



Inverted patterns of genetic diversity in continental and island populations of the heather *Erica scoparia* s.l.

A. Désamoré^{1*}, B. Laenen¹, J.M. González-Mancebo², R. Jaén Molina³, N. Bystriakova⁴, E. Martínez-Klimova⁴, M. A. Carine⁴ and A. Vanderpoorten¹

¹Institute of Botany, University of Liège, B22 Sart Tilman, B-4000 Liège, Belgium,

²Department of Botany, University of La Laguna, Tenerife, Spain, ³Department of Molecular Biodiversity & DNA Bank, Jardín Botánico Canario 'Viera y Clavijo', Las Palmas de Gran Canaria, Spain, ⁴Department of Botany, The Natural History Museum, London, UK

ABSTRACT

Aim Using the heather *Erica scoparia* s.l. as a model, this paper aims to test theoretical predictions that island populations are genetically less diverse than continental ones and to determine the extent to which island and continental populations are connected by pollen- and seed-mediated gene flow.

Location Macaronesia, Mediterranean, Atlantic fringe of Europe.

Methods Patterns of genetic diversity are described based on variation at two chloroplast DNA (cpDNA) loci and one nuclear DNA (nDNA) locus for 109 accessions across the entire distribution range of the species. Global patterns of genetic differentiation were investigated using principal coordinates analysis. Genetic differentiation between island and continental areas, estimations of pollen- and seed-mediated gene flow, and the presence of phylogeographical signal were assessed by means of F_{ST}/N_{ST} (continental scale) and F_{ij}/N_{ij} (local scale). Extant and past distribution ranges of the species were inferred from niche modelling using layers describing present and Last Glacial Maximum (LGM) macroclimatic conditions.

Results The Azores exhibited a significantly higher genetic diversity than the continent. The lowest levels of genetic differentiation were observed between the Azores and the western Mediterranean, and the diversity observed in the Azores resulted from at least two colonization waves. Within the Azores, kinship coefficients showed a significant and much steeper decrease with geographical distance in the cpDNA than in the nDNA. The distribution predicted by LGM models was markedly different from the current potential distribution, particularly in western Europe, where no suitable areas were predicted by LGM models, and along the Atlantic coast of the African continent, where LGM models predicted highly suitable climatic conditions.

Main conclusions The higher diversity observed in Azorean than in continental populations is inconsistent with MacArthur and Wilson's equilibrium model and derived theoretical population genetic expectations. This inverted pattern may be the result of extinction on the continent coupled with multiple island colonization events and subsequent allopatric diversification and lineage hybridization in the Azores. The results highlight the role of allopatric diversification in explaining diversification on islands and suggest that this process has played a much more significant role in shaping Azorean biodiversity than previously thought.

Keywords

Azores, dispersal, endemism, island biogeography, isolation by distance, Macaronesia, spatial autocorrelation, spatial genetic structure, speciation.

*Correspondence: Aurélie Désamoré, Institute of Botany, University of Liège, B22 Sart Tilman, B-4000 Liège, Belgium.
E-mail: adesamore@doct.ulg.ac.be

INTRODUCTION

Population genetic theory predicts that island populations should be genetically depauperate in comparison with source populations because of the founder effect (Barrett, 1996; Franks, 2010). Recently, however, inverted patterns of genetic diversity have been reported, with higher levels of diversity observed on islands (Algar & Losos, 2011; Fernández-Mazuecos & Vargas, 2011). Indeed, if dispersal between islands and continental areas is frequent, owing to either the high mobility of organisms or close geographical proximity, any signature of a bottleneck in present patterns of genetic diversity may be rapidly erased.

The Macaronesian region comprises the volcanic oceanic archipelagos of the Azores, Madeira, the Selvagens, Canaries and Cape Verdes. They are located in the north Atlantic between 1370 km (Azores) and 95 km (Canaries) from the nearest continental land mass. The geographical setting of the Macaronesian archipelagos is likely to have had a profound impact on patterns of diversity. In the initial stages of colonization, the short distances separating at least some of the Macaronesian archipelagos from continental source areas mean that lineages may have experienced multiple introductions into the region. The past existence of 'stepping-stone islands' (Fernández-Palacios *et al.*, 2011), coupled with the southern extension of the Mediterranean climate zone at times during the Pleistocene (Suc, 1984), may have further promoted this in the case of the Canaries. It is notable that Francisco-Ortega *et al.* (2000) found that allozyme diversity in Canarian endemic plants was higher than that in Pacific island endemics, suggesting a less marked bottleneck effect in the Canaries, consistent with this hypothesis.

Molecular systematic analyses suggest that for most Macaronesian lineages, the source area for colonization was Europe or North Africa (e.g. Carine *et al.*, 2004, 2010). These areas experienced marked changes in their climate during the Pliocene and Pleistocene, resulting in the extinctions and migrations of many plant species (Suc, 1984). It has long been recognized that the Macaronesian islands, buffered from the worst effects of continental climate changes by their oceanic climate, may have acted as refugia, harbouring relictual diversity for lineages that became extinct on the continent (Vargas, 2007). Over time, this unique relictual diversity may have been supplemented by the intra-regional diversification of lineages, further enhancing levels of insular diversity. Diversity patterns at the interface between the Macaronesian archipelagos and the near continent may be further complicated if island populations recolonized continental areas, contributing novel genotypes to the continental gene pool and increasing continental diversity levels. Examples of such back-colonization have been increasingly recognized (e.g. Bellemain & Ricklefs, 2008). In Macaronesia, back-colonization has been documented in angiosperm groups such as *Convolvulus*, wherein *Convolvulus fernandesii* P.Silva & Teles, a pinpoint endemic restricted to Cabo Espichel in Portugal, is resolved as a derived member of a clade of Macaronesian endemic taxa

(Carine *et al.*, 2004). Caujapé-Castells (2011) documented other possible cases of back-colonization involving Macaronesian angiosperms, although most remain equivocal. In the fern *Trichomanes speciosum* Willd., chloroplast DNA (cpDNA) haplotype data support dispersal from the archipelagos to continental Europe (Rumsey *et al.*, 1996), and in the liverwort *Radula lindenbergiana* Gottsche ex Hartm., phylogeographical analyses suggest that Macaronesia served as a reservoir of genetic diversity, from which Europe was back-colonized after glaciations (Laenen *et al.*, 2011).

In this paper, we aim to investigate patterns of genetic diversity at the interface between the Macaronesian islands and the continent. Focusing on the *Erica scoparia* L. complex, our goal is to determine the direction and frequency of colonization events and the extent to which lineage diversification within Macaronesia and the effects of Pleistocene climate changes in continental areas have shaped the patterns of diversity observed in the complex. By integrating niche modelling with molecular data from the chloroplast and nuclear genomes, we specifically aim to address the following questions.

1. How is genetic diversity distributed in the *E. scoparia* complex? Is diversity on the Macaronesian islands reduced in comparison with that in continental areas as predicted by population genetic theory? What has been the impact of Pleistocene climate change, and particularly the Last Glacial Maximum (LGM), on continental and island diversity in the species?

2. What is the degree of genetic differentiation between islands and the continent, between archipelagos, and between islands within archipelagos? To what extent are island and continental populations of *E. scoparia* connected by pollen- and seed-mediated gene flow, and at what scale does isolation by distance operate?

MATERIALS AND METHODS

Sampling and molecular protocols

The *E. scoparia* complex is composed of four taxa with a distribution range spanning Macaronesia, south-western Europe and north-western Africa. *Erica scoparia* subsp. *scoparia* is distributed across the western Mediterranean, extending eastwards to west-central Italy, northwards to west-central France, and southwards to the Mediterranean fringe of Morocco and Tunisia (Lundqvist & Nordenstam, 1988). *Erica scoparia* subsp. *azorica* (Hochst.) D.A. Webb (*E. azorica* Hochst.) is endemic to and widespread across all islands of the Azores (Schaefer, 2005). *Erica scoparia* subsp. *maderinicola* D.C. McClint is endemic to the Madeiran archipelago (Press & Short, 1994). Finally, *E. scoparia* subsp. *platycodon* (Webb & Berthel.) A. Hansen & G. Kunkel [*E. platycodon* (Webb & Berthel.) Rivas Mart.] is endemic to the Canary Islands. The species has been reported from Tenerife, La Gomera and El Hierro (Bramwell & Bramwell, 2001), but intensive fieldwork, examination of herbarium material, and discussions with local

botanists suggest that the species is actually restricted to the first two islands.

In total, 109 specimens were sampled from the entire distribution range of *E. scoparia* s.l. (see Appendix S1 in Supporting Information). DNA was extracted using a standard CTAB (cetyl trimethyl ammonium bromide) protocol. Each specimen was genotyped at the cpDNA loci *matK* and *trnH-psbA*, which proved to exhibit intraspecific variation in *Erica* (Désamoré *et al.*, 2011), as well as at the nuclear locus *at103*. The cpDNA loci were amplified using the primers and following the protocols described in Désamoré *et al.* (2011). For *at103*, a specific set of primers (*at103_sco_F*: 5' CAAGCCAAAGTT CATCTT 3'; *at103_sco_R*: 5' AATGTCACATAAA CCTGCAC 3') was designed within conserved regions at the 5' and 3' ends of the marker from an initial alignment of sequences obtained using universal primers from Sang (2002). Polymerase chain reaction (PCR) cycling consisted of denaturation for 2 min at 95 °C, followed by 35 cycles of denaturation for 30 s at 95 °C, annealing for 45 s at 58 °C, extension for 2 min at 72 °C, and finally 7 min at 72 °C. The presence of amplified target DNA fragments was verified visually on agarose gel by staining with ethidium bromide. PCR products were purified with Exosap-it mix (USB Corporation, Cleveland, OH, USA) and sent to MacroGen (MacroGen Inc., Seoul, Korea) for sequencing.

Genetic data analysis

Sequences were aligned and verified by eye using SEQUENCHER 3.1 (Gene Codes Corporation, Ann Arbor, MI, USA). Gaps were inserted where necessary to preserve positional homology in the alignment, and indels were scored as binary characters, regardless of their length. The *at103* locus, unlike ribosomal DNA, is not subject to concerted evolution, and additivity patterns were observed within 25.6% of the sequences obtained (see Appendices S1 and S2). Such infra-genomic polymorphisms were distinguished from sequencing artefacts when they (1) occurred in both reading directions and (2) were observed at polymorphic positions, that is, where they corresponded to the additivity of alleles observed as a single copy in other sequences (see Appendix S3). Additivity patterns revealed by direct sequencing of nuclear DNA (nDNA) have commonly been used to identify instances of hybridization (see Fehrer *et al.*, 2010 for review), and individuals exhibiting such sequence additivity patterns were scored as heterozygotes.

Patterns of genetic diversity and structure were explored at the level of four geographical regions: the Azores, Canary Islands, Madeira, and the western Mediterranean. Diversity indices, including the number of alleles, number of private alleles (*A*) and nucleotide diversity (*N*), were computed for each geographical region and for both the cpDNA matrix and the nDNA matrix. Unbiased haplotype diversity (*H*) and unbiased expected heterozygosity (*H_e*) were computed for the cpDNA and nDNA data matrices, respectively, using GENALEX 6 (Peakall & Smouse, 2006). Heterogeneity in the sample size between regions was further addressed by a resampling procedure: 19 specimens (equivalent to the continental sample

size) were selected at random within the Azorean sample and the corresponding values of *H* and *H_e* were recalculated. The procedure was repeated 99 times, and the continental *H* and *H_e* were compared with the range of resampled *H* and *H_e* values in the Azores.

The structure of the cpDNA and nDNA variation was computed with GENALEX 6 by means of a principal coordinates analysis (PCoA) derived from a matrix of genetic distance among individuals based on Tajima and Nei's substitution model (Tajima & Nei, 1984). Genetic differentiation among geographical regions, estimations of pollen- and seed-mediated gene flow, and presence of phylogeographical signal in the data were assessed by means of comparative analyses employing *F_{ST}* (Wright's fixation index) and *N_{ST}* for each of the cpDNA and nDNA data matrices. *N_{ST}* is a measure of genetic differentiation among populations; it is analogous to *F_{ST}* but takes into account the phylogenetic relationships among alleles (Pons & Petit, 1996). When *N_{ST}* is larger than *F_{ST}*, it means that mutation rates are higher than dispersal rates, generating a phylogeographical pattern. Phylogenetic relationships among alleles were derived from a matrix of mean character differences among alleles, as implemented in PAUP* 4.0b10 (Swofford, 2003). The significance of *F_{ST}* and of *N_{ST}* were tested by constructing the distribution of the null hypothesis by means of 999 random permutations of individuals among populations, as implemented in SPAGeDI 1.3 (Hardy & Vekemans, 2002). The existence of a phylogeographical signal was tested by assessing the significance of the observed difference between *N_{ST}* and *F_{ST}* values by means of 999 random permutations of the mean character difference matrix among alleles.

The significance of isolation by distance in seed- and pollen-mediated gene flow was assessed in the Azores by investigating patterns of genetic differentiation at the scale of individuals along gradients of geographical distance. For that purpose, we computed pairwise kinship coefficients among individuals, *F_{ij}*, using J. Nason's estimator (Loiselle *et al.*, 1995), as well as an *F_{ij}* analogue for ordered alleles, termed *N_{ij}*, taking the phylogenetic relationship among alleles into account. The estimated parameters *F_{ij}* and *N_{ij}* are defined as $F_{ij} \equiv 1 - h_{ij}/\bar{h}$ and $N_{ij} \equiv 1 - v_{ij}/\bar{v}$, where *h_{ij}* is the probability that two gene copies from individuals *i* and *j* carry different alleles (or haplotypes), *v_{ij}* is the phylogenetic distance between the haplotypes carried by individuals *i* and *j* (mean character differences among haplotypes), and \bar{h} and \bar{v} are the averages over all pairs of individuals in the sample of *h_{ij}* and *v_{ij}*, respectively. Both *F_{ij}* and *N_{ij}* were computed from allele frequencies within each geographical region. To test for isolation by distance, the significance of the slope of the regression of *F_{ij}* or *N_{ij}* on the logarithm of spatial distance among individuals, $\ln(d_{ij})$, was tested by means of 999 random permutations of population locations (Mantel test). The mean *F_{ij}* or *N_{ij}* values were also computed over *i, j* pairs separated by predefined geographical distance intervals, *d*, giving *F_(d)* and *N_(d)*. Threshold distance separating intervals were 0, 100, 250, 500 and 750 km. The difference between *N_(d)* and *F_(d)* was tested by means of 999 random permutations of the genetic

distance matrix to test for the presence of a phylogeographic signal at different spatial scales. All computations were performed using SPAGEDI 1.3.

Niche modelling

Erica scoparia collection localities (135) were assembled in a single database. Locality information was obtained from herbarium collections at the Natural History Museum, London, the University of Reading (UK) and the Global Biodiversity Information Facility (GBIF, <http://www.gbif.org>). For specimens without precise latitude and longitude data, coordinates were obtained using locality descriptions and Google Earth 5.2 (<http://earth.google.com>). Current climate data (Hijmans *et al.*, 2005) and LGM climate data were downloaded from the WorldClim website (<http://www.worldclim.org>). To assess modelling uncertainty arising from uncertainty in the LGM climate data, two available climate simulations were used, Community Climate System Model (CCSM) and Model for Interdisciplinary on Climate (MIROC), derived from the PMIP2 database (<http://pmip2.lscce.ipsl.fr/>, last accessed 15 January 2011). For each location point of the dataset, corresponding values of 19 BIOCLIM variables were obtained using ARCVIEW 9.3 (ESRI, Redlands, CA, USA). To model the current distribution of *E. scoparia*, we used MAXENT 3.3.1 (Phillips *et al.*, 2006). The maximum-entropy method of model estimation used in MAXENT can be applied to presence-only data. For training and testing models in MAXENT, the 20:80 split was randomly applied to the available *E. scoparia* presence data, and the default values were accepted for the 'maximum iterations' and 'convergence threshold' parameters. The resulting models were projected onto the LGM climate using MIROC and CCSM simulation results. For each climate scenario (present, LGM/MIROC and LGM/CCSM), five replicate runs were conducted and an average model was produced by MAXENT. Validation of models included examination of the receiver operating characteristic (ROC) curve and its area under the curve (AUC) parameter. The averaged models were then edited in ARCVIEW.

RESULTS

The *matK*, *trnH-psbA* and *at103* genes included 5, 3 and 10 polymorphic sites, respectively. Three indels of 1, 8 and 24 bp were also scored within the *at103* gene. Patterns of genetic diversity in the Azores, Madeira, Canary Islands and the western Mediterranean are given in Table 1. For the chloroplast data, the Azores exhibited a higher diversity than the western Mediterranean, in terms of both unbiased haplotype and nucleotide diversity. At the nuclear locus, the Azores and the western Mediterranean exhibited similar levels of diversity, with a slightly higher expected heterozygosity in the western Mediterranean but a slightly higher nucleotide diversity in the Azores. Madeira exhibited the lowest levels of diversity with both the cpDNA and nDNA. The observed haplotype diversity of the western Mediterranean remained lower than that of each

Table 1 Sampling size (*n*), number of private alleles (*A*), unbiased haplotype diversity (*H*), unbiased expected heterozygosity (*H_e*) and nucleotide diversity (*N*) at the cpDNA loci (*matK* and *trnH-psbA*) and at the nDNA locus (*at103*) in the heather *Erica scoparia* s.l. across its entire distribution range, encompassing the Azores, Madeira, the Canary Islands and the western Mediterranean.

	cpDNA				nDNA			
	<i>n</i>	<i>A</i>	<i>H</i>	<i>N</i>	<i>n</i>	<i>A</i>	<i>H_e</i>	<i>N</i>
Azores	57	4	0.699	0.00170	52	2	0.388	0.0132
Madeira	10	1	0.200	0.00036	4	0	0.000	0.0000
Canary Islands	20	1	0.505	0.00091	13	1	0.077	0.0052
Western Mediterranean	19	0	0.298	0.00076	17	1	0.456	0.0119

of 100 random samples of 19 Azorean specimens in the cpDNA [*H* mean = 0.694; 95% confidence interval (CI) = 0.526–0.861], indicating that the difference observed was not an artifact of sample size. In the nDNA, the expected heterozygosity of the western Mediterranean fell within the confidence interval of the Azorean resampling (*H_e* mean = 0.383; 95% CI = 0.237–0.528).

Nine cpDNA haplotypes, whose distribution is presented in Fig. 1, were resolved. The first two PCoA axes account for 60.1% of the variation in the chloroplast (Fig. 2a). The cpPCoA resolves three groups. Group 1, at the extreme right of the first axis, includes the Canarian endemic haplotype 2 and haplotype 3, found mainly in the Canaries but with one accession in Spain (Galicia). Those haplotypes differ from each other by a single substitution. Group 2, at the upper left of the graph, includes haplotypes 4 and 6, which were predominantly sampled from the Azores. The last group is composed of five haplotypes: haplotype 1, which is the most frequent, was found in the Azores and Madeira and on the continent; haplotypes 8 and 9 were sampled respectively from Santa Maria and Sao Jorge in the Azores; haplotype 5 was found once in Portugal but was otherwise distributed mainly in the Azores; and haplotype 7 was sampled only from Madeira.

Variation at the *at103* gene resolved five alleles, hereafter referred to as alleles 1–5. The first two axes of the PCoA (Fig. 2b) account for 82.8% of the nDNA variation. All but one of the Canarian accessions exhibit the homozygous genotype 3,3 that was sampled only from the Canary Islands and is isolated on the PCoA graph in the upper right corner. A single Canarian accession has a heterozygous combination (4,3) and is thus intermediate between homozygotes 4,4 found on the continent and homozygotes 3,3 found in the Canary Islands. Continental accessions appear in three groups on the PCoA: the homozygous genotypes 4,4 and 5,5, which are restricted to the continent, differ from each other by a single substitution. Heterozygotes 4,1, which were found only on the continent, occupy a central position on the graph. Finally, two continental accessions exhibit the homozygous genotype 1,1,

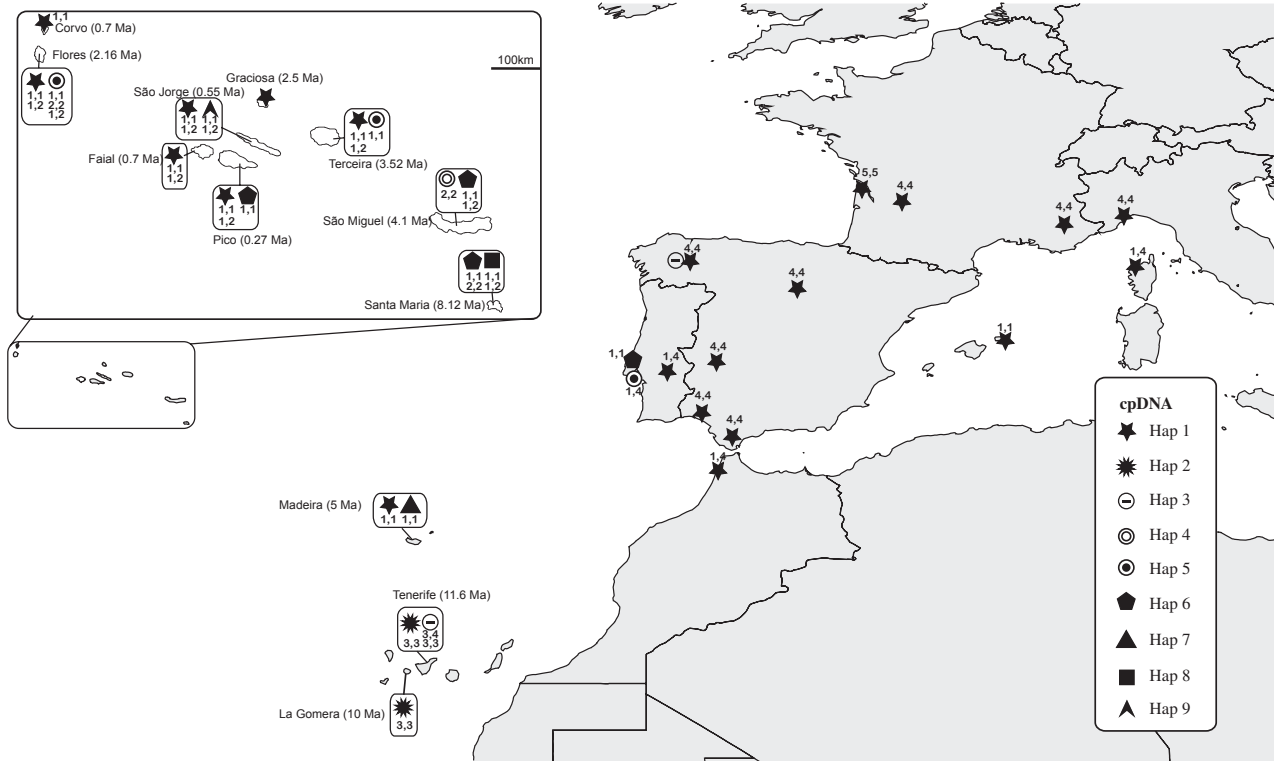


Figure 1 cpDNA (*matK* and *trnH-psbA*; symbols) haplotype and nDNA (*at103*; numbers) genotype distribution in the heather *Erica scoparia* s.l. based on 109 individuals sampled across the entire distribution range of the species.

which is otherwise found among Azorean and Madeiran accessions. The two continental accessions with this genotype exhibit cpDNA haplotypes 1 and 6. Azorean accessions comprise three groups on the PCoA graph, as follows. The homozygous genotype 2,2, which was found only in the Azores, is very close to continental endemic genotypes 4,4 and 5,5, at the lower right of the graph. The heterozygous genotypes 2,1 occupy a central position near the genotype 4,1. Finally, the bulk of Azorean accessions share the homozygous genotype 1,1 and appear at the upper left of the PCoA graph.

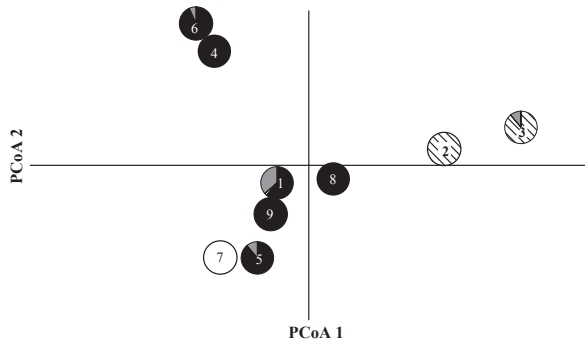
There is a significant partitioning of cpDNA haplotype variation among the Azores, the Canary Islands, Madeira and the western Mediterranean (F_{ST} global = 0.39, $P < 0.001$). The chloroplast data also exhibit a significant phylogeographical signal, with a global N_{ST} of 0.65, which is significantly higher than the global F_{ST} at $P < 0.05$. At the nuclear locus, allelic variation is also significantly geographically structured (F_{ST} global = 0.71, $P < 0.001$), but the global N_{ST} (0.78) is not significantly higher than the global F_{ST} . Pairwise comparisons of F_{ST} and N_{ST} among geographical regions indicate that the Canary Islands are significantly genetically differentiated from all the other regions at both the cpDNA and nDNA loci (Table 2). There is a highly significant difference in cpDNA haplotype frequency, but not in *at103* allele frequency, between the Azores and Madeira ($F_{ST} = 0.43$, $P < 0.001$ and $F_{ST} = 0.09$, $P > 0.05$, respectively). The lowest levels of differ-

entiation in the chloroplast are observed between the Azores and the western Mediterranean, with a comparatively low F_{ST} of 0.11.

Within the Azores, average kinship coefficients between pairs of individuals show a significant decrease with geographical distance for both the cpDNA and nDNA loci (Fig. 3). Average kinship coefficients are, however, significantly different from zero at all classes of distance except one for the cpDNA markers, whereas only the average kinship coefficient derived from variation at the *at103* locus for the largest class of geographical distance among individuals is significant. The slope of the regression analysis is also steeper for the cpDNA markers (-0.51 , $P < 0.01$) than for the nuclear locus (-0.19 , $P < 0.05$), for which it is only marginally significantly different from 0.

In the niche modelling analysis, all models produced by MAXENT predicted the test sites significantly better than random, with AUC values exceeding 0.95. The predicted current distribution of *E. scoparia* is consistent with its current known distribution range (Fig. 4). LGM/MIROC models consistently predicted a distribution area for *E. scoparia* larger than that predicted by LGM/CCSM models. However, little variation was observed between replicates produced for each climate scenario. The distributions predicted by both LGM models were markedly different from the current potential distribution, particularly in western Europe, where no suitable areas were predicted by LGM models, and along the Atlantic

(a) PCoA cpDNA



(b) PCoA nDNA

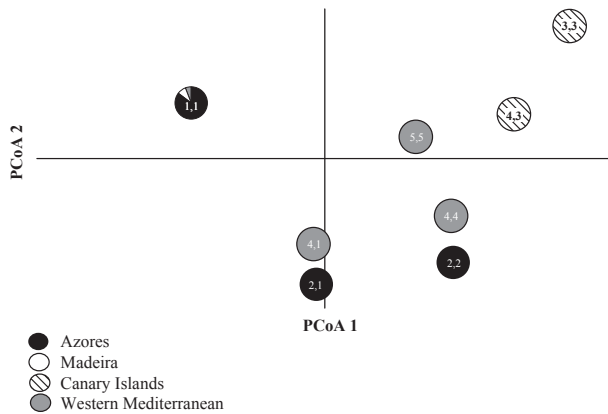


Figure 2 Principal coordinates analysis (PCoA) of the variation of (a) cpDNA (*matK* and *trnH-psbA*) and (b) nDNA (*at103*) in the heather *Erica scoparia* s.l. across its entire distribution range. Pie diagrams indicate the geographical origin of the specimens in proportion to their frequency in the sampling.

coast of the African continent, where LGM models predicted highly suitable climatic conditions (Fig. 4).

DISCUSSION

Levels of chloroplast DNA diversity in Azorean populations of *E. scoparia* are significantly higher than those observed on the continent, but the Azores exhibit similar levels of diversity to the continent for the nuclear locus examined. The genetic diversity patterns observed are therefore inconsistent with MacArthur & Wilson’s (1967) equilibrium model, theoretical expectations derived from population genetic theory (Barrett, 1996), and previous studies on island/continent patterns of genetic diversity (Franks, 2010). They add to a growing body of evidence challenging the traditional assumption that oceanic island populations exhibit lower levels of genetic variation than their continental counterparts (Algar & Losos, 2011; Fernández-Mazuecos & Vargas, 2011).

In *E. scoparia*, this inverted pattern of genetic diversity can be attributed to three main factors. First, palaeoniche recon-

Table 2 Genetic differentiation between populations of *Erica scoparia* s.l. from the Azores, Madeira, the Canary Islands and the western Mediterranean, as measured by F_{ST} and N_{ST} (below and above the diagonal, respectively) derived from a matrix of cpDNA (*matK* and *trnH-psbA*) and a matrix of nDNA (*at103*). The significance test associated with N_{ST} results from the randomization of the genetic distance matrix among alleles and measures whether $N_{ST} > F_{ST}$. The significance test associated with F_{ST} results from the randomization of the alleles among localities and measures whether $F_{ST} > 0$. The asterisks *, **, *** indicate whether the null hypothesis is rejected at the 0.05, 0.01 or 0.001 significance level, respectively.

	Azores	Madeira	Canary Islands	Mediterranean
cpDNA				
Azores	–	0.61	0.62*	0.06
Madeira	0.43***	–	0.85	0.72
Canary Islands	0.38***	0.61***	–	0.68
Mediterranean	0.11*	0.71***	0.59***	–
nDNA				
Azores	–	0.25	0.84	0.37
Madeira	0.09	–	0.98	0.74
Canary Islands	0.70***	0.94**	–	0.78
Mediterranean	0.50***	0.56**	0.70***	–

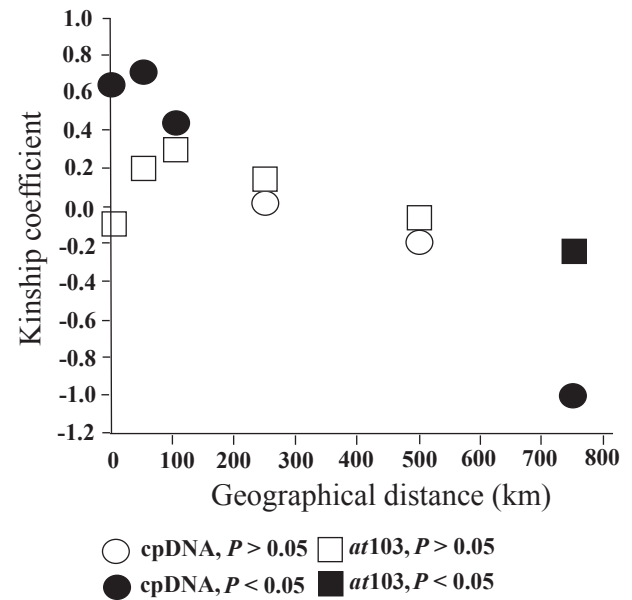


Figure 3 Mean N_{ij} (and associated significance test) resulting from comparisons of the cpDNA haplotype (*matK* and *trnH-psbA*) and nDNA allele at the *at103* locus between pairs of individuals of the heather *Erica scoparia* s.l. in the Azores in relation to the geographical distance separating them.

structions indicate that the Macaronesian archipelagos retained extensive areas of suitable habitat during the LGM, reinforcing the notion that islands, with their buffered climate, might have played a much more important role as refugia than

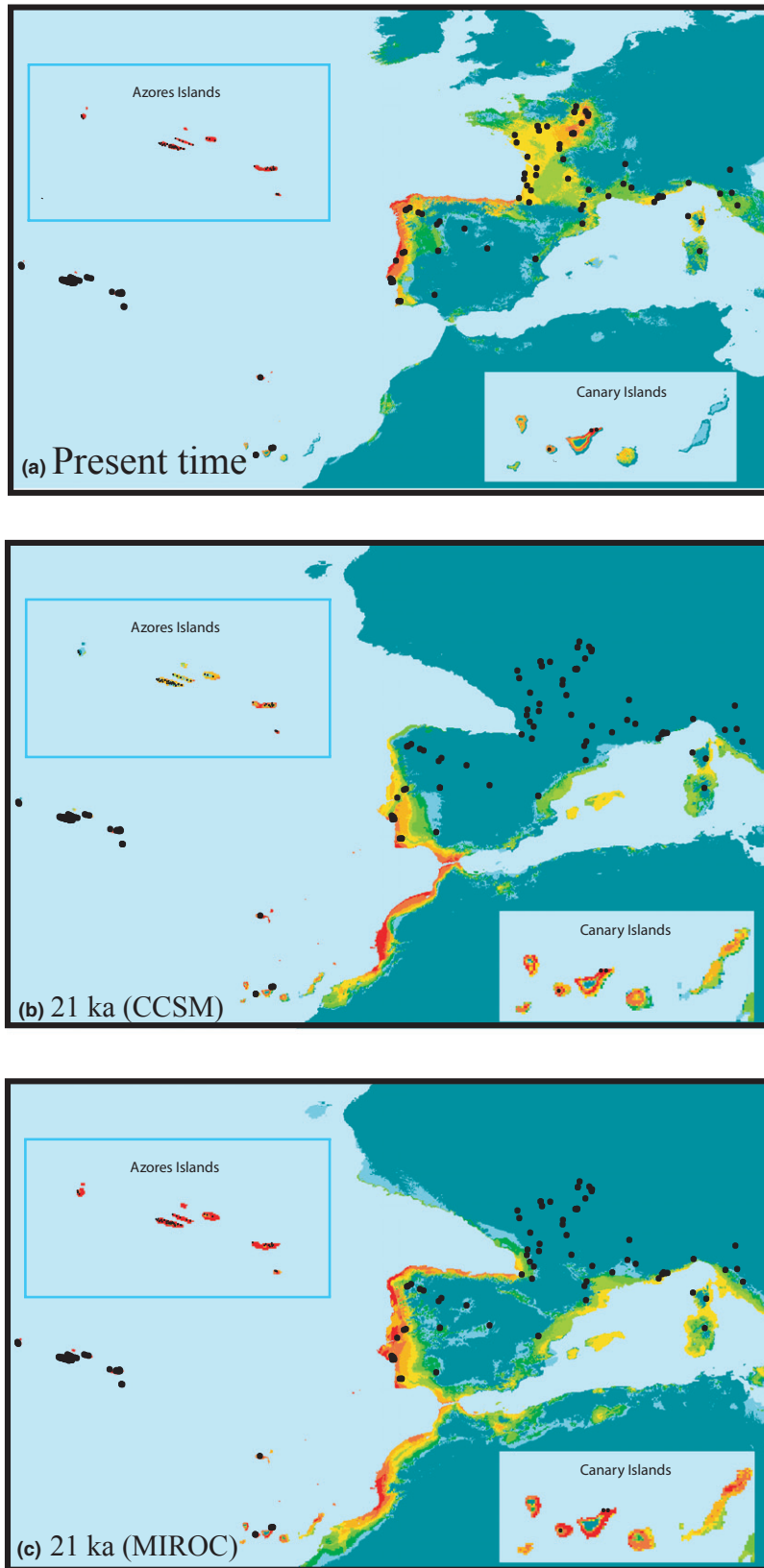


Figure 4 Predicted environmental suitability for *Erica scoparia* s.l. across time stages, according to the niche model fitted to present locations and climate: (a) present, (b) Last Glacial Maximum/CCSM model, (c) Last Glacial Maximum/MIROC model. Dots represent specimens used for extrapolating the niche of *E. scoparia* s.l. Colours are representative of the potential suitability, with a gradient from blue to red, where blue is the less suitable and red the most suitable area for the species.

previously acknowledged (Vargas, 2007; Fernández-Mazuecos & Vargas, 2011; Laenen *et al.*, 2011). In contrast, niche availability on the continent was severely restricted. Fossils of *E. scoparia* from the Gortian interglacial found in Ireland (Nelson, 2009), where this species no longer occurs, further corroborate the hypothesis of marked continental range shifts by *E. scoparia* as a result of climatic changes in the Pleistocene. Much of the continental distribution of the species is therefore likely to be the result of a post-glacial range expansion. The predominance of a single cpDNA haplotype on the continent is fully compatible with this hypothesis.

The second factor is the multiplicity of migration events between the Azores and the continent, as evidenced by the fact that the lowest pairwise comparison of F_{ST} (0.11) in the cpDNA was observed between these two regions. The PCoA of both the nDNA and cpDNA datasets further indicates that multiple colonization events of the Azores contributed to the high levels of diversity observed in the archipelago. In the *at103* gene, the close proximity, along the first two PCoA axes, of the continental 4,4 genotype and the Azorean 2,2 genotype (which differ by a single mutation) points to a recent colonization event. The isolated position of genotype 1,1, which includes the vast majority of Azorean accessions, all of the Madeiran accessions and two continental ones, can be interpreted in several ways. One interpretation is that continental specimens colonized the Azores very recently. This hypothesis, is, however, weakened by the fact that *E. scoparia* actively diversified in the archipelago (see below). Alternatively, the extreme position of genotypes 1,1 and 4,4 or 5,5 on the PCoA graph may reflect an ancient colonization of the Azores by one of the latter genotypes, which may have accumulated divergence either *in situ* within the archipelago or on the continent from a now extinct (or unsampled) genotype. In this scenario, the sharing of the same *at103* genotype 1,1 and two distinct cpDNA haplotypes by the vast majority of Azorean accessions and two continental accessions would be explained by back-colonization of the continent from the Azores. In keeping with the situation in a number of angiosperm groups (Caujapé-Castells, 2011), the molecular data are equivocal with regard to the direction of genetic exchange between the Macaronesian islands and continent in *E. scoparia*, but unambiguously point to several migrations events. The independent colonization of the Azores by different continental ancestors challenges the taxonomic status of *E. azorica*. The polyphyletic origin of *E. azorica*, coupled with the close similarity between the flowers of the continental *E. scoparia* s.s. and *E. azorica*, which differ only in size [1.5–1.75 mm long in the latter versus 2.1–2.3 mm in the former (Bayer, 1993; Schaefer, 2005)], suggest that *E. azorica* should be reduced to synonymy with *E. scoparia*.

The third factor contributing to the inverted pattern of genetic diversity observed in *E. scoparia* is *in situ* diversification in the Azores, which further supplemented diversity and counter-balanced the impact of the founding effect in the archipelago. The significant decrease of cpDNA kinship coefficients between pairs of individuals along a gradient of

geographical distance in the Azores, together with the much steeper slope in the regression analysis of cpDNA kinship coefficients than in the nDNA regression analysis, suggests limited seed-mediated gene flow. This is in line with previous analyses of spatial genetic structure (e.g. Vekemans & Hardy, 2004; Ndiade-Bourobou *et al.*, 2010; Debout *et al.*, 2011; Vanderpoorten *et al.*, 2011). The role of drift in speciation has been questioned because even a small amount of gene flow can delay differentiation by drift (Coyne & Orr, 2004). As a consequence, adaptive diversification has been viewed as a major factor in the diversification of island lineages (see Givnish, 2010 for a review). However, the results presented here reinforce the idea (Comes *et al.*, 2008) that allopatric diversification can play a key role in explaining diversification on islands. Furthermore, the results support the suggestion that allopatric differentiation has played a greater role in shaping Azorean biodiversity than was previously suggested based upon the apparent homogeneity of the biota across the islands of the archipelago (Schaefer *et al.*, 2011). Given the high ecological homogeneity of the Azorean islands (Borges & Hortal, 2009; Cardoso *et al.*, 2010), the main driver of diversification within the archipelago would be isolation-by-distance, a neutral process caused by drift and therefore unlikely to lead to spectacular morphological shifts in the short term.

Although the Canaries and Madeira are closer to potential source areas than the Azores, the data suggest much more limited interchange between these archipelagos and the continent for *E. scoparia*. Indeed, the cpDNA data suggest a single colonization for each of the Canary Islands and Madeira. The presence, in Madeira, of two cpDNA haplotypes, one shared with the Azores and the continent and the other sampled only in Madeira and separated from the former by a single mutation, points to a recent colonization of Madeira from either the Azores or the continent. Morphologically, the flowers of *E. scoparia* subsp. *madeirinicola* are most similar to those of the continental *E. scoparia* s.s., leading Press & Short (1994) to treat the two as the same taxon. The present results are inconsistent with the taxonomic interpretation of *madeirinicola* as a subspecies of the Canarian *E. platycodon* by Rivas-Martínez *et al.* (2002) and with its implicit assumption of a close relationship between the Canarian and Madeiran taxa. The limited diversification in Madeira in comparison with the Azores may reflect the more limited opportunities for allopatric diversification in Madeira in comparison with the Azores, where the archipelago spans nearly 600 km.

The high N_{ST} values between Canarian and continental accessions and the extreme position of Canarian accessions along the PCoA axes of both the cpDNA and nDNA data are consistent with the long isolation of Canarian populations. This isolation may be explained by continental extinction as a result of the Pliocene aridification of North Africa, which resulted in the demise of *E. arborea* in the region (Désamoré *et al.*, 2011). It is notable that whilst most Canarian accessions exhibit a genetically distinct homozygous genotype (3,3), the detection of a heterozygous accession (4,3) points to instances

of recent pollen flow between the Canary Islands and the continent. A more southerly extension of suitable habitat for *E. scoparia* during the LGM was suggested by climate modelling (Fig. 4), and this may have afforded a window of opportunity for some genetic exchange between continental and Canarian populations.

CONCLUSIONS

Globally, the results presented here suggest that multiple colonization events of the Macaronesian region and of the Azores, subsequent *in situ* diversification in the Azores, and range contraction in North Africa and western Europe account for the extant pattern of distribution in *E. scoparia*. This pattern differs strikingly from that exhibited by the congeneric *E. arborea*, for which palaeontological and phylogeographical evidence point to several colonization waves from eastern Africa and the eastern Mediterranean westwards, ultimately resulting in the relatively recent colonization of Macaronesia (Désamoré *et al.*, 2011). The presence of a phylogeographical signal at the scale of the entire distribution range in the cpDNA in *E. scoparia* but not in *E. arborea* further indicates that whilst dispersal rates are higher than mutation rates in *E. arborea*, the converse is true in *E. scoparia*. *Erica arborea* and *E. scoparia* are two conspicuous elements of the Macaronesian vegetation. That they exhibit such markedly different biogeographical histories serves to emphasize the highly complex spatio-temporal relationships of the unique and diverse flora of the Macaronesian region.

ACKNOWLEDGEMENTS

A.D. and B.L. are PhD fellows of the Belgian Industrial and Agricultural Research Funds (FRIA). This research was funded by grants 1.5036.11 and 2.4557.11 from the Belgian Funds for Scientific Research (FSR-FNRS) as well as by a fellowship of the Fonds Léopold III to A.V.

REFERENCES

- Algar, A.C. & Losos, J.B. (2011) Evolutionary assembly of island faunas reverses the classic island–mainland richness difference in *Anolis* lizards. *Journal of Biogeography*, **38**, 1125–1137.
- Barrett, S.C.H. (1996) The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **351**, 725–733.
- Bayer, E. (1993) *Erica*. *Flora Iberica*, Vol. IV (ed. by S. Castroviejo, C. Aedo, C. Gómez Camp, M. Laínz, P. Monteserrat, R. Morales, F. Muñoz Garmandenia, G. Nieto Feliner, E. Rico, S. Talavera and L. Villar), pp. 485–506. Real Jardín Botánico, CSIC, Madrid.
- Bellemain, E. & Ricklefs, R.E. (2008) Are islands the end of the colonization road? *Trends in Ecology and Evolution*, **23**, 461–468.
- Borges, P.A.V. & Hortal, J. (2009) Time, area and isolation: factors driving the diversification of Azorean arthropods. *Journal of Biogeography*, **36**, 178–191.
- Bramwell, D & Bramwell, Z.I. (2001) *Flores silvestres de Las Islas Canarias*, 4th edn. Editorial Rueda, Madrid.
- Cardoso, P., Arnedo, M.A., Triantis, K.A. & Borges, P.A.V. (2010) Drivers of diversity in Macaronesian spiders and the role of species extinctions. *Journal of Biogeography*, **37**, 1034–1046.
- Carine, M.A., Russell, S.J., 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, M.A., Santos Guerra, A., Guma, I.R. & Reyes-Betancort, J.A. (2010) Endemism and evolution of the Macaronesian Flora. *Beyond cladistics* (ed. by D.M. Williams and S.K. Knapp), pp. 101–124. University of California Press, Berkeley, CA.
- Caujapé-Castells, J. (2011) Jesters, red queens, boomerangs and surfers: a molecular outlook to the diversity of the Canarian endemic flora. *The biology of island floras* (ed. by D. Bramwell and J. Caujapé-Castells), pp. 284–324. Cambridge University Press, Cambridge.
- Comes, H.P., Tribsch, A. & Bittkau, C. (2008) Plant speciation in continental island floras as exemplified by *Nigella* in the Aegean archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **B363**, 3083–3096.
- Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer Associates, Sunderland, MA.
- Debout, G.D., Doucet, J.L. & Hardy, O.J. (2011) Population history and dispersal ecology inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). *Heredity*, **106**, 88–99.
- Désamoré, A., Laenen, B., Devos, N., Popp, M., González-Mancebo, J.M., Carine, M.A. & Vanderpoorten, A. (2011) Out of Africa: northwestwards Pleistocene expansions of the heather *Erica arborea*. *Journal of Biogeography*, **38**, 164–176.
- Fehrer, J., Krak, K. & Chrtek, J. (2010) Intra-individual polymorphism in diploid and apomictic polyploid hawkweeds (*Hieracium*, Lactuceae, Asteraceae): disentangling phylogenetic signal, reticulation, and noise. *BMC Evolutionary Biology*, **9**, e239.
- Fernández-Mazuecos, M. & Vargas, P. (2011) Genetically depauperate in the continent but rich in oceanic islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS ONE*, **6**, e17172.
- Fernández-Palacios, J.M., de Nascimento, L., Otto, R., Delgado, J.D., García-del-Rey, E., Arévalo, J.R. & Whittaker, R.J. (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226–246.

- Francisco-Ortega, J., Santos-Guerra, A., Kim, S.C. & Crawford, D.J. (2000) Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany*, **87**, 909–919.
- Franks, S.J. (2010) Genetics, evolution, and conservation of island plants. *Journal of Plant Biology*, **53**, 1–9.
- Givnish, T.J. (2010) Ecology of plant speciation. *Taxon*, **59**, 1326–1366.
- Hardy, O.J. & Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Laenen, B., Désamoré, A., Devos, N., Shaw, A.J., González-Mancebo, J.M., Carine, M.A. & Vanderpoorten, A. (2011) Macaronesia: a source of hidden genetic diversity for post-glacial recolonization of western Europe in the leafy liverwort *Radula lindenberghiana*. *Journal of Biogeography*, **38**, 631–639.
- Loiselle, B.A., Sork, V.L., Nason, J. & Graham, C. (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.
- Lundqvist, J. & Nordenstam, B. (1988) *Index Holmiensis*. VI. *Dicotyledonae D–F*. Swedish Museum of Natural History, Stockholm.
- MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- Ndiade-Bouroubou, D., Hardy, O.J., Favreau, B., Moussavou, H., Nzengue, E., Mignot, A. & Bouvet, J.M. (2010) Long-distance seed and pollen dispersal inferred from the spatial genetic structure at chloroplast and nuclear microsatellite markers in the very low density rainforest tree, *Baillonella toxisperma* Pierre in Central Africa. *Molecular Ecology*, **19**, 4949–4962.
- Nelson, E.C. (2009) *Erica scoparia* and *Erica spiculifolia* (formerly *Bruckenthalia spiculifolia*) in interglacial floras in Ireland and Britain: confused nomenclature leading to misidentification of fossilized seeds. *Quaternary Science Reviews*, **28**, 381–383.
- Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.
- Press, J.R. & Short, M.J. (1994) *Flora of Madeira*. The Natural History Museum, London.
- Rivas-Martínez, S., Díaz-González, T.E., Fernández-González, F., Izco, J., Loidi Arregui, J., Lousã, M. & Penas Merino, A. (2002) Vascular plant communities of Spain and Portugal. *Itinera Geobotanica*, **15**, 5–922.
- Rumsey, F.J., Russell, S.J., Ji, J., Barrett, J.A. & Gibby, M. (1996) Genetic variation in the endangered filmy fern *Trichomanes speciosum* Willd. *Pteridology in perspective* (ed. by J.M. Camus, M. Gibby and R.J. Johns), pp. 161–165. Royal Botanic Gardens, Kew, London.
- Sang, T. (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology*, **37**, 121–147.
- Schaefer, H. (2005) *Flora of the Azores*, 2nd edn. Margraf Publishers, Weikersheim.
- Schaefer, H., Moura, M., Belo Maciel, M.G., Silva, L.F., Rumsey, F.J. & Carine, M.A. (2011) The Linnean shortfall in oceanic island biogeography: a case study in the Azores. *Journal of Biogeography*, **38**, 1345–1355.
- Suc, J.-P. (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, **307**, 429–438.
- Swofford, D.L. (2003) PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*, Version 4. Sinauer Associates, Sunderland, MA.
- Tajima, F. & Nei, M. (1984) Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution*, **1**, 269–285.
- Vanderpoorten, A., Hardy, O.J., Lambinon, J. & Raspé, O. (2011) Two reproductively isolated cytotypes and a swarm of highly inbred, disconnected populations: a glimpse into *Salicornia*'s evolutionary history and challenging taxonomy. *Journal of Evolutionary Biology*, **24**, 630–644.
- Vargas, P. (2007) Are Macaronesian islands refugia of relict plant lineages? A molecular survey. *Phylogeography in southern European refugia: evolutionary perspectives on the origins and conservation of European biodiversity* (ed. by S.J. Weiss and N. Ferrand), pp. 297–314. Kluwer Academic Publishers, Dordrecht.
- Vekemans, X. & Hardy, O.J. (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Voucher information and GenBank accessions numbers.

Appendix S2 Sequence additivity patterns observed at the polymorphic sites of the *at103* gene in the heather *Erica scoparia* s.l.

Appendix S3 Pattern of sequence additivity at the *at103* locus.

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

BIOSKETCHES

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

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

Editor: José María Fernández-Palacios