Faculty of Sciences
Department of Environmental Sciences and Management
Unit of Plant Taxonomy and Conservation Biology

(E. Sérusiaux)

From Engler's relictualism model to Wallace's dynamic interchange model: Genetic structure and gene flow between Macaronesian islands and the neighboring continents in the heathers *Erica arborea* and *Erica scoparia*

Final work defended by Aurélie Désamoré
In order to get the diploma of Master in Biology
Organism Biology and Ecology

Supervisors: A. Vanderpoorten (ULG) and M. Carine (NHM)

Photo: A. Désamoré

Acknowledgment

I want to thank people who helped me all along this work.

Particularly, my supervisor Alain Vanderpoorten:
- For sharing his experiences and for his help with analyses and redaction.
- For fieldwork sampling experience in the Azores and the work experience in Duke University
- For the time spent in help all along the realization of this work.

Mark Carine for his help in collection of sampling, for his welcome in London and his continuous valuable help all along the realization of this work.

John Shaw for his welcome in his laboratory in Duke University, and also Nicolas Devos who spent time in help for laboratory and analyses, but also for accommodations in Duke.

People in Duke, especially Sandy Blake Boles for her help in the Lab, Peter Szövényi for his helpful advises but also B. Shaw, M. Ricca, L. Pokorny, L. Karst, M. Johnson for their welcome and A. Merced-Alejandro, E. Cooper, and P. Sova for their friendship.

Laurent Gohy for his help in the lab in Liège and Mr. Lambinon for his helpful advices.

To all people that provided me herbarium materials.

The Leopold III fund for financial help in the Azores mission.

Finally, my family and especially Benjamin who was always present when uncertainty appears, and helped me all along this work.

Résumé

L'intérêt des îles en biogéographie et en biologie évolutive, lié à leur isolement géographique et leur dynamisme écologique associé à leur activité volcanique, est bien connu depuis Darwin et son étude sur les pinsons des Galápagos. La Macaronésie, qui peut être définie comme un concept biogéographique regroupant quatre archipels de l'Atlantique Nord (les Açores, les Canaries, Madère et les Selvagems) sur base de leurs spectaculaires taux d'endémisme, offre un modèle particulièrement intéressant car, contrairement aux îles Pacifiques, les îles macaronésiennes présentent une histoire volcanique plus complexe et sont plus proches du continent.

Historiquement, l'origine de l'endémisme macaronésien a été attribuée à la théorie relictualiste d'Engler, selon laquelle, la flore macaronésienne serait une relicte d'une flore Tertiaire étendue en Europe et en Afrique du Nord et qui aurait trouvé refuge en Macaronésie suite à la désertification du Sahara et aux évènements glaciaires du Quaternaire. C'est cette théorie qui est testée sur deux espèces prépondérantes de la laurisylve endémique de Macaronésie, *Erica arborea* et *Erica scoparia.*

Erica arborea est un candidat idéal pour tester l'hypothèse selon laquelle la flore macaronésienne peut être regardée comme un « fossile vivant » pour deux raisons majeures. Premièrement, cette espèce présente une distribution disjointe en Macaronésie, en Europe et sur des sommets d'Afrique centrale et orientale suggérant fortement un événement de vicariance ancienne. Deuxièmement, et au contraire de nombreux cas de radiations spectaculaires observés chez de nombreux autres éléments de la flore macaronésienne, cette espèce ne s'est pas diversifiée localement, selon un processus qui rappelle les extinctions massives caractéristiques des lignées des exemples typiques de fossiles vivants. La comparaison avec *Erica scoparia* est intéressante car cette espèce présente une distribution similaire bien qu'on ne la retrouve pas sur le continent Africain, de plus celle-ci présente une apparente diversification morphologique en sous-espèces sur les îles macaronésiennes.

Au cours de la présente étude, les patterns phylogéographiques de ces deux espèces ont été contrastés pour répondre à ces questions principales :

1) Pourquoi *Erica arborea* n'a pas divergé et ne s'est pas diversifiée sur les îles contrairement à la plupart des angiospermes insulaires ? *E. arborea* et *E. scoparia* peuvent–elle réellement être interprétées comme fossiles Tertiaires ?

2) Comment peut-on interpréter le pattern de distribution disjointe caractéristique d'*E. arborea* ? Est-ce dû à la contraction d'une aire ancestrale ou à une récente colonisation suivie d'une expansion ? Les deux espèces présentent-elles la même histoire biogéographique ?

Pour répondre à ces questions, 105 spécimens d'*E. arborea* et 31 d'*E. scoparia* ont été respectivement échantillonnés de manière à couvrir l'entièreté de leur aire de distribution. Chaque échantillon a été génotypé sur quatre marqueurs moléculaires

chloroplastiques, l'espaceur *atp*B-*rbc*L, les régions *trn*H-*psb*A, m*at*K et le gène *rpl*16. Les données ont été analysées en partitionnant la variance des fréquences alléliques dans et entre les populations à l'aide de statistiques F et N, d'analyses phylogénétiques et d'analyses en réseau, et d'une analyse de dispersion/vicariance (DIVA).

Un résultat totalement inattendu est apparu à l'issue du génotypage de certains individus qui présentent, dans leur génome chloroplastique, une copie allélique d'*E. arborea* et une seconde d'*E. scoparia*. Cette observation suggère que l'hétéroplasmie observée chez ces individus est d'origine hybridogène. Elle réfute l'hypothèse traditionnellement admise d'un génome chloroplastique dans lequel les gènes seraient présents en une seule copie d'origine maternelle. Par ailleurs, la présence des deux copies parentales chez trois gènes, mais d'une unique copie issue d'*E. scoparia* dans le gène *rpl*16, suggère que le génome chloroplastique peut recombiner. Ces deux observations battent en brèche deux hypothèses solidement implantées sur les caractéristiques du génome chloroplastique et ont des conséquences majeures pour l'analyse des séquences d'ADN chloroplastique qui sont discutées.

D'un point de vue phylogéographique, les résultats obtenus permettent de voir que, contrairement à l'hypothèse relictualiste d'Engler, *E. arborea* et *E. scoparia* seraient arrivées récemment en Macaronésie à partir de l'Europe. *Erica scoparia* présente une grande diversité sur les Açores, tandis qu'*Erica arborea* ne s'est pas du tout diversifiée en Macaronésie. L'hypothèse la plus plausible de cette non-diversification est une colonisation extrêmement récente de l'espèce. En effet, la reconstruction des aires d'origine de l'espèce et de son patron de recolonisation post-glaciaire suggère que l'espèce a survécu au Quaternaire dans deux refuges au Nord de la Péninsule Ibérique et en Afrique de l'Est. La recolonisation de l'Europe s'est opérée à partir du refuge africain via l'Asie du Sud-Ouest, puis selon une progression d'est en ouest. L'Ouest de l'Afrique du Nord et la Macaronésie sont donc les deux dernières régions touchées par cette vague de recolonisation, ce qui explique que la diversité génétique de l'espèce dans ces deux régions soit minimale. Les résultats suggèrent donc que des phénomènes récents de dispersion puissent imiter un patron de distribution a priori compatible avec les attendus d'une hypothèse relictualiste.

Il apparaît ainsi qu'*E. arborea* et *E. scoparia* sont en pleine expansion et montrent même en Macaronésie le syndrome propre aux espèces invasives, à savoir une diversité génétique très basse et une rapidité d'expansion favorisée par la dégradation des habitats. La laurisylve macaronésienne apparaît dès lors comme une formation composée de relictes Tertiaires, mais également d'un ensemble d'espèces d'origine beaucoup plus récente.

Table of content

I. Introduction

Biogeography is the science that aims at describing the spatial distributions of biota (a pattern) and understanding the means by which these distributions were achieved (a process). Biogeography is a field that existed long before evolutionary biology and indeed helped in founding the evolutionary ideas of Charles Darwin and Alfred Wallace, among others (Humphries and Parenti, 1999). Biogeography and evolutionary biology are therefore at the interface of each other, as the discovery of the mechanisms regulating species distributions involves an understanding of species dispersal ability, evolutionary rates, and diversification mode, which are among the main foci of the sciences of evolution.

Many plant and animal taxa exhibit disjunct distributions such as the Magnolias found in south-east Asia and in America, the springtail *Tetracanthella arctica*, which is common in Iceland and Spitzbergen but is also found in Greenland, Arctic Canada, the Pyrenean mountains (France and Spain) and in the Tatra mountains (Poland, Czech republic) or even the Gorilla found in tropical rainforest in West Africa and also in mountains in East Africa (Cox and Moore, 2005). Two main hypotheses, namely ancient vicariance owing to continental drift and dispersal, were the subject of major debates to explain those distributions. Even today, the debate is not over (McDowall, 2004, de Queiroz, 2005, McGlone, 2005, Waters and Craw, 2006).

Darwin and Wallace did not know the existence of continental drift and hypothesized that disjunct distributions were explained by dispersion from a "center of origin" (Croizat, 1974). This notion is attributed to Darwin but this idea was already accepted by Linné and Buffon (Humphries and Parenti, 1999, Lomolino M.V. et al., 2006). With the theory of continental drift and cladistic studies, vicariant hypotheses emerged, notably with Croizat (Croizat, 1958). Most patterns were then attributed to the disjunction of a formerly continuous range because of the emergence of a barrier, as for example the Wallace line and the Austral domain (Pangea fracturation) (Cecca, 2008). At that time indeed, dispersal was considered to generate random patterns that were not compatible with the regular patterns observed among many taxa and perfectly congruent with the expectations of a continental drift scenario (Croizat, 1974). Because most Gondwanan groups were presumed to be poor dispersers, long-distance dispersal across oceanic barriers was discarded as a valid explanation for the disjunct patterns observed. Moreover, the dispersal hypothesis was considered unfalsifiable and unscientific, (Nelson and Platnick, 1981, Morrone and Crisci, 1995, Craw et al., 1999, Humphries, 2000). If dispersal was conceded for oceanic islands that had never been connected to other landmasses, vicariance was assumed as the more probable explanation for all the other cases of disjunct distribution observed and came to dominate the field of historical biogeography (Wiley, 1988, Morrone and Crisci, 1995).

However, a reappraisal of the dispersal theory has recently gained much attention with the emergence of molecular dating (Renner, 2005, de Queiroz, 2005, Milne, 2006). Although some disjunct distributions can be unambiguously interpreted in terms of ancient vicariance and massive extinctions based on fossil evidence (Tiffney, 1985,

Milne and Abbott, 2002, Xiang et al., 2005), patterns that are spatially congruent with such a hypothesis may, in fact, conceal a complex mixture of relictual distributions and more recent speciation and dispersal events (de Queiroz, 2005, Vanderpoorten et al., 2007) and we are now, in fact, in the middle of a dispersalist counter-revolution (de Queiroz, 2005).

1.1. Why focuses on islands?

The tremendous interest in islands as natural laboratories has long been acknowledged since Darwin's journey around the world and study of the diversification of Galapagos Finches (Emerson and Kolm, 2005). This interest comes from several important features of island ecosystems that make them ideal models for the study of speciation and diversification (Emerson, 2002).

Volcanic oceanic islands are geographical entities that are well defined and separated by oceanic barriers that inhibit gene flow with continental areas. In addition, they are characterized by intense volcanic or erosional activities that permit rapid ecological changes, thereby reducing competition and enhancing the arrival or evolution of new species. Islands also show considerable habitat diversity produced by their topology and humidity gradients (Juan *et al.*, 2000). As a result, oceanic islands typically exhibit high rates of species endemism following spectacular radiations. For example, the Hawaiian silversword alliance has been regarded as the best example of adaptive radiation in plants (Juan et al., 2000, Raven et al., 1992), the Canary Island beetle *Nesotes* form one of the more diverse coleopteran genera in the archipelago (Rees *et al.*, 2001), the *Anolis* lizards include 140 of the 400 species occurring in the Caribbean fauna (Jackman *et al.*, 1999) and the *Drosophila* of the Hawaiian Islands are represented by more than 800 species that have evolved after an initial colonization more than 30 Ma (Kambysellis *et al.*, 1996).

Islands are, consequently, often used as models for the study of evolution, especially the Hawai'i (Wagner and Funk, 1995), the Galapagos (Schilling and Panero, 2002), the Juan Fernández Islands (Sang *et al.*, 1994) and the Canary Islands (Francisco-Ortega *et al.*, 1996).

1.2. The Macaronesian model

The term Macaronesia was used for the first time by the botanist Philip Barker Webb (Sunding, 1979). This term was then adopted by biogeographers such as Engler leading to its acceptance among biologists in general, to circumscribe a biogeographic unit composed of the Azores, Madeira, Selvagems and the Canary islands (Engler, 1879). The Cape Verde was later added to the region by some authors (Bramwell, 1972, Takhtajan, 1969), and in some cases, continental enclaves areas of North Africa and Iberia where species with Macaronesian affinity occur (Sunding, 1979). Macaronesia is situated in the northeast Atlantic ocean and the Canaries islands are situated between 27°37' and 29°25' N and 13°20' and 18°10'W (Juan *et al.*, 2000) ***(Fig. 1)***.

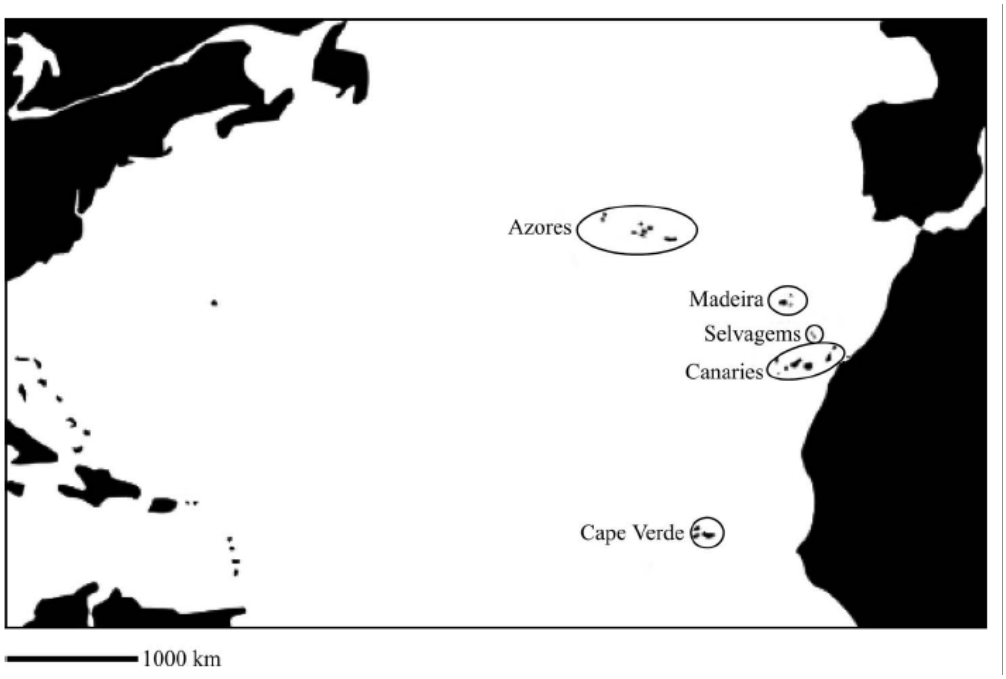


Fig. 1: The Macaronesian archipelagos. Adaptated from (Vanderpoorten et al., 2007)

The islands of the region exhibit considerable variation in age, altitude, climate, ecology and floristic composition (Fernández-Palacios and Dias, 2001). Despite such a great ecological heterogeneity, the Azores, Canary Islands, Madeira, and Cape Verde share a remarkable endemic element that constitutes the basis of Engler's theory. Indeed, in Macaronesia, 20% of angiosperm species are endemic (Gonzáles Martín, 2001., Humphries, 1979., Santos-Guerra, 1999, Shäffer, 2003) and almost 40% in the Canaries (Santos-Guerra, 1999). This high degree of endemism explains why the region is one of the richest hot-spots within the Mediterranean floristic region and one of the most important in terms of conservation.

Macaronesia, and especially the Canary Islands, presents some differences compared to Pacific islands (for example: Hawaii, Galapagos), such as the much more complex volcanic evolution for the Canaries (Juan *et al.*, 2000). However, the main difference is that the Canaries are close to continental areas, 95 km only from North Africa and probably 65 km in the past 20 million years. Moreover, volcanic sea mounts, some of which are 68 Ma (Geldmacher et al., 2001) are located between the islands and the continent and may have facilitated dispersal between the continent and these islands, particularly when sea levels were lower (García-Talavera, 1997). The Canary Islands were formed by volcanic activity that began during Oligocene but the islands began to emerge during the past 21 Ma (20 Ma for Fuerteventura and decrease from east to west to less than 1 Ma for El Hierro) (Juan *et al.*, 2000). Thus, the islands are very recent and represent a good model for the study of adaptive radiation and evolution of endemism.

1.3. From a pattern to a process: two predominant theories underlying the evolutionary origin of Macaronesia

1.3.1. The Engler refugium model

Engler proposed that the Macaronesian flora was, for the most part, a relict of a widespread subtropical flora that covered Southern Europe and Northern Africa during the Tertiary and disappeared because of the quaternary glaciations and the extension of the Sahara desert. This hypothesis was later supported by several other authors including Takhtajan, Bramwell and Cronk (Bramwell, 1972, Cronk, 1992, Takhtajan, 1969). Engler pointed out the existence of two types of endemism: one based on the preservation of ancient forms and the other based on the development of new entirely autochthonous forms. This is, in fact, the distinction between paleoendemics versus neoendemics but, as these terms may be misleading, Cronk (1987, Cronk, 1992) introduced the 'Relictual Series Hypothesis' to describe relicts of different ages from Miocene to Pliocene, due to phases of vegetation changes. The relict theory involves the reduction of range and species extinction, partly as a consequence of climatic changes, as evidenced by the high number of monotypic genera (Cronk, 1992).

1.3.2. The Wallace dynamic interchange model

The idea developed by Wallace (1881) focused on the interchange of species between continental and island areas. This idea was later followed by Matthew, MacArthur and Wilson, Whittaker for example (Whittaker et al., 2008, Wilson, 1969, Mac Arthur and Wilson, 2001). The theory proposed by Wallace focused particularly on islands, whereas the dispersal theory more generally suggests that a species expands and so fragments its range by accidental crossing of a barrier leading to immediate isolation and disjunction (Platnick, 1981). ***(Fig. 2)***

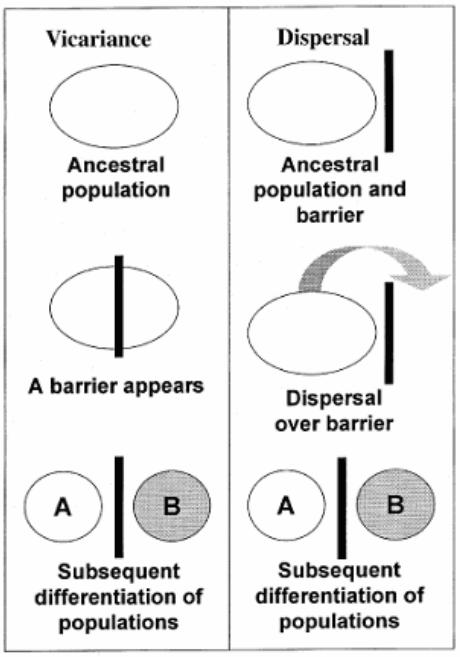


Fig. 2: The two historical explanations (ancient vicariance and dispersion) of the disjunct distribution of two taxa (Crisci, 2001)

Scientists have tended to follow one another of these two apparently dissimilar models. However, the recent development of more and more robust and temporally calibrated phylogenies coupled with more reliable geological reconstructions of oceanic regions, allow a new conceptual paradigm to emerge (Heaney, 2007, Whittaker et al., 2008). This paradigm seems to give more and more importance to the dispersal theory but suggests that the two theories are acceptable with evolutionary histories differing from one plant group to another.

1.3.3. Application to Macaronesia

Taxa that are thought to conform to the Engler model, as paleoendemics that survived continental extinctions on the islands, are hypothesized to have retained a highly conserved morphology for millions years and fossils discovered in continental Europe have been considered conspecific with taxa from Macaronesia (Sunding, 1979, Frahm, 2004). Such data suggest that many groups now found in the laurel forests and sclerophyllous zones of Macaronesia were apparently lost from the flora of Europe (Saporta, 1865, Depape, 1922, Sunding, 1979) at the end of the Tertiary and in the Pleistocene due to climate change (Emerson, 2002). These fossil deposits contain many subtropical genera that are now present in Macaronesia but not in the Mediterranean region such as the genera *Clethra*, *Draceana*, *Ocotea* and *Persea*. However, paleontological records do not provide irrefutable evidence for the Engler refugium model because the fragmentary nature of the fossil record often makes it difficult to assign a fossil to an extant species with confidence (García-Talavera, 1995).

The woody habit of many Macaronesian plants has also been viewed as an ancestral trait that identifies them as a relict fauna. However, it has been shown for numerous shrubs that the arborescent habit is an insular derived trait, for example in the *Bencomia* alliance (Helfgott *et al.*, 2000) and the Sonchinae sub-family (Kim *et al.*, 1996).

Molecular phylogenetic analyses of Macaronesian plants have increasingly allowed a better understanding of the relationships of the Macaronesian endemic flora. Even though it appears that each group has a unique diversification pattern (Francisco-Ortega *et al.*, 2001), some general patterns exist. In the majority of group investigated, Macaronesian endemic taxa are monophyletic (Carine, 2005), which is consistent with the Engler theory. A single colonization has often been followed by a rapid radiation, as for example the spectacular radiations in the genera *Argyranthemum*, that radiated into all habitats of these Atlantic Islands (Juan *et al.*, 2000), *Aeonium* (Webb and Berthelot, 1840, Mort et al., 2002) and *Echium* (Bohle *et al.*, 1996). However, it is important to know that many genera have not radiated in Macaronesia and a study (Carine *et al.*, In press) showed that 25% of the angiosperm surveyed contain only a single taxon .

Multiple congeneric colonizations have also been demonstrated, for example, in the genus *Hedera L.* (Carine *et al.*, 2004) and *Lavatera* (Fuertes-Aguilar *et al.*, 2002). A back-colonization of the continent from Macaronesia is the most plausible explanation

for the pattern of relationships of continental and island taxa *Tolpis* (Mort *et al.*, 2003). One genus for which both multiple colonizations into Macaronesia and back colonization of the continent have been hypothetized is *Convolvulus*. (Carine et al., 2004, Carine, 2005), back colonization from islands to the continent (Cabo-espichel in Portugal) is also consistent for *Aeonium (Carine et al., In press).*

However, it seems that when multiple colonizations occur, radiation is precluded. Silvertown (Silvertown, 2004) showed that 20 genera that are each represented in Macaronesia by unique colonization events had given rise to 269 endemic species, whilst 20 repeated colonization events gave rise to only 38 species (Silvertown *et al.*, 2005). More phylogenetic research will permit a better understanding of the mechanisms of colonization and diversification of the region's flora.

1.4. The *Erica arborea/scoparia* model

Woody Erica (*E. arborea* and *E. scoparia*) are a member of the unique Macaronesian laurisylva forests. In contrast to most laurisylva species, however, Macaronesian *Erica* are not endemic but exhibit disjunct distribution ranges. *Erica arborea* is found in a wide range of habitats, in the Canary Islands and Madeira, in the Mediterranean region where it thrives in the maquis (Mesleard and Lepart, 1991), the Tibesti mountains (in Sahara) where it is present at the top of an upper montane desert steppe vegetation with *Artemisia sp*. around 2000 and 3000m and also on the East African mountains and southwest Asia where it dominates in an *Erica* belt vegetation between 3000 and 4000m with *Philippia sp*. and *Helichrysum sp.* principally. (McGuire and Kron, 2005, Quezel, 1978, Bruneau de Miré and Quezel, 1959, Pichi-Sermolli and Heiniger, 1953, Messerli and Winiger, 1992). *Erica scoparia* has a more restricted geographic distribution that overlaps that of *E. arborea*. It can be found in the Mediterranean region (going up to central France in the western, but no further than North Italy in the Eastern part), North Africa, Macaronesia, but not in East Africa, Southwest Asia or Tibesti. In addition, *E. scoparia* is also found on the Azores where *E. arborea* does not occur (Small, 2001). The similar distributions of *E. arborea* and *E. scoparia* permit the comparison of their biogeographical relationships. In addition, *E. scoparia* contains three subspecies: *E. scoparia* ssp. *platycodon*, *E. scoparia* ssp. *Azorica* and *E. scoparia* ssp. *maderincola* respectively found in Canary Islands (Tenerife, La Gomera, and El Hierro), the Azores and Madeira. In contrast to *E. arborea*, *E. scoparia* therefore shows some diversification on islands. This comparison between the two species will allow us to investigate why diversification has occurred in one species and not the other.

1.5. Goals of the thesis

In this project, we contrast the phylogeographic patterns of *E. arborea* and *E. scoparia* to address the following questions:

1) Why did *E. arborea*, as opposed to most of the members of the Macaronesian laurisylva, fail to diverge and diversify? Are the molecular data showing a genetic variation without morphological differentiation? Can *E. arborea* truly be interpreted as a living fossil? Or, to the reverse, are gene flow between island and continental populations so intense that they prevent genetic isolation and speciation?

2) What is the evolutionary origin of Erica's strikingly disjunct distribution? Is it the result of contraction from a large ancestral area or recent recolonization and range expansion?

In the course of this phylogeographic study, evidence for heteroplasmy, hybridization and recombination within the chloroplast genome were encountered, and these results are presented in the Appendix section.

II. Materiel and methods

2.1. Biological models

The genus *Erica* is composed of nearly 800 species found in a narrow north-south distribution 30°W to 45°E of the Greenwich meridian ranging from the warm wet conditions of western Europe to the dry hot conditions of southern Africa. About 750 species are confined to the south of the Limpopo River in South Africa. They form evergreen shrubs with narrow folded leaves and flowers with 3 bracts, 4 sepals and a corolla. *Erica* does not occur naturally on the American continent, the vast majority of Asia, nor in Australasia but has been naturalized in parts of Australia and New Zealand where it is now viewed as a major weed (Small, 2001). The name *Erica* derives from the Greek word ερεικω meaning to break. It possibly derives from the medieval theory that the plant could dissolve gallstones; alternatively, it may refer to the fact that the stems of some species are easily broken.(Couplan, 2000)

2.1.1. Erica arborea

Fig. 3: Erica arborea flowers. From http://commons.wikimedia.org/wiki/File:Erica_arborea_flors.jpg

Erica arborea is a tall shrub or small tree, usually up to 5 m but sometimes more, often with a well-defined trunk. Young shoots are pubescent with whitish, simple and branched hairs. The leaves have 3 (-4) mm, in whorls of 3 (-4). Flowers are often in a broad panicle. The bracteoles are 2-3, in the lower part of the pedicel. The calyx is saccate at the base. The corolla is 2-2,5 mm, broadly campanulate and white. The anthers have basal appendages. The stigma is prominently exserted (Press, 1994). ***(Fig. 3)*** *E. arborea* presents tetrads of pollen smaller than 40 µm. These small tetrads facilitate pollen dispersal by wind (Rodriguez-Rajo *et al.*, 2005).

The distribution area of *E. arborea* encompasses Madeira and the Canary Islands, the Mediterranean region (from Morocco to Tunisia in the South and from the Iberian

Peninsula to Turkey in the North) and Central Africa (Tibesti mountains), East Africa and South-west Asia ***(Fig. 4)***.

Fig. 4: Distribution of Erica arborea (From: http://www.heathersociety.org.uk/heather_family.html)

2.1.2. Erica scoparia

Fig. 5-6-7: Erica scoparia flowers; Erica scoparia subsp. platycodon; Erica scoparia subsp. azorica
http://botanicavirtual.udl.es/fam/Ericacies/Erica_for.htm#Erica_scoparia
http://www.planetefleurs.fr/Systematique/Ericaceae/Erica_platycodon.htm
http://www.horta.uac.pt/species/plantae/Erica_scoparia_azorica/Erica_scoparia_azorica.htm

Erica scoparia subsp. *scoparia* is a rather untidy bush reaching 2 m, with tiny brownish-green flowers in late spring and early summer, which produce clouds of pollen. This subspecies is found in West Mediterranean from Portugal to North of Italy and growing to middle France and North Africa ***(Fig. 8)***.

Erica scoparia subsp. *platycodon* (Webb& Berthel.) ***(Fig. 6)*** is a tall shrub up to 4 m with stems up to 20 cm or more in diameter. The young twigs are glabrescent. The leaves are 10-12mm, spreading, in whorls of 3-4. The bracteoles are 2, near the middle of the pedicel. The calyx is not saccate. The corolla is 2-3 mm, broadly campanulate and greenish-pink. The anthers are without appendages. The flowering occurs between April and June (Press, 1994). This subspecies is endemic to the Canary Islands only on El Hierro, La Gomera and Tenerife (Banco de datos de Biodiversidad de Canarias, 2004).

Erica scoparia subsp. *maderincola* (D.C Mc clint) is only found in Madeira and Porto Santo. It closely resembles ssp. *platycodon* but has a laxer habit, longer leaves even more spaced and spreading when mature, a shorter calyx and corolla with the stigma more exserted (Small, 2001).

Erica scoparia subsp. *azorica* (Hochst) ***(Fig. 7)*** is an evergreen shrub or small tree up to 6 m. The leaves are simple, needle-like. The flowers are in terminal, lax panicles. The petals are green, often tinged with purple, connate, forming a campanulate corolla of 1,5-1,75 mm. The stigma is exserted. The fruit is a dry capsule. The flowering occurs between April and June. This subspecies is endemic to the Azores. It occurs and is abundant on all islands, on coastal cliffs, dry slopes, young lava flows and in hedges up to 1500 m (Schäfer, 2002).

Fig. 8: Distribution of Erica scoparia (From: http://www.heathersociety.org.uk/heather_family.html)

2.2. Population sampling

387 samples (274 of *E. arborea* and 109 of *E. scoparia*) were collected from herbarium material, silica-gel dried leaves or fresh material collected during a field trip in the Azores. Five islands from the Oriental (Sao Miguel), Central (Faial, Pico and Terceira) and Occidental (Flores) part were visited ***(Fig. 9)***.

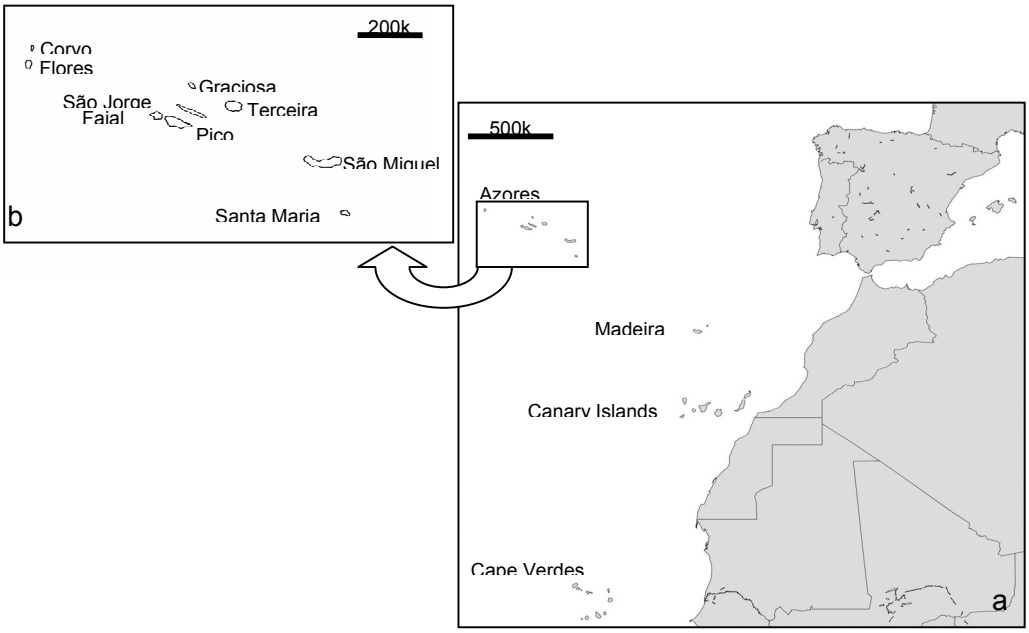


Fig. 9: The Azorean archipelago (From M. Carine: Personal communication)

23 specimens in total of *Erica scoparia* subsp. *azorica* were collected and placed in silica gel for conservation. In Sao Miguel, 13 specimens were collected above 600m in a chamaephytic vegetation dominated by *Erica azorica, Calluna vulgaris* and *Juniperus brevifolia* in three different parts of the island (Pico de Vara, Caldera Seca, Vale de Las Lambadas and Lagoa do Fogo) ***(Fig. 10)***. In most parts of the island, the original laurisylva was replaced by the invasive species *Pittosporum undulatum*, the Australian shrubby fern *Sphaeropteris cooperi* and the Asiatic Zingiberaceae *Edychium gardneranum* ***(Fig. 11)***

Fig. 10: Heathland dominated by Erica azorica at Lagoa do Fogo on the island of Sao Miguel. (photo: A. Désamoré).

In Faial, four specimens were collected in four different parts of the island. One specimen was collected on the island of Pico.

In Flores, the coastal cliff vegetation was dominated by *Erica azorica* and *Myrsine africana.* Three specimens were collected along the Northwest coast in Faja Grande. A further 2 specimens were also collected along the Southwest coast between Mosteiro and Lajedo ***(Fig. 12)***. In Terceira, three specimens were collected in a laurisylva dominated by *Laurus azorica* and *Juniperus brevifolia* (Serrata) and in a coastal moor (Aguavalva) ***(Fig. 13)***.

Fig. 11: Original laurisylva degraded by invasive species, such as the Australian shrubby fern Sphaeropteris cooperi and by the Asiatic Zingiberaceae Edychium gardneranum, on the island of Sao Miguel in the Azores.
(Photo: A. Désamoré)

Fig. 12: Erica azorica and Juniperus brevifolia on the island of Flores; Fig.13: Erica azorica dominating the vegetation on the island of Terceira. (Photos: A. Désamoré)

2.3. Molecular protocols

2.3.1 Extraction

About 10 leaves were taken on one branch from each individual and collected in Eppendorf tubes for extraction. The material was ground to powder in liquid nitrogen with genogrinder 2000. The extraction was performed using a CTAB protocol (Doyle and Doyle, 1987), followed by a purification with Geneclean kit following the manufacturer's instructions.

2.3.2. Loci selection

Since DNA sequencing technology has become available, many comparative studies on cpDNA have been performed (Ritland, 1987, Zurawski and Clegg, 1987). Coding regions were first explored to address phylogenetic issues at high taxonomic levels. Subsequently, the focus has been on non-coding regions, which show less functional constraints than coding regions and provide more variations for phylogenetic analyses, and hence, are more appropriate for low-level taxonomic studies (Gielly and Taberlet, 1994). As a consequence, universal primers to amplify and sequence non-coding cpDNA regions were designed and have successfully been used in plant phylogeographies. I have first selected two cpDNA regions and one from the nuclear genome: the *atp*B-*rbc*L spacer, the 5' half of the chloroplast *mat*K gene and the nrITS, because all three regions were used in a phylogenetic study of the genus *Erica* (McGuire and Kron, 2005). In addition, *Trn*H-*psb*A, the *Trn*S*gcu*-*Trn*G*uuc* spacer and the *Trn*G intron, *Rpl*16 and *psb*B-*psb*H were screened among the 21 noncoding cpDNA regions described by Shaw et al (Shaw *et al.*, 2005).

The nuclear region was abandoned because of the poor quality of the sequence chromatograms. In fact, it has long been recognized that ITS are subject to a phenomenon known as concerted evolution. Concerted evolution occurs when sequence differences among reiterated copies in the genome become homogenized by mechanisms such as high-frequency unequal crossing over or gene conversion. However, mounting evidence suggests that concerted evolution is much less efficient than previously thought and issues of paralogy have increasingly been reported, thereby questioning the actual phylogenetic utility of the ITS, even though this marker is still often used at low-taxonomic level (Alvarez and Wendel, 2003, Bailey et al., 2003).

The *Trn*S*gcu*-*Trn*G*uuc* spacer and the *Trn*G intron were also abandoned because of difficulties of amplification. Finally, *psb*B-*psb*H was not employed because of the lack of polymorphism among sequences. As a result, the four regions *atp*B-*rbc*L spacer, *Trn*H-*Psb*A, *Mat*K and *rpl*16 loci were used for this phylogeographic study. These regions were first amplified using universal primers (Shaw et al., 2005, McGuire and Kron, 2005). Due to the difficulties encountered with the amplification of 3 loci for several accessions, a specific set of primers was designed within conserved regions at the 5' and 3' ends of the molecules. Primers to amplify the *atp*B-*rbc*L spacer were designed from sequences of *Erica arborea* available on Genbank (McGuire and Kron, 2005) (see

Table I in annex). These primers were designed using the online software available at http://www.sigma-genosys.com/calc/DNACalc.asp in order to check for the TM (temperature from which half of the primers are melted), and avoid strong secondary structure that would favor the formation of homodimers and avoid the formation of heterodimers between the forward and reverse primers.

2.3.3. Polymerase chain reaction (PCR) and sequencing

PCR was carried out in a 15µl volume reaction using 1.5µl of 10X reaction buffer, 2.4µl of dNTPs mix (1mM each), 0.6µl of 50mM $MgCl_2$, 0.75 of each primer (10µM), 1.125µl of BSA, 0.3µl of taq DNA polymerase and 1µl of DNA.

The reaction was programmed on a thermocycler with 2 min of denaturation at 95°, followed by 35 cycles of 30s denaturation at 95°, 45s of annealing at 50°, 2 min of extension at 72° and finally 7 min at 72°. This protocol was used for all genes with some modifications depending on the age of samples and amplification easiness. The annealing temperature has been increased to 52° or 54° when smears where present on the gel or diminished to 50° when no band appeared on the gel. The presence of DNA fragment amplified was verified visually on agarose gel ***(Fig. 14)*** by staining with either ethidium bromide, gel green or sybersafe. PCR products were purified with exosap-it mix. The sequencing reaction employed big dye (principally containing Taq polymerase and dideoxynucleotides) and involved 2 min at 96°, 25 cycles of 15s at 96°, 10s at 50°, 4min at 60°. The labeled fragments were then separated by capillary electrophoresis on an ABI-3100 sequencing machine.

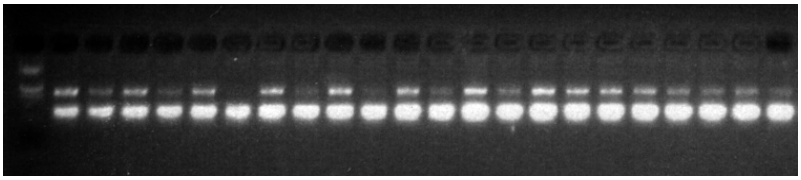


Fig. 14: Photo of an agarose gel for MatK in 22 samples of Erica arborea. Photo taken under UV light to see the presence of amplified segments stained by Ethidium bromide.

2.4. Data analyses

2.4.1. Sequence alignment

The sequences were aligned and verified with Sequencher 3.1 (Schneider, 1998). All polymorphic and monomorphic sites were verified visually. The sequences available for only one gene were deleted for all the analyses. Gaps were inserted where necessary to preserve positional homology in the alignment in Mac Clade 4.0 (Maddison and Maddison, 1989). The four genes where then concatenated in a matrix resulting in 136 samples.

2.4.2. Data description

Individuals exhibiting the same sequence across the 4 loci were grouped within the same haplotype. Those haplotypes were determined from the sequence matrix with the DNAsp program (Rozas and Rozas, 1999, Rozas, 2009). Each individual was assigned to one haplotype. Then, a geographic haplotype map was produced to visualize the haplotype's distribution areas and the occurrence of each haplotype within each of the following geographic regions, namely: 1) Macaronesia (Canary Islands and Madeira), 2) the Iberian Peninsula (Portugal and Spain, including the islands of Mallorca, Menorca and Ibiza), 3) the central and eastern Mediterranean (from France to Turkey comprising Tibesti and also Corsica and Sardinia), 4) North Africa (Morocco, Algeria, Tunisia), 5) Southwest Asia (Arabia and Yemen) and 6) East Africa (Kenya, Ethiopia, Rwanda,Uganda).

For *E. scoparia*, the occurrence of the haplotypes was recorded in each of the following geographic regions: 1) the Azores, 2) Madeira, 3) Canary Islands, 4) Central Mediterranean (France, Corsica, Italy), and 5) the Iberian Peninsula and North Africa (including Portugal, Spain, Algeria, Tunisia, Morocco), which were merged owing to the presence of few samples in North Africa and only one haplotype corresponding to the same present in Iberian Peninsula.

Genetic diversity H was calculated using the Arlequin software for population genetics Version 2.0 (Excoffier *et al.*, 2005). Gene diversity is the probability that two randomly chosen haplotypes are different in the sampling.

$$H = \frac{n}{n-1}(1-\sum_{i=1}^{k} pi^2)$$

n = number of gene copies in the region
k = number of haplotypes
pi = haplotype frequency

Gene diversity is comprised between 0 and 1. This value is close to 0 when there are few haplotype or when one is dominant and the others very rare. The value is close to 1 when there are many haplotypes with nearly identical frequencies. This index permits to see which regions present a high degree of genetic diversity taking haplotype frequencies into account.

2.4.3. Geographic differentiation and phylogeographic signal

Analysis of molecular variance (AMOVA) and Fst were used to measure the genetic structure in the data, determine the amount of genetic variation within and among geographic regions and, indirectly, measure gene flow among regions. These statistics were computed using Arlequin Version 2.0 (Excoffier et al., 2005, Schneider et al., 2000).

Fst is a measure of the variance of haplotype frequencies among regions standardized by the total variance of haplotype frequencies. When all the variation in haplotype frequencies lies among populations Fst=1, which means that population differentiation is maximum. By contrast, an Fst of 0 means that all the variance lies within populations, as if haplotype frequencies had been homogenized by intense dispersal. The significance of the observed Fst value was tested by a permutation test, which randomly shuffles the haplotype distributions. Thus, we produced 1,000 random associations between haplotypes and geographic origin to produce the distribution of the null hypothesis, that is, the value of Fst that would be expected by chance. We then determined whether the observed value was higher than 95% of the randomly generated Fst values.

Presence of phylogeographic signal in the data was explicitly tested by contrasting Nst and Fst values (Pons and Petit, 1996, Weir and Cockerham, 1984) among geographic regions. Nst is a measure of genetic differentiation among populations analogous to Fst but taking into account the phylogenetic relationships between alleles. An interesting property is that Nst > Fst when phylogeographic signal exists; that is, when distinct alleles sampled from within populations are phylogenetically more similar, on average, than alleles sampled from different populations. Nst was computed from a Tajima and Nei distance matrix (Tajima and Nei, 1984) (Table II in annex) among haplotypes obtained with Arlequin. The hypothesis that Nst > Fst was tested by computing the distribution under the null hypothesis by conducting 1000 permutations of rows and columns of the distance matrix among haplotypes, as implemented by SPAGEDI (Hardy and Vekemans, 2002). This hypothesis was accepted if the observed values were higher than at least 95% of the values obtained after permutation of the data.

2.4.4. Phylogenetic reconstruction among Erica species

A phylogenetic analysis of *Erica* species was performed to test the monophyly of *E. arborea* and *E. scoparia*. 42 DNA sequences data for 9 *Erica* species from Europe and 12 *Erica* species from South Africa were downloaded from Genbank (McGuire and Kron, 2005), for two chloroplast DNA regions: the *atp*B-*rbc*L spacer and the 5' half of the chloroplast *Mat*K gene (Table III in annex). One individual of each haplotype from my own data was chosen and added to this dataset. Gaps in the sequences were treated as missing data. Sequences were aligned and verified by eye using Sequencher 3.1 (Schneider, 1998). The data matrix was analyzed under Bayesian inference according to the following procedure:

2.4.4.1. Selection of a nucleotide substitution model

There are 56 nested evolution models and each of these models has different estimated parameters. Three types of parameters exist:
1) Substitution rates: the maximum is 6 rates (for example: transition and transversion correspond to 2 rates);
2) A γ distribution is used to model rate heterogeneity among sites (for example, differences in mutation rate between introns and exons). It is, in fact, impossible to calculate a specific rate for each site since this would dramatically increase the number of parameters to be inferred from the data. Hence, different rates are sampled from a γ distribution. This distribution is characterized by the shape parameter α. When α is small, many sites evolve at a low rate. The shape of the distribution thus permits to determine classes of sites characterized by similar rate values. The distribution is typically divided into four parts characterized by a mean substitution rate. At each site, four likelihood values are calculated for the four different rates, and the overall log-likelihood of the site is obtained by summing the contribution to the likelihood of each rate. For example, the contribution to the likelihood of a fast rate to a monomorphic site will be negligible, whereas the evolution of the site will be best characterized by a low rate value.
3) The proportion of invariable sites I in the locus.

The first model is the simplest one and includes a single parameter, namely the transition rate that is equal among all nucleotides, and subsequent models progressively increase in complexity. The choice of the model is a trade-off between the necessity to include as many parameters as possible to accurately describe the data, and the number of data that are available to assess those parameters. Thus, rate values of a model that is too complex for the data will have high variances, resulting in much phylogenetic uncertainty. The program Modeltest (Posada and Crandall, 1998) offers a useful framework to choose the appropriate model for the data. First, a neighbour-joining (NJ) tree is reconstructed from the data. Then, the program calculates the likelihood of the NJ tree under the assumptions of each of the 56 nested substitution models and determines which model best fits the data based on the Akaike Criterion.

2.4.4.2. Bayesian reconstruction

Once the nucleotide substitution model is selected by Modeltest, the number of parameters of the model is introduced in the MrBayes program. This program will sample trees and model parameters using a Monte Carlo Markov Chain (MCMC). At each generation, a perturbation is introduced in the values of the tree or parameters, and the likelihood of the tree and model parameters is each time calculated. This new state of the chain is accepted or rejected following the Metropolis-Hasting term: if L2>L1 the perturbation is accepted but if L2<L1, the perturbation is accepted if $\frac{L2}{L1} > [0,1]$ (higher than a random value between 0 and 1). The likelihood values increase to a plateau. The group of values before the plateau is called the burnin and is deleted after visualization on a graph and only the plateau values are conserved. Trees and model parameters are subsequently sampled at regular intervals that are sufficiently large to allow for

independence between two successive samplings (in general, 1000 to 10000 generations). The sampled trees and model parameters form what is called the posterior probability distribution. Finally, a 50% majority-rule consensus tree is computed from the trees sampled from the posterior probability distribution; this means that only the branches that appear more than 50% times in the sampled trees are conserved, and support for branches is provided by proportion of sampled trees wherein the branch in question is indeed resolved. This proportion is termed the posterior probability of the branch and was generated using the 'sumt' command of MrBayes.

For the *Erica* dataset, the GTR + I+ γ evolution model was selected by Modeltest 3.7 (Posada and Crandall, 1998). Six parameters plus an invariable sites proportion and a gamma distribution were introduced in MrBayes and were then reconstructed by Bayesian inference using the methods outlined above.

2.4.4.3. Infraspecific haplotype relationships

The phylogenetic relationships among haplotypes within *E. arborea* were also reconstructed by Bayesian inference using the methods outlined above. A GTR+I evolution model was selected by Modeltest 3.7. Six parameters and a proportion of invariable sites were thus introduced in MrBayes. This process allowed obtaining a 50% majority consensus tree and the posterior probabilities associated.

This analysis has been completed with a haplotype network reconstruction for *E. arborea* and *E. scoparia*. In fact, this method is based on distances and is often used in analyses of population genetics, since it permits to visualize topologies that are not possible within a cladistic framework because the latter is constrained by dichotomic topologies.

The haplotype network reconstructions were performed with the TCS program (Clement et al., 2000). The program collapses identical sequences into haplotypes and calculates the frequencies of the haplotypes in the sample. In this case, haplotypes were already defined and were directly introduced in the software. This is a parsimony method that reconstructs the tree with a minimum of mutational steps. Each branch represents one mutational step and each dot between branches represents the non-sampled or extinct haplotypes.

2.5. Dispersal-Vicariance Analysis

Ancestral biogeographic areas and the history of colonizations were reconstructed using DIVA (Dispersal-Vicariance Analysis), a method in which ancestral distributions are inferred based on a cost matrix derived from a simple biogeographic model (Ronquist, 1996). The program computes a general area cladogram that best fits the haplotype phylogeny and the associated distribution matrix specifying for each area whether a species occurs in the area (1) or is absent (0). It also reconstructs the ancestral distributions of the species. This reconstruction is an event-based method where the event are vicariance events (cost=0), sympatric speciation (cost=0), dispersal (cost=1) and

extinction (cost=1). Based on these costs, it is possible to find the minimum-cost reconstruction, which may be considered the most likely explanation for the origin of the pattern being analyzed. The method explicitly favors vicariance events for area change of the species over dispersal events.

III. Results

3.1. Data description

Within the 387 samples collected, 105 specimens of *E. arborea* and 31 of *Erica scoparia* were successfully amplified and sequenced at all the 4 loci: *atp*B-*rbc*L spacer, *Trn*H-*psb*A, *Mat*K and *rpl*16 (Table IV in annex).

The alignment of the *atpB-rbcL* spacer, *Trn*H-*psb*A, *Mat*K and *rpl*16 in *E. arborea* and *E. scoparia* included 1876 base pairs (bp): 253 bp in *atp*B-*rbc*L, 307 in *Trn*H-*psb*A, 584 in *mat*K and 732 in *rpl16*. Of the 1876 bp, 16 were parsimony informative and 5 were autapomorphic for *E. arborea*. For *E. scoparia*, 6 positions were synapomorphic and one autapomorphic. Details for each locus are given in ***Table 1***.

Table 1: Detail of synapomorphic and autapomorphic sites for 105 samples of E. arborea and 31 of E. scoparia for each of the 4 loci and their respective length.

	*Mat*K	*atp*B-*rbc*L	*Trn*H-*psb*A	*rpl*16	total
length (bp)	584	253	307	732	1876
synapomorphic sites- *E. arborea*	6	1	6	3	16
autapomorphic sites - *E. arborea*	0	1	1	3	5
synapomorphic sites- *E. scoparia*	2	1	0	3	6
autapomorphic sites - *E. scoparia*	0	0	0	1	1

Indels provided another 6 synapomorphic characters ***(Table 2)***, 5 for *E. arborea* and 1 for *E. scoparia*. No autapomorphic indels have been recorded.

Table 2: Number of indel per locus and their respective length for the 4 loci in 105 samples of E. arborea and 31 of E. scoparia.

indels	*Mat*K	Bp	*atp*B-*rbc*L	bp	*Trn*H-*psb*A	bp	*rpl*16	bp
E. arborea	***0***	0	***0***	0	***4***	1(for3),10	***1***	8
E. scoparia	***0***	0	***1***	1	***0***	0	***0***	0

Based on those 26 and 8 variable characters in *E. arborea* and *E. scoparia*, 19 and 11 haplotypes were identified respectively for *E. arborea* and *E. scoparia*. ***(Table 3 and 4)***

Erica arborea		
haplotype	nb. 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
19	1	0.95
22	2	1.90
23	2	1.90
24	1	0.95
Total	105	

Erica scoparia		
haplotype	Nb. of samples	frequency
1	6	19.4
2	7	22.6
3	3	9.7
4	1	3.2
5	1	3.2
6	1	3.2
7	4	12.9
9	2	6.5
10	3	9.7
13	1	3.2
14	2	6.5
Total	31	

Table 3 and 4: Haplotype frequency distribution based on variation in four cpDNA regions for a sample of 105 and 31 specimens of the heathers, E. arborea (a) and E. scoparia (b), collected throughout their distribution area.

In *E. arborea*, one haplotype (3) dominates, with a frequency of 64.76%. The frequency of all the other haplotypes ranges between 0.95% and 3.81%. In *E. scoparia*, haplotypes 2, 1 and 7 have a frequency of 22.6, 19.4 and 12.9%, respectively. The frequency of the other haplotypes ranges between 3.2 and 9.7 %.

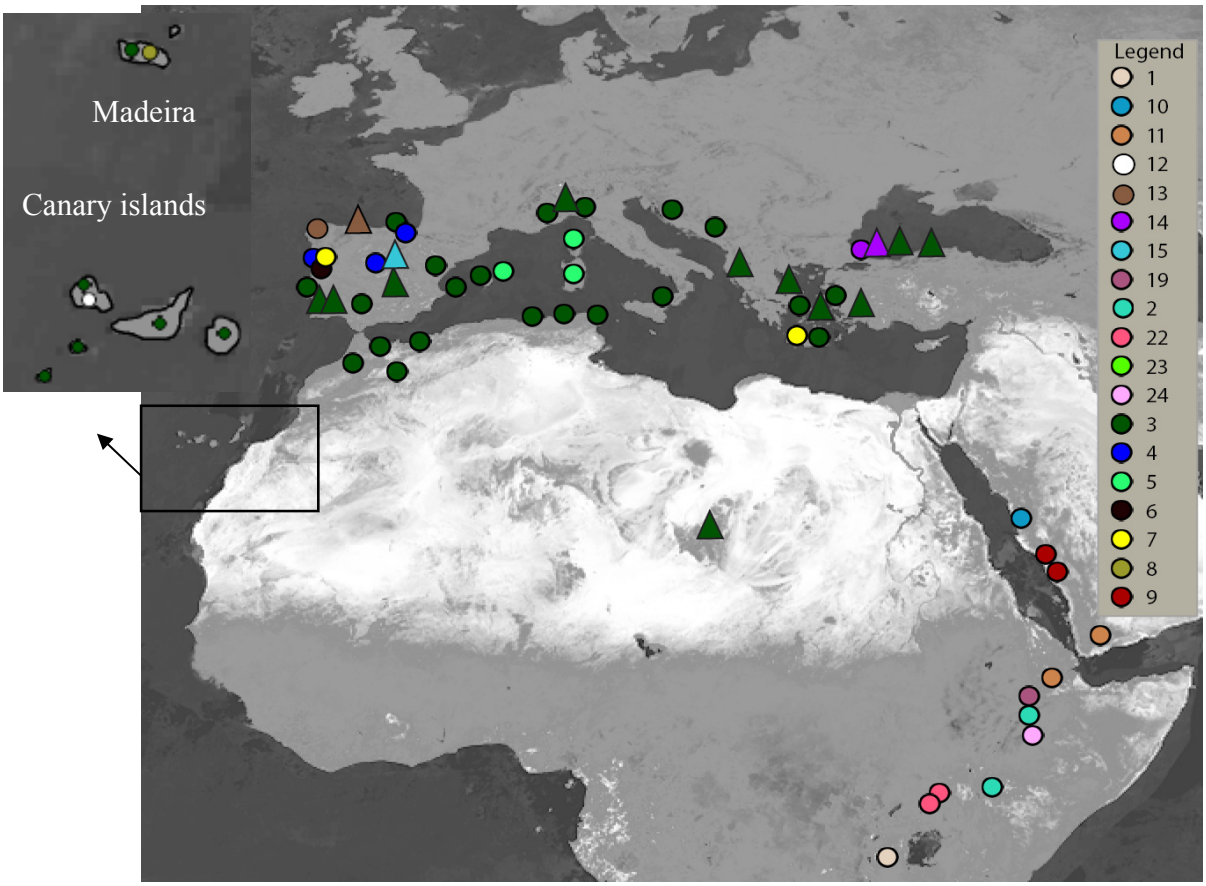
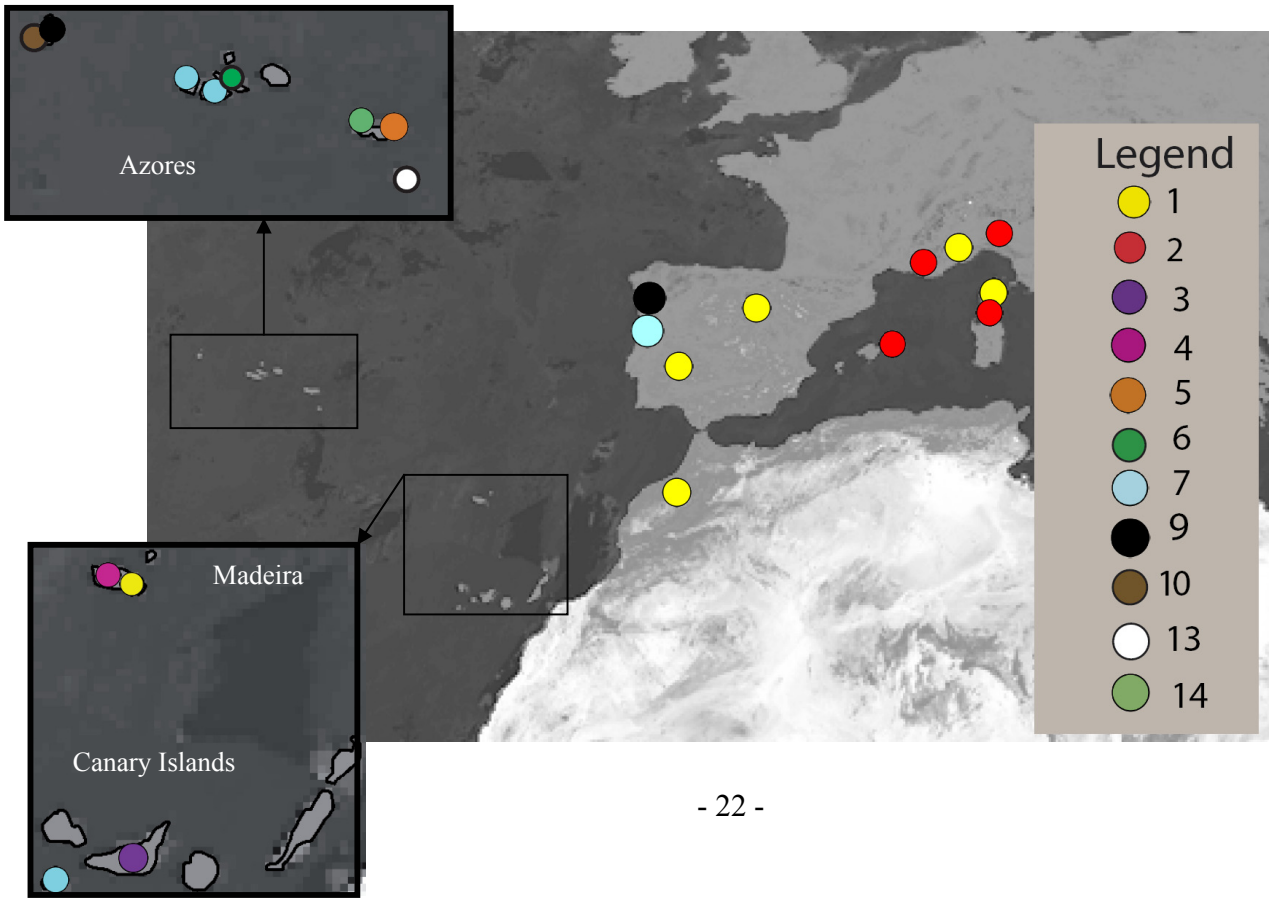


Fig. 15 and 16: 1) Geographic distribution of cpDNA haplotypes of Erica arborea: triangles correspond to the hybrids (corresponding to the haplotype of the same colour) between Erica arborea and Erica scoparia (see Appendix),
2) Geographic distribution of cpDNA haplotypes of Erica scoparia.



In *E. arborea* ***(Fig. 15)***, the bulk of haplotypic diversity is concentrated around East Africa and the northern half of the Iberian Peninsula. Eight different haplotypes, all of which are endemic, are present in East Africa and Southwest Asia. Five different haplotypes, four of which are endemic and one (7) is shared with Crete, are situated in northern Iberian Peninsula (Spain, Portugal). Haplotype 3 is distributed across Europe, North Africa, Tibesti and Macaronesia. Haplotype 5 was only found in the Mediterranean islands (Corsica, Sardinia, Balearic Islands). Haplotype 8 was only found on Madeira and haplotype 12 on La Palma. Haplotype 24 was only found in France.

In *E. scoparia* ***(Fig. 16)***, the bulk of haplotypic diversity is found in the Azores, where seven different haplotypes are found. It is also where most of samples were collected. Haplotypes 5 and 6 were only found on Sao Miguel and haplotype 10 on Flores. Haplotype 13 was found on Santa Maria and 14 on Sao Jorge. Haplotype 9 is shared between Portugal and Flores and haplotype 7 is shared between Faial, Pico, La Gomera and also Portugal. Haplotype 3 was only found on Tenerife and 4 on Madeira. Haplotypes 1 and 2 are distributed in Europe: 1 is found in France, Corsica, Spain, Portugal, Morocco and Madeira; 2 is found in France, Italy, Corsica and Menorca.

For *E. arborea*, the highest gene diversity levels are found in East Africa (0.891), South-west Asia (0.833) and the Iberian Peninsula (0.753). In North Africa, gene diversity is 0 because there is only a single haplotype present in this region. The lowest gene diversities are found in Macaronesia and Europe ***(Table 5)***.

Table 5: Gene diversity per region and number of haplotypes present in each region in 105 samples of the heather E. arborea, throughout its distribution area

E. arborea	gene diversity	Nb of haplotype
Macaronesia	0.227	3
Ib. Peninsula	0.753	7
Europe	0.398	5
N. Africa	0.000	1
SW Asia	0.833	3
E. Africa	0.891	6

3.2. Geographic differentiation and phylogeographic signal

Table 6: Results of the Amova test for 105 samples of E. arborea.

E. arborea	% variation	p-value
Variation among regions	31.13	
Variation within regions	68.87	<0.0001

The Amova test for *E. arborea* shows that there is a significant partitioning of genetic variance among populations and that the genetic variation is partitioned within and among populations at 68.87 and 31.13%, respectively ***(Table 6)***.

Table 7: Global Nst and Fst value for 105 samples of E. arborea, and their comparison by allele permutation test.

	Nst	Fst
value	0.3289	0.2624
P(1-sided test. H1: obs>exp)	<0.0001	<0.0001

allele permutation test	test Nst>Fst
P(1-sided test, H1: obs>exp)	0.3760

Global Nst and Fst values separately are significantly different from zero (p-value <0. 0001). The global Nst (0.3289) is higher than the global Fst (0.2624), but this difference is not significant (p-value= 0.3760) ***(Table 7)***.

Significant Fst values are observed between the Iberian Peninsula and all of the other regions: Europe, SW Asia (0.009), East Africa, North Africa and Macaronesia (<0,0001). Significant Fst values are also observed between East Africa and the other regions (<0.0001) and also between SW Asia and the other regions (<0.0001 and 0.009 with Iberian Peninsula) except the Fst value between East Africa and SW Asia that is not significant.

Significant Nst values are observed for Iberian Peninsula and Europe. Nst value is also significant between East Africa and SW Asia but Fst value is not significant. All the other pairwise Nst values are not significantly different from 0 ***(Table 8)***.

Table 8: Pairwise Fst and Nst values between the six geographical regions for 105 samples of Erica arborea. Significant values are in bold and non significant values are replaced by N.S.

Region	Region	Fst	p-value	Nst	p-value
3	2	0.11	**0.009**	0.37	**0.000**
2	5	0.22	**0.009**	0.37	N.S
6	5	0.11	N.S	0.35	**0.028**
3	4	0.05	N.S	0.04	N.S
3	1	0.02	N.S	0.04	N.S
1	4	0.01	N.S	0.04	N.S
6	2	0.19	**<0.0001**	0.43	N.S
2	4	0.24	**<0.0001**	0.40	N.S
2	1	0.19	**<0.0001**	0.24	N.S
6	3	0.42	**<0.0001**	0.37	N.S
6	1	0.52	**<0.0001**	0.33	N.S
6	4	0.57	**<0.0001**	0.42	N.S
4	5	0.79	**<0.0001**	0.02	N.S
1	5	0.65	**<0.0001**	0.03	N.S
3	5	0.50	**<0.0001**	0.03	N.S

1= Macaronesia
2= Iberian Peninsula
3= Europe
4= North Africa
5= SW Asia
6= East Africa

3.3. Phylogenetic reconstruction among Erica species

3.3.1. Erica phylogeny

The 50% majority-rule consensus from the trees sampled from the posterior probability distribution ***(Fig. 17)*** shows that all accessions of *E. scoparia* form a monophyletic group supported by a posterior probability value of 1. *E. scoparia* is included within a clade of European species supported at 0.71. *E. arborea* is resolved as part of a sister clade comprised of African species and supported at 0.96. The bulk of accessions within *E. arborea* form a monophyletic group supported at 1.00, which is resolved as sister to a clade composed of South African species.

Five accessions obtained from Eastern Africa are resolved outside that *E. arborea* clade. Forcing the MCMCs to sample only trees that are compatible with a monophyletic interpretation of *E. arborea* lead to a significant loss in log-likelihood. Those specimens correspond to sterile material supplied by M. Popp and C. Brochmann, from University of Oslo, which we interpret as specimens from other shrubby Ericaceae. In fact, those species named on the ***Fig.17***, as *Erica sp*. could have been misidentified. Indeed, other shrubby species belonging to the genus *Erica* or *Phyllippia* could be confused with *E. arborea* species, especially when samples are sterile. If those accessions indeed belong to *Phyllipia*, this genus would be paraphyletic and should be included in *Erica* as already suggested by Oliver in 1987 (Oliver, 1987), but this would require further investigations that are beyond the scope of the present work.

3.3.2. Haplotype network

3.3.2.1. E. arborea

Three main clades, clade I, II and III can be identified on the haplotype network ***(Fig. 19)***. Clade I comprises haplotypes from the Iberian Peninsula and Madeira (4, 15, 13, 6, 8), clade II from East Africa (22, 23, 2), clade III from East Africa and Southwest Asia (19, 11) and from Europe, the Iberian Peninsula, Madeira, the Canary Islands and North Africa (3, 5, 7, 12, 14, 24). Haplotypes 9 and 10 are resolved between clade II and III. Macaronesian haplotypes are not resolved as a monophyletic group. One haplotype is found in clade I (8 in La Palma) and the two others in clade III (3 in Canary Islands and Madeira, and 12 in Madeira). In clade III, haplotype 3, which is geographically widespread, occupies a central position from where five other haplotypes radiate (5, 7, 12, 14, 24). The case is nearly the same for clade I, wherein haplotype 4 is central to four other haplotypes (6, 8, 13, 15). This clade is well separated from the two others by 17 mutational steps in the two cases. Inside clade III, 8 mutational steps separate the haplotypes from East Africa and SW Asia regions from those of all the other regions (Iberian Peninsula, Europe, Macaronesia and North Africa).

3.3.2.2. E. scoparia

The haplotype network of *E. scoparia* is shown on ***Fig. 18***. Haplotype 1 occupies a central position, from which all the other haplotypes radiate. In fact, five events of

radiation from haplotype 1 occurred, giving raise to the Macaronesian haplotypes. So, Macaronesian haplotypes are not resolved as a monophyletic group. The mutational step that separates haplotypes 13 and 14 from the others consists of an indel of 10 bp, and is so considered as one mutational event.

3.3.3 Phylogeography of E. arborea derived from the Bayesian analysis

In this tree ***(Fig. 20)***, two sister clades can be identified respectively with 1 and 0.99 of posterior probability value (Pp). Clade I is composed of haplotypes present in Iberian Peninsula and Madeira. Clade IV can be divided into two smaller sister clades II and III. Clade II, with 0.97 Pp, is composed of haplotypes from East Africa. Clade III, with 0.98 Pp, is composed of haplotypes from Europe (comprising Iberian Peninsula), North Africa, Canary Islands and Madeira and is sister to two other haplotypes from East Africa and Southwest Asia (0.98 Pp).

3.4. Dispersal-Vicariance analysis (DIVA)

The DIVA analysis ***(Fig. 21)*** identified two similar competing optimal reconstructions requiring 10 dispersion/extinction events. The ancestral distribution areas for *E. arborea* are the Iberian Peninsula plus Southwest Asia in the first scenario and the Iberian Peninsula, Southwest Asia and East Africa in the second scenario. From the reconstruction of ancestral distributions, two clades are distinguished. The ancestral node of clade I is reconstructed as the Iberian Peninsula; the next node within that clade is also Iberian but the node subtending haplotypes 13, 6 and 8 is reconstructed as Iberian and Macaronesian what implies one dispersion event. The node between haplotypes 6 and 13 is also Iberian. Those reconstructions are the same for both scenarii. The ancestral node of the second clade is reconstructed as SW Asia in the first scenario whereas in the second scenario it is reconstructed as SW Asia and E. Africa. This second clade comprises clades II and III and also haplotypes 9 and 10. This node separating haplotype 10 from the others, imply one dispersal event to East Africa in a case of only a SW Asian reconstruction. The ancestral node of haplotype 9 and all the others is SW Asia and E. Africa. The ancestral node between clades II and III is resolved as E. Africa. Inside clade II, all nodes are resolved as E. Africa in both scenarios. For clade III, the ancestral node is East African. The next node between haplotype 11 and the others implies one or two events of dispersion as ancestral node is resolved as either: Europe plus E. Africa (1 event) or East Africa, Europe and SW Asia (2 events). In the case of one dispersal event, another dispersal event is required to explain the occurrence of haplotype in SW Asia. The next node separating haplotypes 5 and 12 from the rest is resolved as European, with the ancestral nodes for haplotypes 24, 14, 7 (1 dispersal event) and 3 (3 dispersal events) all European in both scenarios. For haplotypes 5 and 12, the ancestral node is either, Europe and Macaronesia or those two plus the Iberian Peninsula implying either one or two dispersal events. In the case of one dispersal event, a second event is required to explain the occurrence of haplotype 5 in the Iberian Peninsula.

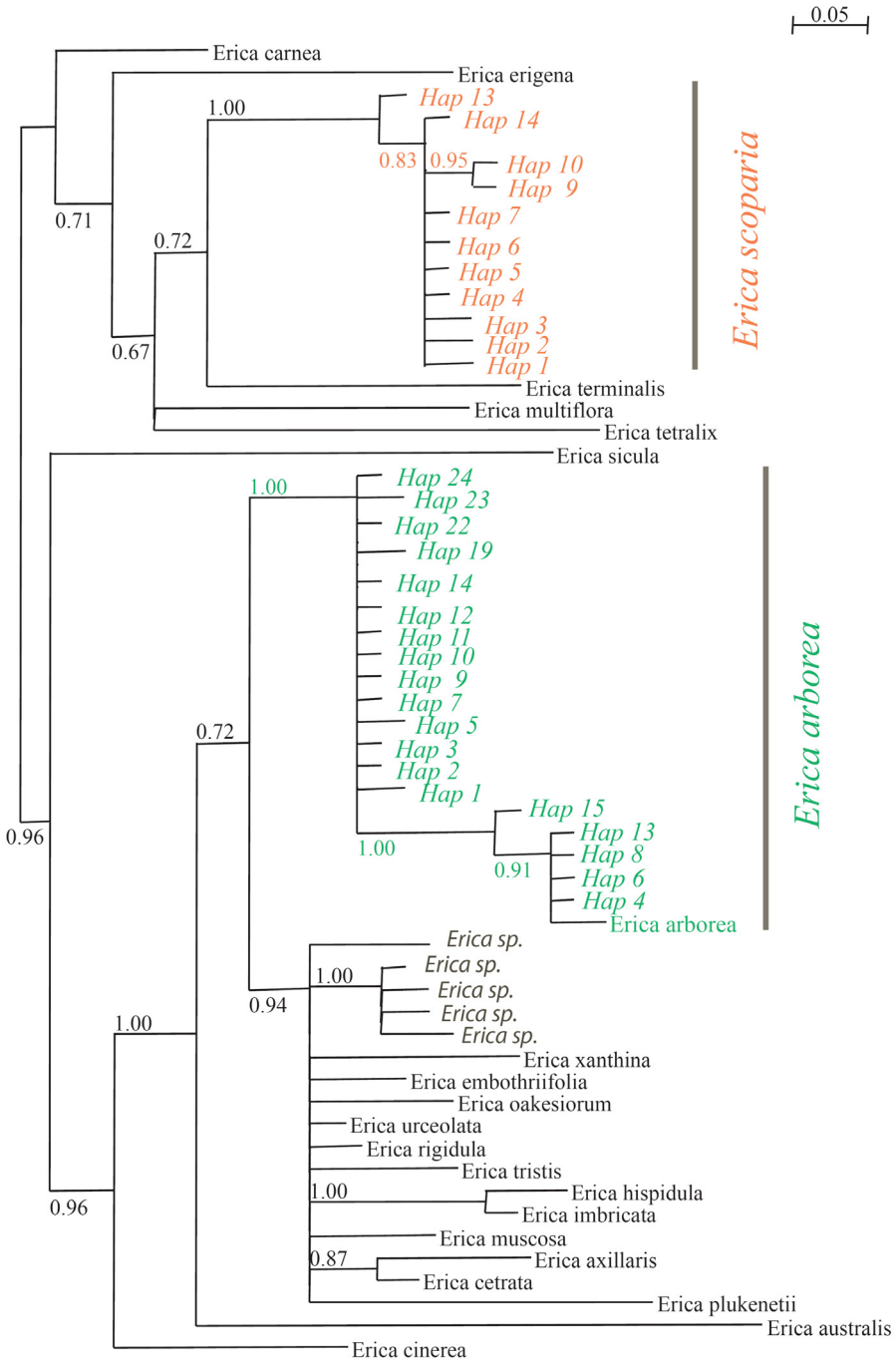


Fig. 17: 50% consensus tree reconstructed in Bayesian inference from 19 and 11 haplotype sequences of E. arborea and E. scoparia and 22 sequences from Genbank for two loci: MatK and atpB-rbcL.

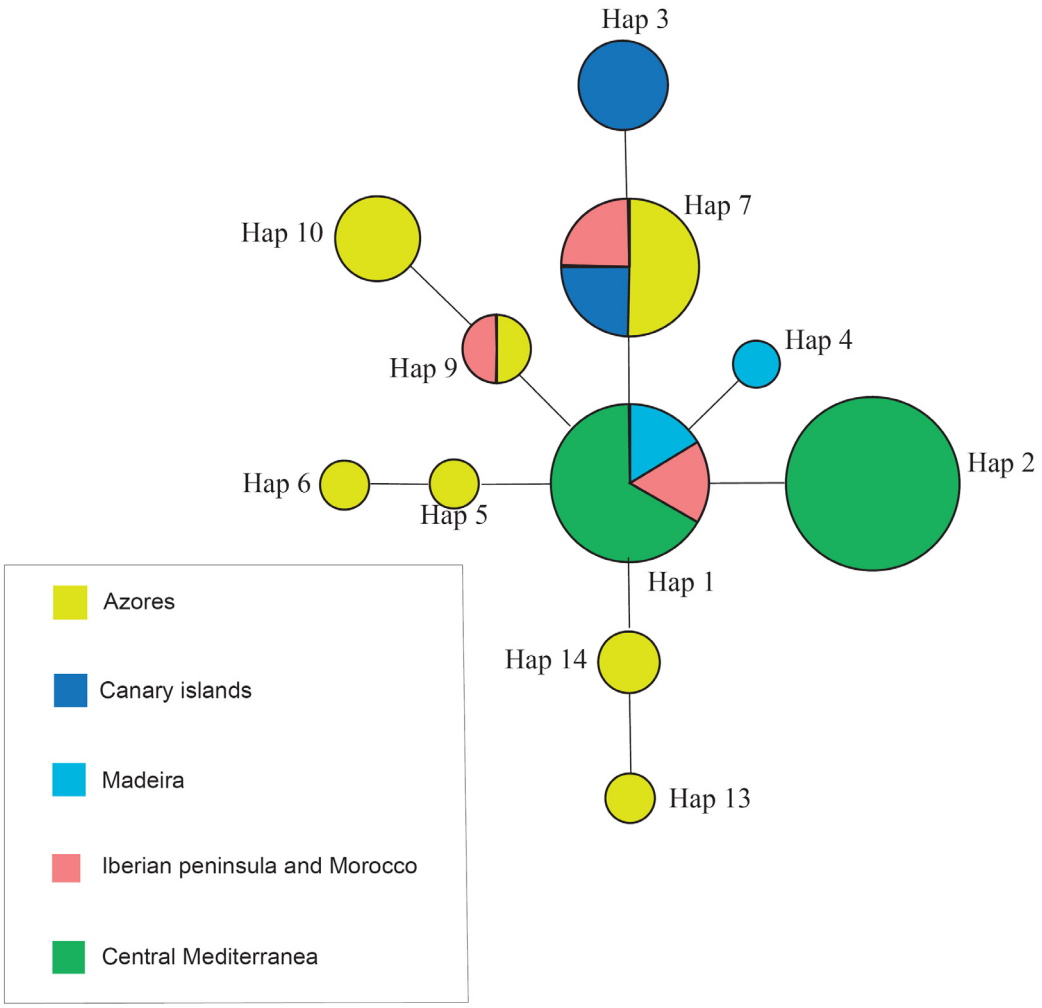


Fig. 18: Erica scoparia network reconstruction based on 11 haplotypes and 31 sequences for the 4 loci: MatK, atpB-rbcL, rpl16 and TrnH-psbA realized with the TCS software. On this tree, circles with haplotype geographical regions (colors) and frequencies (representing by circle size) have been added from the bayesian analysis to clarify the visualization.

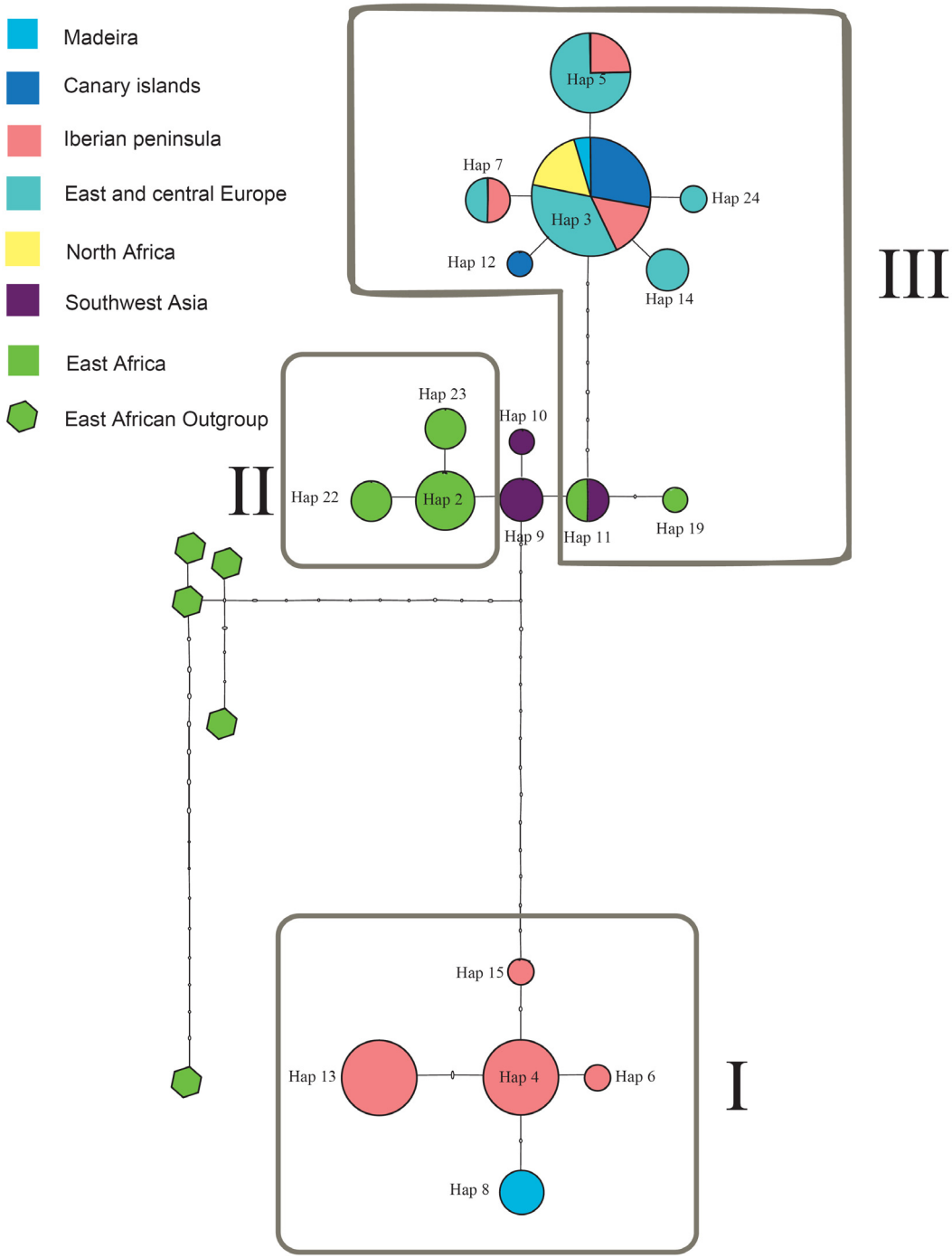


Fig. 19: Erica arborea haplotype network reconstructed from 19 haplotypes and 105 sequences on 4 loci: MatK, atpB-rbcL, rpl16, TrnH-psbA realized with the TCS software. The size of the circle representing each haplotype is proportional to its frequency. Dots represent "missing" haplotypes.

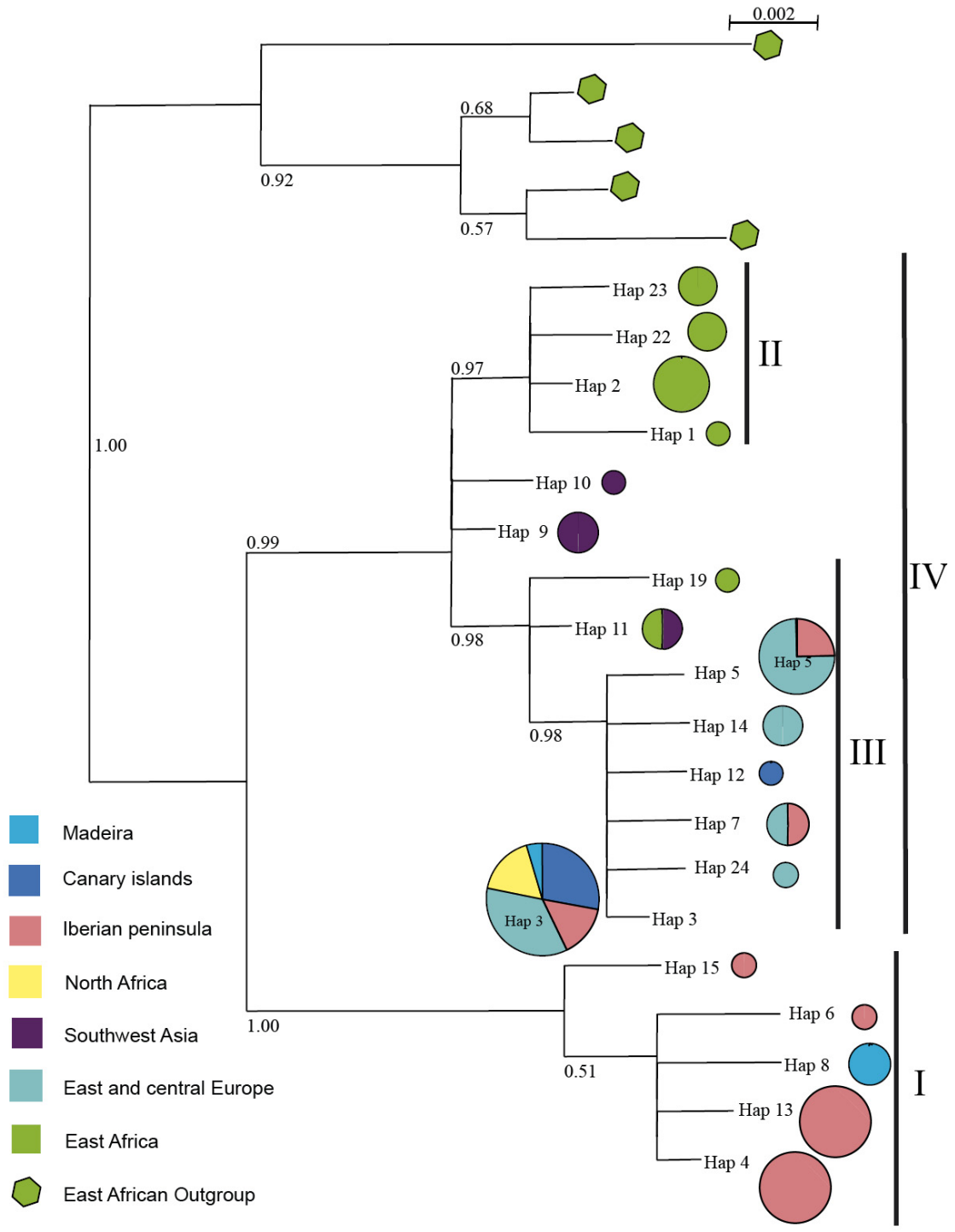


Fig. 20: Erica arborea, 50% consensus tree reconstructed in bayesian inference from 19 haplotypes and 105 sequences for 4 loci: MatK, atpB-rbcL, rpl16, TrnH-psbA. The size of the circle representing each haplotype is proportional to its frequency.

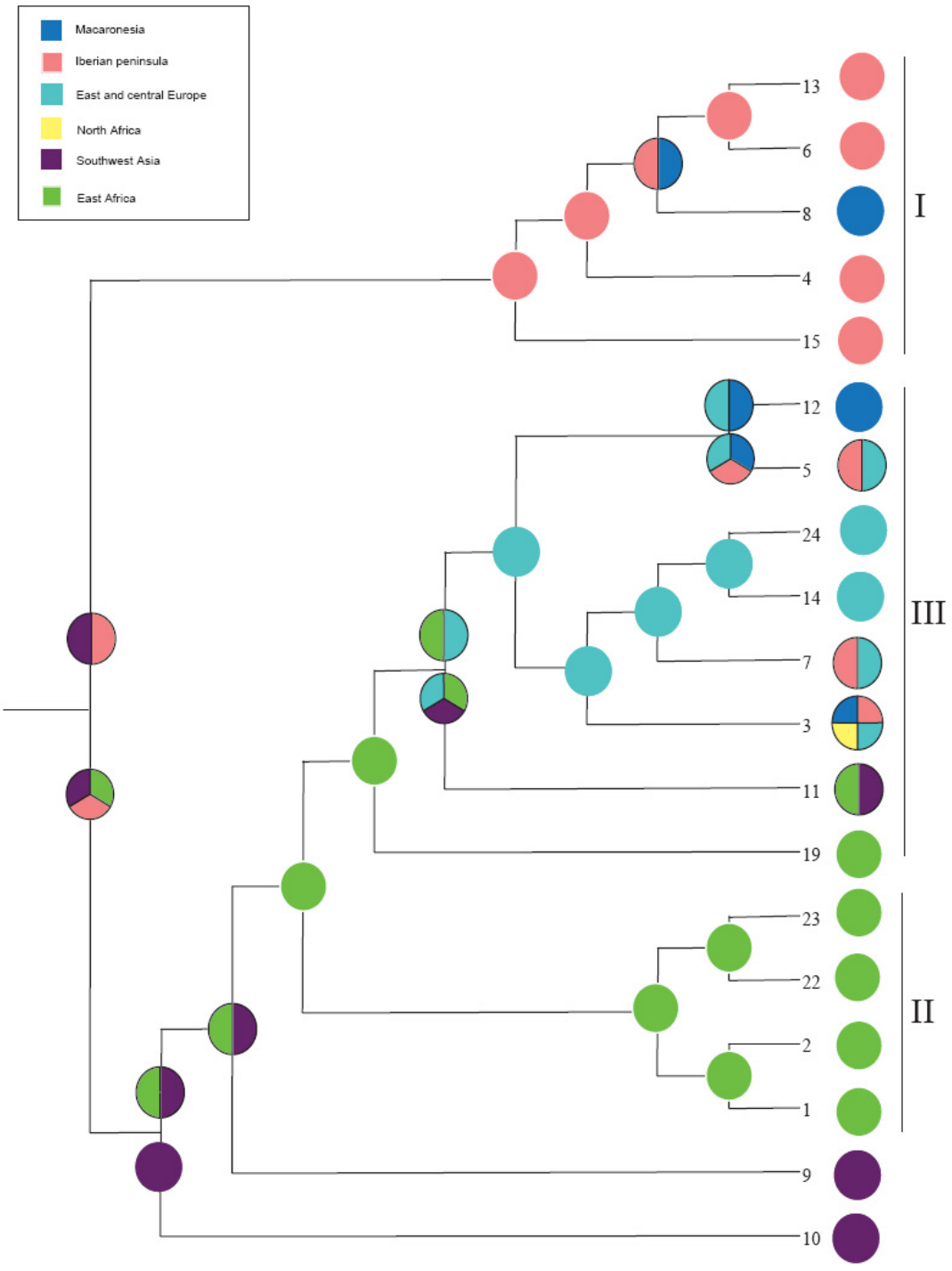


Fig. 21: Area cladogram from the Erica arborea phylogeny showing the most-likely ancestral distributions reconstructed by the dispersal-vicariance analysis (DIVA). The two optimal reconstructions are represented on the figure as two reconstructions on the same node.

IV. Discussion

E. arborea and *E. scoparia* are two morphologically similar elements of the laurel forest of Macaronesia, a unique temperate rainforest dominated by Lauraceae (e.g., *Ocotea foetens, Laurus novocariensis, Persea indica, Appolonias barbujana*), smaller evergreen bushes such as *Clethra arborea* and *Ilex perado* and a rich understory including for example, *Sonchus sp*. and *Argyranthemum sp.* It has been hypothesised that the impact of glaciations in Europe and the aridification of North Africa led to the eradication of continental laurel forests, with their survival only in Macaronesia (Axelrod, 1975) and in small continental refugial areas (Barbero et al., 1981, Arroyo-García et al., 2001, Rodríguez-Sánchez and Arroyo, 2008). The Macaronesian flora has therefore been viewed, for the most part, as a relict of a formerly widespread vegetation (Carine *et al.*, In press). Yet, the data presented here strongly challenge the hypothesis that the two *Erica* species are of relict origin but are rather newcomers from mainland Europe.

4.1. Diversity and diversification patterns

Despite the fact that the same molecular markers were employed for *E. arborea* and *E. scoparia*, the latter shows very little genetic variation. This difference could be due to differences in evolutionary rate, but these two species belong to the same genus and different rates in these species would contradict the classical hypothesis of rates inheritance (Sanderson, 1997). This hypothesis suggests that close branches in a phylogeny are considered to evolve at the same rate because rates are inherited from common ancestor.

If we accept that the evolutionary rate of the two species is similar because they belong to the same genus, an alternative interpretation is that the evolutionary history of *E. scoparia* is more recent. Most of the genetic diversity found in the species lies in the Azores, where several radiating haplotypes are present. Such an observation is, at first sight, consistent with Engler's refugium model and would suggest that, despite apparent morphological homogeneity, the species has undergone a cryptic radiation in the archipelago. Whilst such a cryptic radiation is consistent with the idea that the evolutionary history of *E. scoparia* is recent, it contrasts with the paucity of evolutionary radiations and the widespread distribution of most endemics in Azorean taxa. In fact, it has been hypothesised that the lack of Quaternary climate changes in the Azores, in contrast to the situation in the Canaries, may explain the lack of radiations in the Azores flora (Carine and Schaefer, In Press). However, whilst morphological variation in endemic lineages across the archipelago appears to be limited in most cases, a recent, detailed examination of quantitative floral traits in the widespread endemic *Vaccinium cylindraceum* has revealed subtle differences among island populations (Pereira, 2008) that may be indicative of an ongoing process of inter-island vicariance. Furthermore, ongoing molecular analyses of endemic Azorean plant taxa suggest cryptic allopatric speciation in several lineages (M. Carine pers. comm.)

In *E. arborea*, the global genetic diversity is much higher than in *E. scoparia*. However, within *E. arborea,* genetic diversity within Macaronesia is amongst the lowest compared to other biogeographic regions where the species is present, and only two endemic haplotypes are present in Macaronesia (haplotype 12 on La Palma and haplotype 8 on Madeira). Thus, whilst *E. scoparia* displayed a tendency for cryptic radiation in the Azores, *E. arborea* completely failed to diversify within Macaronesia. In this respect, the species differs markedly from those Canarian taxa, which are typically characterized by spectacular radiations such as *Argyranthemum* (Juan *et al.*, 2000), *Aeonium* (Webb and Berthelot, 1840, Mort et al., 2002) and *Echium* (Bohle *et al.*, 1996) .

Three hypotheses may explain the non-radiation of *E. arborea* in Macaronesia. On island ecosystems, 'windows of opportunities' for colonization of new habitats recurrently occur following geological disturbance through, e.g., volcanic activity or erosion, thereby favoring geographic isolation and ecological specialization and, eventually, adaptive radiation (Whittaker *et al.*, 2008). Radiations are characteristic for species that colonized the islands only once. When multiple colonization occur indeed, competition among populations from the same species, or from sister species, for niche occupancy, likely decreases the possibilities for diversification, a phenomenon known as 'niche pre-emption' (Silvertown, 2004). In the case of *E. arborea*, however, the same haplotype (#3) is found on all islands, which strongly weakens the hypothesis of a 'rain' of colonizing haplotypes from the continents.

A second interpretation of the lack of radiation among Macaronesian haplotypes of *E. arborea* is that the species displays the 'high dispersal' syndrome of bryophytes. In fact, one explanation for the complete failure of the latter to radiate in Macaronesia is that these organisms disperse freely, thereby homogenizing their genetic structure and limiting opportunities for genetic isolation and speciation (Vanderpoorten *et al.*, 2007). Despite the lack of specialized devices such as a wing, the seeds of *E. arborea* are minute (the weight of 100 seeds is $1.3\ 10^{-3}$g) (Molinier and Muller, 1938), which, exactly as spores, may facilitate their long-distance dispersal by wind. This hypothesis of a high dispersability is, however, at odds with the observation that endemic haplotypes from the centers of diversity of the species in North of Iberian Peninsula and eastern Africa failed to disperse.

A third possibility is that the absence of radiation of *E. arborea* in Macaronesia is due to a very recent colonization. This hypothesis refutes completely Engler's refugium model (Engler, 1879) and Axelrod's assumption (Axelrod, 1975) that the species colonized Macaronesia during the Tertiary. The idea of a recent colonization is consistent with the occurrence of only two Macaronesian endemic haplotypes, one in Madeira and one in La Palma, an island that is only 2 Ma old. If *E. arborea* had been in the Canaries for a long time, we would expect additional diversification events on the other islands that are, for the most part, older than La Palma. The fact that haplotype 3 is present in Macaronesia and throughout the Mediterranean also suggests that the colonization of the islands is very recent because no real diversification has appeared yet.

4.2. Centers of origin and refugia

Pleistocene climate fluctuations caused major range shifts for many taxa (Hewitt, 2001). While temperate species contracted their ranges to southern refugia (Taberlet et al., 1998, Hewitt, 1999), the habitat of taxa adapted to cold and dry conditions was larger during cooler periods than today. For wide-ranging species, climate fluctuations may have induced range contraction in some parts of the range and expansion in others. The repeated cycles of range expansions and contractions, which followed distinct patterns in different species, formed the intraspecific genetic structure that we observe today (Avise, 2000, Petit et al., 2003)

In E. *arborea*, the Amova shows that the genetic diversity is principally found within regions (68.87%). In fact, high genetic diversity values were observed in the Iberian Peninsula (0.753), South-West Asia (0.833) and East Africa (0.891). Although most of the diversity lies within regions, significant differences in haplotype frequencies among regions suggest limitations to gene flow from and to these areas. The combination of a high genetic diversity and genetic isolation points to a refugium for those areas. This idea is further reinforced by the identification by the DIVA analysis of either the Iberian Peninsula plus East Africa or these two regions plus South-West Asia as the centers of origin of the species. The identification of those regions considered as refugia is consistent with pollen records for the Near East and the three southern peninsulas of Europe (Balkan, Italian and Iberian), which are sufficiently robust to suggest that trees survived in these southern European regions during the full-glacial period (Willis and Niklas, 2004).

. With seven haplotypes, the Iberian Peninsula is characterized by high gene diversity (0.753). Four of those haplotypes are genetically related and form a fully supported clade well separated from the others by 17 mutational steps. These four haplotypes are endemic to the North of Iberian Peninsula. The significant values of Fst between the Iberian Peninsula and all the other regions show that this region is well separated from the others in terms of gene flows. The significant Nst value between the Iberian Peninsula and the rest of Europe further shows that a phylogeographic signal exists between these two regions, indicating that mutations accumulate at a faster rate than dispersal events. Altogether, these observations strongly reinforce the idea that the North of the Iberian Peninsula provided a refuge for *E. arborea* during the Ice Age. In fact, some areas of North Iberia were identified as refuge by Medail and Diadema (Medail and Diadema, 2009), based on observations made on other typical Mediterranean taxa, notably *Quercus ilex* (Lopez-de-Heredia et al., 2007) as well as other tree species including *Fagus sylvatica* (Lopez-Merino *et al.*, 2008), *Betula pendula* (Palmé *et al.*, 2003), *Coryllus avellana* (Palmé and Vendramin, 2002).

South-west Asia and East African similarly present high genetic diversity values, respectively of 0.833 and 0.891. Significant Nst values among those regions again suggest genetic isolation through faster mutation than dispersal rates. Whilst East Africa has long been identified as a refugium for a suite of species currently disjunct with South Africa (Gehrke and Linder, 2009), our observations on *E. arborea* suggest, together with

other observations made on species that exhibit similar distribution patterns as *E. arborea*, e.g., *Arabis alpine* (Koch *et al.*, 2006) that East Africa may also have played a key refugium role for Mediterranean species. In comparison with the North Iberian refuge, however, fewer mutational steps are present among East African haplotypes than among Iberian ones. In fact, East African haplotypes are separated by one mutational step, while most of North Iberian ones are separated by two. One possible explanation is that the East African region was more heavily affected by the Quaternary glaciations than North Iberia, and that diversification in East Africa is more recent as a result. It has been shown that the peak of the East African mountains was glaciated during the last glacial maximum (Tiercelin *et al.*, 2008), while Iberian Peninsula was potentially isolated by the Pyrenean Mountains in the North as suggested by Turner & Hannon (Turner and Hannon, 1988).

4.3. Patterns of post-glacial recolonization

The haplotype tree shows that the populations of *E. arborea* currently distributed in North Iberia and East Africa share a common ancestor. Those populations are, however, separated by a high mutation number (17 mutational steps). Together with the identification of either the Iberian Peninsula and SW Asia or the two plus East Africa as ancestral distribution areas by the DIVA analysis, these observations are consistent with the idea of an ancient vicariance event. The African refugium, or both the African and South-West Asian refugia, are identified as ancestral areas for the European haplotypes. Two scenarios of colonization of Europe and Macaronesia can therefore be proposed.

The first scenario involves a recent dispersal event from East Africa to Tibesti in Chad. In Tibesti, *E. arborea* occurs at high altitude (2000 to 3000 m) on the peak of the volcano. Pollen of this species was also recorded in Neolithic sediment in areas of North Africa including Touggourt, the Hoggar massif and the diatomites of Borkou (Bruneau de Miré and Quezel, 1959). In this scenario, those regions were colonized from East Africa, and, from there, the species reached North Africa where it is still present today. Such a hypothesis is supported by the fact that the haplotype found in Tibesti is the same as the one found everywhere in North Africa. This hypothesis is also supported by the fact that in the earlier Holocene, climatic conditions in the Sahara were far moister than present. As a result, Mediterranean species such as *Quercus ilex*, *Pistacia* or *Erica sp.* are reported from deposits dating back to 6000 BP (Prentice *et al.*, 2000). This scenario is, however, weakened by the fact that only a single haplotype was found throughout North Africa, whereas haplotypic diversification occurred in other Mediterranean areas. Furthermore, the eastwards circulation of air masses in Africa since 18000 BP (Messerli and Winiger, 1992), also weakens the hypothesis of a westwards migration by air currents.

A second scenario would involve the dispersal of eastern African haplotypes towards south-west Asia before reaching the Mediterranean region. Haplotype 3 would then have dispersed westwards, eventually reaching western North Africa and Macaronesia. The absence of any diversification of *E. arborea* in Macaronesia and western North Africa, by contrast with the evolution of endemic haplotypes in other

Mediterranean regions including the Balearic Islands, Corsica, France, Turkey and Crete, indeed suggests that Macaronesia was the most recently colonized area during the western expansion of the species. A recent colonization of western North Africa and Macaronesia is much consistent with the extremely low genetic diversities observed in those regions that, at least for North Africa, have traditionally been seen as refugia (Medail and Quezel, 1999).

4.4. Range expansion

The phylogeographic scenarii proposed above for *E. arborea* suggest that the species is undergoing a fast range expansion. This is further corroborated by two lines of evidence. First, as documented for other expanding species such as *Sphagnum fimbriatum* (Szovenyi *et al.*, 2006) and *Pogonatum dentatum* (Hassel *et al.*, 2005)), haplotypic variation in *E. arborea* is characterized by many singletons, while the level of genetic differentiation among haplotypes is low, mostly involving a single mutational step. Second, the observation that the global Nst is not significantly higher than the global Fst suggests that, overall, dispersal rates are faster than mutational rates, i.e., the species disperse before diversifying locally, as illustrated by the wide occurrence and abundance of haplotype 3.

In *E. scoparia* similarly, the star-like phylogeography suggests that the species underwent a range expansion. The extremely low genetic diversity as compared to *E. arborea* tends to indicate that this phenomenon is much recent. A fast range expansion of the species would account for the inconsistency found between the present results and the traditional geographically-based sub-species division of the species.

The results of this study therefore suggest that the Macaronesian *Erica* species cannot be considered as relicts of a putative evergreen tropical forest that spanned across Europe and North Africa in the tertiary. Indeed, the two species investigated here present some similarities with invasive species since they appear as recent colonizers and thrive in degraded laurel forests (Anon, 1973, Arevalo and Fernandez-Palacios, 2007, Arevalo et al., 2008). This is also supported by palynological evidence for the presence of the species in the middle Holocene, in Iberian regions where the vegetation was deteriorated by human activities (Lopez-Merino *et al.*, 2008) and also by the ability of this species to thrive after fire (Alvarez *et al.*, 2009). The ericaceous species are considered as pyrophytes to the extent that they are capable of regenerating from the rootstock and that frequent fires favour their survival ('fire dependence') by preventing the establishment of other competing woody species (Mesleard and Lepart, 1991).

In the Macaronesian islands, *Erica* species replace native laurel species following human disturbance (Melville, 1979) and also occurs on ridges where the laurels find it difficult to grow. Macaronesian Erica species therefore appear as opportunistic species that benefit from the degradation or destruction of forest ecosystems, thereby reinforcing the idea that the present-day composition of the Macaronesian laurel forest is a complex mix of ancient relicts and recent colonizers (Aigoin *et al.*, 2009).

V. Bibliography

Adams KL, Qiu YL, Stoutemyer M, Palmer JD. 2002. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **99**: 9905-9912.

Aigoin DA, Devos N, Huttunen S, Ignatov MS, Gonzalez-Mancebo JM, Vanderpoorten A. 2009. And if Engler was not completely wrong? Evidence for multiple evolutionary origins in the moss flora of Macaronesia. *Evolution,* **In press**.

Alvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution,* **29**: 417-434.

Alvarez R, Munoz A, Pesqueira XM, Garcia-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.

Anon. 1973. *Inventario Forestal Nacional,* Santa Cruz de Tenerife, Madrid, Ministerio de Agriculture, Pesca y Alimentacion: Section de inventario y mapas.

Arevalo JR, Fernandez-Palacios JM. 2007. Treefall gaps and regeneration composition in the laurel forest of Anaga (Tenerife): a matter of size? *Plant Ecology,* **188**: 133-143.

Arevalo JR, Peraza MD, Alvarez C, Bermudez A, Delgado JD, Gallardo A, Fernandez-Palacios JM. 2008. Laurel forest recovery during 20 years in an abandoned firebreak in Tenerife, Canary Islands. *Acta Oecologica,* **33**: 1-9.

Arroyo-García R, Martínez-Zapater JM, Fernández Prieto JA, Álvarez-Arbesú R. 2001. AFLP evaluation of genetic similarity among laurel populations (*Laurus L.*). *Euphytica,* **122**: 155-164.

Avise JC. 2000. *Phylogeography : the history and formation of species,* Cambridge, Mass., Harvard University Press.

Axelrod DI. 1975. Evolution and Biogeography of Madrean-Tethyan Sclerophyll Vegetation. *Annals of the Missouri Botanical Garden,* **62**: 280-334.

Bailey CD, Carr TG, Harris SA, Hughes CE. 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution,* **29**: 435-455.

Banco de datos de Biodiversidad de Canarias. 2004. Lista de especies silvestres de Canarias: Hongos, plantas y animales terrestres.

Barbero M, Benabid A, Peyre C, Quézel.P. 1981. Sur la presence au Maroc *de Laurus azorica* (Seub.)Franco. *Anales Jardín Botánico de Madrid,* **37**: 467-472.

Bausher MG, Singh ND, Lee SB, Jansen RK, Daniell H. 2006. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol,* **6**: 21.

Bendall KE, Sykes BC. 1995. Length Heteroplasmy in the First Hypervariable Segment of the Human Mtdna Control Region. *American Journal of Human Genetics,* **57**: 248-256.

Bendich AJ. 1987. Why Do Chloroplasts and Mitochondria Contain So Many Copies of Their Genome. *Bioessays,* **6**: 279-282.

Bensasson D, Zhang DX, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution,* **16**: 314-321.

Birky CW, Jr. 1983. Relaxed cellular controls and organelle heredity. *Science,* **222**: 468-75.

Birky CW, Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci U S A,* **92**: 11331-8.

Bohle UR, Hilger HH, Martin WF. 1996. Island colonization and evolution of the insular woody habit in *Echium L.* (Boraginaceae). *Proceedings of the National Academy of Sciences of the United States of America,* **93**: 11740-11745.

Bramwell D. 1972. Endemism in the flora of the Canary Islands. D. H. Valentine ed. Academic Press, London, UK.

Bruneau de Miré P, Quezel 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.

Burzynski A, Zbawicka M, Skibinski DOF, Wenne R. 2003. Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the Baltic. *Molecular Biology and Evolution,* **20**: 388-392.

Carine MA. 2005. Spatio-temporal relationships of the Macaronesian endemic flora: a relictual series or window of opportunity? *Taxon,* **54**: 895-903.

Carine MA, Russell SJ, Santos-Guerra A, Francisco-Ortega J. 2004. Relationships of the Macaronesian and Mediterranean floras: Molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in *Convolvulus* (Convolvulaceae). *American Journal of Botany,* **91**: 1070-1085.

Carine MA, Santos Guerra A, Guma IR, Reyes-Betancourt JA. In press. Endemism and evolution of the Macaronesian Flora. In: D.M. Williams, Knapp SK eds. *Beyond Cladistics.*

Carine MA, Schaefer H. In Press. The Azores diversity enigma: why are there so few Azorean endemic flowering plants and why are they so widespread? . *Journal of Biogeography*

Cecca F. 2008. La dimension biogéographique de l'évolution de la Vie. *C. R. Palevol*.

Chiu WL, Stubbe W, Sears BB. 1988. Plastid Inheritance in *Oenothera* - Organelle Genome Modifies the Extent of Biparental Plastid Transmission. *Current Genetics,* **13**: 181-189.

Clegg MT. 1993. Chloroplast Gene-Sequences and the Study of Plant Evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **90**: 363-367.

Clegg MT, Gaut BS, Learn GH, Morton BR. 1994. Rates and Patterns of Chloroplast DNA Evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **91**: 6795-6801.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. . *Mol. Ecol.,* **9**: 1657-1659.

Corriveau JL, Coleman AW. 1988. Rapid Screening Method to Detect Potential Biparental Inheritance of Plastid DNA and Results for over 200 Angiosperm Species. *American Journal of Botany,* **75**: 1443-1458.
Couplan F. 2000. *Dictionnaire étymologique de botanique,* Paris, delachaux et niestlé.
Cox CB, Moore PD. 2005. *Biogeography : an ecological and evolutionary approach,* Malden, MA, Blackwell Pub.
Craw R, Grehan JR, Heads MJ. 1999. *Panbiogeography: tracking the history of life,* New York, Oxford University Press.
Crisci JV. 2001. The voice of historical biogeography. *Journal of Biogeography,* **28**: 157-168.
Croizat L. 1958. *Panbiogeography; or, An introductory synthesis of zoogeography, phytogeography, and geology, with notes on evolution, systematics, ecology, anthropology, etc,* Caracas,.
Croizat L, Nelson G., Rosen D.E. 1974. Centers of origin and related concepts. *Systematic Zoology,* **23**: 265-287.
Cronk QCB. 1987. The History of Endemic Flora of St-Helena - a Relictual Series. *New Phytologist,* **105**: 509-520.
Cronk QCB. 1992. Relict floras of Atlantic islands-Patterns assessed., Academic Press Ltd.
de Queiroz A. 2005. The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution,* **20**: 68-73.
Delavault P, Thalouarn P. 2002. The obligate root parasite *Orobanche cumana* exhibits several rbcL sequences. *Gene,* **297**: 85-92.
Demesure B, Sodzi N, Petit RJ. 1995. A Set of Universal Primers for Amplification of Polymorphic Noncoding Regions of Mitochondrial and Chloroplast DNA in Plants. *Molecular Ecology,* **4**: 129-131.
Depape G. 1922. Recherches sur la flore Pliocène de la vallée du Rhone. *Ann. Sci. Nat. Ser,* **10**: 73-265.
Doyle JJ. 1992. Gene Trees and Species Trees - Molecular Systematics as One-Character Taxonomy. *Systematic Botany,* **17**: 144-163.
Doyle JJ, Doyle L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin,* **19**: 11-15.
Dumolin-Lapegue S, Pemonge MH, Petit RJ. 1997. An enlarged set of consensus primers for the study of organelle DNA in plants. *Mol Ecol,* **6**: 393-7.
Emerson BC. 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. (vol 11, pg 951, 2002). *Molecular Ecology,* **11**: 2451-2451.
Emerson BC, Kolm N. 2005. Species diversity can drive speciation. *Nature,* **434**: 1015-1017.
Engler A. 1879. *Versuch einer Entwicklungsgeschichte, insbesondere der Florengebiete seit der Tertiärperiode. I. Die extra-tropischen Gebiete der nördlischen Hemisphäre* Engelmann, Leipzig, Germany

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online,* **1**: 47-50.

Fernández-Palacios JM, Dias E. 2001. Marco Biográphico Macaronésico. In: Esquivel JMF-PaJLM ed. *Naturaleza de las Islas Canarias.* Santa Cruz de Tenerife, Canary islands, Spain, Publicationes Turquesa S.L.,.

Frahm JP. 2004. A new contribution to the moss flora of Baltic and Saxon amber. *Review of Palaeobotany and Palynology,* **129**: 81-101.

Francisco-Ortega J, Jansen RK, Santos-Guerra A. 1996. Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization in the evolution of the Macaronesian flora. *Proceedings of the National Academy of Sciences of the United States of America,* **93**: 4085-4090.

Francisco-Ortega J, Park SJ, Santos-Guerra A, Benabid A, Jansen RK. 2001. Origin and evolution of the endemic Macaronesian Inuleae (Asteraceae): evidence from the internal transcribed spacers of nuclear ribosomal DNA. *Biological Journal of the Linnean Society,* **72**: 77-97.

Fuertes-Aguilar J, Ray MF, Francisco-Ortega J, Santos-Guerra A, Jansen RK. 2002. Molecular evidence from chloroplast and nuclear markers for multiple colonizations of *Lavatera* (Malvaceae) in the Canary Islands. *Systematic Botany,* **27**: 74-83.

García-Talavera F. 1997. Las Canarias Orientales y la vecina costa africana en el Holoceno. *Eres,* **7**: 74-83.

García-Talavera F, L. Sanchez-Pinto, and S.Soccoro. 1995. Vegetales fósiles en el complejo taquítico-sienítico de Gran Canaria. *Revista de la Academia Canaria de Ciencias,* **7**: 77-91.

Gehrke B, Linder HP. 2009. The scramble for Africa: pan-temperate elements on the African high mountains. *Proceedings of the Royal Society B-Biological Sciences,* **276**: 2657-2665.

Geldmacher J, Hoernle.K, Van Den Bogaard P, Zankl G, Garbe-Schöenberg D. 2001. Earlier history of the >70 Ma old Canary hotspot based on the temporal and geochemical evolution of the Selvagen Archipelago and neighbouring seamounts in the eastern North Atlantic. *Journal of Volcanology and Geothermal Research,* **111**: 55–87.

Gielly L, Taberlet P. 1994. The Use of Chloroplast DNA to Resolve Plant Phylogenies - Noncoding Versus Rbcl Sequences. *Molecular Biology and Evolution,* **11**: 769-777.

Gonzáles Martín M, and F. Gonzáles Artiles. 2001. *Naturaleza de las Islas Canarias: Flora introducida.,* S.L., Santa Cruz de Tenerife, Canary Islands, Spain., J. M. Fernández-Palacios and J. L. Martín Esquivel.

Gunther S, Sommer G, von Breunig F, Iwanska A, Kalinina T, Sterneck M, Will H. 1998. Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: Frequency and functional consequences of PCR-introduced mutations. *Journal of Clinical Microbiology,* **36**: 531-538.

Hansen AK, Escobar LK, Gilbert LE, Jansen RK. 2007. Paternal, maternal, and biparental inheritance of the chloroplast genome in Passiflora (Passifloraceae): Implications for phylogenetic studies. *American Journal of Botany,* **94**: 42-46.

Hardy OJ, 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.

Harrison RG, Rand DM, Wheeler WC. 1985. Mitochondrial DNA Size Variation Within Individual Crickets. *Science,* **228**: 1446-1448.

Hassel K, Sastad SM, Gunnarsson U, Soderstrom L. 2005. Genetic variation and structure in the expanding moss *Pogonatum dentatum* (Polytrichaceae) in its area of origin and in a recently colonized area. *American Journal of Botany,* **92**: 1684-1690.

Heaney LR. 2007. Is a new paradigm emerging for oceanic island biogeography? *Journal of Biogeography,* **34**: 753-757.

Helfgott DM, Francisco-Ortega J, Santos-Guerra A, Jansen RK, Simpson BB. 2000. Biogeography and breeding system evolution of the woody Bencomia alliance (Rosaceae) in Macaronesia based on ITS sequence data. *Systematic Botany,* **25**: 82-97.

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Mol Ecol,* **10**: 537-49.

Hewitt MG. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society,* **68**: 87-112.

Hoch B, Maier RM, Appel K, Igloi GL, Kossel H. 1991. Editing of a chloroplast mRNA by creation of an initiation codon. *Nature,* **353**: 178-80.

Howell N, Halvorson S, Kubacka I, Mccullough DA, Bindoff LA, Turnbull DM. 1992. Mitochondrial Gene Segregation in Mammals - Is the Bottleneck Always Narrow. *Human Genetics,* **90**: 117-120.

Huang S, Chiang YC, Schaal BA, Chou CH, Chiang TY. 2001. Organelle DNA phylogeography of *Cycas taitungensis*, a relict species in Taiwan. *Molecular Ecology,* **10**: 2669-2681.

Humphries CJ. 1979. Plants and islands: Endemism and evolution in Macaronesia. D. Bramwell ed., Academic Press.

Humphries CJ. 2000. Form, space and time; which comes first? *Journal of Biogeography,* **27**: 11-15.

Humphries CJ, Parenti LR. 1999. *Cladistic biogeography,* Oxford, Oxford University Press.

Jackman TR, Larson A, de Queiroz K, Losos JB. 1999. Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. *Systematic Biology,* **48**: 254-285.

Jacobs G, Tscholl E, Sek A, Pfreundschuh M, Daus H, Trumper L. 1999. Enrichment polymerase chain reaction for the detection of Ki-ras mutations: relevance of Taq polymerase error rate, initial DNA copy number, and reaction conditions on the emergence of false-positive mutant bands. *Journal of Cancer Research and Clinical Oncology,* **125**: 395-401.

Juan C, Emerson BC, Oromi P, Hewitt GM. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution,* **15**: 104-109.

Kambysellis MP, Ho KF, Craddock EM, Piano F, Parisi M, Cohen J. 1996. Pattern of ecological shifts in the diversification of Hawaiian *Drosophila* inferred from a molecular phylogeny (vol 5, pg 1129, 1995). *Current Biology,* **6**: 213-213.

Kim SC, Crawford DJ, Francisco-Ortega J, SantosGuerra A. 1996. A common origin for woody *Sonchus* and five related genera in the Macaronesian islands:

Molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences of the United States of America,* **93**: 7743-7748.
Koch AM, 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.
Kvist L, Martens J, Nazarenko AA, Orell M. 2003. Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution,* **20**: 243-247.
Lee SB, Kaittanis C, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H. 2006. The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms. *BMC Genomics,* **7**: 61.
Lomolino M.V., Riddle B.R., Brown J.H. 2006. *Biogeography,* Sunderland, Massachussets, Inc. Publishers.
Lopez-de-Heredia U, Carrión JS, 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.
Lopez-Merino L, Lopez-Saez JA, Zapata MBR, Garcia MJG. 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.
M. Metzalff TB, R. Hagemann. 1981. Variations of chloroplast DNA in the genus *Pelargonium* and their biparental inheritance. *Theoretical and Applied Genetics,* **60**: 37-41.
Mac Arthur RH, Wilson EO. 2001. *The theory of island biogeography,* Princeton, Princeton University Press.
Maddison WP, Maddison DR. 1989. Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatol (Basel),* **53**: 190-202.
Marshall HD, Newton C, Ritland K. 2001. Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. *Molecular Biology and Evolution,* **18**: 2136-2138.
Martin W, Herrmann RG. 1998. Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiology,* **118**: 9-17.
McDowall RM. 2004. What biogeography is: a place for process. *Journal of Biogeography,* **31**: 345-351.
McGlone MS. 2005. Goodbye Gondwana. *Journal of Biogeography,* **32**: 739-740.
McGuire AF, Kron KA. 2005. Phylogenetic relationships of European and African Ericas. *International Journal of Plant Sciences,* **166**: 311-318.
Medail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography,* **36**: 1333-1345.
Medail F, Quezel P. 1999. Biodiversity hotspots in the Mediterranean basin: Setting global conservation priorities. *Conservation Biology,* **13**: 1510-1513.
Medgyesy P, Fejes E, Maliga P. 1985. Interspecific chloroplast recombination in a Nicotiana somatic hybrid. *Proc Natl Acad Sci U S A,* **82**: 6960-6964.
Melville R. 1979. *Endangered Islands Flora,* Royal botanical garden Kew England.

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.
Milne RI. 2006. Northern hemisphere plant disjunctions: A window on tertiary land bridges and climate change? *Annals of Botany,* **98**: 465-472.
Milne RI, Abbott RJ. 2002. The origin and evolution of tertiary relict floras. *Advances in Botanical Research, Vol 38,* **38**: 281-314.
Mogensen HL. 1996. The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany,* **83**: 383-404.
Molinier R, Muller P. 1938. La dissémination des espèces végétales. *Rev. Gen. Bot.,* **50**.
Morrone JJ, Crisci JV. 1995. Historical Biogeography - Introduction to Methods. *Annual Review of Ecology and Systematics,* **26**: 373-401.
Mort ME, Crawford DJ, Santos-Guerra A, Francisco-Ortega J, Esselman EJ, Wolfe AD. 2003. Relationships among the Macaronesian members of *Tolpis* (Asteraceae : Lactuceae) based upon analyses of inter simple sequence repeat (ISSR) markers. *Taxon,* **52**: 511-518.
Mort ME, Soltis DE, Soltis PS, Francisco-Ortega J, Santos-Guerra A. 2002. Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Systematic Botany,* **27**: 271-288.
Nelson G, Platnick N. 1981. *Systematics and biogeography: Cladistic and vicariance,* New York, Columbia University Press.
Nesbo CL, Arab MO, Jakobsen KS. 1998. Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (Perca fluviatilis, Acerina cernua, Stizostedion lucioperca). *Genetics,* **148**: 1907-1919.
Oliver EGH. 1987. Studies in the Ericoideae (Ericaceae).VII. The placing of the genus *Phyllipia* into synonymy under *Erica*; the southern African species. *South African Journal of Botany,* **53**: 455-458.
Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M. 2003. Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology,* **12**: 201-212.
Palmé AE, Vendramin GG. 2002. Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology,* **11**: 1769-1779.
Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR. 1988. Chloroplast DNA Variation and Plant Phylogeny. *Annals of the Missouri Botanical Garden,* **75**: 1180-1206.
Pereira MJ. 2008. Reproductive biology of *Vaccinium cylindraceum* (Ericaceae), an endemic species of the Azores archipelago. *Botany-Botanique,* **86**: 359-366.
Petit RJ, 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 JP, Rendell S, Vendramin GG. 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.
Platnick NI. 1981. *The progression rule or progress beyond rules in biogeography,* New York, Columbia University Press.
Pons O, Petit R. 1996. Measuring and testing genetic differentiation with ordered *versus* unordered alleles. *Genetics,* **144**: 1237-1245.
Posada D, Crandall A. 1998. MODELTEST: testing then model of DNA substitution. *Bioinform. Appl.,* **Note 14**: 817-818.
.
Prentice IC, Jolly D, participants B. 2000. Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biogeography,* **27**: 507-519.
Press JR, and Short, M.J. 1994. *Flora of Madeira,* London, Natural history museum.
Quezel P. 1978. Analysis of the flora of Mediterranean and Sahara Africa. *Annals of the Missouri Botanical Garden,* **65**: 479-534.
Rambaut A. 1995. Se–Al: Sequence alignment program. In: University O ed.
Raven PH, Evert RF, Eichhorn SE. 1992. *Biology of plants,* New York, Worth.
Rees DJ, Emerson BC, Oromi P, Hewitt GM. 2001. The diversification of the genus *Nesotes* (Coleoptera : Tenebrionidae) in the Canary Islands: Evidence from mtDNA. *Molecular Phylogenetics and Evolution,* **21**: 321-326.
Renner SS. 2005. Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science,* **10**: 550-558.
Ritland KaMTC. 1987. Evolutionary analysis of plant DNA sequences. *American Naturalist,* **130**: S74-S100.
Rodriguez-Rajo FJ, Mendez J, Jato V. 2005. Airborne ericaceae pollen grains in the atmosphere of Vigo (Northwest Spain) and its relationship with meteorological factors. *Journal of Integrative Plant Biology,* **47**: 792-800.
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.
Ronquist F. 1996. DIVA version 1.1. Computer program and manual available by anonymous FTP from Uppsala University (ftp.uu.se or ftp.systbot.uu.se).
.
Rozas J. 2009. DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol,* **537**: 337-50.
Rozas J, Rozas R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics,* **15**: 174-5.
Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution,* **14**: 1218-1231.
Sang T, Crawford DJ, Kim SC, Stuessy TF. 1994. Radiation of the Endemic Genus *Dendroseris* (Asteraceae) on the Juan-Fernandez-Islands - Evidence from Sequences of the Its Regions of Nuclear Ribosomal DNA. *American Journal of Botany,* **81**: 1494-1501.
Santos-Guerra A. 1999. Origen y evolucíon de la flora de Canarias. In: Fernández-Palacios JM, Bacallado JJ, Belmonte JA eds. *Ecologia y cultura en Canarias.*

Cabildo de Tenerife, Santa Cruz de Tenerife, Spain, Organismo Autónomo de Museos y Centros.
Saporta G. 1865. Etude sur la végétation du sud-est de la France à l'époque Tertiaire. *Ann. Sci. Nat.Ser,* **54**: 5-264.
Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. 1999. Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res,* **6**: 283-90.
Schäfer H. 2002. *Flora of the Azores: a field guide*, Margraf Verlag.
Schilling EE, Panero JL. 2002. A revised classification of subtribe Helianthinae (Asteraceae : Hetiantheae). I. Basal lineages. *Botanical Journal of the Linnean Society,* **140**: 65-76.
Schneider I. 1998. Gene codes sequencher. *Genetic Engineering News,* **18**: 24-24.
Schneider S, Roessli D, Excoffier L. 2000. Arlequin, Version 2.0: a software for population genetic data analysis. Genetics and Biometry Laboratory, Geneva, University of Geneva.
Shadel GS, Clayton DA. 1997. Mitochondrial DNA maintenance in vertebrates. *Annual Review of Biochemistry,* **66**: 409-435.
Shäffer H. 2003. *Chorology and diversity of the Azorean flora,* Borntraeger, Berlin, Germany, Dissertationes Botanicae Band.
Shaw J, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 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.
Shore JS, Mcqueen KL, Little SH. 1994. Inheritance of Plastid DNA in the *Turnera-Ulmifolia* Complex (Turneraceae). *American Journal of Botany,* **81**: 1636-1639.
Silvertown J. 2004. The ghost of competition past in the phylogeny of island endemic plants. *Journal of Ecology,* **92**: 168-173.
Silvertown J, Francisco-Ortega J, Carine M. 2005. The monophyly of island radiations: an evaluation of niche pre-emption and some alternative explanations. *Journal of Ecology,* **93**: 653-657.
Small DaA. 2001. *The Heather Society's Handy Guide to Heathers,* Denbeigh, England, The Heather Society.
Smith SE. 1989. Influence of Parental Genotype on Plastid Inheritance in *Medicago sativa*. *Journal of Heredity,* **80**: 214-217.
Smtih S.E. B, E.T., Fulton, R.W. 1986. Transmission of chlorophyll deficiencies in *Medicago sativa.* Evidence for biparental inheritance of plastids. *Journal of Heredity,* **77**: 35-38.
Soltis D, Soltis, P.S. 1998. *Molecular systematics of plants II - DNA sequencing,* Boston, Kluver Academic Publishers.
Soltis PS, Soltis DE, Chase MW. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature,* **402**: 402-4.
Stern DB, Lonsdale DM. 1982. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature,* **299**: 698-702.
Sunding P. 1979. Origin of the Macaronesian Flora. In: D.Bramwell ed. *Plants and Islands.* London UK, Academic Press.

Szovenyi P, Hock Z, Urmi E, Schneller JJ. 2006. Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *Sphagnum squarrosum* (Bryophyta, Sphagnopsida) in Europe. *New Phytol,* **172**: 784-94.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol,* **7**: 453-64.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal Primers for Amplification of 3 Noncoding Regions of Chloroplast DNA. *Plant Molecular Biology,* **17**: 1105-1109.

Tajima F, Nei M. 1984. Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol,* **1**: 269-285.

Takhtajan A. 1969. Flowering plants: origin and dispersal. Edinburgh, UK., Oliver and Boyd.

Tanabe K, Sakihama N, Farnert A, Rooth I, Bjorkman A, Walliker D, Ranford-Cartwright L. 2002. In vitro recombination during PCR of *Plasmodium falciparum* DNA: a potential pitfall in molecular population genetic analysis. *Molecular and Biochemical Parasitology,* **122**: 211-216.

Taylor JS, Breden F. 2002. The inheritance of heteroplasmy in guppies. *Journal of Fish Biology,* **60**: 1346-1350.

Tiercelin JJ, Gibert E, Umer M, Bonnefille R, Disnar JR, Lezine AM, Hureau-Mazaudier D, Travi Y, Keravis D, Lamb HF. 2008. High-resolution sedimentary record of the last deglaciation from a high-altitude lake in Ethiopia. *Quaternary Science Reviews,* **27**: 449-467.

Tiffney BH. 1985. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. *J. Arnold Arboretum,* **66**: 73-94.

Turner C, Hannon GE. 1988. Vegetational Evidence for Late Quaternary Climatic Changes in Southwest Europe in Relation to the Influence of the North Atlantic Ocean. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences,* **318**: 451-485.

Vanderpoorten A, Rumsey FJ, Carine MA. 2007. Does macaronesia exist? Conflicting signal in the bryophyte and pteridophyte floras. *American Journal of Botany,* **94**: 625-639.

Vijverberg K, Mes THM, Bachmann K. 1999. Chloroplast DNA evidence for the evolution of *Microseris* (Asteraceae) in Australia and New Zealand after long-distance dispersal from Western North America. *American Journal of Botany,* **86**: 1448-1463.

Vogl C, Badger J, Kearney P, Li M, Clegg M, Jiang T. 2003. Probabilistic analysis indicates discordant gene trees in chloroplast evolution. *J Mol Evol,* **56**: 330-40.

Wagner W, Funk V. 1995. *Hawaiian biogeography: Evolution on a Hot Spot Archipelago,* Washington D.C., Smithsonian institution Press.

Waters JM, Craw D. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. *Systematic Biology,* **55**: 351-356.

Webb PB, Berthelot S. 1840. *Histoire naturelle des Iles Canaries: Phytographica Canariensis,* Paris, France, Section 3. Bureaux.

Weir BS, Cockerham CC. 1984. Estimating f-statistics for the analysis of population structure. *Evolution,* **38**: 1358-1370.

Whittaker RJ, Triantis KA, Ladle RJ. 2008. A general dynamic theory of oceanic island biogeography. *Journal of Biogeography,* **35**: 977-994.
Wiley EO. 1988. Vicariance Biogeography. *Annual Review of Ecology and Systematics,* **19**: 513-542.
Wilkinson GS, Chapman AM. 1991. Length and Sequence Variation in Evening Bat D-Loop Mtdna. *Genetics,* **128**: 607-617.
Willis KJ, Niklas KJ. 2004. The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* **359**: 159-172.
Wilson EO. 1969. The species equilibrium. *Brookhaven Symposia in Biology,* **22**: 38-47.
Wolfe AD, Randle CP. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: Implications for plant molecular systematics. *Systematic Botany,* **29**: 1011-1020.
Xiang QY, Manchester SR, Thomas DT, Zhang WH, Fan CZ. 2005. Phylogeny, biogeography, and molecular dating of Cornelian cherries (*Cornus*, Cornaceae): Tracking tertiary plant migration. *Evolution,* **59**: 1685-1700.
Zurawski G, Clegg MT. 1987. Evolution of higher plant chloroplast DNA-coded genes: implications for structure-function and phylogenetic studies. *Annual Review of Plant Physiology,* **38**: 391-418.

Web sites:

http://www.heathersociety.org.uk/heather_family.html
http://commons.wikimedia.org/wiki/File:Erica_arborea_flors.jpg
http://botanicavirtual.udl.es/fam/*Erica*cies/*Erica*_for.htm#*Erica_scoparia*
http://www.planetefleurs.fr/Systematique/*Erica*ceae/*Erica*_platycodon.htm
http://www.horta.uac.pt/species/plantae/*Erica_scoparia*_azorica/*Erica_scoparia*_azorica.htm

Annex

Table I: Primers of MatK, rpl16 and atpB-rbcL loci redesigned from Erica arborea sequences due to difficulty of amplifications in this study.

Primers	Sequence
*Mat*K-*F*	5'AAGACTTCTAGTTCGATTTTT3'
*Mat*K-*R*	5'CGCTTATCTTTTCAGGAGTATA 3'
*rpl*16-*F*	5'CAACTCATCACTTCGTTTTATCTGG3'
*rpl*16-*R*	5'CAGCTCCTCGCGAATAAAAG3'
*Atp*B-*F*	5'TATATTCAAAAAGTCAATATTAGGGCGA3'
*Atp*B-*R*	5'TGAAATAAAAAGCGCCAATGAGATA3'

Table II: Tajima and Nei distance matrix used for Nst calculations for Erica arborea in SPAGEDI.

haplo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	19	22	23
1																		
2	0.00090																	
3	0.00180	0.00055																
4	0.00724	0.00436	0.00383															
5	0.00271	0.00109	0.00055	0.00438														
6	0.00815	0.00490	0.00438	0.00054	0.00493													
7	0.00180	0.00109	0.00055	0.00438	0.00109	0.00493												
8	0.00815	0.00490	0.00438	0.00054	0.00493	0.00108	0.00493											
9	0.00180	0.00054	0.00000	0.00381	0.00055	0.00436	0.00055	0.00436										
10	0.00180	0.00109	0.00055	0.00436	0.00109	0.00490	0.00109	0.00490	0.00054									
11	0.00180	0.00054	0.00000	0.00381	0.00055	0.00436	0.00055	0.00436	0.00000	0.00054								
12	0.00180	0.00109	0.00055	0.00438	0.00109	0.00493	0.00109	0.00493	0.00055	0.00109	0.00055							
13	0.00725	0.00436	0.00383	0.00000	0.00438	0.00054	0.00438	0.00054	0.00381	0.00436	0.00381	0.00438						
14	0.00180	0.00109	0.00055	0.00438	0.00109	0.00493	0.00109	0.00493	0.00055	0.00109	0.00055	0.00109	0.00438					
15	0.00633	0.00328	0.00273	0.00109	0.00328	0.00163	0.00328	0.00163	0.00273	0.00328	0.00273	0.00328	0.00109	0.00328				
19	0.00361	0.00163	0.00109	0.00491	0.00164	0.00545	0.00164	0.00545	0.00109	0.00163	0.00109	0.00164	0.00491	0.00164	0.00383			
22	0.00180	0.00054	0.00109	0.00490	0.00164	0.00545	0.00164	0.00545	0.00109	0.00163	0.00109	0.00164	0.00491	0.00164	0.00383	0.00218		
23	0.00180	0.00054	0.00109	0.00490	0.00164	0.00545	0.00164	0.00545	0.00109	0.00163	0.00109	0.00164	0.00491	0.00164	0.00383	0.00218	0.00109	
24	0.00180	0.00109	0.00055	0.00438	0.00109	0.00493	0.00109	0.00493	0.00055	0.00109	0.00055	0.00109	0.00438	0.00109	0.00328	0.00164	0.00164	0.00164

Table III: Genbank accession number of the 21 Erica species for the two genes MatK and atpB-rbcL used for the reconstruction of the Erica phylogeny.

Species	Genbank accessions number for *Mat*K gene	Genbank accessions number for *atp*B-*rbc*L spacer	Geographic location
Erica tetralix	U61340	AY520778	Europe
Erica erigena	AY517912	AY520766	Europe
Erica carnea	AY517908	AY520762	Europe
Erica sicula	U61341	AY520776	Europe
Erica multiflora	AY517917	AY520771	Europe
Erica terminalis	AY517921	AY520777	Europe
Erica cinerea	AY517909	AY520763	Europe
Erica australis	U61329	AY520761	Europe
Erica arborea	AY517907	AY520760	Europe, East Africa
Erica xanthina	AY517923	AY520779	South Africa
Erica urceolata	AY517928	AY520784	South Africa
Erica plukenetii	AY517920	AY520774	South Africa
Erica muscosa	AY517926	AY520782	South Africa
Erica axillaris	AY517925	AY520781	South Africa
Erica tristis	AY517924	AY520780	South Africa
Erica imbricata	AY517914	AY520768	South Africa
Erica hispidula	AY517913	AY520767	South Africa
Erica cetrata	AY517906	AY520759	South Africa
Erica rigidula	AY517927	AY520783	South Africa
Erica embothriifolia	AY517911	AY520765	South Africa
Erica oakesiorum	AY517918	AY520772	South Africa

Table IV:List of the samples used in this study including: country and date of collection, collector, species and sub-species, the number of gene amplified, the herbarium and the genbank accessions number for the three loci atpB-rbcL, MatK and rpl16

DNA	country	species	sub-species	date	collected by	Herbarium	Gene	*atp*B-*rbc*L	*mat*K	*rpl*16
6	Spain_Menorca	*E.arborea*		7/04/1967	Lorna F.Bowden and Patricia A.Sims	BM	4	GQ401367	GQ401526	GQ415773
7	Croatia	*E.arborea*		25/05/1962	B.E.E Duyfjes, R.Hengeveld and C.W.Van de Voet	BM	3	GQ401368	GQ401541	-
8	Spain_Ibiza	*E.arborea*		31/10/1971	H. Kuhbier and G. Finshow	BM	4	GQ401369	GQ401518	GQ415774
12	Croatia	*E.arborea*		1/07/1983	Doutrelepont	BM	4	GQ401370	GQ401529	GQ415775
13	Italy_Sardinia	*E.arborea*		4/04/1973	Humphries C.J. and I.B.K. Richardson	BM	4	GQ401371	GQ401525	GQ415776
19	Italy	*E.arborea*		7/06/1979	Davis, D and S Sutton	BM	2	GQ401372	GQ401540	-
22	SaudiArabia	*E.arborea*		16/10/1971	G. Popov	BM	4	GQ401373	GQ401581	GQ415777
23	Yemen	*E.arborea*		2/06/1978	J.R.I. Wood	BM	4	GQ401374	GQ401535	GQ415778
28	Tunisia	*E.arborea*		11/05/1975	Davis&Lamond	BM	4	GQ401376	GQ401662	GQ415779
30	Algeria	*E.arborea*		24/04/1971	Davis	BM	4	GQ401377	GQ401568	GQ415780
32	Morocco	*E.arborea*		19/04/1971	Davis	BM	4	GQ401378	GQ401548	GQ415781
43	Ethiopia	*E.arborea*		5/01/1990	G&S Miche	BM	4	GQ401379	GQ401589	GQ415782
44	Ethiopia	*E.arborea*		12/01/1990	G&S Miche	BM	4	GQ401380	GQ401516	GQ415783
47	Spain_LaGomera	*E.arborea*		7/03/1973	Angela E. Aldridge	BM	4	GQ401381	GQ401545	GQ415784
51	Spain_ElHierro	*E.arborea*		12/04/1977	Jarvis&Murphy	BM	4	GQ401382	GQ401546	GQ415785
55	Spain_ElHierro	*E.arborea*		12/04/1977	Jarvis&Murphy	BM	4	GQ401383	GQ401547	GQ415786
57	Portugal_Madeira	*E.arborea*		25/03/1984	J.R. Press and M.J Short	BM	4	GQ401384	GQ401661	GQ415787
59	Portugal_Madeira	*E.arborea*		4/04/1984	J.R.Press and M.J. Short	BM	4	GQ401385	GQ401544	GQ415788
60	Portugal_Madeira	*E.arborea*		8/03/1986	J.F.M., M.J.Cannon	BM	4	GQ401386	GQ401543	GQ415789
66	Algeria	*E.arborea*		16/04/1981	Herbier Robert Roncart	LG	4	GQ401387	GQ401557	GQ415790
68	Greece	*E.arborea*		30/04/1980	Fr Cernoch	LG	3	GQ401388	GQ401590	-
70	Portugal_Madeira	*E.arborea*		10/04/2007	M, Dewald, A, Hambuckers & E, Sérusiaux	LG	4	GQ401389	GQ401614	GQ415791
71	Portugal	*E.arborea*		22/04/1977	Malato-Beliz and J.A.Guerra	LG	4	GQ401390	GQ401610	GQ415792
74	Spain	*E.arborea*		18/05/1992	P.Monserrat	LG	4	GQ401391	GQ401523	GQ415793
78	France_Corsica	*E.arborea*		4/08/1982	J.J. Symoens	LG	4	GQ401393	GQ401527	GQ415794
79	France_Corsica	*E.arborea*		9/05/1980	P.Maquet	LG	2	-	GQ401521	-
82	France	*E.arborea*		15/05/1981	Kai Larsen	LG	4	GQ401394	GQ401522	GQ415795
86	Spain	*E.arborea*		22/06/1991	P.M. Uribe-Echebarria	LG	4	GQ401395	GQ401615	GQ415796
94	Morocco	*E.arborea*		7/05/1992	Fernandez Casas&Julian Molero	LG	2	-	-	GQ415797
110	Rwanda	*E.arborea*		19/02/1972	P.Auquier	LG	3	GQ401396	GQ401587	-
112	Ethiopia	*E.arborea*		9/01/1971	J.J.F.E De Wilde	LG	4	GQ401397	GQ401567	GQ415798
113	Kenya	*E.arborea*		26/07/1975	Windels G.	LG	4	GQ401398	GQ401559	GQ415799
127	Spain	*E.arborea*		27/05/1977	J. Bouharmont	BR	4	GQ401399	GQ401617	GQ415800
138	France_Corsica	*E.arborea*		2/06/1972	C. Vanden Berghen	BR	4	GQ401400	GQ401528	GQ415801
139	France_Corsica	*E.arborea*		4/06/1973	C. Vanden Berghen	BR	2	GQ401401	-	-
140	Spain_Mallorca	*E.arborea*		11/04/2005	C. Vanden Berghen	BR	4	GQ401402	GQ401530	GQ415802
141	Portugal	*E.arborea*		14/08/1959	J. Duvigneaud	BR	4	GQ401403	GQ401612	GQ415803
143	Italy_Sicile	*E.arborea*		5/06/2000	S. Castroviejo	BR	4	GQ401404	GQ401564	GQ415804
147	Montenegro	*E.arborea*		9/08/1976	C. Vanden Berghen	BR	4	GQ401405	GQ401532	GQ415805
148	Greece _Crete	*E.arborea*		1/05/1984	R. Deshamps	BR	4	GQ401406	GQ401531	GQ415806
149	Greece _Crete	*E.arborea*		12/05/1984	R. Deshamps	BR	4	GQ401407	GQ401566	GQ415807
150	Greece _Crete	*E.arborea*		2/04/1980	C. Evrard	BR	4	GQ401408	GQ401533	GQ415808
158	Spain_Tenerife	*E.arborea*		5/03/1994	V.Vasak	BR	4	GQ401409	GQ401534	GQ415810
159	Spain_Tenerife	*E.arborea*		5/07/1977	Cajon. A	BR	4	GQ401410	GQ401538	GQ415811
160	Spain_GranCanaria	*E.arborea*		7/02/1993	V.Vasak	BR	3	-	GQ401537	GQ415812
168	Portugal_Madeira	*E.arborea*		15/03/1984	J.R. Press&M.J Short	BR	4	GQ401411	GQ401618	GQ415813
172	Croatia	*E.arborea*		1/07/1983	Doutrelepont	BR	4	GQ401412	GQ401536	GQ415814
174	Spain	*E.arborea*		16/03/1997	X.Lizaur	BR	4	GQ401413	GQ401539	GQ415815
175	Rwanda	*E.arborea*		3/10/1974	J. Rameloo	BR	3	GQ401414	GQ401588	-
176	Ethiopia	*E.arborea*		4/11/1988	I.Friis, A. Michelsen&Sebsese Demiseu	BR	4	GQ401415	GQ401542	GQ415816
231	Spain_LaPalma	*E.arborea*		16/11/2008	A Santo	TFC	4	GQ401439	GQ401580	GQ415825
238	Spain_LaPalma	*E.arborea*		1/11/2008	Julio Leal	TFC	4	GQ401440	GQ401519	GQ415826
239	Spain_LaPalma	*E.arborea*		1/11/2008	Julio Leal	TFC	4	GQ401441	GQ401520	GQ415827
240	Spain_LaPalma	*E.arborea*		1/11/2008	Julio Leal	TFC	4	GQ401442	GQ401517	GQ415828
244	Spain_LaPalma	*E.arborea*		1/11/2008	Julio Leal	TFC	4	GQ401444	GQ401515	GQ415829
246	Spain_LaGomera	*E.arborea*		1/11/2008	Gonzeles Manchebo	TFC	4	GQ401446	GQ401579	GQ415830
247	Spain_LaGomera	*E.arborea*		1/11/2008	Gonzeles Manchebo	TFC	4	GQ401447	GQ401562	GQ415831
249	Spain_LaGomera	*E.arborea*		1/11/2008	Gonzeles Manchebo	TFC	4	GQ401448	GQ401560	-
300	Spain_GranCanaria	*E.arborea*		27/01/2009	A. Vanderpoorten	LG	4	GQ401459	GQ401578	GQ415841
301	Spain_GranCanaria	*E.arborea*		28/01/2009	A. Vanderpoorten	LG	4	GQ401460	GQ401582	GQ415842
302	Spain_GranCanaria	*E.arborea*		29/01/2009	A. Vanderpoorten	LG	4	GQ401461	GQ401561	GQ415843
303	Spain_GranCanaria	*E.arborea*		30/01/2009	A. Vanderpoorten	LG	4	GQ401462	GQ401563	GQ415844
306	Spain	*E.arborea*		1/01/2009	J.M Gonzalez-Mancebo	TFC	4	GQ401465	GQ401616	GQ415847
307	Spain	*E.arborea*		1/01/2009	J.M Gonzalez-Mancebo	TFC	4	GQ401466	GQ401613	GQ415848
308	Spain_Tenerife	*E.arborea*		7/12/2008	J.M Gonzalez-Mancebo	TFC	4	GQ401467	GQ401570	GQ415849
309	Spain_Tenerife	*E.arborea*		7/12/2008	J.M Gonzalez-Mancebo	TFC	4	GQ401468	GQ401569	GQ415850
310	Sahara	*E.arborea*		1/09/1956	P.Quezel	AIX	4	GQ401469	GQ401592	GQ415851
311	Tchad_EmiKoussi	*E.arborea*		1/09/1958	P.Quezel	AIX	4	GQ401470	GQ401593	GQ415852
312	Tchad_EmiKoussi	*E.arborea*		1/09/1956	P.Quezel	AIX	4	GQ401471	GQ401594	GQ415853
313	Turkey	*E.arborea*		23/04/1963	C.Tobey	E	3	GQ401472	GQ401595	-
315	Turkey	*E.arborea*		24/04/1965	P.Davis	E	4	GQ401473	GQ401565	GQ415854
316	Turkey	*E.arborea*		22/04/1966	P.Davis	E	4	GQ401474	GQ401596	GQ415855
317	Turkey	*E.arborea*		14/07/1962	P.Davis	E	4	GQ401475	GQ401549	GQ415856
320	Turkey	*E.arborea*		15/04/1965	P.Davis	E	3	GQ401476	GQ401597	GQ415857

DNA	country	species	sub-species	date	collected by	Herbarium	Gene	*atp*B-*rbc*L	*mat*K	*rpl*16
321	Greece	*E.arborea*		21/04/1965	M.Coode&B.M.G Jones	E	4	GQ401477	GQ401598	GQ415858
322	Greece	*E.arborea*		2/04/1965	P.Davis	E	3	GQ401478	GQ401599	-
323	Spain	*E.arborea*		11/11/1965	S. Silvestre	RNG	3	GQ401479	GQ401600	GQ415859
324	Spain	*E.arborea*		11/01/1979	J. Rivera	RNG	4	GQ401480	GQ401558	GQ415860
325	Spain	*E.arborea*		12/04/1988	J. Herrera, C. Romero & S. Talavera	RNG	3	GQ401481	GQ401601	GQ415861
327	Spain	*E.arborea*		6/07/1978	P. Harrold & R.J.D. McBeath	RNG	4	GQ401482	GQ401607	GQ415862
328	Spain	*E.arborea*		10/07/1969	P.E. Gibbs	RNG	4	GQ401483	GQ401608	GQ415863
329	Spain	*E.arborea*		12/03/1978	B. Cabezudo	RNG	4	GQ401484	GQ401552	GQ415864
330	Portugal	*E.arborea*		17/03/1986	J. Maxted	RNG	4	GQ401485	GQ401602	GQ415865
333	Portugal	*E.arborea*		17/04/1981	L.A. Grandvauv Barbosa	RNG	4	GQ401486	GQ401603	GQ415866
335	Portugal	*E.arborea*		8/03/1995	H.J.M. Bowen	RNG	4	GQ401487	GQ401550	GQ415867
336	Portugal	*E.arborea*		4/04/2001	H.J.M. Bowen	RNG	2	GQ401488	GQ401609	-
337	France	*E.arborea*		20/04/1971	M. Greenway	RNG	4	GQ401489	GQ401551	GQ415868
339	France	*E.arborea*		1/03/1974	E. Coppejans	RNG	4	GQ401490	GQ401604	GQ415869
340	France	*E.arborea*		7/09/1980	N. Thomson	RNG	4	GQ401491	GQ401575	GQ415870
341	Italy	*E.arborea*		15/06/1997	Optima Iter 8	RNG	4	GQ401492	GQ401576	GQ415871
342	Italy	*E.arborea*		1/06/1997	Optima Iter 8	RNG	4	GQ401493	GQ401577	GQ415872
343	Italy	*E.arborea*		12/06/1997	Optima Iter 8	RNG	4	GQ401494	GQ401574	GQ415873
344	Italy	*E.arborea*		4/06/2003	C. Navarro et al.	RNG	4	GQ401495	GQ401524	GQ415874
345	Greece	*E.arborea*		24/05/1993	L.C. Jury & K. Warren	RNG	3	GQ401496	GQ401573	GQ415875
346	Greece	*E.arborea*		3/04/1972	E. Stamatiadou	RNG	4	GQ401497	GQ401605	GQ415876
347	Greece	*E.arborea*		3/01/1978	H.J.M. Bowen	RNG	4	GQ401498	GQ401572	GQ415877
354	Turkey	*E.arborea*		25/07/1988	R.m. Nesbitt & D. Samule	RNG	4	GQ401504	GQ401606	GQ415882
355	Morocco	*E.arborea*		24/04/1995	S.L. Jury	RNG	4	GQ401505	GQ401571	GQ415883
356	Morocco	*E.arborea*		19/06/1992	Optima Iter V	RNG	4	GQ401506	GQ401553	GQ415884
357	Morocco	*E.arborea*		9/04/1994	J.M. Montserrat & B. Valdes	RNG	4	GQ401507	GQ401554	GQ415885
358	Morocco	*E.arborea*		27/02/1994	S.L. Jury et al	RNG	4	GQ401508	GQ401556	GQ415886
359	Algeria	*E.arborea*		28/04/1976	D.A. & S.J. Sutton	RNG	2	GQ401509	-	-
371	Spain_LaGomera	*E.arborea*		1/12/2008	P.N. Garajonay	LPA	3	GQ401510	GQ401555	GQ415887
376	Ethiopia	*E.arborea*		7/01/2008	M. Popp and C. Brochmann	O	4	GQ401511	GQ401583	GQ415888
380	Ethiopia	*E.arborea*		7/01/2008	M. Popp and C. Brochmann	O	3	GQ401512	GQ401584	GQ415889
383	Uganda	*E.arborea*		7/01/2008	M. Popp and C. Brochmann	O	3	GQ401513	GQ401585	GQ415890
384	Kenya	*E.arborea*		7/01/2008	M. Popp and C. Brochmann	O	3	GQ401514	GQ401586	GQ415891
180	France	*E.scoparia*		9/06/1992	D.Podlech	LG	4	GQ401417	GQ401626	GQ415817
182	France	*E.scoparia*		22/05/1972	D. Podlech	LG	2	GQ401418	-	-
188	Italy	*E.scoparia*		19/05/1972	A.Dieterle	LG	2	GQ401420	GQ401628	-
189	Spain	*E.scoparia*		23/03/1985	A. Segura-zubizarretta	LG	4	GQ401421	GQ401629	GQ415818
190	Morocco	*E.scoparia*		16/05/1985	F.Damblon	LG	2	GQ401422	-	-
191	Portugal_Madeira	*E.scoparia*	*maderincola*	10/04/2007	M.Dewald	LG	4	GQ401423	GQ401646	GQ415819
192	Tunisia	*E.scoparia*		19/05/1983	J.W.A Jansen	BR	2	GQ401424	-	-
194	Morocco	*E.scoparia*		10/07/1988	F.Damblon	BR	4	GQ401425	GQ401634	GQ415820
196	Portugal	*E.scoparia*		12/07/1977	J.E De Langhe	BR	3	GQ401426	GQ401663	GQ415821
199	France_Corsica	*E.scoparia*		16/09/1980	J.Lambinon	BR	4	GQ401427	GQ401620	GQ415822
204	France_Corsica	*E.scoparia*		2/07/1970	J. Léonard	BR	2	GQ401430	-	-
208	Spain_Tenerife	*E.scoparia*	*platycodon*	3/01/1972	J.Duvigneaud	BR	2	GQ401431	-	-
210	Spain_Tenerife	*E.scoparia*	*platycodon*	8/03/1994	V.Vasak	BR	2	GQ401432	GQ401638	-
214	France	*E.scoparia*		1/06/1956	L.Delvosalle	BR	3	GQ401434	GQ401591	GQ415823
216	Spain_Menorca	*E.scoparia*		14/08/1967	J. Duvigneaud	BR	3	GQ401435	GQ401633	-
218	Spain	*E.scoparia*		22/04/1957	J. Duvigneaud	BR	2	GQ401436	-	GQ415824
252	Spain_Tenerife	*E.scoparia*	*platycodon*	1/11/2008	Gonzeles Manchebo	TFC	4	GQ401449	GQ401637	GQ415832
256	Portugal_Azores	*E.scoparia*	*azorica*	23/04/1972	B. Goncalves	BM	3	GQ401450	GQ401627	GQ415833
259	Portugal_Azores	*E.scoparia*	*azorica*	12/08/1972	B. Goncalves	BM	2	GQ401451	GQ401632	-
273	Portugal_Madeira	*E.scoparia*	*maderincola*	29/07/1981	R.J Hampshire	BM	3	GQ401452	GQ401631	GQ415834
280	Portugal_Azores	*E.scoparia*	*azorica*	6/01/2009	A.Désamoré	LG	3	GQ401453	GQ401652	GQ415835
288	Portugal_Azores	*E.scoparia*	*azorica*	8/01/2009	A.Désamoré	LG	3	GQ401454	GQ401641	GQ415836
291	Portugal_Azores	*E.scoparia*	*azorica*	10/01/2009	A.Désamoré	LG	3	GQ401455	GQ401659	GQ415837
294	Portugal_Azores	*E.scoparia*	*azorica*	11/01/2009	A.Désamoré	LG	3	GQ401456	GQ401644	GQ415838
298	Portugal_Azores	*E.scoparia*	*azorica*	20/01/2009	A.Désamoré	LG	3	GQ401457	GQ401621	GQ415839
299	Portugal_Azores	*E.scoparia*	*azorica*	20/01/2009	A.Désamoré	LG	3	GQ401458	GQ401619	GQ415840
304	Spain_LaGomera	*E.scoparia*	*platycodon*	11/01/2009	J.M Gonzalez-Mancebo	TFC	3	GQ401463	GQ401649	GQ415845
305	Spain_LaGomera	*E.scoparia*	*platycodon*	11/01/2009	J.M Gonzalez-Mancebo	TFC	3	GQ401464	GQ401650	GQ415846
348	Spain	*E.scoparia*	*platycodon*	26/04/1994	Optima Iter VI	RNG	3	GQ401499	GQ401648	GQ415878
350	Spain	*E.scoparia*	*platycodon*	11/01/1978	J. Rivera	RNG	3	GQ401500	GQ401647	GQ415879
351	Spain	*E.scoparia*		4/01/1987	H.J.M. Bowen	RNG	4	GQ401501	GQ401624	GQ415880
352	Portugal	*E.scoparia*		12/04/1967	E.V. Watson	RNG	2	GQ401502	GQ401636	-
353	Portugal	*E.scoparia*		7/03/1995	H.J.M. Bowen	RNG	3	GQ401503	GQ401635	GQ415881

Appendix

Heteroplasmy, recombination and hybridization in the chloroplast DNA of the heather Erica arborea

Introduction

Despite its relatively small size (mean = 150 kbp (Clegg et al., 1994)) and comparatively slow rate of evolution (Clegg, 1993, Clegg et al., 1994), chloroplast DNA (cp DNA) remains a primary source of data for phylogenetic analysis in plants (Soltis et al., 1999, Shaw et al., 2005). Chloroplast genes exist as single copies (Soltis, 1998) and the amplification and sequencing of cpDNA loci is relatively straightforward thanks to the existence of universal primers (Taberlet et al., 1991, Demesure et al., 1995, Dumolin-Lapegue et al., 1997). It is also widely accepted that cpDNA exhibits uniparental inheritance (Birky, 1995, Corriveau and Coleman, 1988), and non-recombination (Palmer et al., 1988, Doyle, 1992, Vogl et al., 2003, Clegg, 1993, Wolfe and Randle, 2004) and that, consequently, it is immune from the problems that recombination and paralogy may pose for phylogeny reconstruction.

Whilst there are many chloroplasts within each plant cell, with each containing several copies of the cp genome (for instance, as many as 900 genomes per chloroplast and 50,000 such genomes per cell (Bendich, 1987)), studies to date indicate that in most plant lineages, progeny exhibit only one of the two parental chloroplast haplotypes because of selection against deleterious mutations, vegetative sorting, or differential disintegration of organelles or organellar DNA in the zygote (Wolfe and Randle, 2004). Chloroplastic heteroplasmy, the condition in which cells have more than one chloroplast type, may arise from independent mutations in organelle genomes or when organellar transmission is 'leaky', allowing occasional biparental inheritance and it has been reported in several flowering plants (*Silene*, *Medicago*, *Coreopsis* and *Cynomorium*) and in the conifer (*Chamaecyparis)*(Wolfe and Randle, 2004). However, where it does occur, heteroplasmy is normally an unstable condition. (Wolfe and Randle, 2004)

Evidence for recombination of cpDNA is extremely limited. It can be induced *in vitro* by means of interspecific somatic cell fusion, resulting in viable cytoplasmic hybrid lines (Medgyesy et al., 1985), but in the wild, it has been reported in only two species of conifer (Huang et al., 2001, Marshall et al., 2001) although it is also suggested by genomic rearrangements in Asteraceae (Vijverberg et al., 1999). There has therefore been little evidence to date that challenges the view of the cp genome as a non-recombining, uniparentally-inherited molecule.

In this paper, we document extensive infra-individual polymorphism within the chloroplast genome of the tree heather *Erica arborea* (Ericaceae) and provide evidence for recombination between the cpDNA genomes of *E. arborea* and the related *E. scoparia.* Our observations suggest that heteroplasmy and recombination within the chloroplast genome of angiosperms might be more common than previously thought and

suggest that the consequences of this phenomenon for phylogenetic reconstruction demand greater attention.

Material and methods

In total, 136 individual samples, including 105 accessions for *E. arborea* and 31 for *E. scoparia,* were obtained from the entire distribution range of each species *(Table IV in annex)*. Genomic DNA was isolated from dried leaf tissue, using a CTAB protocol (Doyle and Doyle, 1987) and further purified with Geneclean Kit following the manufacturer's protocol. Each specimen was genotyped at three cpDNA loci, namely the *atp*B-*rbc*L intergenic spacer, the 5' end of *Mat*K, (McGuire and Kron, 2005) and *rpl*16 (Shaw et al., 2005). *Mat*K was amplified and sequenced using whether published primers (McGuire and Kron, 2005) or specific primers designed for the study (*Mat*K-F: 5'AAGACTTCTAGTTCGATTTTT3'; *Mat*K-R: 5'CGCTTATCTTTTCAGGAGTATA 3'). The *rpl*16 locus was amplified using primers published by Shaw et al (Shaw et al., 2005) or the primer pair *rpl*16-F: 5'CAACTCATCACTTCGTTTTATCTGG3' and *rpl*16-R: 5'CAGCTCCTCGCGAATAAAAG3' designed for this study.
Primers to amplify the *atp*B-*rbc*L spacer were designed from sequences of *Erica arborea* available on Genbank (McGuire and Kron, 2005):
AtpB-F: 5'TATATTCAAAAAGTCAATATTAGGGCGA3';
AtpB-R: 5'TGAAATAAAAAGCGCCAATGAGATA3'.

Polymerase chain reactions (PCR) were carried out in 15µl volumes reaction using 1.5µl of 10X reaction buffer, 2.4µl of dNTPs mix (1mM each), 0.6µl of 50mM $MgCl_2$, 0.75 of each primer (10µM), 1.125µl of BSA, 0.3µl of taq DNA polymerase and 1µl of DNA. . After initial denaturation at 95° for 2 min, 35 amplification cycles were performed comprising 30s denaturation at 95°, 45s of annealing at 50°, 2min of extension at 72° followed by 7min at 72°. This protocol was used for all genes with some modifications depending on the age of samples and amplification easiness. Sequences were edited using Sequencher 3.1. Contigs were constructed from single-stranded forward and reverse sequences using Sequencher 3.1. Sequences were aligned manually using Se-al.2.0a11(Rambaut, 1995) and gaps were inserted where necessary to preserve positional homology.

Chromatograms were examined for evidence of double peaks of approximately equal height, indicative of the superposition of gene copies. Where these were observed, PCR cloning was used to identify individual gene copies. PCR products were cloned using the TOPO-TA kit following the manufacturer's instructions. In order to duplicate manipulation to avoid problems of contamination, samples were reextracted from the original material and sequenced again.

Results

Nineteen (A1-A19) and 11 (S1-S11) haplotypes were identified in *E. arborea* and *E. scoparia*, respectively. The two species differed by 31 positions: 9 in *Mat*K, 8 in the *atp*B-*rbc*L spacer and 14 in *rpl*16. Eighteen individuals were found that displayed a clear

pattern of sequence additivity at positions in the *Mat*K region that distinguished *E. arborea* and *E. scoparia* (Fig. 1). The presence of two alleles, one of *E. arborea* (allele A1 in 14 individuals and allele A2 in 4 individuals; (see Table 1) and one of *E. scoparia* (allele S1), was confirmed by cloning. The 14 individuals with *mat*K haplotype A1 also displayed sequence additivity between *E. scoparia* and *E. arborea* at all 8 sites in *atp*B-*rbc*L that segregated between the two species The presence of two *atp*B-*rbc*L alleles in these individuals was again confirmed by cloning. In contrast, the 4 individuals that exhibited *Mat*K haplotype A2 did not show evidence of polymorphism at any of the 8 segregating sites of *atp*B-*rbc*L between *E. arborea* and *E. scoparia,* with all four individuals exhibiting the *E. arborea* A2 haplotype sequence (Table 1.) In *rpl*16, none of the segregating sites between *E. arborea* and *E. scoparia* exhibited conflicting base calls.

Discussion

Multiple copies of the cpDNA *mat*K gene and *atp*B-*rbc*L intergenic spacer were discovered in individual genomes of the heather *E. arborea* sampled from the wild. Although such observations are almost never reported in plant phylogenetic studies, the presence of multiple copies of chloroplast genes can occur following gene duplication and transfer towards the nucleus, or heteroplasmy. The case of mitochondrial genes transferred into the nucleus is often reported in eukaryotes (Bensasson et al., 2001); indeed most regulatory cytoplasmic genes were relocated to the nucleus in their early evolutionary history (Martin and Herrmann, 1998) and in vertebrates, more than 100 proteins necessary for mitochondrial function are encoded in the nucleus (Shadel and Clayton, 1997). In *Arabidopsis*, eight mitochondrial genes have been transferred from the mitochondrion to the nucleus (Adams et al., 2002) while in the holoparasitic *Orobanche cumana,* the chloroplast *rbcL* gene is present in both the chloroplast and nucleus (Delavault and Thalouarn, 2002). Gene duplication and transfer is probably enhanced in plants owing to the presence of three genomic compartments and transfers from the chloroplast to the mitochondrion have also been reported (Hoch et al., 1991, Stern and Lonsdale, 1982).

In the case of *Erica arborea,* the presence of two copies of *Mat*K and *atp*B-*rbc*L in individuals of *E. arborea* would appear to be the result of heteroplasmy owing to biparental inheritance through interspecific hybridization. Duplication and transfer of *Mat*K and *atpb*B-*rbc*L to another compartment seems highly unlikely. The two loci are well separated on the cpDNA genome (Bausher et al., 2006, Lee et al., 2006, Sato et al., 1999) meaning that multiple duplication and transfer events or an almost complete transfer of the whole cpDNA genome into the nucleus would be necessary to explain the pattern. Furthermore, the two copies found precisely correspond to those found in *E. scoparia* and *E. arborea*. The presence of two distinct *Mat*K haplotype combinations within *E. arborea*, (i.e., A1/S1 and A2/S1) indicates that at least two hybridization events have occurred.

Chloroplast inheritance has long been thought to be maternal (Birky, 1995, Corriveau and Coleman, 1988) although several studies have documented cases of paternal inheritance (*Medicago* (Smith, 1989), *Oenothera* (Chiu et al., 1988, M. Metzalff,

1981) *Pelargonium* (M. Metzalff, 1981), *Turnera* (Shore et al., 1994)) and biparental cpDNA inheritance is believed to occur, at least sometimes, in nearly a third of the embryo of the angiosperms surveyed (Smith, 1989). In most lineages, however, only one of the two parental copies remains at maturity, following sorting out mechanisms (Birky, 1983, Smtih S.E., 1986, Mogensen, 1996). However, the persistence of the two parental copies within genomes of *E. arborea*, along with similar observations in *Passiflora* (Hansen et al., 2007) and *Oenothera* (Chiu et al., 1988), suggest, that those sorting mechanisms might not be as efficient as previously thought. Chloroplast heteroplasmy can be a stable state as is the case in mitochondrial heteroplasmy in animals (Burzynski et al., 2003) (Kvist et al., 2003, Taylor and Breden, 2002, Harrison et al., 1985, Bendall and Sykes, 1995, Howell et al., 1992, Wilkinson and Chapman, 1991, Nesbo et al., 1998).

Failure to detect heteroplasmy may be because cpDNA is assumed to be uniparentally inherited and non-recombining so that any polymorphisms observed are generally assumed to be 'noise' owing to sequencing error and are coded as ambiguous. Distinguishing haplotype polymorphism from Taq DNA polymerase replication error when using PCR cloning techniques to examine sequence divergence of this nature (Gunther et al., 1998, Jacobs et al., 1999, Tanabe et al., 2002) can be problematic unless the likely source of the polymorphisms is known, as is the case in *Erica* where all polymorphisms occur at the segregating sites between the two parental species of the hybrids.

Whilst eighteen individuals exhibited heteroplasmy in the *Mat*K gene, only fourteen were polymorphic for the *atp*B-*rbc*L spacer, and none of the sequences of the *rpl*16 displayed any signal for the presence of two copies in the chromatograms. Whilst the possibility that a potential second copy failed to be amplified by PCR cannot be completely ruled out, the observation does suggest recombination within the cpDNA genome. Whilst cpDNA recombination has been successfully achieved in vitro through cell fusion between two different species of *Nicotinia* (Medgyesy et al., 1985), until now, it was documented in only two spermatophytes *Pinus contorta* (Marshall et al., 2001) and *Cycas* (Huang et al., 2001).

Our observations suggest that heteroplasmy owing to biparental cpDNA inheritance and subsequent recombination might be much more common in the wild than previously acknowledged. This has long been evident to animal systematists employing mtDNA. The assumption that recombination and heteroplasmy are not of concern if sequence data are obtained from cpDNA loci, as Wolfe & Randle (Wolfe and Randle, 2004) emphasized, is untenable.

Fig. 1: Electropherogram of a portion of MatK at two of the 9 diagnostic sites between E. arborea and E. scoparia. (a) typical E. arborea sequence. (b) typical E. scoparia sequence. (c) and (d): Sequential additivity at the two diagnostic sites within 18 E. arborea individuals (arrows).

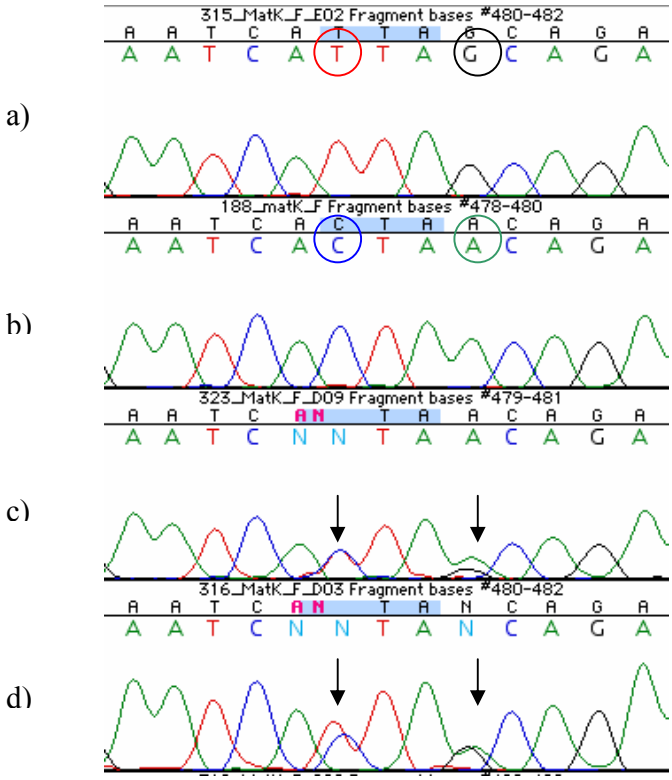


Table 1: Summary of the cpDNA polymorphic sites.
K= GT; Y=CT ; R=AG ; W=AT ; M=AC; S=CG; from IUPAC ambiguity code, and # is chosen when an insert (-) and an other base (A or T) are present.
E. arborea-A1 = A1 haplotype of E. arborea; E. arborea-A2= A2 haplotype of E. arborea; Hybrid 1= hybrid between haplotype A1 of E. arborea and E. scoparia; Hybrid 2= haplotype A2 of E. arborea and E. scoparia

	MatK (585bp)	AtpB-rbcL spacer (253bp)
E. arborea-A1	G C C A A A A T G T G C C	A T C - - - - - - G G A
E. arborea-A2	. . T T . . . G T . . .	C . . - - - - - - - . . .
E. scoparia	T T . . G T C . . C A T G	C G A T T T T T A T C G
Hybrid 1	**K Y** C A **R W M** T G **Y R Y S**	**M K M # # # # # # K S R**
Hybrid 2	**K Y Y W R W M K K Y R Y S**	C . . - - - - - - - . . .

References

Adams KL, Qiu YL, Stoutemyer M, Palmer JD. 2002. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **99**: 9905-9912.

Bausher MG, Singh ND, Lee SB, Jansen RK, Daniell H. 2006. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol,* **6**: 21.

Bendall KE, Sykes BC. 1995. Length Heteroplasmy in the First Hypervariable Segment of the Human Mtdna Control Region. *American Journal of Human Genetics,* **57**: 248-256.

Bendich AJ. 1987. Why Do Chloroplasts and Mitochondria Contain So Many Copies of Their Genome. *Bioessays,* **6**: 279-282.

Bensasson D, Zhang DX, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution,* **16**: 314-321.

Birky CW, Jr. 1983. Relaxed cellular controls and organelle heredity. *Science,* **222**: 468-75.

Birky CW, Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci U S A,* **92**: 11331-8.

Burzynski A, Zbawicka M, Skibinski DOF, Wenne R. 2003. Evidence for recombination of mtDNA in the marine mussel *Mytilus trossulus* from the Baltic. *Molecular Biology and Evolution,* **20**: 388-392.

Chiu WL, Stubbe W, Sears BB. 1988. Plastid Inheritance in *Oenothera* - Organelle Genome Modifies the Extent of Biparental Plastid Transmission. *Current Genetics,* **13**: 181-189.

Clegg MT. 1993. Chloroplast Gene-Sequences and the Study of Plant Evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **90**: 363-367.

Clegg MT, Gaut BS, Learn GH, Morton BR. 1994. Rates and Patterns of Chloroplast DNA Evolution. *Proceedings of the National Academy of Sciences of the United States of America,* **91**: 6795-6801.

Corriveau JL, Coleman AW. 1988. Rapid Screening Method to Detect Potential Biparental Inheritance of Plastid DNA and Results for over 200 Angiosperm Species. *American Journal of Botany,* **75**: 1443-1458.

Delavault P, Thalouarn P. 2002. The obligate root parasite *Orobanche cumana* exhibits several rbcL sequences. *Gene,* **297**: 85-92.

Demesure B, Sodzi N, Petit RJ. 1995. A Set of Universal Primers for Amplification of Polymorphic Noncoding Regions of Mitochondrial and Chloroplast DNA in Plants. *Molecular Ecology,* **4**: 129-131.

Doyle JJ. 1992. Gene Trees and Species Trees - Molecular Systematics as One-Character Taxonomy. *Systematic Botany,* **17**: 144-163.

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

Dumolin-Lapegue S, Pemonge MH, Petit RJ. 1997. An enlarged set of consensus primers for the study of organelle DNA in plants. *Mol Ecol,* **6**: 393-7.

Gunther S, Sommer G, von Breunig F, Iwanska A, Kalinina T, Sterneck M, Will H. 1998. Amplification of full-length hepatitis B virus genomes from samples from patients with low levels of viremia: Frequency and functional consequences of PCR-introduced mutations. *Journal of Clinical Microbiology,* **36**: 531-538.

Hansen AK, Escobar LK, Gilbert LE, Jansen RK. 2007. Paternal, maternal, and biparental inheritance of the chloroplast genome in Passiflora (Passifloraceae): Implications for phylogenetic studies. *American Journal of Botany,* **94**: 42-46.

Harrison RG, Rand DM, Wheeler WC. 1985. Mitochondrial DNA Size Variation Within Individual Crickets. *Science,* **228**: 1446-1448.

Hoch B, Maier RM, Appel K, Igloi GL, Kossel H. 1991. Editing of a chloroplast mRNA by creation of an initiation codon. *Nature,* **353**: 178-80.

Howell N, Halvorson S, Kubacka I, Mccullough DA, Bindoff LA, Turnbull DM. 1992. Mitochondrial Gene Segregation in Mammals - Is the Bottleneck Always Narrow. *Human Genetics,* **90**: 117-120.

Huang S, Chiang YC, Schaal BA, Chou CH, Chiang TY. 2001. Organelle DNA phylogeography of *Cycas taitungensis*, a relict species in Taiwan. *Molecular Ecology,* **10**: 2669-2681.

Jacobs G, Tscholl E, Sek A, Pfreundschuh M, Daus H, Trumper L. 1999. Enrichment polymerase chain reaction for the detection of Ki-ras mutations: relevance of Taq polymerase error rate, initial DNA copy number, and reaction conditions on the emergence of false-positive mutant bands. *Journal of Cancer Research and Clinical Oncology,* **125**: 395-401.

Kvist L, Martens J, Nazarenko AA, Orell M. 2003. Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution,* **20**: 243-247.

Lee SB, Kaittanis C, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H. 2006. The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms. *BMC Genomics,* **7**: 61.

M. Metzalff TB, R. Hagemann. 1981. Variations of chloroplast DNA in the genus *Pelargonium* and their biparental inheritance. *Theoretical and Applied Genetics,* **60**: 37-41.

Marshall HD, Newton C, Ritland K. 2001. Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. *Molecular Biology and Evolution,* **18**: 2136-2138.

Martin W, Herrmann RG. 1998. Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiology,* **118**: 9-17.

McGuire AF, Kron KA. 2005. Phylogenetic relationships of European and African Ericas. *International Journal of Plant Sciences,* **166**: 311-318.

Medgyesy P, Fejes E, Maliga P. 1985. Interspecific chloroplast recombination in a Nicotiana somatic hybrid. *Proc Natl Acad Sci U S A,* **82**: 6960-6964.

Mogensen HL. 1996. The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany,* **83**: 383-404.

Nesbo CL, Arab MO, Jakobsen KS. 1998. Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (Perca fluviatilis, Acerina cernua, Stizostedion lucioperca). *Genetics,* **148**: 1907-1919.

Palmer JD, Jansen RK, Michaels HJ, Chase MW, Manhart JR. 1988. Chloroplast DNA Variation and Plant Phylogeny. *Annals of the Missouri Botanical Garden,* **75**: 1180-1206.

Rambaut A. 1995. Se–Al: Sequence alignment program. In: University O ed.

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. 1999. Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res,* **6**: 283-90.

Shadel GS, Clayton DA. 1997. Mitochondrial DNA maintenance in vertebrates. *Annual Review of Biochemistry,* **66**: 409-435.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 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.

Shore JS, Mcqueen KL, Little SH. 1994. Inheritance of Plastid DNA in the *Turnera-Ulmifolia* Complex (Turneraceae). *American Journal of Botany,* **81**: 1636-1639.

Smith SE. 1989. Influence of Parental Genotype on Plastid Inheritance in *Medicago sativa*. *Journal of Heredity,* **80**: 214-217.

Smtih S.E. B, E.T., Fulton, R.W. 1986. Transmission of chlorophyll deficiencies in *Medicago sativa.* Evidence for biparental inheritance of plastids. *Journal of Heredity,* **77**: 35-38.

Soltis D, Soltis, P.S. 1998. *Molecular systematics of plants II - DNA sequencing,* Boston, Kluver Academic Publishers.

Soltis PS, Soltis DE, Chase MW. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature,* **402**: 402-4.

Stern DB, Lonsdale DM. 1982. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature,* **299**: 698-702.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal Primers for Amplification of 3 Noncoding Regions of Chloroplast DNA. *Plant Molecular Biology,* **17**: 1105-1109.

Tanabe K, Sakihama N, Farnert A, Rooth I, Bjorkman A, Walliker D, Ranford-Cartwright L. 2002. In vitro recombination during PCR of *Plasmodium falciparum* DNA: a potential pitfall in molecular population genetic analysis. *Molecular and Biochemical Parasitology,* **122**: 211-216.

Taylor JS, Breden F. 2002. The inheritance of heteroplasmy in guppies. *Journal of Fish Biology,* **60**: 1346-1350.

Vijverberg K, Mes THM, Bachmann K. 1999. Chloroplast DNA evidence for the evolution of *Microseris* (Asteraceae) in Australia and New Zealand after long-distance dispersal from Western North America. *American Journal of Botany,* **86**: 1448-1463.

Vogl C, Badger J, Kearney P, Li M, Clegg M, Jiang T. 2003. Probabilistic analysis indicates discordant gene trees in chloroplast evolution. *J Mol Evol,* **56**: 330-40.

Wilkinson GS, Chapman AM. 1991. Length and Sequence Variation in Evening Bat D-Loop Mtdna. *Genetics,* **128**: 607-617.

Wolfe AD, Randle CP. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: Implications for plant molecular systematics. *Systematic Botany,* **29**: 1011-1020.