**, 562–576 (1993).
36. Heitz, E. Über multiple und aberrante Chromosomenzahlen. *Abh. Naturw. Ver. Hamburg* **21**, 47–57 (1927).
37. Pericàs, J. & Roselló, J. Sobre la falsa presència de *Cymbalaria hepaticifolia* (Poiret) Wettst. a Balears. *Boll. Soc. Hist. Nat. Balears* **27**, 195–199 (1983).
38. Favarger, C. & Contandriopoulos, J. Essai sur l'endémisme. *Psychol. Bull.* **71**, 384–408 (1961).
39. Bonin, A. *et al.* How to track and assess genotyping errors in population genetics studies. *Mol. Ecol.* **13**, 3261–3273 (2004).
40. Stull, G. W., D de Stefano, R., Soltis, D. E. & Soltis, P. S. Resolving basal lamiid phylogeny and the circumscription of Icacinaeaceae with a plastome-scale data set. *Am. J. Bot.* **102**, 1794–1813 (2015).
41. Carvalho-Sobrinho, J. *et al.* Revisiting the phylogeny of Bombacoideae (Malvaceae): novel relationships, morphologically cohesive clades, and a new tribal classification based on multilocus phylogenetic analyses. *Mol. Phylogenet. Evol.* **101**, 56–74 (2016).
42. Puglisi, C., Yao, T. L., Milne, R., Möller, M. & Middleton, D. J. Generic recircumscription in the Loxocarpaceae (Gesneriaceae), as inferred by phylogenetic and morphological data. *Taxon* **65**, 277–292 (2016).
43. Španiel, S., Marhold, K., Zozomová-Lihová, J. & Thiv, M. A new circumscription of *Alyssum montanum* ssp. *montanum* and *A. montanum* ssp. *gmelinii* (Brassicaceae) in Central Europe: molecular and morphological evidence. *Bot. J. Linn. Soc.* **169**, 378–402 (2012).
44. Kuzmanović, N. *et al.* Genetic, cytological and morphological differentiation within the Balkan-Carpathian *Sesleria rigida* sensu Fl. Eur. (Poaceae): a taxonomically intricate tetraploid-octoploid complex. *Taxon* **62**, 458–472 (2013).
45. Magauer, M., Schönswetter, P., Jang, T. & Frajman, B. Disentangling relationships within the disjunctly distributed *Alyssum ovirens*/*A. wulfenianum* group (Brassicaceae), including description of a novel species from the north-eastern Alps. *Bot. J. Linn. Soc.* **176**, 486–505 (2014).
46. Ronikier, M. & Zalewska-Gałosz, J. Independent evolutionary history between the Balkan ranges and more northerly mountains in *Campanula alpina* s.l. (Campanulaceae): genetic divergence and morphological segregation of taxa. *Taxon* **63**, 116–131 (2014).
47. Kellner, A., Ritz, C. M. & Wissmann, V. Low genetic and morphological differentiation in the European species complex of *Rosa sherardii*, *R. mollis* and *R. villosa* (*Rosa* section *Caninae* subsection *Vestitae*). *Bot. J. Linn. Soc.* **174**, 240–256 (2014).
48. Van Son, G. A proposal for the restriction of the use of the term subspecies. *The Lepid. News* **9**, 1–3 (1995).
49. Hamilton, C. W. & Reichard, S. H. Current practice in the use of subspecies, variety, and forma in the classification of wild plants. *Taxon* **41**, 485–498 (1992).
50. Stuessy, T. *Plant taxonomy: the systematic evaluation of comparative data*. (Columbia University Press, New York, 2009).
51. Rodríguez, J. J. Additions a la flore de Minorca. *Bull. Soc. Bot. Fr.* **25**, 238–242 (1978).
52. Chevalier, A. Les espèces élémentaires françaises du genre *Cymbalaria*. *Bull. Soc. Bot. Fr.* **83**, 638–653 (1936).
53. IUCN. 2012. IUCN Red List Categories and Criteria: Version 3.1. Second edition. Gland and Cambridge.
54. Crawford, D. J., *et al.* Genetic diversity in Asteraceae endemic to oceanic islands: Baker's Law and polyploidy. (eds Funk, V., Susanna, A., Stuessy, T. & Bayer, R.) 139–151 (International Association for Plant Taxonomy, Vienna, 2009).
55. Vitales, D. *et al.* The explosive radiation of *Cheirilophus* (Asteraceae, Cardueae) in Macaronesia. *BMC Evol. Biol.* **14**, 118 (2014).
56. Luikart, G., Allendorf, F. W., Cornuet, J. M. & Sherwin, W. B. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **89**, 238–247 (1988).
57. Kuhlmann, J. *et al.* Würmian maximum glaciation in Corsica. *Austrian J. Earth Sc.* **97**, 68–81 (2005).
58. Reille, M. *et al.* The late-glacial at Lac de Creno (Corsica, France): a key site in the western Mediterranean Basin. *New Phytol.* **135**, 547–559 (1997).
59. Salvi, D., Bisconti, R. & Canestrelli, D. High phylogeographical complexity within Mediterranean islands: insights from the Corsican fire salamander. *J. Biogeogr.* **43**, 192–203 (2016).
60. Lambeck, K., Antonioli, F., Purcell, A. & Silenzi, S. Sea-level change along the Italian coast for the past 10,000 yr. *Quat. Sci. Rev.* **23**, 1567–1598 (2004).
61. Lambeck, K. & Purcell, A. Sea-level change in the Mediterranean Sea since the LGM: model predictions for tectonically stable areas. *Quat. Sci. Rev.* **24**, 1969–1988 (2005).
62. Falchi, A. *et al.* Phylogeography of *Cistus creticus* L. on Corsica and Sardinia inferred by the *trnL*-F and *rpl32-trnL* sequences of cpDNA. *Mol. Phylogenet. Evol.* **52**, 538–543 (2009).
63. Schönswetter, P., Elven, R. & Brochmann, C. Trans-Atlantic dispersal and large-scale lack of genetic structure in the circumpolar, arctic-alpine sedge *Carex bigelowii* s.l. (Cyperaceae). *Am. J. Bot.* **95**, 1006–1014 (2008).
64. Burnier, J., Buerki, S., Arrigo, N., Küpfer, P. & Alvarez, N. Genetic structure and evolution of Alpine polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Mol. Ecol.* **18**, 3730–3744 (2009).
65. Currat, M., Ruedi, M., Petit, R. J. & Excoffier, L. The hidden side of invasions: massive introgression by local genes. *Evolution* **62**, 1908–1920 (2008).
66. Gràcia, F., Clamor, B., Landreth, R., Vicens, D. & Watkinson, P. Evidències geomorfològiques dels canvis del nivell marí. *Monogr. Soc. Hist. Nat. Balears* **9**, 91–119 (2001).

67. Nathan, R. *et al.* Mechanisms of long-distance seed dispersal. *Trends Ecol. Evol.* **23**, 638–647 (2008).
68. Gillespie, R. G. *et al.* Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol. Evol.* **27**, 47–55 (2012).
69. Losos, J. B. & Malher, L. D. *Evolution after Darwin: the First 150 Years. Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation.* (eds Bell, M. A., Eanes, W. F., Futuyma, D. J. & Levinton, J. S.) 381–420 (Sinauer Associates, Sunderland, 2010).
70. Roquet, C., Boucher, F. C., Thuiller, W. & Lavergne, S. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* **40**, 1874–1886 (2013).
71. Guzmán, B. & Vargas, P. Long-distance colonization of the Western Mediterranean by *Cistus ladanifer* (Cistaceae) despite the absence of special dispersal mechanisms. *J. Biogeogr.* **36**, 954–968 (2009).
72. Dixon, C. J., Schönswetter, P., Suda, J., Wiedermann, M. M. & Schneeweiss, G. M. Reciprocal Pleistocene origin and postglacial range formation of an allopolyploid and its sympatric ancestors (*Androsace adfinis* group, Primulaceae). *Mol. Phylogenet. Evol.* **50**, 74–83 (2009).
73. Escudero, M., Valcárcel, V., Vargas, P. & Luceño, M. Bipolar disjunctions in *Carex*: long-distance dispersal, vicariance, or parallel evolution? *Flora* **205**, 118–127 (2010).
74. Santos-Gally, R., Vargas, P. & Arroyo, J. Insights into Neogene Mediterranean biogeography based on phylogenetic relationships of mountain and lowland lineages of *Narcissus* (Amaryllidaceae). *J. Biogeogr.* **39**, 782–798 (2011).
75. Piñeiro, R. *et al.* Circumarctic dispersal and long-distance colonization of South America: the moss genus *Cinclidium*. *J. Biogeogr.* **39**, 2041–2051 (2012).
76. Linder, H. P. & Barker, N. P. Does polyploidy facilitate long-distance dispersal? *Ann. Bot.* **113**, 1175–1183 (2014).
77. Bussell, J. D., Waycott, M. & Chappill, J. A. Arbitrarily amplified DNA markers as characters for phylogenetic inference. *Perspect. Plant Ecol. Evol. Syst.* **7**, 3–26 (2005).
78. Corriveau, J. L. & Coleman, A. W. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *Am. J. Bot.* **75**, 1443–1458 (1988).
79. Soltis, D. E., Soltis, P. S. & Doyle, J. *Molecular systematics of plants II: DNA sequencing.* (Kluwer Academic, Boston, 1998).
80. Degnan, J. H. & Rosenberg, N. A. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340 (2009).
81. Pelsner, P. B. *et al.* Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. *Am. J. Bot.* **97**, 856–873 (2010).
82. Hilpold, A. *et al.* Phylogeny of the *Centaurea* group (*Centaurea*, Compositae) - Geography is a better predictor than morphology. *Mol. Phylogenet. Evol.* **77**, 195–215 (2014).
83. Sun, M. *et al.* Deep phylogenetic incongruence in the angiosperm clade Rosidae. *Mol. Phylogenet. Evol.* **83**, 156–166 (2015).
84. Mckinnon, G. E., Jordan, G. J., Vaillancourt, R. E., Steane, D. A. & Potts, B. M. Glacial refugia and reticulate evolution: the case of the Tasmanian eucalypts. *Philos. Trans. R. Soc. B* **359**, 275–284 (2004).
85. Lorenz-Lemke, A. P. *et al.* Diversity and natural hybridization in a highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. *Mol. Ecol.* **15**, 4487–4497 (2006).
86. Koopman, W. J. M. Phylogenetic signal in AFLP data sets. *Syst. Biol.* **54**, 197–217 (2005).
87. El-Rabey, H. A., Badr, A., Schäfer-Pregl, R., Martin, W. & Salamini, F. Speciation and species separation in *Hordeum* L. (Poaceae) resolved by discontinuous molecular markers. *Plant Biol.* **4**, 567–575 (2002).
88. Semerikov, V., Zhang, H., Sun, M. & Lascoux, M. Conflicting phylogenies of *Larix* (Pinaceae) based on cytoplasmic and nuclear DNA. *Mol. Phylogenet. Evol.* **27**, 173–184 (2003).
89. Álvarez, I. & Wendel, J. F. Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* **29**, 417–434 (2003).
90. Pelsner, P. B. & Gravendeel, B. & Van der Meijden, R. Phylogeny reconstruction in the gap between too little and too much divergence: the closest relatives of *Senecio jacobaea* (Asteraceae) according to DNA sequences and AFLPs. *Mol. Phylogenet. Evol.* **29**, 613–628 (2003).
91. Doyle, J. & Doyle, J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15 (1987).
92. Tel-Zur, N., Abbo, S., Myslabodski, D. & Mizrahi, Y. Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera *Hylocereus* and *Selenicereus* (Cactaceae). *Plant Mol. Biol.* **17**, 249–254 (1999).
93. Vos, P. *et al.* AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**, 4407–4414 (1995).
94. Schönswetter, P., Solstad, H., García, P. E. & Elven, R. A combined molecular and morphological approach to the taxonomically intricate European mountain plant *Papaver alpinum* s. l. (Papaveraceae) — taxa or informal phylogeographical groups? *Taxon* **58**, 1326–1343 (2009).
95. Rešetnik, I., Frajman, B., Bogdanović, S., Ehrendorfer, F. & Schönswetter, P. Disentangling relationships among the diploid members of the intricate genus *Knautia* (Caprifoliaceae, Dipsacoidae). *Mol. Phylogenet. Evol.* **74**, 97–110 (2014).
96. Arrigo, N., Tuszynski, J. W., Ehrlich, D., Gerdes, T. & Alvarez, N. Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics* **10**, 33 (2009).
97. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (2015).
98. Vekemans, X., Beauwens, T., Lemaire, M. & Roldan-Ruiz, I. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molec. Ecol.* **11**, 139–151 (2002).
99. Nei, M. & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* **76**, 5269–5273 (1979).
100. Van de Peer, Y. & de Wachter, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.* **10**, 569–570 (1994).
101. Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. & RoyChoudhury, A. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932 (2012).
102. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
103. Rambaut, A. & Drummond, A. Tracer 1.6, <http://beast.bio.ed.ac.uk/Tracer> (2013).
104. Bouckaert, R. & Heled, J. DensiTree 2: seeing trees through the forest. *bioRxiv* 012401 (2014).
105. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010).
106. Drummond, A. J., Suchard, M., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–73 (2012).
107. Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).
108. Nylander, S. *et al.* On the biogeography of *Centipeda*: a species-tree diffusion approach. *Syst. Biol.* **63**, 178–191 (2014).
109. Rambaut, A. & Drummond, A. Tracer 1.6, <http://beast.bio.ed.ac.uk/Tracer> (2013).
110. Bielejec, F., Rambaut, A., Suchard, M. A. & Lemey, P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27**, 2910–2912 (2011).
111. Hartigan, J. & Wong, M. A k-means clustering algorithm. *Appl. Stat.* **28**, 100–108 (1979).
112. Arrigo, N. *et al.* Origin and expansion of the allotetraploid *Aegilops gemiculata*, a wild relative of wheat. *New Phytol.* **187**, 1170–1180 (2010).

113. Eldenäs, P., Källersjö, M. & Anderberg, A. Phylogenetic placement and circumscription of tribes Inuleae s. str. and Plucheeae (Asteraceae): evidence from sequences of chloroplast gene *ndhF*. *Mol. Phylogenet. Evol.* **13**, 50–58 (1999).
114. Kim, K. & Jansen, R. *ndhF* sequence evolution and the major clades in the sunflower family. *Proc. Natl. Acad. Sci.* **92**, 10379–10383 (1995).
115. Shaw, J., Lickey, E., Schilling, E. E. & Small, R. L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.* **94**, 275–288 (2007).
116. Galbany-Casals, M., Carnicero-Campmany, P., Blanco-Moreno, J. M. & Smissen, R. D. Morphological and genetic evidence of contemporary intersectional hybridisation in Mediterranean *Helichrysum* (Asteraceae, Gnaphalieae). *Plant Biol.* **14**, 789–800 (2012).
117. Tamura, K., Stecher, G. & Peterson, D. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
118. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–9 (2000).
119. Aldrich, J., Cherney, B. W., Merlin, E. & Christopherson, L. The role of insertions/deletions in the evolution of the intergenic region between *psbA* and *trnH* in the chloroplast genome. *Curr. Genet.* **14**, 137–146 (1988).
120. Swofford, D. *PAUP*: phylogenetics analysis using parsimony (*and other methods), version 4.0 beta 10* (Sinauer Associates, 2002).
121. Simmons, M. P. & Ochoterena, H. Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.* **49**, 369–381 (2000).
122. Müller, K. *SeqState. Appl. Bioinformatics* **4**, 65–69 (2005).
123. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
124. Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
125. Sáez, L. & Crespo, M. B. A taxonomic revision of the *Linaria verticillata* group (Antirrhineae, Scrophulariaceae). *Bot. J. Linn. Soc.* **148**, 229–244 (2005).
126. Jolliffe, I. T. *Principal Component Analysis*, 2nd edn. (Springer Series in Statistics, New York, 2002).
127. Webb, D. A. 1972. *Flora Europaea*, Vol. 3, *Cymbalaria* Hill. (eds Tutin, T.G., Heywood, V. H., Burges, N. A., Moore, D. M., Valentine, D. H., Walters, S. M. & Webb, D. A.) 236–238 (Cambridge University Press, Cambridge, 1972).
128. Pignatti, S. 1982. *Flora d'Italia*. Vol. 2 (Edagricole, Bologna).
129. Payne, W. Glossary on plant hair terminology. *Brittonia* **30**, 239–255 (1978).
130. Abrámov, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. *Biophotonics Int.* **11**, 36–41 (2004).
131. Koutceky, P. MorphoTools: a set of R functions for morphometric analysis. *Plant Syst. Evol.* **301**, 1115–1121 (2015).
132. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
133. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project, <http://qgis.osgeo.org> (2018).

Acknowledgements

We thank all colleagues who assisted during fieldwork or provided plant material. M. Magauer, M. Pacary, D. Pirkebner and R. Vilatersana provided valuable help with lab work. We also thank B. Frajman, M. Magauer and I. Rešetnik for assistance in the BEAST continuous phylogeography and K-means clustering analyses, and E. Závěská for assistance in SNAPP analysis. J.M. Blanco-Moreno contributed to the morphometric analyses with valuable comments. This research was funded in part by the Spanish government (CGL2010-18631/BOS and the *Flora iberica* project, CGL2011-28613-C03-01) as well as the Catalan government (2014-SGR 514). Pau Carnicero benefited from the support of a PIF PhD student fellowship from the Universitat Autònoma de Barcelona and an Erasmus + traineeship mobility fellowship, which supported a research stay at the Department of Botany of the University of Innsbruck.

Author Contributions

P.C., N.G., L.S. and M.G. conceived the project. P.C., P.F., L.S. and M.G. conducted the field samplings. P.C., M.G. and P.S. designed and conducted the analyses. All authors contributed to the interpretation of results. P.C., P.S. and M.G. drafted the manuscript. All authors revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36412-1>.

Competing Interests: The authors declare no competing interests.

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



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

© The Author(s) 2018