MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 1381–1395 doi: 10.1111/j.1365-294X.2010.04549.x

Ecological rather than geographical isolation dominates Quaternary formation of Mediterranean *Cistus* species

M. FERNÁNDEZ-MAZUECOS and P. VARGAS

Real Jardín Botánico, CSIC, Plaza de Murillo 2, 28014 Madrid, Spain

Abstract

The lack of a comparative approach makes it impossible to determine the main factors influencing colonization and evolution in plants. Here we conducted the first comparative study of a characteristic Mediterranean lineage (white-flowered *Cistus*) taking advantage of its well-known phylogenetic relationships. A two-scale approach was applied to address the hypothesis of higher levels of isolation in mountain than in lowland species. First, a time-calibrated phylogeny using plastid sequences of Cistaceae suggested that the origin of *Cistus* species postdated both the refilling of the Mediterranean Sea (5.59–5.33 Ma) and the onset of the Mediterranean climate (3.2 Ma). Two hundred and sixty-three additional, plastid sequences from 111 populations showed different numbers of haplotypes in *C. laurifolius* (7), *C. monspeliensis* (2) and *C. salviifolius* (7). Although haplotype sharing among disjunct populations was observed in all species, phylogeographic analyses revealed haplotype lineages exclusive to Europe or Africa only in the mountain species (*C. laurifolius*). Isolation by either geographical distance or sea barriers was not significantly supported for the lowland species (*C. monspeliensis*; *C. ladanifer* from a previous study). The same is true for the less habitat-specific species of the lineage (*C. salviifolius*). Comparative phylogeography of the *Cistus* species leads us to interpret a general pattern of active colonization surpassing Mediterranean barriers. In contrast, ecological conditions (precipitation, temperature, soils) appear to have determined the distribution of the *Cistus* species of Mediterranean mountains. This study further provides molecular evidence for multiple colonization patterns in the course of successful adaptation of *Cistus* species to Mediterranean habitats.

Keywords: disjunctions, mountains, phylogeny, phylogeographic patterns, plastid haplotypes, sea barriers

Received 29 September 2009; revision received 11 December 2009; accepted 2 January 2010

Introduction

Since high altitude creates an island-like isolation, mountains provide an excellent geographical system to elucidate historical colonization patterns. Numerous studies have indicated that the distribution patterns of European mountain plants are largely determined by interactions between the species characteristics and the climatic events of the Pleistocene (Comes & Kadereit 1998; Schönswetter *et al.* 2005; Ronikier *et al.* 2008). Indeed, historical processes of isolation have been more easily reconstructed in mountain than in lowland taxa (Zhang *et al.* 2001; Comes & Kadereit 2003; Arafeh & Kadereit 2006). The alternation of glacial and interglacial periods subjected plants to fundamental distribution changes. The disjunct distributions of mountain plants in Central Europe are primarily explained by: (1) survival in local refugia; (2) stepwise colonization from other areas, particularly from southern Europe; and (3) long-distance dispersal (Taberlet *et al.* 1998; Comes & Kadereit 2003; Tribsch & Schönswetter 2003; Schönswetter *et al.* 2005).

Phylogeography is concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species (Avise 2000, 2009). Emerging

Correspondence: Mario Fernández-Mazuecos,
Fax: +34914200157; E-mail: mfmazuecos@rjb.csic.es

comparative phylogeographic studies are helpful for testing non-exclusive colonization hypotheses. Some studies have suggested that species of the same genus do not necessarily display comparable colonization patterns (Després *et al.* 2003; Dixon *et al.* 2007). Similarly, species of different plant genera sharing similar alpire ecological conditions appear not to follow the same patterns (Taberlet 2002). Survival and colonization processes in the high Mediterranean mountains are expected to be complex given the higher biological diversity of the region and the more limited effects of the Last Glacial Maximum mass extinction event in the southern latitudes (Hewitt 2000; Vargas 2003; Kropf *et al.* 2008). It has been proposed that plant genera with similar ecological requirements display more dissimilar colonization patterns in the Mediterranean than in the Central European mountains (Vargas 2003; Kropf *et al.* 2008). The question remains as to whether historical colonization is shared by Mediterranean species of the same genus with different habitat requirements but sharing a recent common ancestor.

The genus *Cistus* L. (Cistaceae) consists of 21 species distributed primarily in the Mediterranean. Most *Cistus* species are widespread with a few narrow endemics, patterns resulting from colonization processes after species formation and post-glaciation population survival (Guzmán *et al.* 2009). Long-distance colonizations have been documented in *Cistus* despite its capsules and dry seeds that display no obvious capability for extensive dispersal. A phylogenetic hypothesis has revealed a parallel colonization pattern of the two major lineages of *Cistus* (white and purple-flowered species) across the Mediterranean (Guzmán & Vargas 2005). The colonization potential of the purple-flowered lineage was further manifested by the presence of a sublineage of five Canary Island endemics resulting from a single introduction. On the other hand, colonization of the white-flowered lineage has been successful as evidenced by the presence of *C. monspeliensis* in most of the Canary Islands. In addition, at least two long-distance colonization events of the white-flowered *C. ladanifer* across the Strait of Gibraltar have been identified by geographical distribution of haplotypes (Guzmán & Vargas 2009b).

Although relatively active dispersal has been inferred for *Cistus* lineages, a comparative phylogeographic study of some *Cistus* species is needed to identify general colonization patterns and the main factors influencing them. Here we used white-flowered species of the same *Cistus* lineage (Guzmán *et al.* 2009) to address explicit hypotheses about geographical (Europe/Africa; eastern/western Mediterranean) and ecological (altitude) colonizations. Four particular white-flowered species are subject to this study. *Cistus laurifolius* presents a fragmented distribution in the Mediterranean mountains, leading us to hypothesize a significant altitude-associated isolation. Less isolation is expected for two lowland species: *C. ladanifer* (previously analysed by Guzmán & Vargas 2009b) and *C. monspeliensis*. Finally, even less isolation is expected for *C. salviifolius*, given its successful distribution in numerous habitats from sea level to timberline. Two comparisons were accordingly used to contrast the phylogeographic patterns of (i) the mountain *C. laurifolius* with the lowland *C. ladanifer* and *C. monspeliensis*; and (ii) three altitude limited species (*C. laurifolius*, *C. ladanifer* and *C. monspeliensis*) with the most widespread species of *Cistus*, *C. salviifolius*.

To explore the hypothesis of a higher degree of genetic isolation in mountains, we applied a two-scale approach based on the analysis of sequences of the plastid genome (cpDNA), which is structurally stable, haploid, non-recombinant, and maternally inherited in *Cistus* (Guzmán & Vargas 2009b). First, a relaxed molecular clock analysis of Cistaceae was performed to estimate divergence times of *Cistus* species and relate them to biogeographic features of the Mediterranean region. Then, combined analyses of species distributions, haplotype polymorphisms, and haplotype clade relationships were carried out in *Cistus laurifolius*, *C. monspeliensis*, and *C. salviifolius* populations (1) to reconstruct the phylogeographic relationships and historical colonization patterns of each species; (2) to contrast the colonization histories of the three species, plus *C. ladanifer*, at the temporal and spatial scales; and (3) to investigate the net contribution of geographical and ecological barriers in the species distributions across the Mediterranean.

Materials and methods

Study species

The four species chosen to perform a comparative phylogeographic analysis are widespread and belong to the white-flowered lineage of *Cistus*, i.e. they share a recent common ancestor (Guzmán & Vargas 2005).

Cistus laurifolius L. is a mountain shrub typically found in the understory and successional scrub of oak and pine woodlands. A dense scrub is formed on acidic substrates of Mediterranean mountains from 800 to 1900 m (supramediterranean vegetation belt) (Martín & Guinea 1949; Demoly & Montserrat 1993). Disjunct distribution of this species in the western and eastern Mediterranean basin, and a relict population in central Italy, has been interpreted as the result of the contraction of a wider pre-glacial area (Dansereau 1939; Pignatti 1982).

Cistus ladanifer L. is a shrub occurring on acidic soils of the western Mediterranean from sea-level to 1000 m (mesomediterranean vegetation belt) (see Guzman & Vargas 2009 for details of this species).

Cistus monspeliensis L. is a lowland shrub displaying a rather continuous distribution in the Mediterranean basin, even though it becomes scarcer eastwards. It also occurs in the Canary Islands. Dense *C. monspeliensis* formations are found on poor soils from sea-level to 600–800 m (thermomediterranean vegetation belt) on both calcareous and acidic substrates. This species forms dense scrub when holm oak, cork oak and pine woodlands are degraded (Juhren 1966).

Cistus salviifolius L. is a subshrub occurring on sandy soils of a wide range of habitats (Demoly & Montserrat 1993). Unlike the two preceding species, this species does not form dense scrub, and has a patchy distribution in a wide altitudinal range from sea level to 1800 m. *Cistus salviifolius* has a circum-Mediterranean distribution, from Portugal and Morocco to Palestine and the eastern coast of the Black Sea and also occurs in the south of the Eurosiberian floristic region. It has also been recorded in Macaronesia (Madeira), where it is probably an introduced species (Short 1994).

Artificial crossings between four species (*C. parviflorus*, *C. laurifolius*, *C. libanotis*, *C. ladanifer*) of the white-flowered lineage revealed maternal inheritance of plastid haplotypes (Guzmán & Vargas 2009b). Accordingly, we interpret that phylogeographic reconstruction of plastid haplotypes reflects plant colonization by seeds.

Sampling strategy and DNA sequencing

In order to estimate divergence times among *Cistus* lineages and plastid sequences of the study species, we obtained GenBank sequences of the *trn*L–*trn*F spacer (501 bp) and the *mat*K gene (5′ end of 486 bp) from 20 *Cistus* species and 10 additional Cistaceae (*Fumana*, *Helianthemum*, *Tuberaria*, *Halimium*) (Guzmán *et al.* 2009; Guzmán & Vargas 2009a). We employed sequences of *Hopea wightiana* (Dipterocarpaceae) as outgroup (Gamage *et al.* 2003, 2006).

For the inference of intraspecific patterns of cpDNA variation, a total of 111 populations of *C. laurifolius*, *C. monspeliensis* and *C. salviifolius* were sampled to cover their entire distribution in the Mediterranean basin (Table S1, Supporting Information). Forty-seven populations of *C. ladanifer* sampled and sequenced by Guzmán & Vargas (2009b) were also included in the comparative analyses (see below). Thirty-three individuals of *C. laurifolius*, 26 of *C. monspeliensis* and 52 of *C. salviifolius* were collected in the field and dried in silica gel, except for some individuals obtained from herbarium collections (see Table S1, Supporting Information). First, a pilot study using three to six samples of each species from distant geographical areas was performed to find the most variable sequences among 17 plastid DNA regions previously used in phylogenetic and phylogeographic analyses (Table S2, Supporting Information).

Procedures used for amplification and sequencing of DNA regions followed Guzmán & Vargas (2005, 2009b), except for minor details. Total genomic DNA was extracted using DNeasy Plant Mini Kit (QIAGEN Inc., California). DNA regions were amplified in an Eppendorf Mastercycler Epgradient S (Westbury, NY) or a MJ Research PTC-200 (Massachusetts) thermal cycler, with standard primers (Table S2, Supporting Information). After 1–2 min pretreatment at 94–95 °C, PCR conditions were: 30–40 cycles of 45 s—1 min at 94–96 °C, 1–2 min at 48–59 °C (Table S2, Supporting Information) and 1–2 min at 72 °C. In certain reactions, a volume of 1 μL of bovine serum albumine (BSA) at 1 mg mL^{-1} was included in each 25 mL reaction to improve the efficiency of the amplification. Amplified products were cleaned using spin filter columns (UltraClean PCR Clean-Up Kit, MO BIO Laboratories Inc., CA, USA) or treated with ExoSAP-IT (USB Corporation, OH, USA) following the manufacturers' protocols. Sequencing of cleaned products followed Guzmán & Vargas (2005), and resulting sequence data were assembled and edited using the SeqEd software (Applied Biosystems, CA, USA).

After identifying DNA regions with sequence variation (*trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC} and *trn*H^{GUG}-*trn*K^{UUU} for *C. laurifolius*; *trn*S-*trn*G and *psb*K-*trn*S for *C. monspeliensis*; and *trn*S-*trn*G and *trn*H^{GUG}-*trn*K^{UUU} for *C. salviifolius*), we extended the sequencing to every population (one individual per population). Sequences from two additional individuals were obtained from certain populations (population 38 and population 50 of *C. salviifolius*) to check for potential artefacts or haplotype diversity in particular geographical areas. One internal primer (5′-CCCATGTCAACCAATAACCA-3′) was designed to facilitate the amplification of a short region in the *trn*H^{GUG}-*trn*K^{UUU} spacer because of fragment length. It was in this short region where sequence variation was found for two species (*C. laurifolius*, *C. salviifolius*). The same DNA regions were also sequenced from the other nine species of the white-flowered lineage (Guzmán & Vargas 2005), plus two purple-flowered species as the outgroup, to reconstruct phylogenetic relationships of the cpDNA haplotypes (see below).

Estimation of divergence times

All *trn*L-*trn*F and *mat*K sequences were aligned using ClustalX 2.0.10 (Larkin *et al.* 2007), and further adjustments were made by visual inspection. To determine the simplest model of sequence evolution that best fits the sequence data (GTR+G for both data sets), the AIC criterion was implemented in each data set using jModeltest 0.1.1 (Posada 2008; Guindon & Gascuel

2003). Consequently both matrices were combined for further analyses.

The combined matrix was analysed using a relaxed Bayesian approach as implemented in BEAST v.1.4.8 (Drummond & Rambaut 2007; Drummond *et al.* 2006). This software simultaneously estimates tree topology and node ages, and uncertainty is incorporated to both aspects of the analysis. A relaxed molecular clock was used, and the substitution rate variation was modelled using an uncorrelated lognormal distribution. A Birth-Death speciation process (Gernhard 2008) was employed as tree prior, assuming constant rates of speciation and extinction per lineage. Two MCMC analyses were run for 10 million generations, sampling every 1000th generation. Analysis with Tracer 1.4 (Rambaut & Drummond 2007) confirmed convergence among chains and adequate sample size. Both chains were combined using LogCombiner 1.4.8 after discarding the first 10% of sampled generations as burn-in, and trees were summarized in a maximum clade credibility tree obtained in TreeAnotator 1.4.8. This tree was finally visualized using FigTree 1.1.2.

Topological constraints were imposed in those nodes where a particular resolution was needed for subsequent calibrations. Thus, constriction of three monophyletic groups (Cistaceae; *Helianthemum-Tuberaria-Halimium-Cistus*; and *Tuberaria-Halimium-Cistus*) was based on previous phylogenetic results (Guzmán & Vargas 2009a). For temporal calibration of the tree, we employed fossil data and molecular estimates. The divergence age between Dipterocarpaceae and Cistaceae obtained by Wikström *et al.* (2001) and the macrofossil *Cistinocarpum roemeri* Conw (a Cistaceae reproductive structure from the German Middle Oligocene; Palibin 1909) were used to constrain the basal node at a minimum age of 28 Ma and a maximum age of 39 Ma. *Helianthemum* pollen found in Upper Miocene formations in France (Naud & Suc 1975) and *Tuberaria* pollen from German Pliocene formations (Menke 1976) led to establishing minimum ages of 11 Ma and 5.3 Ma respectively for the divergence of these genera. All calibrations were implemented using uniform priors.

Haplotype data analyses

For the intraspecific analyses, all sequences obtained from each species were combined in a matrix and manually aligned given the low number of gaps. Sequences of the species most closely related to *C. laurifolius*, *C. monspeliensis* and *C. salviifolius* according to previous results (Guzmán & Vargas 2005; Guzmán *et al.* 2009) were included as the outgroup.

Genealogical relationships among haplotypes were inferred using two different coalescence algorithms for construction of haplotype networks: statistical parsimony and reduced median network (Posada & Crandall 2001). The statistical parsimony algorithm (Templeton *et al.* 1992), implemented in TCS 1.21 (Clement *et al.* 2000), was used to infer relationships based on nucleotide substitutions. A haplotype analysis was conducted for each species plus one or more outgroup samples depending on previously described phylogenetic relationships (Guzmán *et al.* 2009). The maximum number of differences resulting from single substitutions among haplotypes was calculated with 95% confidence limits, treating gaps as missing data. The reduced median network (RMN) method, implemented in Network 4.5.0.2 (http://www.fluxus-engineering.com), allowed for the analysis of both nucleotide substitutions and indels. However, indels are generally considered to be less reliable characters than substitutions, and homology is especially uncertain in poly-A/T microsatellites (Kelchner 2000). Therefore, each type of sequence mutation was given a different weight in the analysis: 10 for substitutions, 5 for several-nucleotide indels and 1 for differences in number of repeats within microsatellites (see Besnard *et al.* 2007; Naciri & Gaudeul 2007). The analysis was rerun using different weighting (4/2/1 and 3/2/1) to assess potential effects in the resulting haplotype network. As required by the reduced median network algorithm, all characters were coded as binary data. Several-nucleotide indels, most likely resulting from a single event, were coded following the 'simple indel coding' approach (Simmons & Ochoterena 2000), while each mononucleotide repeat in microsatellites was coded independently as a character, following Naciri & Gaudeul (2007). The reduction threshold was set to its default value ($r = 2$) in all the analyses. A maximum parsimony (MP) calculation was finally performed to eliminate nonparsimonious links and median vectors (putative haplotypes) from the resulting network, as recommended in the Network 4.5.0.2 manual.

Phylogenetic relationships were assessed using Maximum Parsimony (MP) and Bayesian Inference (BI). Analyses were conducted combining sequences of the four DNA regions (*trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC}, *psb*K-*trn*S and a fragment of *trn*H^{GUG}-*trn*K^{UUU}) representing substitution-based haplotypes of the three study species plus representatives of the other nine white-flowered and two purple-flowered *Cistus* species. MP analyses were performed using PAUP* 4.0b10 (Swofford 2002), with the following parameters for the heuristic search: 1000 random addition replicates holding 100 trees per replicate, tree-bisection-reconnection (TBR) branch swapping and the options Multrees and Steepest Descent in effect. Robustness of clades was estimated using 1 000 000 bootstrap replicates (fast stepwise-addition, Mort *et al.* 2000). BI was implemented using two identi-

cal searches with 10 million generations each (chain temperature = 0.2; sample frequency = 100). The simplest model of sequence evolution that best fits the sequence data (GTR+G) was determined under the Akaike information criterion (AIC) in jModeltest 0.1.1. Probabilities converged on the same stable value after *c.* 20 000 generations in both runs. A 50% majority rule consensus tree was calculated to obtain the final Bayesian estimate of phylogeny (see Guzmán & Vargas 2005 for details).

Genetic differentiation analysis

The nearest-neighbour statistic (S_{nn}) was calculated to assess genetic differentiation in *Cistus* due to isolation by sea barriers and geographical distance. This statistic is a measure of how often the 'nearest neighbours' (in sequence space) of sequences are from the same locality in geographic space (Hudson 2000). S_{nn} is expected to approach one when two partitions (localities) of a dataset form highly differentiated populations, and one-half when they are part of a single panmictic population. Thus, higher values of S_{nn} associated to biogeographic barriers would be expected in mountain than in lowland and wide-range species according to our hypothesis. To detect genetic differentiation associated to sea barriers, we used the plastid datasets of *C. laurifolius* and *C. salviifolius*. In addition, haplotypes of *C. ladanifer* obtained by Guzmán & Vargas (2009b) were equally analysed. The *C. monspeliensis* dataset was not analysed due to lack of nucleotide variation (see below). Each of the three datasets was divided into two partitions (European-Turkish vs. North African samples) separated by the Mediterranean Sea. Island populations were excluded.

To assess genetic differentiation associated with isolation by geographical distance, the *C. laurifolius* and *C. salviifolius* datasets were partitioned into Iberian vs. Balkan-Anatolian samples, separated by no less than 1,600 km. Populations from other areas were excluded from the analysis. In all cases, S_{nn} was calculated using DnaSP v5 (Librado & Rozas 2009) in datasets both with and without indel coding. Permutation tests with 1000 replicates were performed to evaluate significance of the obtained values.

Bayesian hypothesis testing

Bayes factors (BF) allow testing of alternative hypotheses in a Bayesian framework (Kass & Raftery 1995). They quantify the support for one hypothesis versus another given the data. Using this approach, implemented in BEAST 1.4.8 (Drummond & Rambaut 2007; Drummond *et al.* 2006) and Tracer 1.4 (Rambaut & Drummond 2007), we tested alternative phylogeographic hypotheses using the sequence datasets of *C. laurifolius*, *C. ladanifer* and *C. salviifolius*.

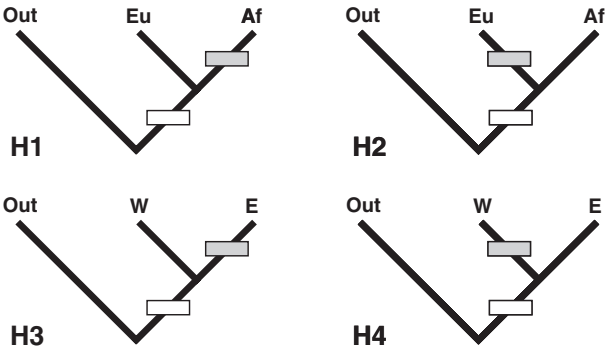

Fig. 1 Topological constraints of phylogeographic hypotheses tested using Bayes factors. H1: single colonization of Africa from Europe; H2: single colonization of Europe from Africa; H3: single eastward colonization; H4: single westward colonization. Grey bars indicate monophyly constraints associated to single colonization events. White bars indicate monophyly of the populations of each species. Out, outgroup; Eu, Europe-Turkey; Af, Africa; W, western Mediterranean; E, eastern Mediterranean.

In each dataset, several hypotheses (tree topologies, Fig. 1) were tested against an unconstrained analysis, all of them involving single colonization events across a sea barrier or a long geographical distance. Thus, different monophyly constraints were set in each BEAST analysis under the assumption that a single colonization event of an area involves monophyly of its extant individuals. Monophyly of the populations of each species was also constrained based on phylogenetic results (see below). The two hypotheses involving single colonization events across the Mediterranean sea barrier were contrasted from Europe to Africa (H1) and vice versa (H2). Single colonization events between the western and eastern Mediterranean are analysed in hypotheses H3 (eastward colonization) and H4 (westward colonization). Under our main hypothesis, single colonization events across barriers would be more easily rejected in lowland or wide-range species than in the mountain species.

Markov Chain Monte Carlo (MCMC) analyses were run for 30 million generations, sampling every 1000th generation. The outgroup (sister species) was included in all the analyses, and a Yule tree prior was implemented. In all cases, the root node was constrained to a maximum age defined by the calculated divergence times between the study species and its sister taxon (see below). Substitution models were determined using jModeltest 0.1.1 under the AIC criterion. We used a strict molecular clock in order to minimize the number of parameters to be estimated following Dixon *et al.*

(2009). Stationarity was determined using Tracer. A second run was performed to confirm convergence among chains, and both runs were combined for hypothesis testing after discarding the first 10% of sampled generations as burn-in.

Marginal likelihoods, their standard error (estimated using 1000 bootstrap replicates) and BFs were calculated in Tracer to measure the support for each hypothesis (Newton & Raftery 1994; Suchard *et al*. 2001). As proposed by Kass & Raftery (1995), we considered $2 \times \ln BF(H \text{ vs. } H') > 10$ as decisive support for hypothesis H vs. hypothesis H'. Likewise, $2 \times \ln BF(H \text{ vs. } H') < -10$ was considered as decisive evidence against hypothesis H vs. hypothesis H'.

Results

Divergence times

Values of standard deviation of the uncorrelated log-normal relaxed clock (0.693) and coefficient of variation (0.754) for rate heterogeneity within our *trn*L-*trn*F and *mat*K data indicate that a relaxed molecular clock is adequate for the analysis. According to the parameter analysis in Tracer, number of MCMC iterations was sufficient, with values of effective sample size (ESS) above 500 in all cases and plots showing equilibrium after discarding burn-in. The maximum clade credibility tree with 95% highest posterior density intervals (HPD) for the divergence time estimates of relevant nodes (also listed in Table 1) is shown in Fig. 2. The chronogram is congruent with the divergence of the purple- and white-flowered *Cistus* lineages around the Pliocene (1.69–7.52 Ma). Pleistocene divergence of *C. laurifolius*, *C. monspeliensis* and *C. salviifolius* lineages is clearly inferred, with speciation times postdating the Messinian Salinity Crisis (5.59–5.33 Ma; Krijgsman *et al*. 1999): 0.08–1.68 Ma for the split of *C. salviifolius* and *C. ladanifer*; 0.22–2.18 Ma for that of *C. laurifolius*, *C. parviflorus* and *C. psilosepalus*; and 0.24–2.41 Ma for that of *C. monspeliensis* and *C. populifolius*.

Table 1 Mean posterior estimated ages and 95% highest posterior density (HPD) intervals for the constrained (A–C) and unconstrained relevant nodes (E–G) based on relaxed molecular-clock analysis of Cistaceae *trn*L-*trn*F and *mat*K sequences in BEAST. Nodes are named as in Fig. 2

Node	Estimated age (Ma)	95% HPD interval
A	32.79	28.00–38.27
B	25.80	11.02–25.97
C	9.55	5.30–14.80
D	4.25	1.69–7.52
E	0.73	0.08–1.68
F	1.07	0.22–2.18
G	1.13	0.24–2.41

Haplotype analysis of Cistus laurifolius

Length of the DNA regions after sequencing the 33 populations was 653 bp for *trn*S-*trn*G, 695 bp for $trnC^{GCA}$-$trnD^{GUC}$ and 85 bp for the amplified stretch of the $trnH^{GUG}$-$trnK^{UUU}$ spacer. The other 14 plastid regions sequenced for the pilot study yielded no variation or we failed to obtain PCR amplifications. Molecular variation was distributed as shown in Table 2.

Statistical parsimony analysis of the combined matrix allowed recognition of four substitution-based haplotypes, distributed in Morocco (two haplotypes), Spain (two haplotypes), Southern France (one haplotype), central Italy (one haplotype) and Turkey (one haplotype). Haplotypes A and B are exclusive to African populations, whereas C and D to European (plus Turkish) populations. No haplotype is shared between the two sides of the Strait of Gibraltar. In contrast, haplotype D is widely distributed from northern Iberia to the eastern Mediterranean. The network constructed by TCS (Fig. 3b) shows three *C. laurifolius* haplotype clades connected to an ancestral, missing haplotype (extinct or not found). This haplotype is connected to the sister species *C. psilosepalus* through five additional mutation steps. One haplotype clade corresponds to European-Turkish populations (subsp. *laurifolius*), whereas the other two clades to African populations (subsp. *atlanticus*). Two different haplotypes coexist in the Moroccan Rif (A and B) and the Spanish Sistema Central (C and D). When microsatellites and other indels were considered in the analysis, we obtained seven haplotypes (Table 2): four in Africa and three in European and Turkish populations (Fig. 3a). The new haplotypes are closely related to the substitution-based ones (A_1, B_1, C_1, D), and are found in the Middle Atlas (A_2 and B_2) and the SE Iberia (C_2). In fact, the RMN algorithm constructed a network (Fig. 3c) showing a topology congruent with that based on nucleotide substitutions. The same network was obtained with different character weighting. The central loop is considered to be related to homoplasy affecting haplotypes A_1, A_2 and C_2 based on one microsatellite of the *trn*S-*trn*G spacer.

Haplotype analysis of Cistus monspeliensis

Sixteen of the 17 plastid regions sequenced for the pilot study showed no variation across the Mediterranean. The choice of the *trn*S-*trn*G spacer (624 bp) for the 26 samples was based on the finding of a 4-bp indel, while the sequencing of *psb*K-*trn*S (431 bp) was extended to

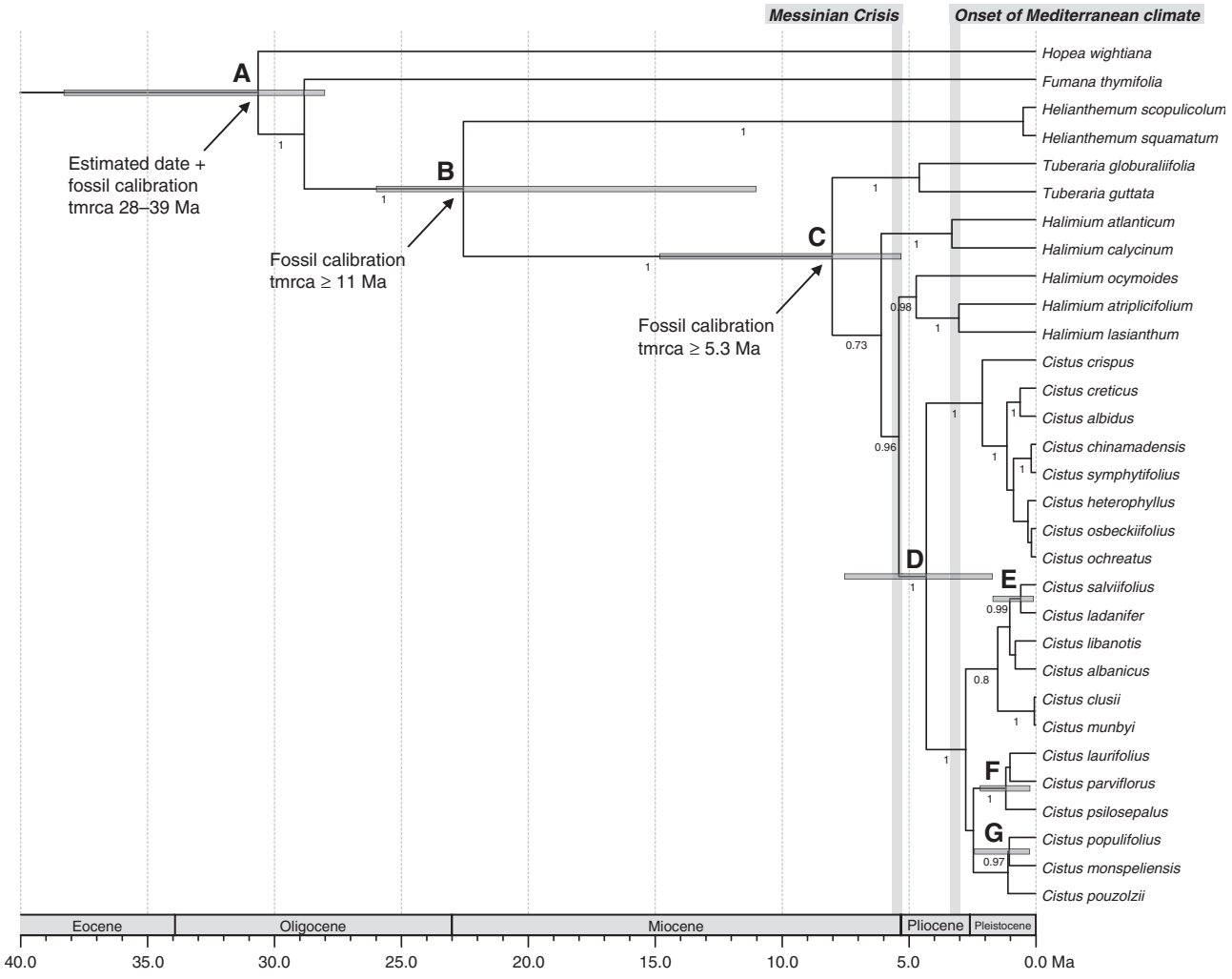

Fig. 2 Maximum clade credibility tree produced by relaxed molecular-clock analysis of Cistaceae *trn*L-*trn*F and *mat*K sequences in BEAST. Node bars represent the 95% highest posterior density intervals for the divergence time estimates. Values below branches indicate Bayesian posterior probabilities. Dates for molecular and fossil calibrations, as well as the Messinian Salinity Crisis (ca. 5.59–5.33 Ma) and the onset of the Mediterranean climate (c. 3.2 Ma) are indicated. Tmrca, Time to most recent common ancestor. A–G, relevant nodes listed in Table 1.

every sample because of the high variation found in the Canary Islands (M. Fernández-Mazuecos & P. Vargas, unpublished results). The combined analysis of the two regions with no indel coding yielded a single haplotype (A) distributed across the Mediterranean basin, from the Iberian Peninsula and Maghreb to Greece and Cyprus. In the TCS analysis, this haplotype was connected to the outgroup through the *C. parviflorus* haplotype by nine steps (result not shown). When indels were used in the RMN analysis (Table 2), the rare haplotype A_1 found in a single Iberian population (Cornalvo reservoir, Badajoz, Spain) was shown to be ancestral (Fig. 4b), while A_2 remained widespread in the Mediterranean basin (Fig. 4a). The same network was obtained with different character weighting.

Haplotype analysis of Cistus salviifolius

Length of the DNA region after sequencing the 52 populations was 550 bp for the *trn*S-*trn*G spacer and 103 bp for the amplified stretch of the $trnH^{GUG}$-$trnK^{UUU}$ spacer. The other 15 plastid regions sequenced for the pilot study yielded no variation, except for a nucleotide substitution found in a single Turkish population (50) of the $trnK^{UUU}$-$trnK^{UUU}$ spacer.

We found only two substitution-based haplotypes (A, B), which are widely distributed and separated by a single substitution. Haplotype A was found in most (47) of the populations across the Mediterranean. Populations of southern Spain, eastern Morocco, Tunisia, Sicily, Italian Peninsula and Turkey (Bursa population)

Table 2 Haplotypes found in 33 *C. laurifolius*, 26 *C. monspeliensis* and 52 *C. salviifolius* populations. Variable sites of the sequenced plastid regions (*trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC}, *psb*K-*trn*S and a fragment of *trn*H^{GUG}-*trn*K^{UUU}) are shown. Mononucleotide repeats (microsatellites) are represented in brackets. Several-nucleotide indels are shown as presence/absence (1/0) data in square brackets. An 'N' in square brackets represents an inapplicable character following the simple indel coding metodology (Simmons & Ochoterena 2000)

		DNA region									
		*trn*S-*trn*G (553 pb)			*trn*C^{GCA}-*trn*D^{GUC} (695 pb)		*trn*H^{GUG}-*trn*K^{UUU} (85 pb)				
Haplotype	Number of populations	58	(130)	394	22	218	[29–34][1]	[31–34][2]	[33–34][3]	44	70
Cistus laurifolius											
A_1	2	A	(T)	G	T	C	[1]	[1]	[0]	A	(T)
A_2	1	A	(T)	G	T	C	[0]	[N]	[N]	A	(T)
B_1	2	C	(-)	G	G	C	[1]	[1]	[1]	G	(T)
B_2	1	C	(-)	G	G	C	[1]	[0]	[N]	G	(T)
C_1	7	C	(-)	G	G	A	[1]	[1]	[1]	A	(T)
C_2	3	C	(T)	G	G	A	[1]	[1]	[1]	A	(T)
D	17	C	(-)	T	G	A	[1]	[1]	[1]	A	(-)

		*trn*S-*trn*G (624 pb)
		[451–454] [4]
Cistus monspeliensis		
A_1	1	[0]
A_2	25	[1]

		*trn*S-*trn*G (549 pb)			*trn*H^{GUG}-*trn*K^{UUU} (103 pb)
		(13–14)	(39)	69	[27–31][5]
Cistus salviifolius					
A_1	3	(T-)	(-)	T	[1]
A_2	22	(–)	(-)	T	[1]
A_3	21	(T-)	(-)	T	[0]
B_1	3	(T-)	(-)	G	[1]
B_2	1	(–)	(-)	G	[1]
B_3	1	(TT)	(T)	G	[1]
B_4	1	(T-)	(-)	G	[0]

[1]ATATAT; [2]ATAT; [3]AT; [4]AAAT; [5]TATAT.

harboured haplotype B. Given the lack of geographical structure of haplotypes and the presence of more than one haplotype in the same area, three individuals of the Turkish and Italian populations were sampled. However, all of them shared the same haplotype (B). The TCS network connected the sister species *C. ladanifer* to haplotype A through five mutational steps (Fig. 5b). When coding indels, we found seven haplotypes (Table 2), two of which (A_2, A_3) are widely distributed in western and eastern Mediterranean populations: A_2 in Morocco, western Iberia, Balkan Peninsula, Crete and Cyprus and A_3 in Iberia, South France, Corsica, Sardinia, Italian Peninsula, Balkan Peninsula and Turkey (Fig. 5a). Haplotype B_1 is also widely distributed (three populations of Morocco, Tunisia and Sicily). Haplotype A_1 was found nearby the Strait of Gibraltar (Andalusia and northern Morocco), while B_2 and B_3 in single populations (Andalusia and Turkey). The RMN analysis built a network with no loops using character weighting of 10/5/1 and 4/2/1 and showed the *C. salviifolius* clade connected to *C. ladanifer* through haplotype A_1 by nine mutational steps (Fig. 5c). The two most widely distributed haplotypes (A_2 and A_3) are connected to A_1 by a single mutation step. A group formed by the B subhaplotypes (B_1, B_2, B_3, B_4) connected populations distributed across the Mediterranean. Using a lower substitution weight (3/2/1), an additional connection between haplotypes B_4 and A_3 was found

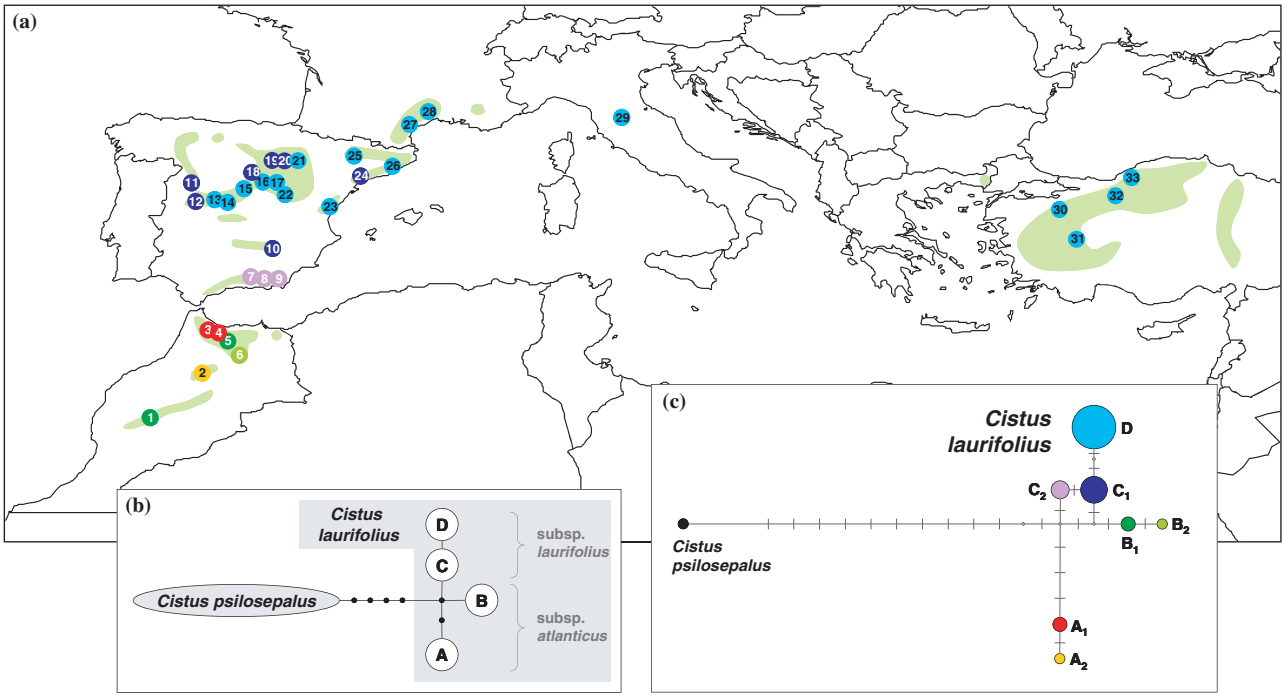

Fig. 3 (a) *Cistus laurifolius* distribution map, sampled populations (numbers), and geographical range of the seven cpDNA haplotypes (colours) obtained from *trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC} and *trn*H^{GUG}-*trn*K^{UUU} sequences as a result of coding nucleotide substitutions and indels. (b) Statistical parsimony network of plastid haplotypes (indicated by letters) based on nucleotide substitutions in *C. laurifolius* and the sister species *C. psilosepalus*. Lines represent a single nucleotide substitution, and dots indicate absent haplotypes (extinct or not found). (c) Reduced median network of *C. laurifolius* haplotypes obtained after coding indels. Each haplotype is shown as a coloured circle with a size proportional to the number of populations. Colours are the same as in the map; median vectors are shown as small dots; and mutational steps are indicated as small lines perpendicular to links between haplotypes.

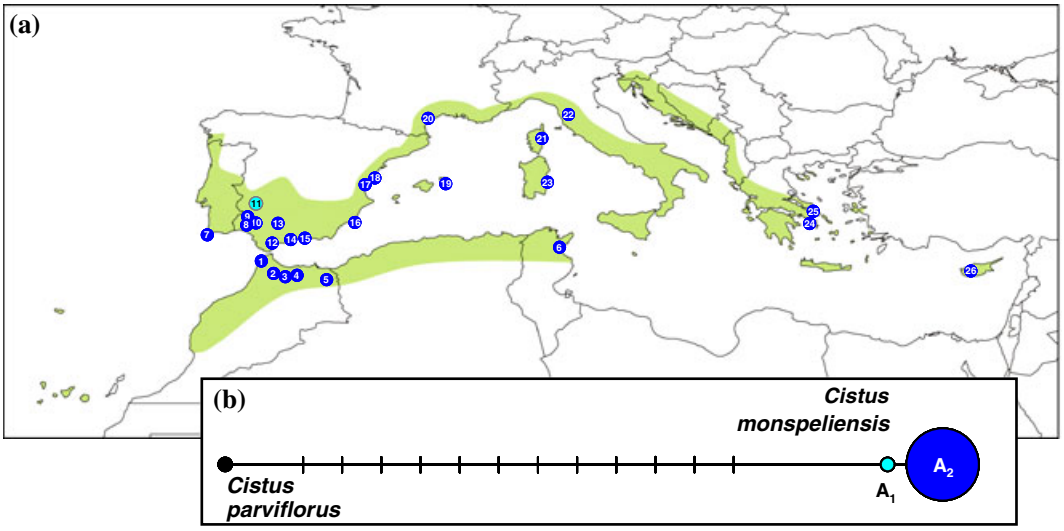

Fig. 4 (a) *Cistus monspeliensis* distribution map, sampled populations (numbers), and geographical range of the two cpDNA haplotypes obtained from *trn*S-*trn*G sequences as a result of coding nucleotide substitutions and indels. (b) Reduced median network of the two *C. monspeliensis* haplotypes. Colours are the same as in the map; and mutational steps are indicated as small lines perpendicular to links between haplotypes.

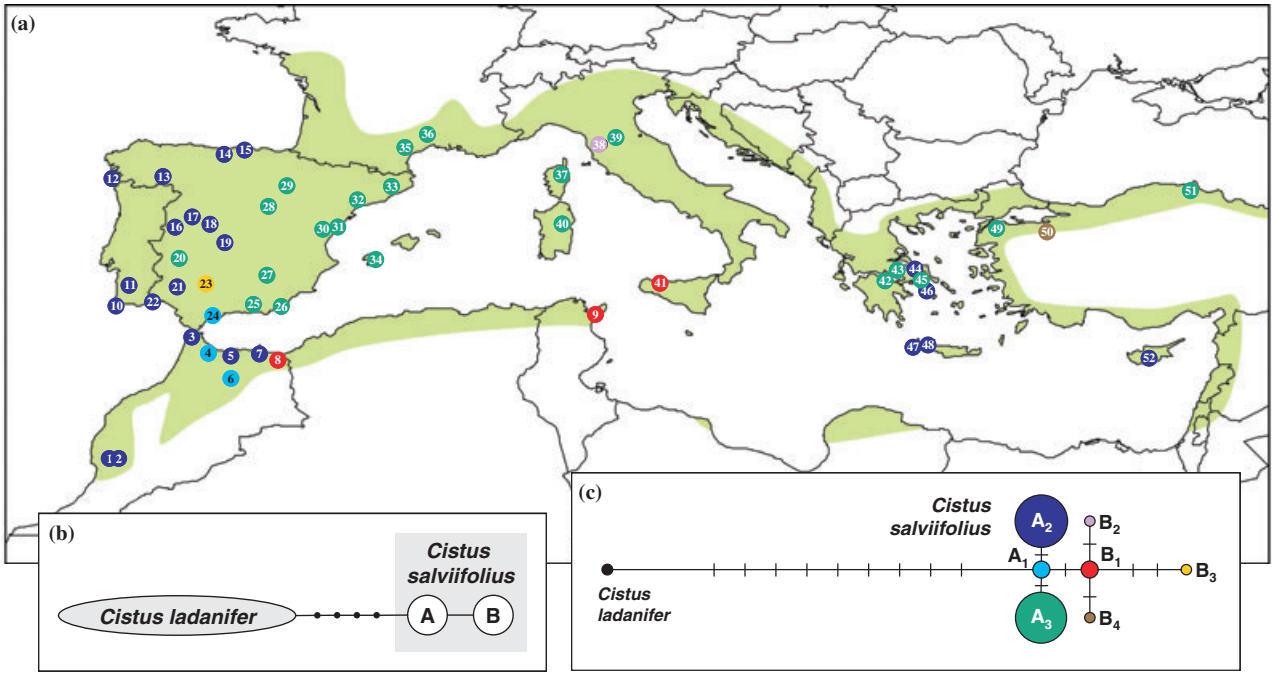

Fig. 5 (a) *Cistus salviifolius* distribution map, sampled populations (numbers), and geographical range of the seven cpDNA haplotypes (colours) obtained from *trn*S-*trn*G and *trn*H^{GUG}-*trn*K^{UUU} sequences as a result of coding nucleotide substitutions and indels. (b) Statistical parsimony network of plastid haplotypes (indicated by letters) based on nucleotide substitutions in *C. salviifolius* and the sister species *C. ladanifer*. Lines represent a single nucleotide substitution, and dots indicate absent haplotypes (extinct or not found). (c) Reduced median network of *C. salviifolius* haplotypes obtained after coding indels. Each haplotype is shown as a coloured circle with a size proportional to the number of populations. Colours are the same as in the map; and mutational steps are indicated as small lines perpendicular to links between haplotypes.

because they share a 5-nucleotide deletion in the *trn*H^{GUG}-*trn*K^{UUU} spacer.

Phylogenetic relationships of haplotypes

Combination of *trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC}, *trn*H^{GUG}-*trn*K^{UUU} and *psb*K-*trn*S sequences of 14 *Cistus* species (including those of the seven substitution-based haplotypes of the three study species) resulted in an aligned length of 1990 bp. Seventy-three of the 118 variable sites from the matrix were phylogenetically informative. MP analysis generated 198 trees of 140 steps with a consistency index (CI) of 0.88 and a retention index (RI) of 0.85. The strict consensus tree (Fig. 6) recognizes *C. laurifolius* (BS 91%) and *C. salviifolius* (BS 100%) populations as monophyletic. The BI phylogeny depicted a congruent topology, with posterior probabilities (PP) of 1 for the haplotype sequences of these two species. Phylogenetic relationships among haplotypes of the study species are congruent with those retrieved in network analyses. African (A, B) and European-Turkish (C, D) haplotypes of *C. laurifolius* form sister groups (BS 63%; 0.98 and 0.97 PP).

Genetic differentiation analysis

Values of S_{nn} (Table 3) show a strong and significant genetic differentiation (1.00) across the Mediterranean Sea barrier for *C. laurifolius*. A lower but still highly significant differentiation was found in *C. ladanifer* (0.85, 0.86), whereas no highly significant differentiation was retrieved for *C. salviifolius*. When assessing isolation by geographical distance in *C. laurifolius* and *C. salviifolius*, no significant genetic differentiation was found in either of the datasets between Iberian and Balkan-Anatolian samples.

Bayesian hypothesis testing

Marginal likelihood and $2 \times \ln BF$ values for the tested phylogeographic hypotheses are shown in Table 4. All hypotheses involving single colonization events across the Mediterranean Sea (H1, H2) and geographical distance (H3, H4) barriers are strongly rejected for both *C. ladanifer* and *C. salviifolius* ($2 \times \ln BF < -20$ in all cases). In *C. laurifolius*, only a single westward colonization (H4) is strongly rejected ($2 \times \ln BF < -30$),

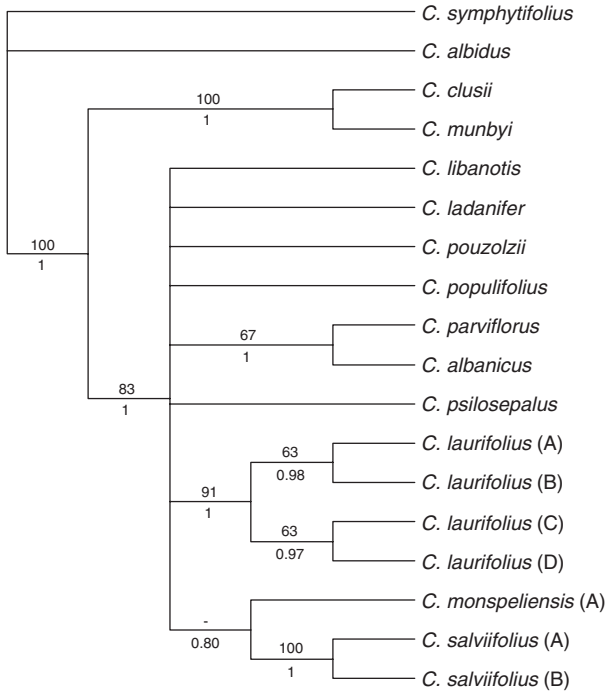

Fig. 6 Strict consensus tree of the 198 shortest trees of 140 steps (CI = 0.88; RI = 0.85) from the combined analysis of *trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC}, *trn*H^{GUG}-*trn*K^{UUU}, and *psb*K-*trn*S sequences. Numbers above branches are bootstrap values. Numbers below branches are Bayesian posterior probabilities (PP). A hyphen (–) indicates no bootstrap support over 50%.

which supports a single colonization of Turkey from the western Mediterranean and a single colonization across the Strait of Gibraltar. Direction of the latter event is not clearly revealed by our data.

Discussion

Pleistocene origin of Cistus species

A recent origin of the four *Cistus* species is suggested by the haplotype identity (*C. monspeliensis*) and the low nucleotide variation (*C. laurifolius*, *C. salviifolius*, *C. ladanifer*) of highly variable DNA regions (Shaw *et al*. 2005, 2007). Phylogenetic analyses and molecular dating estimates were consistent with those of previous studies obtained using different methods (Guzmán & Vargas 2009b; Guzmán *et al*. 2009): the diversification of the white-flowered *Cistus* lineages post-dated the refilling of the Mediterranean basin (5.59–5.33 Ma) and the onset of the Mediterranean climate (3.2 Ma) (Fig. 2). Although interspecific hybridization in *Cistus* has been widely described (Dansereau 1939; Demoly & Montserrat 1993), the monophyly or identity of plastid (Fig. 6) and nuclear (Guzmán *et al*. 2009) sequences of white-flowered species leads us to rule out reticulation as a factor significantly influencing the overall patterns of DNA variation. Thus, the common ancestry assumption needed in phylogeography was fulfilled by the populations of all three newly analysed species, as also reported for *C. ladanifer* (Guzmán & Vargas 2009b). Furthermore, the shared phylogenetic, temporal and geographical framework provides an appropriate model for comparative analyses of species with contrasting ecological requirements.

Isolation by sea barriers and geographical distance

Different levels of isolation, due to geographical distance or sea barriers, were observed for *Cistus*. Neither haplotype nor haplotype lineage sharing were observed

Table 3 Nearest-neighbour statistic (S_{nn}) values calculated partitioning the *C. laurifolius*, *C. ladanifer* and *C. salviifolius* datasets in order to evaluate genetic differentiation associated to isolation by sea barriers and geographical distance. Number of colonizations across the same barriers inferred from S_{nn} values, Bayesian hypothesis testing (Table 4) and network analyses (Fig. 3–5) are also included for comparison

	S_{nn}		
	Only substitutions	Substitutions + indels	No. of colonizations
Isolation by sea barriers (European-Turkish vs. North African populations)			
C. laurifolius	1.00000 (P = 0.0000***)	1.00000 (P = 0.0000***)	1
C. ladanifer	0.85051 (P = 0.0000***)	0.86411 (P = 0.0000***)	2
C. salviifolius	0.65745 (P = 0.5410 ns)	0.73282 (P = 0.0120*)	3–4
Isolation by geographical distance (Iberian vs. Balkan-Anatolian populations)			
C. laurifolius	0.73913 (P = 0.0540 ns)	0.73913 (P = 0.1850 ns)	1
C. salviifolius	0,60741 (P = 1.0000 ns)	0.69597 (P = 0.7410 ns)	3

ns, not significant; *, 0.01 < P < 0.05; **, 0.001 < P< 0.01; ***, P < 0.001.

Table 4 Marginal likelihoods (with standard error) and test statistic 2 × lnBF for the unconstrained analyses (H0) and the tested phylogeographic hypotheses (H1–H4) in *C. laurifolius*, *C. ladanifer* and *C. salviifolius*. H1: single colonization of Africa from Europe; H2: single colonization of Europe from Africa; H3: single eastward colonization; H4: single westward colonization (see Fig. 1). The compared hypotheses (*H* and *H*′) are arranged in rows and columns, respectively. 2 × lnBF*H* vs. *H*′ > 10 is considered as decisive support for *H*, while 2 × lnBF*H* vs. *H*′ < −10 is considered as decisive support for *H*′ (Kass & Raftery 1995). NA = non-applicable

		2xlnBF				
		Hypothesis (*H*′)				
Hypothesis (*H*)	Marginal likelihood ± SE	H0	H1	H2	H3	H4
Cistus laurifolius						
H0	−1925.017 ± 0.047	—	0.328	−0.652	−0.378	36.58
H1	−1925.181 ± 0.046	−0.328	—	−0.98	−0.706	36.252
H2	−1924.691 ± 0.04	0.652	0.98	—	0.276	37.234
H3	−1924.828 ± 0.047	0.378	0.706	−0.276	—	36.958
H4	−1943.307 ± 0.047	−36.58	−36.252	−37.234	−36.958	—
Cistus ladanifer						
H0	−2873.25 ± 0.075	—	65.81	64.372	NA	NA
H1	−2906.15 ± 0.085	−65.81	—	−1.44	NA	NA
H2	−2905.43 ± 0.084	−64.372	1.44	—	NA	NA
Cistus salviifolius						
H0	−828.352 ± 0.052	—	22.116	22.056	22.19	21.922
H1	−839.41 ± 0.056	−22.116	—	−0.058	0.074	−0.194
H2	−839.381 ± 0.055	−22.056	0.058	—	0.134	−0.134
H3	−839.447 ± 0.055	−22.19	−0.074	−0.134	—	−0.268
H4	−839.313 ± 0.058	−21.922	0.194	0.134	0.268	—

between the European and African populations of *C. laurifolius*, suggesting no transmaritime connection. Genetic differentiation values and phylogeographic hypothesis testing provided further evidence for the isolation of *C. laurifolius* populations by the Mediterranean barrier (Tables 3, 4). In contrast, at least two colonizations over the Mediterranean Sea have been documented for *C. ladanifer* (Guzmán & Vargas 2009b), which are responsible for the comparatively lower genetic differentiation encountered when re-analysing the haplotype distribution of this species (Table 3). Active colonization was found in *C. salviifolius*, as suggested by its even lower genetic differentiation and the presence of the same haplotypes and haplotyope lineages in Europe and Africa, which resulted from at least three intercontinental colonizations. Bayesian hypothesis testing supported that *C. ladanifer* and *C. salviifolius*, but not *C. laurifolius*, followed the predominant pattern of multiple colonizations across the Strait of Gibraltar (14 km), as previously proposed in a study of 18 plant groups (Rodríguez-Sánchez *et al.* 2008).

Isolation by longer geographical distances, such as that between the extremes of the Mediterranean, was only tested for *C. salviifolius* and *C. laurifolius* because *C. ladanifer* is exclusively distributed in the western Mediterranean. Our results suggested that full isolation was not accomplished, since a significant genetic differentiation was not found in *Cistus* populations separated by over 1600 km between the Iberian and Balkan Peninsulas (Table 3). There was little isolation by geographical distance in *C. salviifolius*, since there were no exclusive haplotype lineages between the Iberian and Balkan-Anatolian populations and the possibility of a single migration event was clearly rejected by our hypothesis testing (see also Farley & McNeilly 2000). In contrast, our phylogeographic reconstruction and hypothesis testing of *C. laurifolius* indicated a single, eastward migration event. Western Europe may have been the source area for a recent dispersal of this species, as indicated by the higher diversity of haplotypes and haplotype lineages within this area. Indeed, a recent (Pleistocene) European route of colonization is consistent with a single haplotype connection between Iberia and Anatolia through France and Italy (but see Dansereau 1939; Pignatti 1982). The fact that *C. monspeliensis* did not display plastid DNA variation *per se* indicated its rapid dispersal across the Mediterranean after species formation in the Pleistocene. Recent colonization between the Mediterranean extremes has also been documented for lowland species such as *Pinus pinea* (Vendramin *et al.* 2008) and *Olea europaea* (Besnard *et al.* 2009).

Elevation requirements as a determinant factor

The reproductive structures (naked capsules, smooth seeds), which are shared by every *Cistus* species, do not

explain their different levels of colonization across the Mediterranean basin. Ecological requirements, rather than dispersal characteristics, are thought to be crucial factors in the distribution of *Cistus* species (Guzmán & Vargas 2009b). As predicted by our ecological hypothesis, the phylogeographic patterns of species with lowland distributions (*C. ladanifer*) or no specific altitude requirements (*C. salviifolius*) were less geographically structured than that of the mountain species (*C. laurifolius*). Mediterranean mountains exhibiting island-like characteristics may have been an important determinant of the genetic isolation detected in *C. laurifolius* populations.

The results of haplotype distribution were consistent with limited geographical distribution of *C. laurifolius* not only in Mediterranean mountains but also within the Iberian Peninsula and within northern Africa (Fig. 3). This relatively long-term isolation pattern is reflected in the morphological differences between European-Turkish (subsp. *laurifolius*) and African (subsp. *atlanticus*) populations (Dansereau 1939). Therefore, the geographical structure of the genetic variation of *Cistus* indicates that various abiotic factors (water bodies, geographical distance, temperature, precipitation) were at play, favouring differentiation between European and African populations. In addition, strict edaphic conditions (acidic soils) and low germination rates (Tárrega *et al.* 2001; Ramos *et al.* 2006) may have also contributed to the relatively high isolation levels of *C. laurifolius*.

In summary, long geographical distances and the presence of water bodies across the Mediterranean basin have constituted permeable barriers for the distribution of *Cistus* since species formation in the Pleistocene. We argue that the ecological conditions in the acidic soils of medium-altitude mountains limited the colonization history of *C. laurifolius* through the European mountains. Additionally, barriers such as the Strait of Gibraltar may have played important roles in the differentiation of *C. laurifolius* populations. In contrast, the lowland (*C. monspeliensis*, *C. ladanifer*) and the widely distributed (*C. salviifolius*) species considered here displayed less isolation. Further insights into the net contributions of abiotic and biotic factors are required. This could be achieved by examining additional characteristics previously studied only in certain *Cistus* species, such as self-compatibility (Bosch 1992), pollinator effectiveness (Herrera 1987; Talavera *et al.* 1993), adaptation to vegetation disturbances, and competitiveness in new areas (Tárrega *et al.* 1997; Luis-Calabuig *et al.* 2000; Tárrega *et al.* 2001).

Acknowledgements

The authors thank Emilio Cano and Fátima Durán for laboratory assistance; Alan Forrest, Beatriz Guzmán and Luis Valente for advice on some analyses and comments that improved the quality of the manuscript; R. G. Page, B. Guzmán, V. Valcárcel, J. Arroyo, O. Filippi, J. J. Aldasoro, M. Escudero, M. Luceño, J. Martínez and the MA herbarium for plant material. This research is supported by the Spanish Dirección General de Investigación y Técnica (DGICYT) through project CGL2005-06017-C02 and a grant (JAE-Intro, CSIC) and PhD scholarship to M.F.

References

Arafeh R, Kadereit JW (2006) Long-distance seed dispersal, clone longevity and lack of phylogeographical structure in the European distributional range of the coastal *Calystegia soldanella* (L.) R. Br. (Convolvulaceae). *Journal of Biogeography*, **33**, 1461–1469.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3–15.

Besnard G, Rubio de Casas R, Vargas P (2007) Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea*). *Journal of Biogeography*, **34**, 736–752.

Besnard G, Rubio de Casas R, Christin PA, Vargas P (2009) Phylogenetics of *Olea* (Oleaceae) based on plastid and nuclear ribosomal DNA sequences: Tertiary climatic shifts and lineage differentiation times. *Annals of Botany*, **104**, 143–160.

Bosch J (1992) Floral biology and pollinators of three co-occurring *Cistus* species (Cistaceae). *Botanical Journal of the Linnean Society*, **109**, 39–55.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432–438.

Comes HP, Kadereit JW (2003) Spatial and temporal patterns in the evolution of the flora of the European Alpine System. *Taxon*, **52**, 451–462.

Dansereau PM (1939) Monographie du genre *Cistus*. *Boissiera*, **4**, 1–90.

Demesure B, Sodzi N, Petit J (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129–134.

Demoly JP, Montserrat P (1993) *Cistus*. In: *Flora Iberica*, *Vol. 3* (eds Castroviejo S, Aedo C, Cirujano S *et al.*), pp. 319–337. Consejo Superior de Investigaciones Cientifícas, Madrid.

Després L, Gielly L, Redoutet B, Taberlet P (2003) Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. *Molecular Phylogenetics and Evolution*, **27**, 185–196.

Dixon CJ, Schönswetter P, Schneeweiss GM (2007) Traces of ancient range shifts in a mountain plant group (*Androsace halleri* complex, Primulaceae). *Molecular Ecology*, **16**, 3890–3901.

Dixon CJ, Schönswetter P, Vargas P, Ertl S, Schneeweiss GM (2009) Bayesian hypothesis testing supports long-distance Pleistocene migrations in a European high mountain plant

(*Androsace vitaliana*, Primulaceae). *Molecular Phylogenetics and Evolution*, **53**, 580–591.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Ho SYW, Phyllips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Dumolin-Lapegue S, Pemonge MH, Petit J (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.

Farley RA, McNeilly T (2000) Diversity and divergence in *Cistus salvifolius* (L.) populations from contrasting habitats. *Hereditas*, **132**, 183–192.

Gamage DT, de Silva MP, Yoshida A, Szmidt AE, Yamazaki T (2003) Molecular phylogeny of Sri Lankan Dipterocarpaceae in relation to other Asian Dipterocarpaceae based on chloroplast DNA sequences. *Tropics*, **13**, 79–87.

Gamage DT, de Silva MP, Inomata N, Yamazaki T, Szmidt AE (2006) Comprehensive molecular phylogeny of the sub-family Dipterocarpoideae (Dipterocarpaceae) based on chloroplast DNA sequences. *Genes and Genetic System*, **81**, 1–2.

Gernhard T (2008) The conditioned reconstructed process. *Journal of Theoretical Biology*, **253**, 769–778.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Guzmán B, Vargas P (2005) Systematics, character evolution, and biogeography of *Cistus* L. (Cistaceae) based on ITS, *trnL-trnF*, and *matK* sequences. *Molecular Phylogenetics and Evolution*, **37**, 644–660.

Guzmán B, Vargas P (2009a) Historical biogeography and character evolution of Cistaceae (Malvales) based on analysis of plastid *rbc*L and *trn*L-*trn*F sequences. *Organisms, Diversity & Evolution*, **9**, 83–99.

Guzmán B, Vargas P (2009b) Long-distance colonization of the Western Mediterranean by *Cistus ladanifer* (Cistaceae) despite the absence of special dispersal mechanisms. *Journal of Biogeography*, **36**, 954–968.

Guzmán B, Lledó MD, Vargas P (2009) Adaptive radiation in Mediterranean *Cistus* (Cistaceae). *PLoS ONE*, **4**, e6362.

Hamilton M (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**, 521–523.

Herrera J (1987) Flower and fruit biology in Southern Spanish Mediterranean shrublands. *Annals of the Missouri Botanical Garden*, **74**, 69–78.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hudson RR (2000) A new statistic for detecting genetic differentiation. *Genetics*, **155**, 2011–2014.

Johnson LA, Soltis DE (1994) *mat*K DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany*, **19**, 143–156.

Juhren MC (1966) Ecological observations on *Cistus* in the mediterranean vegetation. *Forest Science*, **12**, 415–426.

Kass RE, Raftery AE (1995) Bayes factors and model uncertainty. *Journal of the American Statistical Association*, **90**, 773–795.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*, **87**, 482–498.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**, 652–655.

Kropf M, Comes HP, Kadereit JW (2008) Causes of the genetic architecture of south-west European high mountain disjuncts. *Plant Ecology & Diversity*, **1**, 217–228.

Larkin MA, Blackshields G, Brown NP *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Luis-Calabuig E, Tárrega R, Valbuena L (2000) Ten years of recovery of *Cistus ladanifer* after experimental disturbances. *Israel Journal of Plant Sciences*, **48**, 271–276.

Martín M, Guinea E (1949) Jarales y jaras (Cistografía Hispánica). *Boletín del Instituto Forestal de Investigaciones y Experiencias*, **49**, 1–228.

Menke B (1976) Pliozäne und ältestquartäre Sporen- und Pollenflora von Schleswig-Holstein. *Geologisches Jahrbuch Reihe A*, **32**, 3–197.

Mort ME, Soltis PS, Soltis DE, Mabry M (2000) Comparison of three methods for estimating internal support on phylogenetic trees. *Systematic Biology*, **49**, 160–171.

Naciri Y, Gaudeul M (2007) Phylogeography of the endangered *Eryngium alpinum* L. (Apiaceae) in the European Alps. *Molecular Ecology*, **16**, 2721–2733.

Naud G, Suc JP (1975) Contribution à l'étude paléofloristique des Coirons (Ardèche). Première analyses polliniques dans les alluvions sous-basaltique et interbasaltiques de Mirabel (Miocène supérieur). *Bulletin de la Société Géologique de France*, **17**, 820–827.

Newton MA, Raftery AE (1994) Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*, **56**, 3–48.

Olmstead RG, Reeves PA (1995) Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbc*L and *ndh*F sequences. *Annals of the Missouri Botanical Garden*, **82**, 176–193.

Palibin IV (1909) Cistaceae. In:*Flora Caucasica Critica* (eds Kuznetsov N, Bush N, Fomin A). pp. 117–123, Pensoft Publisher, Tartu.

Pignatti S (1982) *Flora d'Italia, vol. II*. Edagricole, Bologna.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, **16**, 37–45.

Rambaut A, Drummond AJ (2007) *Tracer v1.4*. Available from http://beast.bio.ed.ac.uk/Tracer

Ramos ME, Robles AB, Castro J (2006) Efficiency of endozoochorous seed dispersal in six dry-fruited species (Cistaceae): from seed ingestion to early seedling establishment. *Plant Ecology*, **185**, 97–106.

Rodríguez-Sánchez F, Pérez-Barrales 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.

Ronikier M, Cieślak E, Korbecka G (2008) High genetic differentiation in the alpine plant *Campanula alpina* Jacq. (Campanulaceae): evidence for glacial survival in several Carpathian regions and long-term isolation between the Carpathians and the Alps. *Molecular Ecology*, **17**, 1763–1775.

Schönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547–3555.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**, 275–288.

Short MJ (1994) Cistaceae. In:*Flora of Madeira* (eds Press JR, Short MJ). pp. 226–227, The Natural History Museum, HMSO, London.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, **49**, 369–381.

Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001–1013.

Swofford D (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer, Sunderland, MA.

Taberlet P (2002) Comparative phylogeography of alpine plants. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds Weiss S, Ferrand N), p. 16. International Symposium, Vairão (Portugal), March 11–15. – 15. Universidade do Porto, Vairão [Abstract].

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three noncoding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Talavera S, Gibbs PE, Herrera J (1993) Reproductive biology of *Cistus ladanifer* (Cistaceae). *Plant Systematics and Evolution*, **186**, 123–134.

Tárrega R, Luis-Calabuig E, Alonso I (1997) Space-time heterogeneity in the recovery after experimental burning and cutting in a *Cistus laurifolius* shrubland. *Plant Ecology*, **129**, 179–187.

Tárrega R, Luis-Calabuig E, Valbuena L (2001) Eleven years of recovery dynamic after experimental burning and cutting in two *Cistus* communities. *Acta Oecologica*, **22**, 277–283.

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.

Tribsch A, Schönswetter P (2003) Patterns of endemisms and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon*, **52**, 477–497.

Vargas P (2003) Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon*, **52**, 463–476.

Vendramin GG, Fady B, González-Martínez SC *et al.* (2008) Genetically depauperate but widespread: the case of an emblematic Mediterranean pine. *Evolution*, **62**, 680–688.

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 1–10.

Zhang LB, Comes HP, Kadereit JW (2001) Phylogeny and Quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *American Journal of Botany*, **88**, 2331–2345.

M.F.-M. is a PhD student working on the evolution, biogeography and conservation of Mediterranean plants, in particular species of the genus *Cistus* and *Linaria*. P.V. leads projects on the evolution and molecular systematics of primarily Mediterranean and Macaronesian plant groups.

Supporting Information

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

Table S1 Cistaceae taxa and populations sequenced for plastid regions (*trn*S-*trn*G, *trn*C^{GCA}-*trn*D^{GUC}, *psb*K-*trn*S, *trn*H^{GUG}-*trn*K^{UUU}, *trn*L-*trn*F and *mat*K). Taxonomy follows that of Guzmán & Vargas (2005), except for *C. laurifolius* subspecies and *C. albanicus* (formely called *C. sintenisii*)

Table S2 Primer pairs used in the pilot study performed to search for variable regions in the plastidial genome of *C. laurifolius*, *C. monspeliensis* and *C. salviifolius*. A hyphen (–) in the annealing temperature indicates that the region could not be amplified

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.