Journal of Biogeography (J. Biogeogr.) (2011) 38, 164–176

ORIGINAL ARTICLE

Out of Africa: north-westwards Pleistocene expansions of the heather *Erica arborea*

Aurélie Désamoré[1]*, Benjamin Laenen[1], Nicolas Devos[1], Magnus Popp[2], Juana Maria González-Mancebo[3], Mark A. Carine[4] and Alain Vanderpoorten[1]

[1]*Institute of Botany, University of Liège, Liège, Belgium,* [2]*National Centre for Biosystematics, Natural History Museum, University of Oslo, Oslo, Norway,* [3]*Department of Botany, University of La Laguna, Tenerife, Canary Islands, Spain,* [4]*Department of Botany, The Natural History Museum, London, UK*

*Correspondence: Aurélie Désamoré, Institut de Botanique, B22, Sart Tilman, Université de Liège, B4000, Liège, Belgium.
E-mail: adesamore@doct.ulg.ac.be

ABSTRACT

Aim The heather *Erica arborea* L. is a dominant element of the circum-Mediterranean region. Its broad, disjunct distribution, ranging from Macaronesia to eastern Africa, is consistent with the fragmentation of the evergreen tropical and subtropical forests that dominated Europe and North Africa in the Tertiary. This study aims to investigate phylogeographical patterns in *E. arborea* and to determine whether the current disjunct distribution of the species is a relict of a once wider distribution, or a recent range expansion in response to the establishment of suitable conditions.

Location Mediterranean, Macaronesia, North and eastern Africa.

Methods A total of 105 samples were collected across the species' distribution range and sequenced at four cpDNA loci (*atp*B–*rbc*L, *mat*K, *trn*H–*psb*A and *rpl*16). Phylogenetic reconstructions, molecular dating techniques and Bayesian ancestral area reconstructions were used in combination with population genetic statistics (haplotype diversity, N_{ST}, F_{ST}, Fu's F_S) to describe the pattern of present genetic diversity in *E. arborea* and infer its biogeographical history.

Results Haplotype diversity in Macaronesia and the east and central Mediterranean is much lower than that observed in eastern Africa/Arabia and the western Mediterranean. Bayesian ancestral area reconstructions and molecular dating suggest that *E. arborea* colonized the Mediterranean westwards from eastern Africa/Arabia at least twice during a time period ranging between the upper Miocene and the upper Pleistocene.

Main conclusions The phylogeography of *E. arborea* involves a complex history of range expansions and contractions, which has resulted in a pattern of distribution that mimics that expected for a Tertiary vicariance event. Despite the presence of a late Tertiary refugium in the Iberian Peninsula, the current distribution of the species throughout the Mediterranean is explained by a Pleistocene expansion originating from eastern Africa. One explanation for the isolation of the Iberian refugium is the rapidity of the most recently identified colonization wave, as inferred by the absence of global phylogeographical signal in the data and significantly negative values of Fu's F_S statistic for European populations. Macaronesia was colonized during each of these two expansion waves, confirming that the laurisilva (laurel forest flora) is a complex entity including both ancient relicts and recent colonizers.

Keywords
Eastern Africa, *Erica arborea*, heather, laurisilva, Macaronesia, Mediterranean, phylogeography, refugium, relictualism.

www.blackwellpublishing.com/jbi
doi:10.1111/j.1365-2699.2010.02387.x

Journal of Biogeography

INTRODUCTION

With about 50% of endemism at the species level, comparable with that of many tropical areas, the circum-Mediterranean region is one of the world's majors centres of plant diversity (Myers *et al.*, 2000). The onset of the summer-dry Mediterranean climate during the early Pleistocene (Suc, 1984) resulted in the rapid proliferation of many herbaceous lineages in the region (Axelrod, 1973; Valente *et al.*, 2010), with fire as a significant factor in diversification (Cowling *et al.*, 1996). In contrast, most woody Mediterranean taxa are thought to have originated in the late Cretaceous or early Tertiary as part of the evergreen tropical and subtropical laurel forests (also named laurisilva) that dominated the region during that period (Axelrod, 1973).

These forests became impoverished during the Pliocene as summer rainfall decreased, and largely disappeared in the early Pleistocene as summer rainfall was further reduced and as winter temperatures lowered (Mai, 1989; Svenning, 2003; Rodríguez-Sánchez & Arroyo, 2008). Today, they persist only in Macaronesia, where the oceanic location of the archipelagos mean that a sufficiently humid and mild climate has been maintained (e.g. Bramwell, 1976; Sunding, 1979; Whittaker & Fernández-Palacios, 2007), with some elements also persisting in humid refugia within the Mediterranean Basin.

The climate-driven fragmentation of the Mediterranean laurel forests is supported by ecological niche modelling and chloroplast DNA (cpDNA) sequence data analyses for *Laurus* L. (Rodríguez-Sánchez *et al.*, 2009). Hedberg (1970) and Bramwell (1976) also suggested that a suite of taxa that have distributions disjunct between Macaronesian/Mediterranean and eastern Africa are relicts of this process and that their distributions in the Tertiary would have been more widespread, possibly extending from southern Africa via the eastern African highlands and the mountains of the central Sahara (Hoggar, Tibesti) into the western part of North Africa and Macaronesia, and with a subsequent extension northwards into the Mediterranean (for a review see Andrus *et al.*, 2004). Indeed, the existence of spectacular range disjunctions and sister-group relationships between Macaronesian/Mediterranean and eastern/southern Africa have long been interpreted as a signature of the ancient fragmentation of the Tertiary forests by many authors (Sunding, 1979; Bramwell, 1985; Cronk, 1987, 1992; Mies, 1995, 1998; Hjertson, 1997; Olmstead & Palmer, 1997; Ghebrehiwet, 2000; Bohs & Olmstead, 2001).

However, molecular phylogenetic data have challenged hypothesized disjunctions between Macaronesia/western Mediterranean and eastern Africa in many groups (Andrus *et al.*, 2004). Furthermore, it has been shown that disjunct patterns can also be established by dispersal rather than strict vicariance. Consequently, some vicariance-like patterns appear to be a complex mix involving both vicariance and dispersal events (for a review see Mansion *et al.*, 2008).

The heather *Erica arborea* L. has a disjunct distribution encompassing Macaronesia, the Mediterranean and eastern Africa. In Macaronesia, it is a member of thermophilous forest communities, dominant on dry and shallow soils such as wind-prone mountain crests, southern aspects and forest margins, or transition zones to pine forests and in secondary woodlands (Del Arco Aguilar *et al.*, 2009). Outside Macaronesia, it occurs throughout the Mediterranean, although its distribution is not continuous. It occurs in maquis vegetation (Mesleard & Lepart, 1991) and also as understorey in different forest types in both Mediterranean and Atlantic areas. It is also known from the Tibesti Mountains (Chad) in Sahara, where it is present at the top of an upper montane desert steppe vegetation between 2000 and 3000 m a.s.l., and in eastern Africa and Arabia (including Yemen), where it frequently dominates between 3000 and 4000 m a.s.l. (Pichi-Sermolli & Heiniger, 1953; Bruneau de Miré & Quézel, 1959; Quézel, 1978; Messerli & Winiger, 1992; McGuire & Kron, 2005). Evidence from mesofossils suggests that the range of *E. arborea* extended as far north as Germany in the Miocene (Van Der Burgh, 1987), whilst palaeobotanical studies based on pollen cores indicate that, in response to recent climate changes (specifically from the Last Glacial Maximum to the present day), the species persisted in refugia in the Mediterranean (Carrión *et al.*, 2000). In the Saharan region, its distribution appears to have been more widespread as recently as the Neolithic (Bruneau de Miré & Quézel, 1959).

In this paper, we investigate geographical patterns of diversity in *E. arborea* using cpDNA sequence data. Our goal is to determine the extent to which the processes of vicariance, fragmenting a once continuous distribution, and dispersal through rapid range expansion, explain the current distribution of *E. arborea.*

MATERIALS AND METHODS

Sampling and molecular protocols

A total of 105 specimens of *E. arborea* from across its entire distribution range (Fig. 1) were sampled from herbarium material or from silica gel-dried material (see Appendix S1 in the Supporting Information). Preliminary phylogenetic analyses suggested that *E. arborea* is monophyletic and sister to the South African *E. trimera* (Engl.) Beentje (Appendix S2). Five specimens of the latter were consequently used as the outgroup. Approximately 10 leaves from each specimen were ground to powder in liquid nitrogen with a Genogrinder 2000 (BT&C/OPS Diagnostics, Lebanon, NJ, USA).

DNA extraction was performed using a cetyl trimethyl ammonium bromide (CTAB) protocol (Doyle & Doyle, 1987), followed by purification with Geneclean Kit (Qbiogene Inc., Solon, OH, USA) following the manufacturer's instructions. Each specimen was genotyped at four cpDNA loci: the *atp*B–*rbc*L intergenic spacer; the 5′ end of *mat*K; *trn*H–*psb*A; and *rpl*16. For *trn*H–*psb*A, the primers from Shaw *et al.* (2005) were used. Owing to difficulties in amplification for a number of accessions, specific sets of primers were designed for the three other loci (Table 1). These were designed within conserved regions at the 5′ and 3′ ends of the markers from an initial alignment of sequences obtained using universal primers from McGuire & Kron (2005) and Shaw *et al.* (2005).

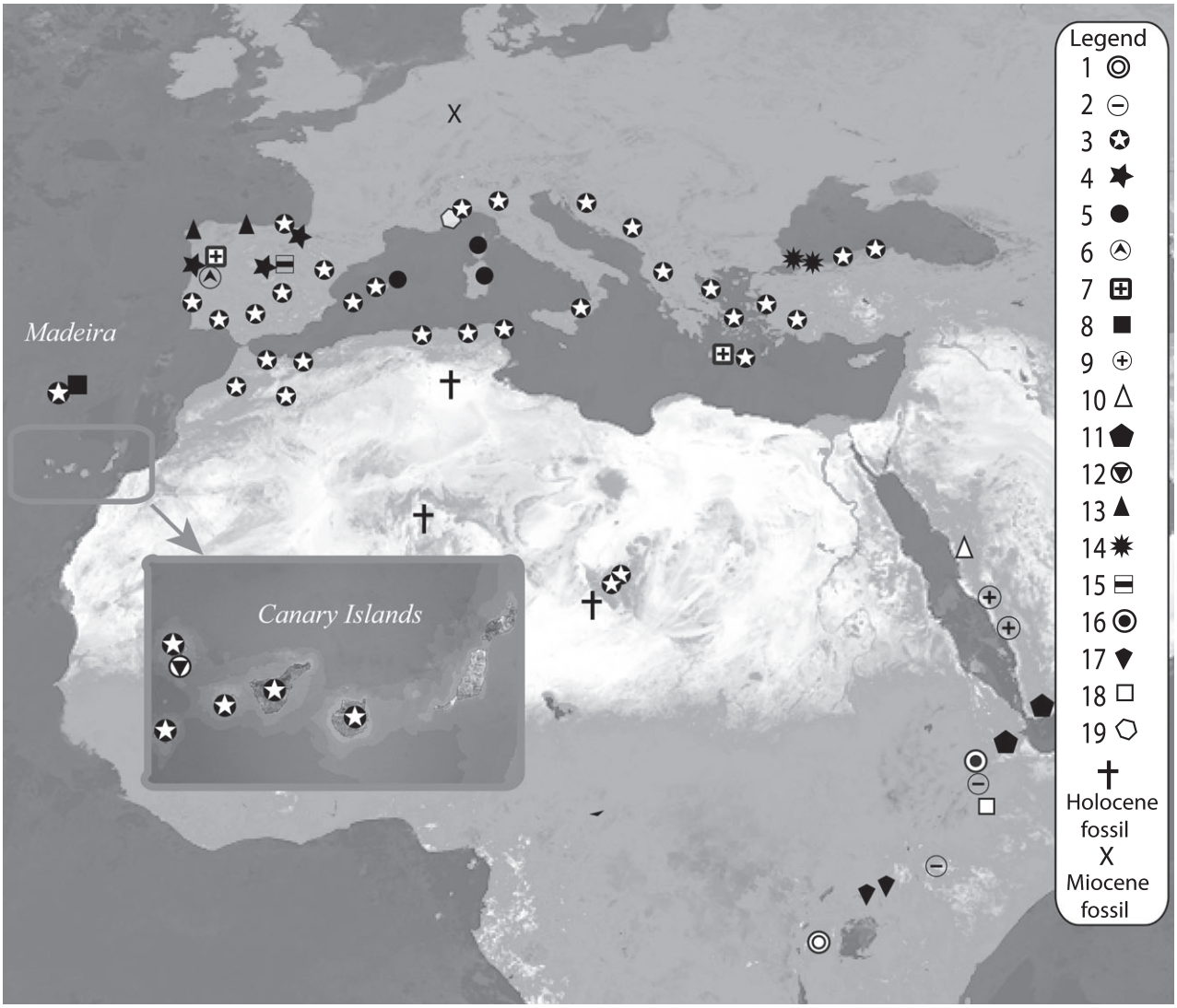

Figure 1 Haplotype distributions based on 105 specimens of *Erica arborea* collected throughout its distribution range and sequenced at *mat*K, *atp*B–*rbc*L, *rpl*16 and *trn*H–*psb*A (see Appendix S1 for sampling details). Where two or more specimens from a similar area share the same haplotype, only one is indicated for clarity. Fossil data are from Bruneau de Miré & Quézel (1959) and Van Der Burgh (1987).

Table 1 Specific primer pairs for the *rpl*16 and *mat*K genes and the *atp*B–*rbc*L intergenic spacer regions in *Erica arborea.*

Primers	Sequence (5′–3′)
*Mat*K–F	AAGACTTCTAGTTCGATTTTT
*Mat*K–R	CGCTTATCTTTTCAGGAGTATA
*rpl*16–F	CAACTCATCACTTCGTTTTATCTGG
*rpl*16–R	CAGCTCCTCGCGAATAAAAG
*atp*B–F	TATATTCAAAAAGTCAATATTAGGGCGA
*atp*B–R	TGAAATAAAAAGCGCCAATGAGATA

Polymerase chain reaction (PCR) was carried out in a 15 μL volume reaction using 1.5 μL 10× reaction buffer, 2.4 μL dNTPs mix (1 mM each), 0.6 μL 50 mM $MgCl_2$, 0.75 μL of each primer (10 μM), 1.125 μL BSA, 0.3 μL taq DNA polymerase and 1 μL DNA. PCR cycling consisted of 2 min denaturation at 95 °C, followed by 35 cycles of 30 s denaturation at 95 °C, 45 s annealing at 50 °C, 2 min extension at 72 °C and finally 7 min at 72 °C. The presence of amplified target DNA fragments was verified visually on agarose gel by staining with ethidium bromide. PCR products were purified with Exosap-it mix (USB Corporation, Cleveland, OH, USA). The sequencing reaction involved 2 min at 96 °C, 25 cycles of 15 s at 96 °C, 10 s at 50 °C, 4 min at 60 °C. The labelled fragments were then separated by capillary electrophoresis on an ABI-3100 sequencing machine (Applied Biosystems, Carlsbad, CA, USA).

Data analyses

Haplotype identification, variation and differentiation

Sequences were aligned and edited using Sequencher 3.1 (Schneider, 1998). Gaps were inserted where necessary to

preserve positional homology in the alignment. The four genes were concatenated to a single matrix in MacClade 4.0 (Maddison & Maddison, 1989).

Individuals exhibiting the same sequence across the four loci were grouped within the same haplotype using DnaSP (Rozás & Rozás, 1999; Rozás, 2009). The dataset was partitioned into four geographical regions. Two regions were recognized in the Mediterranean, and these correspond to the two main biogeographical regions identified in an analysis of *Erica* species distributions in Europe in relation to environmental variables: western Mediterranean (from Portugal to France plus Morocco, corresponding to region IIb in Ojeda *et al.*, 1998) and the central and eastern Mediterranean (hereafter C.E. Mediterranean), corresponding to regions Ia and Ib in Ojeda *et al.*, 1998, in which we also included the Tibesti massif in Chad. The other two areas recognized correspond to regions that are disjunct from the Mediterranean area: Macaronesia (the Canary Islands and Madeira) and eastern Africa/Arabia (Kenya, Ethiopia, Rwanda, Uganda, Arabia and Yemen).

Each region was characterized by its haplotype diversity (H), which was computed with Arlequin 3.11 (Schneider *et al.*, 2000; Excoffier *et al.*, 2005). The genetic variation was partitioned within and among regions by an analysis of molecular variance (AMOVA) with Arlequin 3.11. Variations in haplotype frequency among regions were estimated through pairwise F_{ST}, with significance tested by means of 10,000 random haplotype permutations among regions, as implemented by SPAGeDi 1.3 (Hardy & Vekemans, 2002). Presence of phylogeographical signal in the data was explicitly tested by contrasting N_{ST} and F_{ST} values among geographical regions. N_{ST} is a measure of genetic differentiation among populations analogous to F_{ST} but taking into account the phylogenetic relationships between alleles (Pons & Petit, 1996). Here, N_{ST} values were computed from a Tajima and Nei distance matrix (Tajima & Nei, 1984). An interesting property is that $N_{ST} > F_{ST}$ when phylogeographical signal exists; that is, when distinct alleles sampled from within populations are phylogenetically closer, on average, than alleles sampled from different populations (Pons & Petit, 1996). The hypothesis that $N_{ST} > F_{ST}$ was tested by computing the distribution under the null hypothesis by conducting 1000 permutations of rows and columns of the distance matrix among haplotypes. A signature of past demographic events in the present patterns of genetic variation was tested by means of Fu's F_S statistic (Fu, 1997). The F_S statistic was calculated using frequency distribution of alleles for each biogeographical region separately using Arlequin 3.11.

Phylogeographical relationships and reconstruction of ancestral distribution areas

A haplotype network was constructed using tcs (Clement *et al.*, 2000). Gaps were treated as missing data, and each indel was scored as a single binary character irrespective of its length, in an appended binary character matrix. Relationships between haplotypes were also inferred using MrBayes 3.1 (Ronquist & Huelsenbeck, 2003). A GTR+I DNA substitution model was selected based on the Akaike information criterion as implemented by Modeltest 3.7 (Posada & Buckley, 2004). Indel evolution was described using a model employing identical forward and backward transition rates (Lewis, 2001), which implements Felsenstein's correction (Felsenstein, 1992) for biased data matrices without constant characters. We thus applied the 'variable' coding option of MrBayes. Four Markov chain Monte Carlo (MCMC) simulations were run independently for 10,000,000 generations with MrBayes. Trees and model parameters were sampled every 10,000 generations. Convergence of the MCMCs was estimated in three ways. First, the standard deviation of split frequencies was < 0.01 after 10,000,000 generations. Second, visual inspection of the plot of the log-likelihood score at each sampling point suggested that the four chains reached stationarity. Third, the posterior probability plots of all splits for paired MCMC runs showed high correlation, which diagnoses convergence among the four chains (Nylander *et al.*, 2008). The trees of the burn-in for each run were excluded from the tree set, and the remaining trees from each run were combined to form the full sample of trees assumed to be representative of the posterior probability distribution.

The trees from the posterior probability distribution were used to reconstruct ancestral distribution areas. Each haplotype was assigned to one of five categories: (1) endemic to Macaronesia, (2) endemic to the western Mediterranean, (3) endemic to C.E. Mediterranean, (4) endemic to eastern Africa/Arabia, or (5) distributed in more than one region, which we refer to as combined distributions. The latter category was included because, as Humphries & Parenti (1999) noted, 'widespread taxa are uninformative by comparison to area relationships derived from endemic taxa' and 'introduce complications into biogeographical analysis because they obscure resolution and introduce redundancy by representing areas more than once on the terminals of any given cladogram'. The probabilities of change in a branch were calculated by estimating the instantaneous rates of transitions among all possible pairs of states. We used the 'global' approach, wherein model parameters are first fixed and then used to derive the set of most likely ancestral character states (Pagel, 1999). We examined the impact of the choice of a range of model parameters within and among trees by using the Markov chain model implemented by BayesTraits 1.0 (Pagel & Meade, 2004) to estimate the posterior probability distributions of ancestral states and rate coefficients. The latter were sampled from flat, uniform prior distributions ranging between 0 and 100. The rate at which parameters are changed ('ratedev') was set at the beginning of each run so that the acceptance rate of the proposed changes globally ranges between 20 and 50%. The chain was run for 10,000,000 generations, and rate parameters and probabilities of ancestral distribution areas were sampled every 1000 generations. In order to circumvent issues associated with the fact that not all the trees necessarily contain the internal nodes of interest, reconstructions were performed using a 'most recent common

ancestor' approach. This approach identifies, for each tree, the most recent common ancestor to a group of haplotypes and reconstructs the state at the node, then combines this information across trees (Pagel & Meade, 2004). This offers several advantages, including the incorporation of branch length information to determine the probabilities of change, and the integration of both mapping and phylogenetic uncertainty in the estimates, an advantage over parsimony-based approaches (Kodandaramaiah, 2009). In keeping with most recently developed programs for inferring historical distribution ranges (Ree & Smith, 2008; Sanmartín *et al.*, 2008), this approach is parametric; it is most similar to Sanmartín *et al.*'s (2008) model in the absence of information on island carrying capacity and dispersal rates.

Molecular dating

We determined whether sequence evolution conformed to the assumptions of the molecular clock by running maximum-likelihood analyses successively with and without a molecular clock enforced. The analyses employed heuristic searches implementing the GTR+I substitution model with 100 random addition replicates and tree bisection–reconnection (TBR) branch swapping to find the most likely trees in PAUP* 4.0b10 (Swofford, 2002). The likelihood of the trees obtained with and without the clock enforced was compared using a hierarchical likelihood ratio test. The likelihood ratio, which equals twice the difference in log-likelihood returned by the two models, is asymptotically chi-square distributed with the degrees of freedom equal to the number of operational taxonomic units (OTUs) minus two (Zhang, 1999). The trees with the highest likelihoods with and without the clock enforced differed by 7.3 log units. Since the critical value of a chi-square distribution with 17 d.f. is 27.59, the test suggests that enforcing the clock did not significantly decrease the log-likelihood.

From the 50% majority-rule consensus of the MRBAYES analysis, the MCMCs implemented by BEAST 1.5.2 (Drummond & Rambaut, 2007) were used to sample rate parameters and divergence times depending on their posterior probabilities. A normal distribution, with a mean and standard deviation of 5.0×10^{-10} and 10^{-10} substitutions per site per year, respectively, which corresponds to the average absolute substitution rate of cpDNA across land plants and largely encompasses their variation range, was used as a prior on the absolute rates of evolution for the four cpDNA regions combined (for reviews see Huttunen *et al.*, 2008; Aigoin *et al.*, 2009). The MCMC implemented the GTR+I substitution model and was run for 10,000,000 steps. Parameter values were sampled every 1000 generations during the 10,000,000 MCMC steps. Convergence and acceptable mixing of the parameters sampled was checked using the program TRACER v1.4 (Rambaut & Drummond, 2007), and the burn-in steps were discarded to obtain an estimation of the posterior probability distribution of divergence dates at the ancestral nodes.

RESULTS

Genetic diversity

The total alignment of the four regions was 1876 base pairs (bp) long and included a total of 21 base substitutions within *E. arborea*. The length and number of substitutions per locus was as follows: *atp*B–*rbc*L, 253 bp long, two substitutions; *trn*H–*psb*A, 307 bp long, seven substitutions; *mat*K, 584 bp long, six substitutions; *rpl*16, 732 bp long, six substitutions. Indels provided another five variable characters. Based on the 26 variable characters, 19 haplotypes were identified. One haplotype (no. 3) is characterized by having a strikingly high frequency of occurrence in the sample (64.8%). The frequency of all the other haplotypes ranged between 0.95 and 3.81% (Table 2).

The bulk of haplotype diversity is concentrated around eastern Africa/Arabia and the northern half of the Iberian Peninsula (Fig. 1). Eight haplotypes, all of which are endemic, are present in eastern Africa/Arabia. Five haplotypes, four of which are endemic and one of which is shared with Crete (no. 7), are situated in northern Iberian Peninsula. Haplotype 3 is distributed across Europe, North Africa, Tibesti and Macaronesia. Haplotype 5 was found only in the Mediterranean islands (Corsica, Sardinia, Balearic Islands). Two haplotypes were found only in Macaronesia: haplotype 8 in Madeira and haplotype 12 in La Palma. Haplotype 19 was found in a single sample from France. The highest haplotype diversity levels

Table 2 Haplotype frequency within a sample of 105 individuals of *Erica arborea* collected throughout its distribution range and sequenced at *rpl*16, *mat*K, *trn*H–*psb*A and the *atp*B–*rbc*L intergenic spacer.

Haplotype*	Number of samples	Frequency
1	2	1.90
2	3	2.86
3	68	64.76
4	4	3.81
5	4	3.81
6	1	0.95
7	2	1.90
8	2	1.90
9	2	1.90
10	1	0.95
11	2	1.90
12	1	0.95
13	4	3.81
14	2	1.90
15	1	0.95
16	1	0.95
17	2	1.90
18	2	1.90
19	1	0.95
Total	105	

*Haplotype numbers correspond to those used in Fig. 1.

Table 3 Number of haplotypes per region, gene diversity (H) and Fu's F_S coefficient per biogeographical region within a sample of 105 individuals of *Erica arborea* collected throughout its distribution range and sequenced at *rpl*16, *mat*K, *trn*H–*psb*A and the *atp*B–*rbc*L intergenic spacer.

Region	Number of haplotypes per region	H	F_S
Macaronesia	3	0.227	n.s.
Western Mediterranean	7	0.593	n.s.
Central and eastern Mediterranean	4	0.377	−2.979**
Eastern Africa/Arabia	8	0.924	n.s.

n.s., $P > 0.05$; **$P < 0.01$.

were found in eastern Africa/Arabia (0.924) and the western Mediterranean (0.593). The lowest haplotype diversities were found in Macaronesia (0.227) and central and eastern Mediterranean region (0.377; Table 3).

Geographical differentiation, phylogeographical signal and neutrality tests

Genetic variance among biogeographical regions accounts for 20.8% of the total genetic variance, and this differentiation is significant at the 0.001 level. The global N_{ST} value (0.22) is also significant ($P < 0.001$) and is the same as the global F_{ST} (0.22). Significant pairwise F_{ST} values were observed between all regions, except for the comparisons between C.E. Mediterranean, Macaronesia and western Mediterranean (Table 4). Fu's F_S statistics for the biogeographical regions are provided in Table 3. Significantly negative statistics were observed only for the C.E. Mediterranean region ($F_S = -2.979$, $P < 0.01$).

Phylogeography

Three main groups were identified in the 50% majority-rule consensus tree from the Bayesian analysis (Fig. 2) and these were also represented in the haplotype network (Fig. 3). Group I, which comprises haplotypes from the Iberian Peninsula (haplotypes 4, 6, 13, 15) and a single haplotype endemic to Madeira (haplotype 8), is supported with a posterior probability (hereafter, PP) of 1, and is resolved as sister to group IV. The latter, which includes the remainder of the haplotypes, is supported with a PP of 0.99. Group IV includes haplotypes 9 and 10 forming a polytomy with groups II and III, which are also strongly supported (PP of 0.97 and 0.98, respectively). Group II is composed of endemic haplotypes from eastern Africa. Group III includes haplotypes from the western Mediterranean, C.E. Mediterranean and Macaronesia. The three haplotypes found in Macaronesia are thus not resolved as a monophyletic group, as one is found in group I (haplotype 8 in Madeira) whereas the other two are resolved in group III (the widespread haplotype 3 in the Canary Islands and Madeira, and haplotype 12 in Madeira). In the haplotype network, it is noticeable that within group III, the geographically widespread haplotype 3 is central to a radiation of five other haplotypes (5, 7, 12, 14, 19). Within group I, haplotype 4 is similarly central to a radiation of four other haplotypes (6, 8, 13, 15). Group I is well separated from groups II and III by 17 mutational steps in each case. Within group III, eight mutational steps separate the haplotypes from eastern Africa/Arabian region from those of all the other regions.

The Bayesian reconstruction of ancestral distribution areas unambiguously identifies eastern Africa/Arabia as the most likely ancestral distribution area of the species with a PP of 0.97 (Fig. 2). The western Mediterranean region was identified as the ancestral distribution area for group I with a PP of 0.94, whilst the eastern Africa/Arabia region is identified as the most likely ancestral distribution area for groups II, III and IV with a PP of 0.98, 0.86 and 0.98, respectively.

Molecular dating

Ninety-five per cent confidence intervals (CI) of time estimates from the BEAST analysis are presented in Fig. 2. The divergence between groups I and IV took place between 5.82 and 1.49 million years ago (Ma), with a mean of 3.25 Ma. This period encompasses the upper Miocene to upper Pliocene. The node separating groups II and III is dated between 3.38 and 0.78 Ma with a mean of 1.56 Ma, corresponding to late Pliocene/early Pleistocene. The node separating the eastern Africa/Arabia haplotypes 11 and 16 from the rest of group III is dated between 1.6 and 0.3 Ma, with a mean of 0.83 Ma. This period corresponds to the early Pleistocene.

Table 4 Pairwise F_{ST} values between the four geographical regions defined across the distribution range of *Erica arborea* from a sample of 105 specimens sequenced at *rpl*16, *mat*K, *trn*H–*psb*A and the *atp*B–*rbc*L intergenic spacer.

	Macaronesia	Western Mediterranean	Central and eastern Mediterranean	Eastern Africa/Arabia
Macaronesia	0			
Western Mediterranean	0.08*	0		
Central and eastern Mediterranean	n.s.	n.s.	0	
Eastern Africa/Arabia	0.47***	0.26***	0.39***	0

n.s., $P > 0.05$; *$P < 0.05$; ***$P < 0.001$.

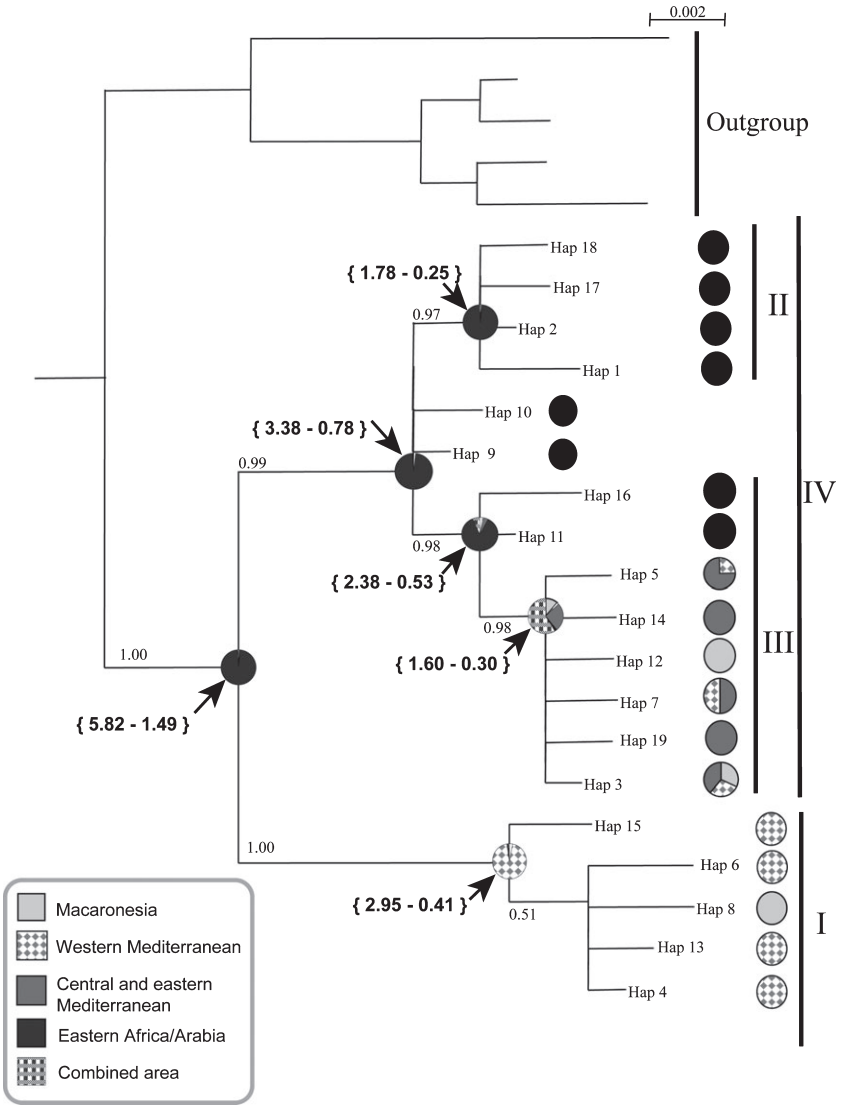

Figure 2 Fifty per cent majority-rule consensus tree with branch lengths averaged across 3920 trees sampled after convergence of four Markov chain Monte Carlo (MCMC) simulations implementing a GTR+I nucleotide substitution model from the analysis of *mat*K, *atp*B–*rbc*L, *rpl16* and *trn*H–*psb*A sequences represented by 19 haplotypes in *Erica arborea*. Posterior probabilities >0.5 are indicated on the branches. Pie diagrams at terminal nodes indicate the geographical distribution of the correspondent haplotype in the western Europe, central and eastern Europe, Macaronesia, eastern Africa/Arabia, and combined area, which corresponds to haplotypes present in more than one region. Pie diagrams at ancestral nodes represent the average posterior probabilities of ancestral areas derived from a Bayesian analysis implementing an MCMC visiting the space of trees and rate parameters depending on their posterior probabilities. Numbers in parentheses correspond to the 95% confidence intervals of the age (in Ma) inferred at internal nodes from a Bayesian analysis implementing an MCMC visiting a prior distribution of absolute nucleotide substitution rates documented across land plants and absolute divergence times in proportion to their posterior probability (see text for details). The tree was rooted with *E. trimera* but accessions of the latter were pruned in the ancestral distribution area reconstructions and molecular dating analyses. The scale bar represents the expected substitutions per site.

DISCUSSION

Evolutionary origin and late Tertiary range fragmentation of *Erica arborea*

The reconstruction of ancestral areas identifies eastern Africa/Arabia as the centre of origin of *E. arborea*, with a divergence time between clades I and IV of 5.82–1.49 Ma (mean = 3.25 Ma). This suggests an early range disjunction separating the Iberian clade I from the eastern African/Arabian clade IV, and broadly coincides with the onset of the Mediterranean climate in the course of the Pliocene.

The reconstruction of ancestral distribution areas could be interpreted in terms of either dispersal from eastern Africa/Arabia to North Iberia, or vicariance due to the fragmentation of a widespread distribution as a result of the cooler and drier climates that were established in the Upper Miocene. Mesofossils from Germany suggest that *E. arborea* was more widespread during the Miocene, with a distribution extending much further north than at present (Van Der Burgh, 1987). Climate-induced vicariance, with the survival of *E. arborea* in two refugial areas (north of Iberian Peninsula and eastern Africa/Arabia), therefore appears to be the most plausible explanation for the disjunct pattern observed. It is notable that clades I and IV (the latter comprising clades II and III), separated by 17 steps on the haplotype network, are the two most divergent groups in *E. arborea*. In the genus *Olea* L., range disjunctions have similarly been interpreted as a consequence of the desertification of the Sahara following the expansion of the East Antarctic ice sheet and the uplift of the eastern African mountains (Besnard *et al.*, 2007, 2009). Relict elements have been identified in the floras of both eastern Africa/Arabia and the Iberian Peninsula. For example, the giant lobelias of the eastern African mountains are considered isolated relicts (Knox & Palmer, 1998). In Iberia, the differentiation observed is comparable with that of other taxa of late Tertiary origin, such as *Frangula alnus* Mill. (Hampe *et al.*, 2003) and *Hedera* L. (Valcárcel *et al.*, 2003). Subsequent periods of glacial/interglacial fragmentations during the Quaternary have further shaped the patterns of genetic

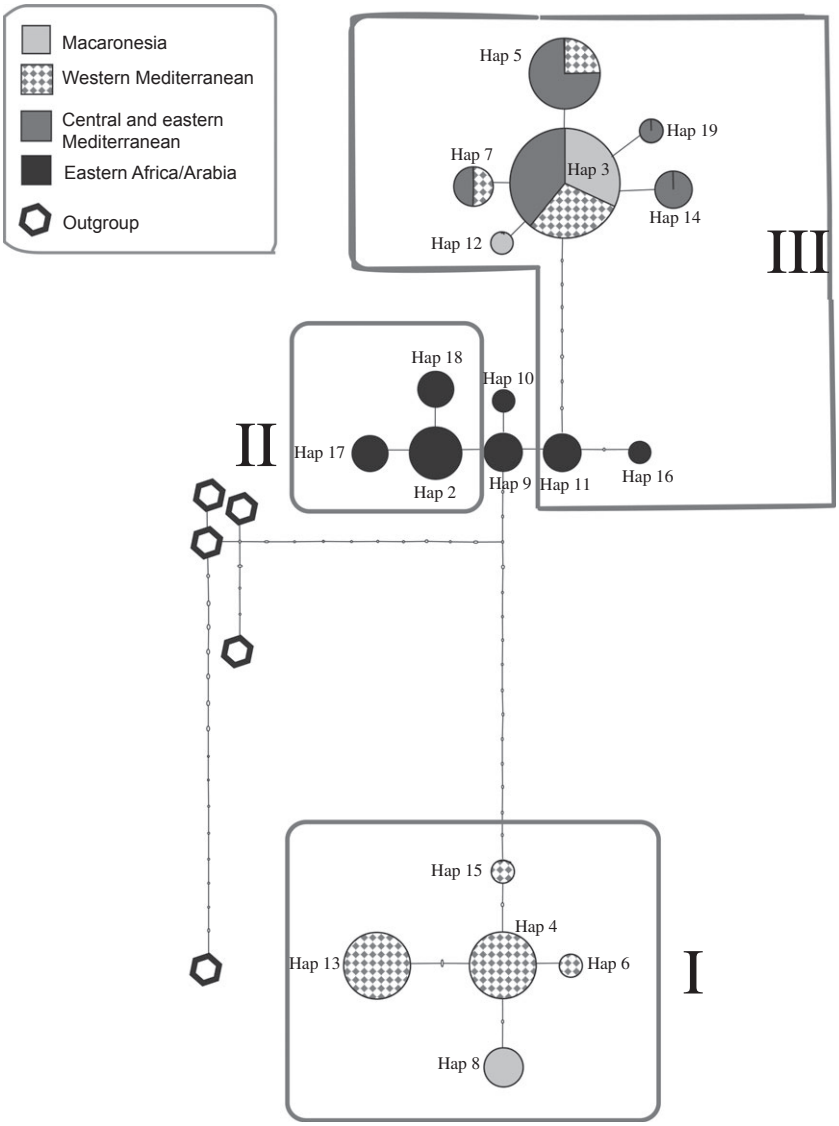

Figure 3 *Erica arborea* haplotype network reconstructed from sequences of *mat*K, *atp*B–*rbc*L, *rpl*16 and *trn*H–*psb*A from across the distribution range of the species. The size of the circle representing each haplotype is proportional to its frequency. Dots represent extinct or unsampled haplotypes.

variation, for example in the fern *Culcita macrocarpa* Presl., which is restricted to a few Iberian localities and Macaronesia (Bañares *et al.*, 2009), and in *Reseda* L. (Martín-Bravo *et al.*, 2010).

Changes in climate during the Upper Miocene precipitated the demise of the continental subtropical forests more generally (for a review see Mansion *et al.*, 2008) and it is notable that a pattern of transcontinental disjunct refugia similar to that observed in *E. arborea* has also been proposed for several elements of the laurel forest flora, notably in *Canarina* L. and the dragon tree group of *Dracaena* Vandelli ex L. (Bramwell, 1972). However, the pattern in both of these taxa differs from that of *E. arborea* in that the western refugium in each case is hypothesized to occur in Macaronesia rather than in the Iberian Peninsula.

Pleistocene range expansion and fragmentation

The reconstruction of ancestral areas, combined with the molecular dating analysis, suggests that Mediterranean haplotypes in group III (Fig. 2) form a monophyletic group that became isolated from the eastern Africa/Arabia refugium during the Pleistocene (2.38–0.53 Ma). The late branching position of the Mediterranean haplotype clade with a paraphyletic grade of eastern African/Arabian haplotypes suggests colonization of the Mediterranean from the eastern Africa/Arabia refugium. Thus, in contrast to other taxa that share with *E. arborea* a disjunct distribution between Europe and eastern Africa (e.g. *Lychnis* L., Popp *et al.*, 2008; *Arabis alpina* L., Assefa *et al.*, 2007; Ehrich *et al.*, 2007; Koch *et al.*, 2006), eastern Africa has served as a source area for recolonization of Europe rather than as a sink. The recolonization of Europe from an eastern Africa/Arabia refugium contrasts with the typical scenario of post-glacial recolonization of Europe from the Mediterranean peninsulas (Hewitt, 1999, 2004; Petit *et al.*, 2003) and the eastern Mediterranean (for a review see Mansion *et al.*, 2010). It further raises the question of why the Mediterranean basin was not recolonized from the geographically closer Iberian refugium, which has acted as a refugium for animal (for a review see Centeno-Cuadros *et al.*, 2009) and

plant species including *Quercus ilex* L. (López de Heredia *et al.*, 2007), *Fagus sylvatica* L. (López-Merino *et al.*, 2008), *Betula pendula* Roth. (Palmé *et al.*, 2003) and *Corylus avellana* L. (Palmé & Vendramin, 2002; for a review see Médail & Diadema, 2009). One possibility is that the colonization wave from the eastern African/Arabia refugium was extremely rapid, possibly preventing subsequent establishment by other haplotypes. The C.E. Mediterranean region is characterized by the occurrence of one main haplotype (haplotype 3) and by a strong reduction in haplotype diversity, a signature expected in regions that were rapidly recolonized when climatic conditions improved (Kerdelhué *et al.*, 2009).

Three independent lines of evidence support this interpretation. First, the branching patterns in group III of the haplotype network exhibit a star-like polytomy that is typically inferred as evidence for demographic expansion (Slatkin & Hudson, 1991; Szövényi *et al.*, 2006). Second, the observation that global N_{ST} is not significantly higher than global F_{ST} indicates that dispersal has occurred at a higher rate than mutation, that is, the species has dispersed before diversifying locally. Third, it is notable that haplotype variation within *E. arborea* is characterized by many singletons, and that the level of genetic differentiation among haplotypes is low, mostly involving a single mutational step. Such an excess of rare mutations in comparison with the expectations under a neutral coalescent in a constant sized population leads to a significantly negative F_S for the C.E. Mediterranean region. Whilst such a result can be interpreted in terms of either selection or population demographic expansion (Ramos-Onsins & Rozás, 2002), we favour the latter, given that most of the nucleotide variation observed in *E. arborea* was restricted to non-coding regions. *Erica arborea* is an effective colonizer, producing many small, easily dispersed seeds that remain viable in the seed bank for a long time (Fernández-Palacios & Arévalo, 1998). It thrives following disturbance (Anon., 1973; Arévalo & Fernández-Palacios, 2007; Arévalo *et al.*, 2008; Álvarez *et al.*, 2009), and Holocene pollen records from northern Morocco show peaks of abundance of *E. arborea* coinciding with historical periods of high disturbance (Ojeda *et al.*, 2008). Climate-induced disturbance during the Pleistocene may therefore have facilitated rapid range expansion in this species.

In North Africa, the distribution of *E. arborea* is currently restricted to the Maghreb region and the Tibesti massif in Chad, where it occurs at high elevation (2000–3000 m). However, pollen of this species is also known from Neolithic sediment in areas of North Africa, including Touggourt, the Hoggar massif and the diatomites of Borkou (Bruneau de Miré & Quézel, 1959). Palaeoecological data suggest that favourable conditions for the existence of broadleaved evergreen trees in the present Sahara desert could have existed for millennia with climatic oscillations on the Milankovitch time-scale, forcing disjunctions during some phases and leading to mixing of floras during others (Prentice *et al.*, 2000). Certainly, earlier in the Holocene, conditions in the Sahara region were far moister than at present, and the Saharan mountains supported typically Mediterranean taxa such as *Quercus ilex* L. and *Pistacia* L. species, as well as *Erica* species, as recently as 6000 BP (Jolly *et al.*, 1998; Prentice *et al.*, 2000). In *E. arborea*, the lack of haplotype differentiation between disjunct populations in Tibesti and Maghreb is entirely consistent with the idea of a rapid colonization of North Africa followed by more recent range fragmentation induced by Holocene aridification.

The colonization of Macaronesia

Both the haplotype network and the reconstruction of ancestral distribution areas suggest that *E. arborea* colonized Macaronesia at least twice independently. Haplotype 8, which is endemic to Madeira, is derived from a haplotype that is restricted to the north Iberian refugium. The most recent common ancestor of this group is dated between 2.95 and 0.41 Ma, which is fully compatible with the emergence of Madeira, 5.2 Ma. Haplotype 12, which is endemic to La Palma, is derived from the widespread haplotype 3 (Fig. 3), which also occurs in Madeira. Haplotype 3, together with the other Mediterranean haplotypes in group III, are hypothesized to have diverged between 1.6 and 0.3 Ma, compatible with the emergence of La Palma, 2 Ma. The recent origin of haplotypes distributed in Macaronesia suggests that *E. arborea* is not a Tertiary relict, and further challenges the idea that range disjunctions between Macaronesia/Mediterranean and eastern Africa are the result of vicariance, driven by changes in climate during the Pliocene (for a review see Andrus *et al.*, 2004).

CONCLUSIONS

Patterns of molecular diversity in *E. arborea*, together with palaeobotanical evidence, suggest a complex explanation for the extant distribution of this species, in which both vicariance and range expansion have played a major role. Two waves of range expansion and contraction are hypothesized. The first involves dispersal from the eastern Africa/Arabia centre of origin of the species, across northern Africa and Europe. This was followed by a dramatic range contraction dated from the late Miocene/Pliocene that restricted the species to two refugia located in northern Iberia and eastern Africa/Arabia, respectively. A second, rapid range expansion westwards from the eastern Africa/Arabia refugium occurred during the Pleistocene and led to the recolonization of the Mediterranean region, Macaronesia and North Africa. Finally, there was a more recent range fragmentation in North Africa, most likely driven by the expansion of the Sahara. Altogether, these observations reinforce the idea of a highly heterogeneous origin of the Mediterranean flora (Mansion *et al.*, 2008).

ACKNOWLEDGEMENTS

We are grateful to the National Botanical Garden of Belgium, the Muséum d'histoire naturelle d'Aix-en-Provence, Ruth Jaen Molina (Jardín Botánico Canario 'Viera y Clavijo'), David Harris (Royal Botanic Garden, Edinburgh), and Stephen Jury (The University of Reading) for the loan of collections. Thanks

are also due to Jon Shaw (Duke University) for providing laboratory facilities, Sandy Boles and Peter Szövényi for assistance, and two anonymous referees for their constructive comments on the manuscript. This study was supported by FSR-FNRS grant F.4505.07 and the Fond Léopold III of Belgium.

REFERENCES

Aigoin, D.A., Devos, N., Huttunen, S., Ignatov, M.S., González-Mancebo, J.M. & Vanderpoorten, A. (2009) And if Engler was not completely wrong? Evidence for multiple evolutionary origins in the moss flora of Macaronesia. *Evolution*, **63,** 3248–3257.

Álvarez, R., Muñoz, A., Pesqueira, X.M., García-Duro, J., Reyes, O. & Casal, M. (2009) Spatial and temporal patterns in structure and diversity of Mediterranean forest of *Quercus pyrenaica* in relation to fire. *Forest Ecology and Management*, **257,** 1596–1602.

Andrus, N.J., Trusty, J., Santos-Guerra, A., Jansen, R.K. & Francisco-Ortega, J. (2004) Using molecular phylogenies to test phytogeographical links between East/South Africa, Southern Arabia and the Macaronesian islands: a review, and the case of *Vierea* and *Pulicaria* section *Vieraeopsis* (Asteraceae). *Taxon*, **53,** 333–346.

Anon. (1973) *Inventario forestal nacional*. Ministerio de Agriculture, Pesca y Alimentación: Section de inventario y mapas, Santa Cruz de Tenerife, Madrid.

Arévalo, J.R. & Fernández-Palacios, J.M. (2007) Treefall gaps and regeneration composition in the laurel forest of Anaga (Tenerife): a matter of size? *Plant Ecology*, **188,** 133–143.

Arévalo, J.R., Peraza, M.D., Alvárez, C., Bermúdez, A., Delgado, J.D., Gallardo, A. & Fernández-Palacios, J.M. (2008) Laurel forest recovery during 20 years in an abandoned firebreak in Tenerife, Canary Islands. *Acta Oecologica*, **33,** 1–9.

Assefa, A., Ehrich, D., Taberlet, P., Nemomissa, S. & Brochmann, C. (2007) Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity*, **99,** 133–142.

Axelrod, D.I. (1973) History of the mediterranean ecosystem in California. *Mediterranean-type ecosystems* (ed. by F. di Castri and H.A. Mooney), pp. 225–277. Springer, New York.

Bañares, Á., Blanca, G., Güemes, J., Moreno, J.C. & Ortiz, S. (2009) *Atlas y Libro Rojo de la flora vascular amenazada de España. Adenda 2008*. Dirección General de Medio Natural y Política Forestal y Sociedad Española de Biología de la Conservación de Plantas, Madrid.

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, P.-A. & 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.

Bohs, L. & Olmstead, R.G. (2001) A reassessment of *Normania* and *Triguera* (Solanaceae). *Plant Systematics and Evolution*, **228,** 33–48.

Bramwell, D. (1972) Endemism in the flora of the Canary Islands. *Taxonomy, phytogeography and evolution* (ed. by D.H. Valentine), pp. 141–159. Academic Press, London.

Bramwell, D. (1976) The endemic flora of the Canary Islands. *Biogeography and ecology in the Canary Islands* (ed. by G. Kunkel), pp. 207–240. Junk, The Hague.

Bramwell, D. (1985) Contribución a la biogeografía de las Islas Canarias. *Botanica Macaronesica*, **14,** 3–34.

Bruneau de Miré, P. & Quézel, P. (1959) Sur la présence de la bruyère en arbre (*Erica arborea* L.) sur les sommets de l'Emi Koussi (Massif du Tibesti). *Comptes Rendus de la Société Biogéographique*, **34,** 66–70.

Carrión, J.S., Navarro, C., Navarro, J. & Munuera, M. (2000) The distribution of cluster pine (*Pinus pinaster*) in Spain as derived from palaeoecological data: relationships with phytosociological classification. *The Holocene*, **10,** 243–252.

Centeno-Cuadros, A., Delibes, M. & Godoy, J.A. (2009) Phylogeography of Southern Water Vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? *Molecular Ecology*, **18,** 3652–3667.

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

Cowling, R.M., Rundel, P.W., Lamont, B.B., Arroyo, M.K. & Arianoutsou, M. (1996) Plant diversity in mediterranean-climate regions. *Trends in Ecology and Evolution*, **11,** 362–366.

Cronk, Q.C.B. (1987) The history of endemic flora of St Helena – a relictual series. *New Phytologist*, **105,** 509–520.

Cronk, Q.C.B. (1992) Relict floras of Atlantic islands: patterns assessed. *Biological Journal of the Linnean Society*, **46,** 91–103.

Del Arco Aguilar, M., Pérez de Paz, P.L., Acebes Ginovés, J.R., Fernández López, A.B. & Wildpret de la Torre, W. (2009) La vegetación del Parque Nacional de Garajonay. *Parque Nacional de Garajonay, Patrimonio Mundial* (ed. by A.B. Fernández López), pp. 273–337. O. A. Parques Nacionales, Serie Técnica, Ministerio de Medio Ambiente, Rural y Marino, Madrid.

Doyle, J.J. & Doyle, L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19,** 11–15.

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

Ehrich, D., Gaudeul, M., Assefa, A., Koch, A.M., Mummenhoff, K., Nemomissa, S. & Brochmann, C. (2007) Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology*, **16,** 2542–2559.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Felsenstein, J. (1992) Phylogenies from restriction sites: a maximum-likelihood approach. *Evolution*, **46,** 159–173.

Fernández-Palacios, J.M. & Arévalo, J.R. (1998) Regeneration strategies of tree species in the laurel forest of Tenerife (the Canary Islands). *Plant Ecology*, **137,** 21–29.

Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147,** 915–925.

Ghebrehiwet, M. (2000) Taxonomy, phylogeny and biogeography of *Kickxia* and *Nanorrhinum* (Scrophulariaceae). *Nordic Journal of Botany*, **20,** 655–689.

Hampe, A., Arroyo, J., Jordano, P. & Petit, R.J. (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, **12,** 3415–3426.

Hardy, O.J. & Vekemans, X. (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2,** 618–620.

Hedberg, O. (1970) Evolution of the Afroalpine flora. *Biotropica*, **2,** 16–23.

Hewitt, G.M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68,** 87–112.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359,** 183–195.

Hjertson, M. (1997) Systematics of *Lindenbergia* and *Campylanthus* (Scrophulariaceae). *Acta Universitatis Upsaliensis*, **331,** 1–46.

Humphries, C.J. & Parenti, L.R. (1999) *Cladistic biogeography*, 2nd edn. Oxford University Press, Oxford.

Huttunen, S., Hedenäs, L., Ignatov, M.S., Devos, N. & Vanderpoorten, A. (2008) Origin and evolution of the northern hemisphere disjunction in the moss genus *Homalothecium* (Brachytheciaceae). *American Journal of Botany*, **95,** 720–730.

Jolly, D., Harrison, S.P., Damnati, B. & Bonnefille, R. (1998) Simulated climate and biomes of Africa during the late Quaternary: comparison with pollen and lake status data. *Quaternary Science Reviews*, **17,** 629–657.

Kerdelhué, C., Zane, L., Simonato, M., Salvato, P., Rousselet, J., Roques, A. & Battisti, A. (2009) Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9,** 220.

Knox, E.B. & Palmer, J.D. (1998) Chloroplast DNA evidence on the origin and radiation of the giant lobelias in eastern Africa. *Systematic Botany*, **23,** 109–149.

Koch, M.A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C. & Mummenhoff, K. (2006) Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology*, **15,** 825–839.

Kodandaramaiah, U. (2009) Use of dispersal–vicariance analysis in biogeography – a critique. *Journal of Biogeography*, **37,** 3–11.

Lewis, P.O. (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, **50,** 913–925.

López de Heredia, U., Carrión, J.S., Jiménez, P., Collada, C. & Gil, L. (2007) Molecular and palaeoecological evidence for multiple glacial refugia for evergreen oaks on the Iberian Peninsula. *Journal of Biogeography*, **34,** 1505–1517.

López-Merino, L., López-Sáez, J.A., Zapata, M.B.R. & García, M.J.G. (2008) Reconstructing the history of beech (*Fagus sylvatica* L.) in the north-western Iberian Range (Spain): from Late-Glacial refugia to the Holocene anthropic-induced forests. *Review of Palaeobotany and Palynology*, **152,** 58–65.

Maddison, W.P. & Maddison, D.R. (1989) Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatologica*, **53,** 190–202.

Mai, D.H. (1989) Development and regional differentiation of the European vegetation during the Tertiary. *Plant Systematics and Evolution*, **162,** 79–91.

Mansion, G., Rosenbaum, G., Schoenenberger, N., Bacchetta, G., Rosselló, J.A. & Conti, E. (2008) Phylogenetic analysis informed by geological history supports multiple, sequential invasions of the Mediterranean basin by the angiosperm family Araceae. *Systematic Biology*, **57,** 269–285.

Mansion, G., Selvi, F., Guggisberg, A. & Conti, E. (2010) Origin of Mediterranean insular endemics in the Boraginales: integrative evidence from molecular dating and ancestral area reconstruction. *Journal of Biogeography*, **36,** 1282–1296.

Martín-Bravo, S., Valcárel, V., Vargas, P. & Luceño, M. (2010) Geographical speciation related to Pleistocene range shifts in the western Mediterranean mountains (*Resada* sect. *Glaucoreseda*, Resedaceae). *Taxon*, **59,** 466–482.

McGuire, A.F. & Kron, K.A. (2005) Phylogenetic relationships of European and African ericas. *International Journal of Plant Sciences*, **166,** 311–318.

Médail, F. & Diadema, K. (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36,** 1333–1345.

Mesleard, F. & Lepart, J. (1991) Germination and seedling dynamics of *Arbutus unedo* and *Erica arborea* on Corsica. *Journal of Vegetation Science*, **2,** 155–164.

Messerli, B. & Winiger, M. (1992) Climate, environmental change, and resources of the African mountains from the Mediterranean to the equator. *Mountain Research and Development*, **12,** 315–336.

Mies, B.A. (1995) On the comparison of the flora and vegetation of the island groups of Socotra and Macaronesia. *Boletim do Museu Municipal do Funchal*, **4,** 455–471.

Mies, B.A. (1998) The phytogeography of Socotra: evidence for disjunctive taxa, especially with Macaronesia. *Conservation and sustainable use of biodiversity of Socotra Archipelago* (ed.

by H.J. Dumont), pp. 83–105. United Nations Development Programme, New York.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403,** 853–858.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L. & Swofford, D.L. (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, **24,** 581–583.

Ojeda, F., Arroyo, J. & Marañón, T. (1998) The phytogeography of European and Mediterranean heath species (Ericoideae, Ericaceae): a quantitative analysis. *Journal of Biogeography*, **25,** 165–178.

Ojeda, F., Arroyo, J. & Marañón, T. (2008) Ecological distribution of four co-occurring Mediterranean heath species. *Ecography*, **23,** 148–159.

Olmstead, R.G. & Palmer, J.D. (1997) Implications for the phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Systematic Botany*, **22,** 19–29.

Pagel, M. (1999) The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology*, **48,** 612–622.

Pagel, M. & Meade, A. (2004) A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Systematic Biology*, **53,** 571–581.

Palmé, A.E. & Vendramin, G.G. (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology*, **11,** 1769–1779.

Palmé, A.E., Su, Q., Rautenberg, A., Manni, F. & Lascoux, M. (2003) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology*, **12,** 201–212.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Muller-Starck, G., Demesure-Musch, B., Palme, A., Martin, J.P., Rendell, S. & Vendramin, G.G. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300,** 1563–1565.

Pichi-Sermolli, R. & Heiniger, H. (1953) Adumbratio Florae Aethiopicae II – Ericaceae. *Webbia*, **9,** 9–48.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144,** 1237–1245.

Popp, M., Gizaw, A., Nemomissa, S., Suda, J. & Brochmann, C. (2008) Colonization and diversification in the African 'sky islands' by Eurasian *Lychnis* L. (Caryophyllaceae). *Journal of Biogeography*, **35,** 1016–1029.

Posada, D. & Buckley, T.R. (2004) Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53,** 793–808.

Prentice, I.C., Jolly, D. & BIOME 6000 participants (2000) Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biogeography*, **27,** 507–519.

Quézel, P. (1978) Analysis of the flora of Mediterranean and Sahara Africa. *Annals of the Missouri Botanical Garden*, **65,** 479–534.

Rambaut, A. & Drummond, A.J. (2007) *Tracer version 1.4* (Computer program). Available at: http://beast.bio.ed.ac.uk/Tracer.

Ramos-Onsins, S. & Rozás, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19,** 2092–2100.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57,** 4–14.

Rodríguez-Sánchez, F. & Arroyo, J. (2008) Reconstructing the demise of Tethyan plants: climate-driven range dynamics of *Laurus* since the Pliocene. *Global Ecology and Biogeography*, **17,** 685–695.

Rodríguez-Sánchez, F., Guzmán, B., Valido, A., Vargas, P. & Arroyo, J. (2009) Late Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. *Journal of Biogeography*, **36,** 1270–1281.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19,** 1572–1574.

Rozás, J. (2009) DNA sequence polymorphism analysis using DnaSP. *Methods in Molecular Biology*, **537,** 337.

Rozás, J. & Rozás, R. (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15,** 174–175.

Sanmartín, I., van der Mark, P. & Ronquist, F. (2008) Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. *Journal of Biogeography*, **35,** 428–449.

Schneider, I. (1998) Thermal cycler innovations: expanded applications and increased performance. *Genetic Engineering News*, **18,** 24.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin, version 2.0: a software for population genetic data analysis*. University of Geneva, Genetics and Biometry Laboratory, Geneva.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W.S., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E. & Small, R.L. (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.

Slatkin, M. & Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129,** 555–562.

Suc, J.P. (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, **307,** 429–432.

Sunding, P. (1979) Origin of the Macaronesian flora. *Plants and islands* (ed. by D. Bramwell), pp. 13–48. Academic Press, London.

Svenning, J.C. (2003) Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters*, **6,** 646–653.

Swofford, D.L. (2002) *PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10*. Sinauer Associates, Sunderland, MA.

Szövényi, P., Hock, Z., Urmi, E. & Schneller, J.J. (2006) Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *Sphagnum squarrosum* (Bryophyta, Sphagnopsida) in Europe. *New Phytologist*, **172,** 784–794.

Tajima, F. & Nei, M. (1984) Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution*, **1,** 269–285.

Valcárcel, V., Fiz, O. & Vargas, P. (2003) Chloroplast and nuclear evidence for multiple origins of polyploids and diploids of *Hedera* (Araliaceae) in the Mediterranean basin. *Molecular Phylogenetics and Evolution*, **27,** 1–20.

Valente, L.M., Savolainen, V. & Vargas, P. (2010) Unparalleled rates of species diversification in Europe. *Proceedings of the Royal Society B: Biological Sciences*, **277,** 1489–1496.

Van Der Burgh, J. (1987) Miocene floras in the lower Rhenish Basin and their ecological interpretation. *Review of Palaeobotany and Palynology*, **52,** 299–366.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography: ecology, evolution, and conservation*, 2nd edn. Oxford University Press, Oxford.

Zhang, J.Z. (1999) Performance of likelihood ratio tests of evolutionary hypotheses under inadequate substitution models. *Molecular Biology and Evolution*, **16,** 868–875.

SUPPORTING INFORMATION

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

Appendix S1 Voucher information and GenBank accession numbers for *atp*B–*rbc*L, *mat*K, *trn*H–*psb*A and *rpl*16 sequenced for 105 specimens of *Erica arborea* sampled across its entire distribution range, and for *atp*B–*rbc*L and *mat*K sequenced for 13 specimens of *E. scoparia* and a specimen of *E. lusitanica* used in the tree presented in Appendix S2.

Appendix S2 Fifty per cent majority-rule consensus tree resulting from the Bayesian analysis of *atp*B–*rbc*L and *mat*K gene sequenced for a sample of 24 *Erica* species.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Aurélie Désamoré and **Benjamin Laenen** are PhD students at the University of Liège, Belgium. Aurélie Désamoré is interested in plant biogeography and population genetics, and is currently working on biogeographical patterns in European bryophytes. Benjamin Laenen currently focuses on the influence of mating systems on the evolutionary history of bryophytes.

Author contributions: A.V. and M.C. conceived the project. M.P. and J.M.G.-M. provided material and assisted with the manuscript. A.D., B.L., N.D. and A.V. collected and analysed the data. A.D., B.L., A.V. and M.C. wrote the manuscript. A.D. and B.L. contributed equally to this paper as senior authors, while M.C. and A.V. contributed equally as junior authors.

Editor: José María Fernández-Palacios