MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 1702–1726 doi: 10.1111/j.1365-294X.2012.05495.x

Divergence and biogeography of the recently evolved Macaronesian red *Festuca* (Gramineae) species inferred from coalescence-based analyses

A. J. DÍAZ-PÉREZ*†, M. SEQUEIRA‡, A. SANTOS-GUERRA§ and P. CATALÁN*

*Departamento de Agricultura (Botánica), Escuela Politécnica Superior de Huesca, Universidad de Zaragoza, Ctra. Cuarte km 1, 22071 Huesca, Spain, †Instituto de Genética and Centro de Investigaciones en Biotecnología Agrícola (CIBA), Facultad de Agronomía, Universidad Central de Venezuela, 2101, Maracay, Aragua, Venezuela, ‡Centro de Ciências da Vida, Universidade da Madeira, Alto da Penteada, 9000 Funchal, Portugal, §Unidad de Botánica (ICIA), Retama 2, 38400 Puerto de la Cruz, Tenerife, Spain

Abstract

Studying the biogeography and the phylogeography of the endemic Macaronesian red *Festuca* species (Loliinae, Poaceae) is of prime interest in understanding the speciation and colonization patterns of recently evolved groups in oceanic archipelagos. Coalescence-based analyses of plastid *trn*LF sequences were employed to estimate evolutionary parameters and to test different species-history scenarios that model the pattern of species divergence. Bayesian IM estimates of species divergence times suggested that ancestral lineages of diploid Macaronesian and Iberian red fescues could have diverged between 1.2 and 1.57 Ma. When empirical data were compared to coalescence-based simulated distributions of discordance and *p-distance* statistics, two species-history models were chosen in which the first branching lineage derived in Canarian *Festuca agustinii*. Its sister lineage could have involved a recent polytomy leading to the Madeiran *Festuca jubata*, the Azorean *Festuca francoi* + *Festuca petraea* and the continental *Festuca rivularis* lineages (*Canarian* model) or the sequential branching of lineages leading to *F. jubata* and finally to the sister clades of *F. rivularis* and *F. francoi* + *F. petraea* (*Sequential* model). Nested clade phylogeographic analysis (NCPA) and a first adapted host–parasite co-evolutionary ParaFit method were used to detect the phylogeographic signal. NCPA inferred long-distance colonizations for the entire diploid red *Festuca* complex, but allopatric-fragmentation and isolation-by-distance (IBD) patterns were inferred within archipelagos. In addition, the ParaFit method suggested a generalized pattern of a stepping-stone model at all hierarchical levels. Maximum-likelihood-based dispersal-extinction-cladogenesis (DEC) models were superimposed on the *Sequential* model species tree. The three-independent-colonization (3IC) model was the best supported biogeographic scenario, concurring with previous analysis based on multilocus AFLP data.

Keywords: biogeography, coalescence, oceanic island plants, phylogeography, plastid *trn*LF gene, species divergence time

Received 1 May 2010; revision received 4 January 2012; accepted 13 January 2012

Introduction

Oceanic islands have been recognized as optimal natural testbeds for studying the speciation processes of terrestrial plants (Kreft *et al.* 2008; Kier *et al.* 2009).

Correspondence: Antonio Díaz-Pérez, Fax: +58 243 5507322; E-mail: antonio.dz.pz@gmail.com

Island populations are more prone to experience accelerated divergences in these isolated environments than those of their continental relatives that are usually interconnected through gene flow (Matsumura *et al.* 2009). Speciation events in oceanic archipelagos have usually occurred in recent times (Emerson 2003), considering the young age and the short lifespans attributed to most of these volcanic islands (Emerson 2003; Whittaker & Fernández-Palacios 2007). The Macaronesian archipelagos constitute one of the major hotspots of plant diversity in oceanic islands (Caujapé-Castells *et al.* 2010); different studies have demonstrated that most of their rich endemic flora derived from geographically close NW Africa–SW European Mediterranean ancestors (Carine *et al.* 2004; Kim *et al.* 2008). Alternative hypotheses on single colonizations followed by islands radiations vs. multiple colonizations have been proposed to explain the speciation processes of the Macaronesian flora and the concurrence of highly diversified and less diversified groups in the same geographic settings and environments (Silvertown 2004; Carine 2005; and references therein; Díaz-Pérez *et al.* 2008). Despite the accumulation of a large number of phylogenetic studies of Macaronesian plants, few of them have documented and dated the origins and the dispersal patterns of their lineages. Recently, Kim *et al.* (2008) reported three discrete windows of colonization spanning the Miocene and early Pliocene to explain the within- and among-archipelagos radiation of five eudicot groups in Macaronesia, apparently derived from single continental colonizations. However, no other studies have attempted to date the timing and tempo of speciation of Macaronesian plants with respect to those of their close continental relatives, testing at the same time hypothetical multiple and bi-directional colonizations involving the origins of the species within a chronological scenario.

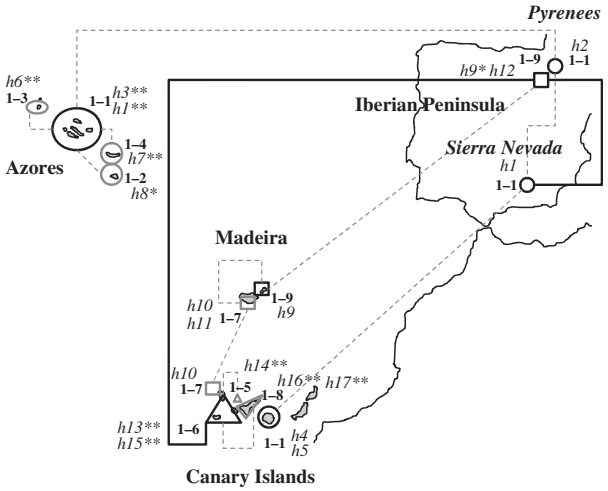


Fig. 1. Geographic location of haplotypes (preceded by the letter h) and nested clade phylogeographic analysis clades of five diploid red *Festuca* species (see Tables S1 and S2, Supporting information and Fig. 2a for further details). Species (geographic distribution) = *Festuca agustinii* (Canary Islands), *Festuca jubata* (Madeiran archipelago), *Festuca francoi* and *Festuca petraea* (Azores) and *Festuca rivularis* (Iberian peninsula). Circle, triangle and square symbols correspond to 2-step *Azorean–Iberian* (2-1), *Canarian* (2-2) and *Madeiran-Iberian* (2-3) clades of the statistical parsimony network of Fig. 2a, respectively. Black and grey symbols represent interior and tip 1-step clades, respectively. Solid black lines connect 2-step clades. Dashed grey lines connect 1-step clades within the same 2-step clade. Asterisks indicate significant associations between haplotypes and geographic locations according to ParaFit method (Legendre *et al.* 2002) (Table 3).

The Macaronesian red fescues

The Macaronesian red fescues (*Festuca* Poaceae) constitute a model group of recently diverged Macaronesian species (Fig. 1) with close extant relatives in the SW European continent (Díaz-Pérez *et al.* 2008; Inda *et al.* 2008). The group comprises four related species, the Canarian *Festuca agustinii* Linding, the Madeiran *Festuca jubata* Lowe, and the Azorean *Festuca francoi* Fern.Prieto, C.Aguiar, E.Dias & M.I.Gut and *Festuca petraea* Guthnick ex Seub (Díaz-Pérez *et al.* 2008), distributed in three of the four Macaronesian archipelagos. All species grow in relatively high-altitude cliffs (from 400 to 2000 m. a. s. l.) in the central-western Canary Islands (Gomera, Gran Canaria, El Hierro, La Palma, Tenerife), Madeira and the Azores, respectively. The only exception is *F. petraea*, which is found in coastal halophytic soils of the Azores. These diploid lineages (Sequeira *et al.* 2009), together with the single Iberian continental diploid relative *Festuca rivularis* (Catalán 2006), were resolved as earlier branching lineages with respect to the more recently evolved clade of high polyploid red *Festuca* taxa within the crown of the fine-leaved *Festuca* sect. *Aulaxyper* s. l. clade (Catalán 2006; Inda *et al.* 2008). Diploid Macaronesian and Iberian red *Festuca* grasses were found to be of relatively recent Pliocene origin, as they collapsed in a basal polytomy in the *Festuca* sect. *Aulaxyper* clade, dated at 2.5 ± 0.9 Ma using plastid *trn*TF and nuclear ITS data (Catalán 2006; Inda *et al.* 2008). Independent evidence from AFLP markers supported this scenario (Díaz-Pérez *et al.* 2008), where few private markers were found for each of the four island species. Genetic differentiation among the Macaronesian fescues was mainly attributed to differences in allelic frequencies in all marker loci, pointing towards the existence of putative ancient polymorphism, which is theoretically associated with recent divergence times (Rosenberg 2003).

AFLP markers were also used to characterize the genetic structure, phylogenetic relationships and biogeographic patterns of the Macaronesian red fescues (Díaz-Pérez *et al.* 2008). Aside from the practical benefit of employing a large number of hypervariable AFLP markers, their usefulness was supported on the assumption that the sharing of homologous fragments reflected a common ancestry, even if individual loci might not all reconstruct the same coalescent history. Using a set of individuals that largely overlapped the current set of samples (all except the continental *F. rivularis* and one *F. jubata* population that were not included in the study of Díaz-Pérez *et al.* (2008)), AFLP markers were used to subdivide each species according to Bayesian intraspecific groups (sensu Díaz-Pérez *et al.* 2008). Those optimal groups were generated through the F model of the STRUCTURE version no.2.2 program (Pritchard *et al.* 2000; Pritchard & Wen 2004; Falush *et al.* 2007). Each grouping represented pools of individuals highly related to each other, which probably originated from recent common ancestry from a few number of founder individuals. Given that the Bayesian Priority Consensus Tree (PCT; Luo *et al.* 2007; Díaz-Pérez *et al.* 2008) and the F model showed that each species was represented by a monophyletic assemblage of individuals, it was assumed that Bayesian groups subdivided the genetic pool of a species into groups of populations sharing a more recent intraspecific ancestor. The F model identified two Bayesian groups of geographically contiguous populations that subdivided *F. agustinii* into the eastern Gran Canaria, Tenerife and Fagus9 (La Gomera) group (coded as *Ae* in Tables 1, 4 and Table S1, Supporting information) and the western La Palma, El Hierro and Fagus8 (La Gomera) group (*Aw*), in addition to a possible genetic bridge between both groups represented by the western population Fagus10 located in the northern section of La Palma. The Azorean species *F. francoi* was subdivided into the eastern Sao Miguel populations (*Fe*) and the geographically contiguous central and western subarchipelagos populations (*Fw*). In *F. petraea*, the Bayesian analyses identified a group that followed a southern axis along the three subarchipelagos of Azores, formed by Santa Maria, Faial and Flores populations (*Ps*), and a northern group located exclusively in the northern section of the central subarchipelago of Azores, represented by Graciosa, Pico and Sao Jorge populations (*Pn*). The Madeiran *F. jubata* showed a homogeneous genetic structure, except for the Fjuba6 population that diverged as a separate Bayesian group. Nevertheless, in this study, we decided to unite Fjuba6 individuals to the

Table 1 Molecular diversity estimates of *trn*LF sequences of diploid red *Festuca* species: *Festuca agustinii* (Fagus), *Festuca francoi* (Ffran), *Festuca petraea* (Fpetr), *Festuca jubata* (Fjuba) and *Festuca rivularis* (Frivu), *Ntrn* designation refers to nucleotide polymorphism; *Ntrn + G* includes three additional multiresidual gaps (see Tables S1 and S2 for details, Supporting information)

			Ntrn			*Ntrn + G*			
Population	Locality	*N*	*Ss*	*nh*	θπ	*Ss*	*nh*	Θπ	Gap frequency G1/G3/G6
Festuca agustinii	Canary Islands	**45**	**9**	**8**	**1.9313 (0.20–5.11)**	**10**	**8**	**2.1677 (0.38–5.51)**	**0.13/1/1**
Ae		22	7	5	2.1558 (0.31–5.58)	8	5	2.5714 (0.45–6.79)	0.27/1/1
Aw		23	3	3	0.9881 (0.00–2.96)	3	3	0.7981 (0.00–2.94)	0/1/1
Festuca francoi	Azores	**23**	**5**	**3**	**1.7075 (0.17–4.61)**	**5**	**3**	**1.7075 (0.17–4.67)**	**1/1/1**
Fe		6	0	1	0	0	1	0	1/1/1
Fw		17	3	2	1.1471 (0.00–3.21)	3	2	1.1471 (0.00–3.38)	1/1/1
Festuca petraea	Azores	**24**	**4**	**3**	**1.0978 (0.00–3.27)**	**6**	**4**	**1.7174 (0.23–4.60)**	**1/0.88/0.75**
Pn		13	0	1	0	1	2	0.5385 (0.00–1.85)	1/1/0.54
Ps		11	4	3	1.9636 (0.18–5.42)	5	3	2.4000 (0.18–6.40)	1/0.73/1
Festuca jubata	Madeira	**15**	**2**	**3**	**0.7810 (0.00–2.95)**	**2**	**3**	**0.7810 (0.0–2.50)**	**0/1/1**
Jm		13	1	2	0.5128 (0.0–1.82)	1	2	0.5128 (0.00–1.82)	0/1/1
Jp		2	0	1	0	0	1	0	0/1/1
Festuca rivularis	Iberian Peninsula	**23**	**2**	**3**	**0.3874 (0.00–1.49)**	**3**	**4**	**0.8300 (0.00–2.64)**	**0.70/1/1**
Rs		12	0	1	0	0	1	0	1/1/1
Rp		11	2	3	0.6909 (0.00–2.29)	3	3	1.2000 (0.00–3.49)	0.36/1/1
Total		130	18	15	1.9751 (0.33–5.15)	21	17	2.6112 (0.55–6.65)	

N is the sample size. Haplotype frequency parameters: number of segregating sites (*Ss*), number of distinct haplotypes (*nh*), molecular diversity θπ estimates, and gap frequency. *Ae*, *Aw*, *Fe*, *Fw*, *Pn*, *Ps* denote Bayesian groups (*F. agustinii* east, *F. agustinii* west, *F. francoi* east, *F. francoi* west, *F. petraea* north, *F. petraea* south) following Díaz-Pérez *et al.* (2008) (see text for details); *Jm* and *Js* represent a pool of *F. jubata* samples from Madeira and Porto Santo Islands, respectively; *Rs* and *Rp* represent *F. rivularis* populations from Sierra Nevada and the Pyrenees, respectively. Within brackets, the 95% confidence intervals of θπ generated through 10 000 θ-based simulations under the coalescence model using the program DNAsp version no. 5.10.01 (Librado & Rozas 2009). Species estimations, highlighted in boldface, were obtained after pooling all individuals of each species.

Table 2 Estimates of the number of shared mutation (*shm*; above diagonal) and the average number of nucleotide differences (*d*; below diagonal) of five diploid red *Festuca* species. Analyses were conducted with DNAsp version no. 5.10.01 (Librado & Rozas 2009)

	Festuca agustinii	*Festuca jubata*	*Festuca rivularis*	*Festuca francoi*	*Festuca petraea*
F. agustinii	—	1	0	0	0
F. jubata	3.06	—	0	0	0
F. rivularis	2.00	1.62	—	0	0
F. francoi	2.82	2.44	1.26	—	3
F. petraea	2.40	2.03	0.84	1.50	—

remaining Madeiran individuals to increase the sample size for Bayesian analyses (coded as *Jm* in Tables 1, 4 and Table S1, Supporting information).

The Canarian *F. agustinii* was the most divergent species of the group according to the F model and the PCT, whereas the Madeiran *F. jubata* was placed at an intermediate position between *F. agustinii* and the group of closely related Azorean species *F. francoi* and *F. petraea* (Díaz-Pérez *et al.* 2008). Interspecific analyses based on neutral correlations of standardized residuals (Nicholson *et al.* 2002) among Bayesian groups generated by the nonadmixture model of STRUCTURE were indicative of independent history among the archipelagos' species, which was further interpreted as three independent colonization events (*3IC* model) from the continent to Canary Islands, Madeira and Azores, respectively (Díaz-Pérez *et al.* 2008). According to this model, the Azorean species probably evolved from a single founder event. Furthermore, the paraphyletic nature of *F. francoi* with respect to *F. petraea* in the PCT suggested that *F. petraea* probably originated from *F. francoi* by *in situ* speciation, following ecological adaptation.

Biogeographic scenarios

In this study, biogeographic scenarios were proposed according to floristic and evolutionary evidence in the red fescues and in other Macaronesian groups of plants. These scenarios were connected to species-history models, in which lineage divergence was modelled according to species-tree branching patterns. The independent evolution of the ancestral lineages (*Radiation* species-history model; Fig. 3a), derived from the small time elapsed since the radiation of the *Festuca* sect. *Aulaxyper* (Inda *et al.* 2008), suggests the existence of a three-independent-colonization (*3IC*) biogeographic scenario, where different mainland ancestors gave rise to the current 3-archipelagos geographic ranges of the Macaronesian red fescues. This scenario is compatible with relationships recovered from other Macaronesian endemic flora, for which the Canarian and the Madeiran archipelagos, and to a lesser extent the Azorean archipelago, have been highly influenced by different continental sources (Schäffer 2003; Carine *et al.* 2004; Comes 2004; Carine & Schäffer 2010). An alternative biogeographic scenario to the *3IC* model was modelled by the *Canarian* model. This model involves the Canarian archipelago being colonized by an ancient lineage followed by an independent more-recent sister lineage radiation that gave rise to the Madeiran and Azorean lineages. Earlier isolation of the Canarian *F. agustinii* from the remaining taxa has been suggested from the AFLP data (Díaz-Pérez *et al.* 2008).

We also propose, as another alternative model, the *Sequential* model, which implies the Canarian archipelago being colonized by an ancient lineage, followed by the successive divergences of more recent ancestral lineages that subsequently and independently colonized the other archipelagos originating the Madeiran and the Azorean lineages. In addition to independent colonization scenarios, other models involving stepping-stone colonizations among archipelagos could be compatible with phylogenetic analyses of several Macaronesian angiosperm groups that have suggested colonizations from the Canary Islands to Madeira (e.g. Kim *et al.* 1996, 2008; Panero *et al.* 1999; Francisco-Ortega *et al.* 2002; Trusty *et al.* 2005; Meimberg *et al.* 2006). Also, connections between the Azores and another Macaronesian archipelago have been inferred (e. g. Panero *et al.* 1999; Moore *et al.* 2002), possibly explaining why 10.2% of the Azorean endemic taxa show phyto-geographic relationships to Madeira + Canary Islands (Schäffer 2003; Carine & Schäffer 2010). In this sense, the *Sequential* model could also be interpreted as a stepping-stone colonization model of the common ancestor from the Canary Islands to Madeira and then to Azores, which could also explain the intermediate genetic AFLP composition of Madeiran *F. jubata* between those of Canarian *F. agustinii* and Azorean *F. francoi* and *F. petraea* (Díaz-Pérez *et al.* 2008). Additionally, colonizations from the Canary Islands to Madeira coupled with an independent colonization of Azores from other continental source areas could have been also possible (*Dichotomous* model). In all the analysed biogeographic scenarios, the addition of the close continental relative *F. rivularis* favoured the reconstruction of the historical distribu-

Table 3 Parameters of the phylogeographic signal analysis using the host–parasite co-evolutionary model implemented in ParaFit (Legendre *et al.* 2002). F1 estimators tested global and specific association between haplotype and geographic location. Geography was represented by sampled locations (see Fig. 1 for details)

NGtrn2 haplotype	Population	F1	Significance level
Global test		4.7917	0.0000**
h1	Ffran3	0.2894	0.0001**
h1	Ffran4	0.3173	0.0001**
h1	Ffran5	0.3024	0.0002**
h1	Ffran6	0.2951	0.0001**
h1	Fpetr4	0.3192	0.0001**
h1	Fpetr5	0.3065	0.0001**
h1	Fpetr6	0.3056	0.0001**
h1	Frivu1	−0.2940	1.0000ns
h2	Frivu2	0.0411	0.2859ns
h3	Fpetr2	0.2840	0.0012**
h3	Fpetr3	0.2826	0.0012**
h4	Fagus1	−0.1896	0.9999ns
h5	Fagus2	−0.1880	0.9999ns
h6	Ffran7	0.7861	0.0003**
h6	Fpetr7	0.7858	0.0006**
h7	Ffran1	0.2721	0.0114*
h7	Ffran2	0.2827	0.0096**
h8	Fpetr1	0.1983	0.0322*
h9	Fpetr5	0.0406	0.1243ns
h9	Frivu2	0.0704	0.0282*
h10	Fagus13	0.0974	0.0797ns
h10	Fjuba7	0.0571	0.2061ns
h10	Fjuba5	0.0587	0.2008ns
h11	Fjuba1	0.0876	0.2283ns
h11	Fjuba7	0.0850	0.2379ns
h11	Fjuba5	0.0877	0.2318ns
h11	Fjuba6	0.0874	0.2313ns
h12	Frivu2	0.1342	0.0101*
h13	Fagus8	0.2206	0.0030**
h13	Fagus15	0.2244	0.0020**
h14	Fagus5	0.3459	0.0078**
h15	Fagus12	0.2940	0.0088**
h15	Fagus14	0.2977	0.0068**
h15	Fagus10	0.2944	0.0076**
h15	Fagus11	0.2907	0.0072**
h16	Fagus9	0.3371	0.0029**
h16	Fagus3	0.3401	0.0031**
h16	Fagus7	0.3424	0.0037**
h17	Fagus4	0.3940	0.0055**

Significance level was based on 9999 permutations. ***0.1%, **1%, *5%, ns, nonsignificant.

tions of the Macaronesian red fescues within a larger co-evolving continental–oceanic geographic setting.

These biogeographic scenarios were tested through maximum-likelihood-based dispersal-extinction-cladogenesis (DEC) analyses (Ree & Smith 2008) on the species-history model most supported by current plastid data. A species-history model is a summary of evolutionary events represented by a tree-like structure. Four species-history models that mirrored the alternative biogeographic scenarios were tested. The *Radiation* model was related to *3IC*, because it included a polytomy at the root of the evolutionary tree simulating the basal radiation of *Festuca* sect. *Aulaxyper* (cf. Inda *et al.* 2008). The *Canarian* model was constructed to fit the ancient divergence of the *F. agustinii* lineage in the Canary Islands (cf. Díaz-Pérez *et al.* 2008). By contrast, the *Sequential* model was represented by a ladder-shaped tree, not only simulating stepping-stone colonizations from the Canary Islands to Madeira and from Madeira to Azores, but also simulating a 3IC scenario where colonizations occurred at different evolutionary times. Finally, a putative colonization of Madeira from the Canary Islands was represented by a close connection between *F. agustinii* and *F. jubata* in the *Dichotomous* model, being characterized by a four-taxon symmetric tree, further connecting *F. rivularis* + the Azorean species.

Alternative species-history models are reflected in different branching pattern of gene trees (Knowles & Maddison 2002). Consequently, discordance between species-histories and gene trees could be employed to evaluate the relative support of alternative models (Knowles 2001, 2004; Knowles & Maddison 2002; Bailey *et al.* 2005; Carstens *et al.* 2005; Steele & Storfer 2006; Knowles *et al.* 2007). A common measure of discordance is the Slatkin & Maddison (1989)'s *S* statistics, which measures the number of additional parsimony steps in a given tree that are needed for the gene tree to correspond to the population/species history in which it is contained (Knowles 2001; Bailey *et al.* 2005). Gene trees simulated by a neutral coalescent process under a species-tree topology, or the null model, are used to derive an expected distribution of the test statistic against which to compare the observed value (Hickerson *et al.* 2010). In addition, as an alternative test to discordance analysis, distance-based statistics can be used to take into account a direct relationship between pairwise distances and coalescence times of genes (Slatkin & Hudson 1991). In this case, an expected distribution of pairwise genetic distances can be generated from simulated sequences for each species-history model and tested in the same way as discordance statistics (Carstens & Richards 2007).

The aims of this study were (i) to test the alternative *Canarian*, *Radiation*, *Dichotomous* and *Sequential* models of divergence to infer the history and timing of species divergence events in the diploid Macaronesian and Iberian red *Festuca* group using the plastid *trn*LF data set; (ii) to compare the times of divergence of the studied Macaronesian and Iberian red fescues, estimated from current coalescence-based approaches, with those obtained from previous phylogenetic-based Bayesian relaxed-clock methods; (iii) to detect phylogeographic

Table 4 IM highest posterior probability estimates of model mutation ($\theta = 2N_e\mu$), splitting time ($t = \mathbf{t}\mu$) and migration rates ($m = \mathbf{m}/\mu$) for five Macaronesian and Iberian diploid red fescues based on the analysis of plastid *trn*LF sequences. Model and demographic parameters are highlighted in italics and boldface, respectively. Description of Bayesian group and species polymorphism is given in Table 1 and Table S1 (Supporting information)

	N	θ	$\mathbf{N_e}$ ($\times10^{-3}$)	*t*	**t** ($\times10^{-3}$)	$m(\times10^{6})$	$\mathbf{N_em}$	LRT
Between Bayesian groups								
Ae vs.	22	**1.35** (0.45–3.62)	**769** (255–2056)	**1.33** (udf)	**756** (udf)	**0.004**	**0.002**	**0**
(Jm + Jp)	15	**0.24** (0.03–1.40)	**139** (15–797)			**0.009**	**<0.001**	**0**
Ae vs.	22	**1.24** (0.32–3.72)	**707** (180–2117)	**1.38** (udf)	**784** (udf)	**0.005**	**0.002**	**0**
Fw	17	**0.31** (0.01–1.98)	**178** (7–1128)			**0.009**	**<0.001**	**0**
Ae vs.	22	**1.06** (0.22–3.18)	**603** (125–1814)	**1.15** (udf)	**658** (udf)	**0.009**	**0.003**	**0.01**
Rp	11	**0.46** (0.04–3.02)	**260** (22–1721)			**0.007**	**<0.001**	**0**
(Jm + Jp) vs.	15	**0.13** (0.01–1.11)	**73** (6–635)	**0.44** (udf)	**248** (udf)	**0.035**	**0.001**	**<0.01**
Fw	17	**0.38** (0.02–2.22)	**214** (10–1263)			**0.009**	**<0.001**	**0**
(Jm + Jp) vs.	17	**0.14** (0.01–1.38)	**79** (6–785)	**0.52** (udf)	**296** (udf)	**1.465**	**0.058**	**0.22**
Rp	11	**0.26** (0.02–4.92)	**145** (9–2801)			**0.009**	**<0.001**	**0**
Rp vs.	11	**0.37** (0.03–6.54)	**208** (14–3724)	**0.08** (udf)	**44** (udf)	**0.009**	**<0.001**	**0**
Fw	17	**0.22** (0.01–3.36)	**123** (7–1917)			**3.484**	**0.214**	**0.17**
Between species								
Festuca agustinii vs.	45	**1.67** (0.64–3.89)	**950** (367–2215)	**1.00** (udf)	**572** (udf)	**1.021**	**0.49**	**1.18**
Festuca jubata	15	**0.23** (0.03–1.22)	**128** (16–694)			**0.009**	**<0.001**	**0**
F. agustinii vs.	45	**1.83** (0.77–4.08)	**1043** (434–2321)	**1.07** (udf)	**609** (udf)	**0.005**	**0.003**	**0**
(*Festuca petraea* + *Festuca francoi*)	47	**1.03** (0.34–2.73)	**584** (192–1553)			**0.003**	**<0.001**	**0**
F. agustinii vs.	45	**1.77** (0.74–3.99)	**1009** (422–2275)	**0.92** (udf)	**523** (udf)	**0.007**	**0.003**	**0.02**
Festuca rivularis	23	**0.33** (0.05–1.55)	**186** (26–883)			**0.004**	**<0.001**	**0**
F. jubata vs.	15	**0.13** (0.01–1.04)	**72** (6–592)	**1.00** (udf)	**566** (udf)	**1.042**	**0.037**	**0.19**
(*F. petraea* + *F. francoi*)	47	**1.02** (0.29–2.62)	**583** (166–1492)			**0.005**	**0.002**	**0.01**
F. jubata vs.	15	**0.15** (0.01–1.87)	**84** (5–1067)	**.92** (udf)	**521** (udf)	**2.127**	**0.089**	**0.78**
F. rivularis	23	**0.17** (0.01–1.64)	**97** (7–933)			**0.011**	**<0.001**	**0**
F. rivularis vs.	23	**0.36** (0.03–2.24)	**207** (19–1276)	**0.51** (udf)	**309** (udf)	**0.011**	**0.001**	**<0.01**
(*F. petrea* + *F. francoi*)	47	**0.86** (0.18–2.82)	**492** (101–1603)			**4.591**	**1.130**	**0.42**

N is the number of sequences sampled per Bayesian group or per species. Values within brackets indicate 90% highest posterior density (HPD) bounds for posterior distributions. Estimates of the effective number of gene copies $\mathbf{N_e}$ ($\times10^3$) and the time since population splitting **t** ($\times10^3$) in generations were obtained assuming a per locus per generation mutation rate μ of 1.7549×10^{-6}. The population migration rate was obtained as $\mathbf{N_em} = \theta m/2$. LRT is a likelihood ratio test against the null hypothesis (H_0) of $m = 0$, in which LRT < 1 was taken as indicative of no rejection of H_0. Bayesian group and population codes are given in Table 1. udf, undefined.

signals using the classical nested clade phylogeographic analysis (NCPA) and a firstly proposed host–parasite co-evolution method; and (iv) to reconstruct the most likely biogeographic colonization scenario according to a coalescent-based supported species-history model of species divergence.

Material and methods

Collections and geographic sampling

The sample sizes of 34 Macaronesian (*F. agustinii*: Canary Islands, *F. jubata*: Madeira, and *F. francoi*, *F. petraea*: both in the Azores) and two continental (*F. rivularis*: Iberian Peninsula) diploid red *Festuca* populations and their geographic localities are shown in Fig. 1 and Table S1 (Supporting information). One hundred and thirty samples from all Macaronesian islands where endemic red fescues have been reported were included. *F. rivularis* samples were obtained from two geographically separated mountain ranges of Spain, the southernmost Sierra Nevada and northernmost Pyrenean ranges. Additional data from 13 species of the genera *Festuca* and *Vulpia* that cover at least four taxonomic sections of the fine-leaved Loliinae (Inda *et al.* 2008) plus the close outgroup *Poa infirma* were incorporated to the analysis. The additional samples were used to reconstruct the evolutionary relationships of these diploid red fescues within the *Festuca* sect. *Aulaxyper* clade and to estimate the rates of nucleotide substitution in their plastid *trn*LF sequences.

DNA isolation and sequencing

DNA isolation was carried as described in the study by Díaz-Pérez *et al.* (2008). The plastid *trn*LF region was

PCR-amplified using the c-f primer combination (Taberlet *et al.* 1991) following Torrecilla *et al.* (2004) with minor modifications. PCR products were purified using QIAgen purification kit, sequenced using the ABI PRISM Dye Terminator Cycle Sequencing kit (Qiagen) and analysed in an automated ABI PRISM 377 sequencer. Chromatograms from direct and reverse directions were corrected and assembled using SEQUENCHER. Clustal W (Thompson *et al.* 1994) was used to obtain a matrix of aligned sequences that was subsequently edited by hand. A data matrix of 751 *trn*LF aligned nucleotide positions plus a set of eight mono- and six multiresidual unambiguously aligned gaps (multiresidual gaps were coded as G1 to G6, ranging from 4 to 7 nucleotide sites) were created for this study. Gaps were coded as binary characters. The newly generated *trn*LF sequences of *F. agustinii*, *F. francoi*, *F. jubata*, *F. petraea* and *F. rivularis* were submitted to GenBank under Accession nos HM068379 to HM068506 (Table 1).

The set of other Loliinae sequences included in the studies by Catalán *et al.* (2006) and Inda *et al.* (2008) were downloaded from GenBank including *Festuca* sect. *Aulaxyper*: *F. rivularis* 1 (2×; GenBank AF478512), *F. rivularis* 2 (2×; EF593000), *F. ampla* (2×-4×-6×; EF592952), *F. iberica* (6×; AF478516), *F. nevadensis* (10×; AF478514), *F. rubra* subsp. *pruinosa* (6×; EF592995), *F. rubra* subsp. *arctica* (6×; EF592997), *F. rubra* subsp. *rubra* (6×; AY118098), *F. rubra* subsp. *megastachys* (6×-8×; AY118088); *Festuca* sect *Festuca*: *F. alpina* (2×; AF478522), *F. ovina* (2×; AF533063); *Festuca* subsect. *Exarate*: *F. borderei* (2×; AF478510), *F. capillifolia* (2×; AF478511); *Festuca* sect. *Eskia*: *F. eskia* (2×; AF478508); *Vulpia* sect. *Vulpia*: *V. bromoides* (2×; AF487616), *V. muralis* 1 (2×; AF478526), *V. muralis* 2 (2×; AY118102); and outgroup *Poa infirma* (2×; AF488773).

Polymorphism of trnLF haplotypes

Within-species and within-population molecular diversity was estimated through the average number of pairwise nucleotide differences between DNA sequences ($\theta\pi$, Tajima 1983), the number of haplotypes (*nh*) and the number of segregating sites (*Ss*). Although *Ss* is not a good statistic for estimating DNA polymorphism as it depends on the sample size (Tajima 1989), it is included here given its high sensitivity to the presence of rare haplotypes in the population. Genetic diversity among species was quantified through the average number of nucleotide differences (*d*) and the number of shared mutations (*shm*) among haplotypes of pairwise species comparisons. All analyses were performed including or excluding gaps using the program DNAsp version no. 5.10.01 (Librado & Rozas 2009). Binary gaps were recoded as two arbitrarily nucleotide-based characters.

Evolutionary relationships of trnLF haplotypes

Phylogenetic relationships of diploid Macaronesian and Iberian red *Festuca* and other representatives of *Festuca* sect. *Aulaxyper* and of Loliinae were reconstructed using the classical maximum parsimony, maximum-likelihood and Bayesian inference methods (Huelsenbeck & Ronquist 2001; Felsenstein 2004) (Fig. S1; Supporting information). With these sets of analyses, we intended to recover a more resolved topology for the basal polytomy of *Festuca* sect. *Aulaxyper*, now with a greater number of samples from the diploid red fescues than that originally employed in the wide-scope biogeographic study of Loliinae (Inda *et al.* 2008). In addition, to handle any putative lack of phylogenetic resolution originated by low levels of diversity usually found at intraspecific levels (Woolley *et al.* 2008), we represented the genealogical relationships of haplotypes through statistical parsimony analysis (Templeton *et al.* 1992). Data matrices with gaps (G1, G3 and G6; Table S2, Supporting information), re-coded using two arbitrarily chosen nucleotide bases, or without gaps were used in this analysis. Patristic distances among haplotypes obtained from the statistical parsimony network were computed with TCS (Clement *et al.* 2000).

Phylogeographic signal detection

We performed the classical NCPA (Templeton *et al.* 1995; Posada *et al.* 2000; Templeton 2004) for detecting associations of haplotypic variation with its geographic distribution (Table S3, Supporting information).

In addition, we introduced a new approach to detect phylogeographic signal directly adapted from a method developed for the detection of host–parasite co-evolution and implemented in ParaFit (Legendre 2001; Legendre *et al.* 2002). Here, co-evolution is characterized as the extent to which host and parasites occupy corresponding positions in the phylogenetic trees. Consequently, a direct relationship between host and parasite patristic distances is expected for a group of organisms showing co-evolution. To exploit such condition, ParaFit combines three types of information that are jointly necessary: a matrix of patristic distances among hosts (H), a matrix of patristic distances among parasites (L) and a binary matrix of presence/absence of host–parasite associations (A). Matrices H and L are then transformed into the matrices of principal coordinates (eigenvectors) Hc and Lc, respectively. Using the multivariate function D = Hc A′ Lc, a matrix that represented the H-L associations (D) is generated, in which the sum of squares of the values of D represented the trace of the matrix D′D, or F1, a global statistic to test the hypothesis of association. Specific F1 estimators that measure each of the

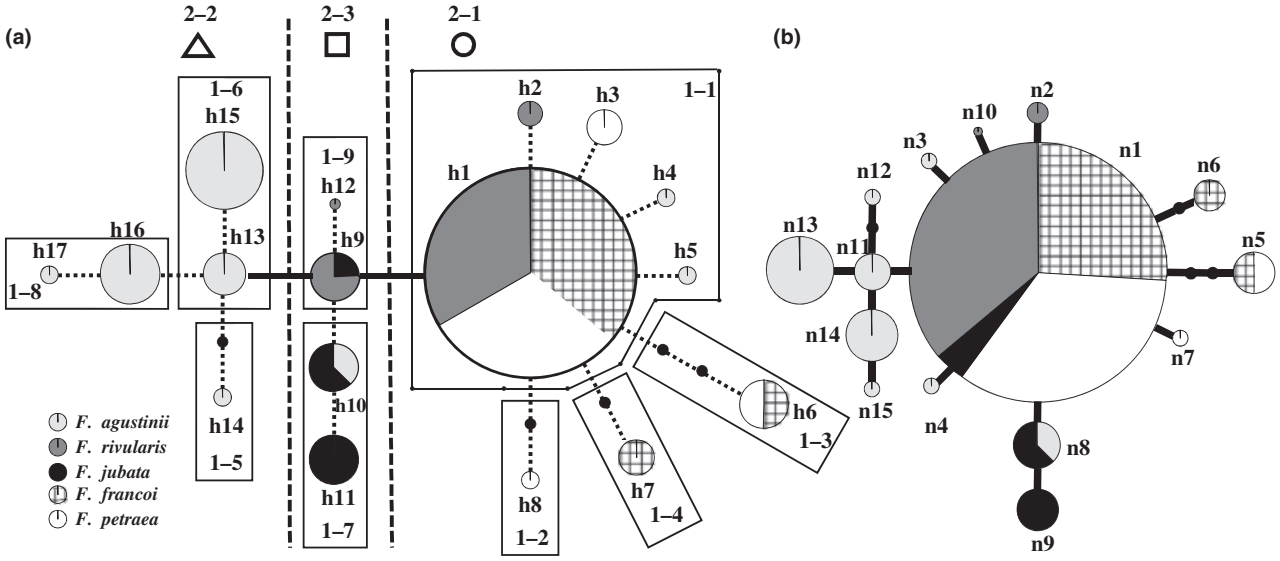


Fig. 2 (a) Statistical parsimony network and associated nested layout for haplotypes of five diploid red *Festuca* species. Haplotypes (see Table S2, Supporting information) are represented by circles with sizes approximately proportional to their relative frequencies; fill pattern areas are proportional to species assignations. Solid-line polygons and dashed lines enclose 1-step and 2-step clades, respectively. Solid black and dotted black lines connect 2-step clades and 1-step clades within the same 2-step, respectively. Δ = *Canarian* (2-2), □ = *Madeiran–Iberian* (2-3), ○ = *Azorean–Iberian* (2-1) 2-step clades. Geographic localities of haplotypes, 1-step and 2-step clades are shown in Fig. 1. (b) Statistical parsimony network for haplotypes without considering gaps. Circle sizes and fill pattern assignations are the same as in Fig. 2a.

individual associations between host and parasites are also computed. Permutation of the elements of A 10 000 times generates a distribution of F1 values under the null hypothesis of randomness of the H-L associations, whose rejection is achieved using a 5% one-tailed test.

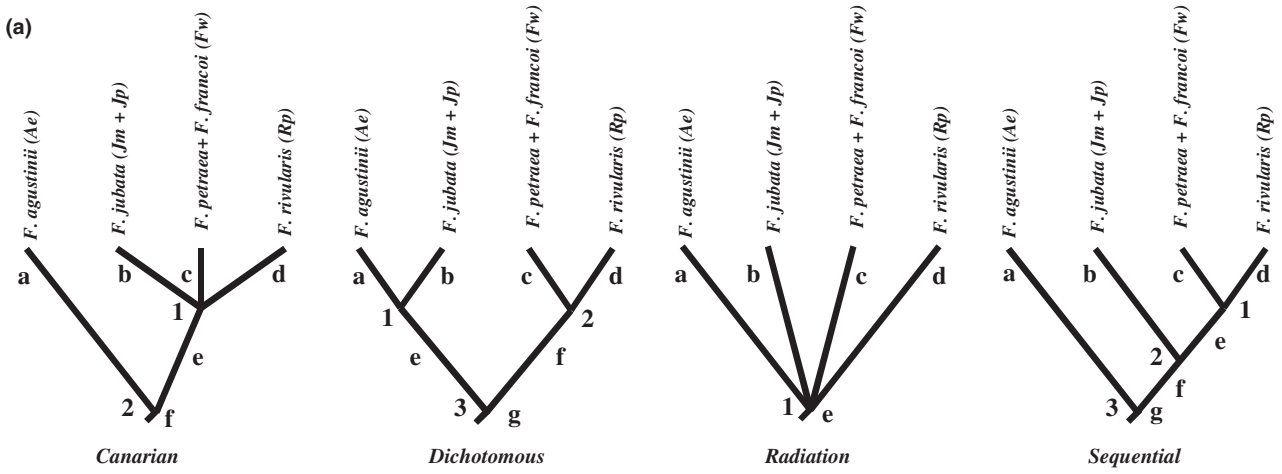
In our case, hosts and parasites were replaced by localities and haplotypes, respectively. Patristic distances among haplotypes were extracted from the statistical parsimony network (Fig. 2); however, the patristic distances of hosts were replaced by a linear geographic distance matrix among localities (L). Nonindependence of H-L associations was taken as a direct relationship between geographic and patristic distances, which is an indicator of restricted gene flow by isolation-by-distance (IBD) or by a related variant, like the stepping-stone model (SSM).

Bayesian estimation of demographic parameters

In order to place divergence events on a chronological framework, pairwise divergence times were computed using an 'isolation-with-migration' model that was fitted to nucleotide data according to the Bayesian method of Hey & Nielsen (2004), as implemented in the program IM (revision of 3 May 2007). Despite the IM model is designed to estimate evolutionary parameters between a pair of closely related populations without migration to any other third population, simulation studies have shown that it can give unbiased estimates of divergence times, even though they are sampled from a larger population data set (Hey 2010). In this case, both populations should not violate the assumptions of the IM model, that is, absence of gene flow in the larger data set and no population size changes. We found that the diploid red fescues did not violate these assumptions. For example, we did not detect evidence of population expansion at the Bayesian group level or at the species level through the Fu's Fs test (Fu 1997; data not shown), nor we detect any conclusive indications of existence of gene flow using IM models (see Table 4). Consequently, we calculated averaged values of individual estimates in different pairwise analyses to construct initial parameters (N^0T^0) for each of the species-history models that were further used in subsequent simulations (See next section for details; Figs 3b and 4).

Although recent simulations have shown that the isolation-with-migration model is relatively robust to another violation of the model, that is, some departure from random mating or substructure (Strasburg & Riesenberg 2010), we chose to work with single Bayesian groups, thus minimizing the effect of genetic structure within species. For this, Bayesian groups were considered as genetic units because they represented a natural choice to pooling individuals in order to increase sample sizes and genetic variability. We assumed similar genetic relationships between any pair

(a)



(b)

					Canarian				*Dichotomous*						*Radiation*		*Sequential*					
	a	**b**	**c**	**d**	**e**	**f**	**1**	**2**	**e**	**f**	**g**	**1**	**2**	**3**	**e**	**1**	**e**	**f**	**g**	**1**	**2**	**3**
$PpN^{0}T^{0}$	6.9	1	1.7	2.0	3.9	2.7	2	7.3	–	5.7	2.8	5	0.4	5	3.3	4.7	5.7	3.0	2.7	0.4	2.7	7.3
$PpN^{2}T^{0}$	21	8	19	17	19	21	2	7.3	–	19	21	5	0.4	5	21	4.7	19	19	21	0.4	2.7	7.3
$PpN^{1}T^{2}$	8	2	2	2	6	6	3	10	6	6	6	7	0.5	10	6	10	6	6	6	0.5	3	10
$PpN^{1}T^{3}$	8	2	2	2	6	6	3	15	6	6	6	7	0.5	15	6	15	6	6	6	0.5	3	15
$PpN^{-1}T^{1}$	2	1	1	1	1	2	2	8	–	1	2	5	0.5	5	2	5	1	1	2	0.5	3	8
$PpN^{-1}T^{2}$	2	1	1	1	1	2	3	10	2	1	2	7	0.5	10	2	10	1	1	2	0.5	3	10
$PpN^{-1}T^{3}$	2	1	1	1	1	2	3	15	2	1	2	7	0.5	15	2	15	1	1	2	0.5	3	15
$PpN^{-2}T^{1}$	1	0.5	0.5	0.5	0.5	1	2	8	–	0.5	1	5	0.5	5	1	5	0.5	0.5	1	0.5	3	8
$SpN^{0}T^{0}$	10	1	5.5	1.6	4.1	0.8	4.7	5.7	–	6.1	1.8	5.5	3.1	5.5	2.4	5.2	6.1	3	0.8	3.1	5.4	5.7
$SpN^{-1}T^{3}$	3.7	0.1	1	0.1	1	3.7	3	15	3.7	1	3.7	7	0.5	15	3.7	15	1	1	3.7	0.5	3	15

Fig. 3 Species-trees history models tested using coalescent modelling. (a) Four basic models were tested: *Canarian, Dichotomous, Radiation and Sequential*. Letters a to g represent branch widths ($N \times 10^5$), and nodal numbers 1-3 indicate nodal ages ($T \times 10^5$), where N is the effective population size and T is the number of generation elapsed backwards from present times. *Ae*, (*Jm + Jp*), *Fw* and *Rp* Bayesian groups were selected as representatives of *Festuca agustinii*, *Festuca jubata*, the Azorean species and *Festuca rivularis*, respectively. (b) Parameters for coalescent simulations under different scenarios. Simulations at the Bayesian group level (scenarios with Pp prefixes) and at the species level (Sp prefixes) were performed using the program Mesquite version no. 2.5 (Maddison & Maddison 2008) and SIMCOAL version no. 1 (Excoffier *et al.* 2000), respectively. Sp scenarios included a unidirectional migration rate of $m = 1 \times 10^{-6}$ or 1×10^{-3} events per gene per generation from *F. jubata* to *F. agustinii* (Table 6). In all cases, the length of the root was fixed at 10 000 generations.

of Bayesian groups from different species based on the observed monophyly and close relationships of these red Festuca species (Díaz-Pérez *et al.* 2008; Inda *et al.* 2008). This assumption is of special relevance for most of the coalescent analyses performed in this study, because under the assumption of a molecular clock, we would not expect appreciable differences in splitting time estimations whatever set of pair of groups is used.

We also repeated all IM analysis, but merging now all populations of a species into a single genetic unit in the second run to increase sample sizes (Table 4), thereby minimizing any putative violation of the model by this cause. The Azorean species *F. petraea* and *F. francoi* were treated as a single genetic unit (Table 4), considering that they showed a sympatric distribution over a large number of islands and presented a close genetic relationship based on AFLP data (Díaz-Pérez *et al.* 2008).

IM estimates the posterior probability density for parameters of the isolation-with-migration (IM) model, which assumes that an ancestral population gave rise to two daughter populations that could have experienced gene exchange between them. Using the infinite site model, six population parameters (in italics) scaled by μ, the neutral mutation rate, were estimated directly by

the model (bold letters represent associated demographic parameters): the three populations sizes ($\mathbf{N_A}$ of the ancestral population and $\mathbf{N_1}$ and $\mathbf{N_2}$ of the descendant populations), represented by $\theta = 2\mathbf{N_e}\mu$ (twice the number of gene copies in the population), two migration rates (from $\mathbf{N_1}$ to $\mathbf{N_2}$ and vice versa), represented by $m = \mathbf{m}/\mu$ (the rate of migration per gene copy, per mutation event), and the population divergence time, represented by $t = \mathbf{t}\mu$. The population migration rate was obtained by $M = \theta m/2 = \mathbf{N_e m}$. Considering that uniform priors were used, the posterior probabilities were proportional to the likelihood of the parameters given the data, and consequently the mode of the posterior distribution provided maximum-likelihood estimates of these parameters, which are also known as the highest posterior probability (HPP) estimates. Three simulations (runs) were performed to assess convergence of the MCMC chain(s). In the first preliminary run (from 10×10^6 to 90×10^6 iterations and a burn-in of 10×10^6), wide uniform priors were assigned to set appropriate limits for the priors of the two subsequent independent longer runs (from 70×10^6 to 170×10^6 iterations and a burn-in of 10×10^6). In addition, updated accepting rates and effective sample sizes (ESS) estimates were used for assessing good mixing of the chain(s). Running parameters were adjusted to reach at least 5% of updating rates and ESS > 50.

A likelihood ratio test (LRT) was computed to discard the absence of migration ($m = 0$) between a pair of species or populations (Nielsen & Wakeley 2001; Hey 2010). From the posterior distribution of the m parameter, LRT was defined as $LRT = -2*\log(L_0/L_{max})$, where L_{max} is the maximum-likelihood value at the HPP estimate and L_0 is the likelihood at $m = 0$, that is, the likelihood when the posterior distribution intersects the ordinate for $m = 0$. The L_0 hypothesis was tested against a mixture of distributions, $\frac{1}{2}\,\chi^2_0 + \frac{1}{2}\,\chi^2_1$ (Nielsen & Wakeley 2001; Hey 2010). Given that this test is conservative at lower values, we considered that LRT values lesser than 1 ($P < 0.15$) suggested that the hypothesis of $m = 0$ could not be rejected (Rasmus Nielsen, personal communication). LRT values slightly higher than 1 but lower than 2.74 ($P < 0.05$) were taken as inconclusive because the LRT statistic apparently does not fit well theoretic expectation until higher values of LRT are achieved (Hey 2010).

Population parameter estimates can be transformed into demographic units ($\mathbf{N_e}$, $\mathbf{m}$ and $\mathbf{t}$; see above for estimation) if the μ estimate and the generation time are available. Two years per generation was assumed to be an approximate value for perennial herbaceous red fescues based on the minimum generation time hypothesis suggested by Gaut *et al.* (1992) (Catalán *et al.* 2006). The mutation rate of the *trn*LF region in the ingroup species was estimated using 23 sequences of the *Festuca* sect. *Aulaxyper* clade, a taxonomic group that has shown relatively homogeneous rates of molecular evolution in this plastid region (Catalán *et al.* 2006), and 15 sequences from the Macaronesian red fescues. The overall average number of nucleotide substitution per site (K) was estimated following the Tajima & Nei (1984) model implemented in the software MEGA version 3.1 (Kumar *et al.* 2004). The neutral mutation rate was estimated as $\mu = K/2T_D$, where the averaged K was 0.005842 and T_D was set according to the divergence time estimated for the basal radiation of *Festuca* sect. *Aulaxyper* (2.5 Ma) by Inda *et al.* (2008).

Testing species-history models

Within a recent evolutionary framework, coalescence-based methods constitute the most appropriated approaches to investigate the evolution and speciation of the Macaronesian red fescues, because they integrate incomplete lineage sorting in the estimation of parameters. Single-gene data, like the plastid *trn*LF gene, can provide estimates of evolutionary parameters and comparative information concerning relative support for some species-history scenarios over others. These maternal-line-based evolutionary scenarios could be contrasted with the biparental ones deduced from AFLP markers (Díaz-Pérez *et al.* 2008), aiming to compare the levels of congruence between the two historical reconstructions and to reveal further potential intraspecific multigenomic sources (e. g. distinct plastid haplotype donors) that could have passed undetected within the AFLP data.

Gene trees simulated by a neutral coalescent process under a species/population tree topology (the null model) were used to derive an expected distribution of the test statistic against which the observed value could be compared (Hickerson *et al.* 2010). This approach is similar to that of Steele & Storfer (2006) and Carstens & Richards (2007) where competing population-history hypotheses are tested using the discordance (*S*) and *distance-based* statistics, respectively. Following a species/population tree, the coalescent simulations of the matrices and the derived gene trees were performed under the assumption that variation among/within gene trees was attributable to genetic drift and migration events during the genealogic process.

Our approach for testing species-history models first established a prior distribution of scenarios according to IM highest posterior density (HPD) bounds. The exploration of scenarios within the range was not uniform, but embedded into discrete jumps in the parameter space following a three-criterion condition (see below; Fig. 4). In this sense, these scenarios comprised

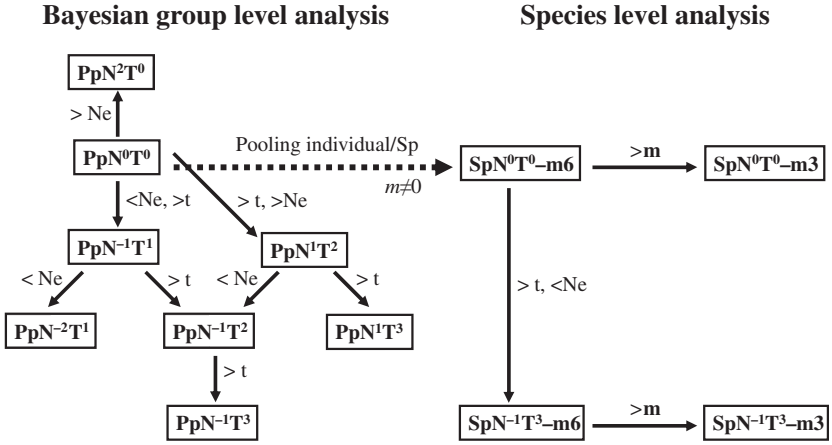


Fig. 4 Jumping/Selection scheme of demographic scenarios according to the three-criterion condition (see text for details). Simulations were generated using Mesquite version no. 2.5 (Maddison & Maddison 2008) and SIMCOAL version no. 1 (Excoffier *et al.* 2000), respectively. Each box represents a demographic scenario (Fig. 3b). Starting with the N^0T^0 scenarios, effective population size (N), migration rate (*r*) and/or nodal ages (T) were decreased '<'or increased '>'following specification in Fig. 3b. A broken arrow connects two related demographic scenarios that represented independent starting point for the jumping/selection scheme performed at the Bayesian group and at the species level.

a subset of parameter combinations selected from the large number of possible combinations for branch lengths and population sizes. Given that the optimal scenario had to fulfil the three-criterion condition better than others, exploration and selection of the optimal scenario was performed simultaneously in the same process. The *Canarian*, *Dichotomous*, *Radiation* and *Sequential* species-history models (Fig. 3a) were finally selected under the optimal scenario according to *S* and *p-distance* statistics. It must be noted that this approach was not aimed at estimating evolutionary parameters because the data, limited to a one-locus gene genealogy, were insufficient for such objective (Fan & Kubatko 2011). However, the optimal scenario could shed light about approximate values of parameters within the wide HPD bounds, representing supplementary information to IM results.

In order to avoid extensive sampling of the parameter space, we started explorations within HPD bounds with the N^0T^0 scenario (Fig. 4), where N^0 is a vector of average effective population sizes (branch width) and T^0 is a vector of average nodal ages of the tree. Each element of the vector was computed from individual IM HPP values of all pairwise comparisons comprising species or groups placed at each descending branch of a given node. From this scenario, values of N and T were simultaneously moved at discrete jumps (Fig. 4) towards the lower (N^{-2}) or upper (N^2) limits of the HPD bounds. Nevertheless, given that HPD bounds for T were undefined (see *t* or **t** in Table 4), the maximum value of the root of the trees was fixed at 1.5×10^6 (T^3) generations ago, which was defined from the upper limit (2.5 + 0.9 Ma = 1.7×10^6 generations ago; assuming 2 years per generation) of the time for the basal radiation of Festuca sect. *Aulaxyper*, as estimated by Inda *et al.* (2008). The lower limit of the root of the trees was fixed at T^0, because more recent times generated unlike scenarios that lacked power to discriminate among models.

The three-criterion condition favoured jumps in one direction of the parameter space (see arrows in Fig. 4) if the new parameter values generated a better fit to this condition. A better fit was firstly met (first criterion) when a new scenario included at least one rejected model or one nonrejected model. If all models were simultaneously rejected or accepted, it indicated unrealistic demographic parameters or lack of power of the test by insufficient DNA variability, respectively. Based on this, a model was rejected if the observed statistic *S* or *p-distance* fell outside the 95% interval of the simulated distribution (see below). Second (second criterion), a better fit was also met if a new scenario comprised a lesser number of nonrejected models. And third (third criterion), a better fit was attained if the nonrejected model(s) showed probability values of the observed statistic close to 0.5, placing them near the centre of the simulated distribution.

Given that the starting point N^0T^0 scenario was conditioned by IM HPD bounds and HPP values, we recreated coalescent-based scenarios that matched the same individuals used both at the population-level and at the species-level analyses (Table 4). Therefore, we performed two sets of simulations. The first set comprised the Bayesian groups *Ae* (*F. agustinii*), *Jm* + *Jp* (*F. jubata*), *Fw* (*F. francoi* + *F. petraea*) and *Rp* (*F. rivularis*). Simulations were performed with the program Mesquite ver-

sion 2.5 (Maddison & Maddison 2008) following a similar approach employed by Maddison & Knowles (2006). Within each population tree, gene trees were simulated using the Mesquite's Neutral Coalescent module. Along each simulated gene tree, sequence evolution was simulated with the Mesquite's Genesis module, generating a nucleotide data matrix of identical sample size and sequence length as the empirical data set. Sequence evolution was modelled approximating the F81 model (Felsenstein 1981) selected as optimal model by MODELTEST version 3.7 (Posada & Crandall 1998) for Macaronesian red fescues *trn*LF data. Scaling factors were adjusted to values ranging between 9.5×10^{-10} and 8×10^{-9} in order to generate sequence divergences comparable to empirical data, according to average θ and number of segregating sites.

The second set of simulations comprised all populations per species as a single genetic unit and treating *F. petraea* + *F. francoi* as a single species (Azores) (Table 4). In order to include unidirectional migration events from *F. jubata* to *F. agustinii* at a rate of $m = 1 \times 10^{-6}$ per gene per generation, simulations were performed using the program SIMCOAL version no. 1 (Excoffier *et al*. 2000) because Mesquite does not handle unidirectional migration rates. Although migration events were suggested by an IM LRT > 1 (=1.18) (see results; Table 4), its close proximity to 1 makes it inconclusive as indication of migration. A higher migration rate of $m = 1 \times 10^{-3}$ per gene per generation was also simulated to explore the performance of the two statistics under higher though unreal migration rates. The generation of data matrices followed the same criteria and parameters than those used for Mesquite simulations, but assuming a Kimura-2-parameter mutation model (Kimura 1980), a transition bias = 1 and a gamma shape parameter $\alpha = 0$. The mutation rate was adjusted to generate sequence divergences comparable to empirical data, according to average θ and number of segregating sites.

To compute the simulated distribution for the Slatkin and Maddison's *S* statistic (Slatkin & Maddison 1989) under a given species-history tree and parameter combination (Fig. 3), 1000 simulated sequence matrices were subjected to a Basic PAUP* (Swofford 2002) Tree Search following default scripts generated by the Batch Architect Package of Mesquite version 2.5 (Maddison & Maddison 2008). An expected distribution of *S* values was generated using the simulated gene trees through the Bar & Line Chart for Trees of Mesquite (Maddison & Maddison 2008). This test statistic considers the defined species as categorical variables, and its value could be considered as a measure of the minimum number of sorting events between species, following a specified gene tree. The value of the *S* statistic of the observed gene tree, obtained in a similar way as simulated data, is compared to the expected distribution in order to determine the significance of discord between the simulated gene trees and the species divisions reflected in each model (Steele & Storfer 2006). If the observed value falls outside the 95% of the simulated distribution, the model is rejected.

We also included a *p-distance* statistic that is expected to be directly related to coalescence times of gene trees. For each species tree–parameter combination, 500 sequence matrices were subjected to *p-distance* comparisons as implemented in PAUP* (Swofford 2002). For each matrix, an average *p-distance* was obtained from all pairwise comparisons involving sequences from different species (*F. francoi* + *F. petraea*, *F. jubata* and *F. rivularis*) or Bayesian groups (*Fw*, *Jm* + *Jp* and *Rp*), except those involving *F. agustinii* or *Ae* because they were not resolutive enough to discriminate between some models. A distribution of 500 average *p-distances* was generated and the observed *p-distance* was compared following the same criteria as that of the *S* statistics.

Biogeography analysis

Once a species-history model (*Sequential*) was chosen according to *S* and *p-distance* statistics, we selected one dispersal-extinction-cladogenesis (DEC)-based model, as developed by Ree & Smith (2008), to infer ancestral ranges and range historical inheritance scenarios of the diploid Macaronesian and Iberian red *Festuca* species. The highest global likelihood of the DEC model was the criterion for selection. The *Sequential* model was chosen because (i) it was not rejected by discordance or distance-based analyses in the optimal scenario selected in the previous section (see Fig. 4 and Table 5); (ii) it summarized better than the other models the intermediate status of *F. jubata* between *F. agustinii* and the Azorean *F. francoi* and *F. petraea*, as suggested by previous AFLP results (Díaz-Pérez *et al*. 2008); and (iii) it is compatible with a completely resolved species tree, as required by DEC models. Two contrasting demographic scenarios for the *Sequential* model were evaluated: T^0 and T^3 (Fig. 5), in order to account for uncertainty of IM estimation of divergence times.

Historical distributions of species lineages were evaluated through the likelihood method implemented in LAGRANGE version no. 2.0.1 (Ree & Smith 2008). Lineage geohistory was modelled both along phylogenetic internodes (anagenetic change) and at lineage branching events (cladogenetic change) (Ree *et al*. 2005). Following Ree & Smith (2008), the range of a species evolved by dispersal between areas and by local extinction within areas in the absence of lineage divergence, whereas lineage divergence from a widespread ancestor

Table 5 *P*-values for species-history models under demographic scenarios of Fig. 4. Four basic models were tested: *Canarian* (C), *Dichotomous* (D), *Radiation* (R) and *Sequential* (S). Bayesian group and species designations are shown in Table S1 (Supporting information)

	p-distance				Slatkin and Maddison's *S*			
Scenario	*C*	*D*	*R*	*S*	*C*	*D*	*R*	*S*
Bayesian group level								
PpN^0T^0	0.51	0.34	0.28	0.64	0.50	0.70	0.42	0.64
PpN^2T^0	0.41	0.39	0.35	0.44	0.999*	1.00*	0.997*	0.998*
PpN^1T^2	0.48	0.18	0.12	0.51	0.46	0.48	0.17	0.68
PpN^1T^3	0.57	0.10	0.05*	0.58	0.43	0.41	0.08	0.68
$PpN^{-1}T^1$	0.53	0.18	0.10	0.43	0.32	0.20	0.09	0.25
$PpN^{-1}T^2$	**0.50**	**0.04***	**0.018***	**0.64**	0.23	0.07	0.011*	0.33
$PpN^{-1}T^3$	0.63	0.03*	0.004*	0.39	**0.30**	**0.10**	**0.004***	**0.39**
$PpN^{-2}T^1$	0.43	0.08	0.022*	0.31	0.20	0.06	0.03*	0.09
Species level								
SpN^0T^0-m6	0.34	0.25	0.23	0.41	0.54	0.49	0.47	0.56
$SpN^{-1}T^3$-m6	**0.28**	**0.014***	**0.002***	**0.5**	0.008*	0.014*	0.002*	0.012*
SpN^0T^0-m3	0.31	0.26	0.21	0.31	0.72	0.69	0.66	0.75
$SpN^{-1}T^3$-m3	0.008*	0.014*	0.002*	0.012*	**0.19**	**0.67**	**0.001***	**0.69**

P-values indicate the probability of a simulated *p-distance* being lower than the empirical *p-distance* or the probability of a simulated Slatkin & Maddison (1989)'s *S* being equal or higher than the empirical *S* estimate for each of the four species-history models. An asterisk denotes a rejected model whose expected one-tailed 95% distribution did not contain the empirical value. Bold numbers in a line indicate the optimal scenario according to the jumping/selection scheme based on the three-criterion condition (Fig. 5).

occurred either within a single area or between a single area and the rest of the range.

Four biogeographic areas were assumed in the analysis (Table 6). Area model parameters were adjusted for 17 alternative area-models (Table 6) assuming that the ancestral lineage of the Macaronesian red fescues was spread (fixed) in the biogeographic area O (Continent): (i) the free-independent-colonization (FIC) model, where dispersal rates were left to vary freely among biogeographic areas; (ii) the 3IC model, where independent dispersal events were constrained from O to C (Canary Islands), M (Madeira) or A (Azores); and (iii) 15 additional constrained area-models that included single or two independent colonizations. Fixing O at the root was considered a reasonable biogeographic scenario, considering that other alternatives would imply extinction of the ancestral lineage in the continental area O after an initial colonization of C, M or A and subsequent retro-colonization of O. Likelihoods of alternative range inheritance scenarios at internal nodes of the *Sequential* tree model were compared once the O biogeographic area was fixed at the root node.

Results

Polymorphism of trnLF haplotypes

Haplotype variation and frequency and molecular diversity indexes of the diploid red *Festuca trn*LF sequences are shown in Table 1, Tables S1 and S2 (Supporting information). *Festuca agustinii* was shown to be the most variable species among diploid red fescues, having twice the number of haplotypes (*nh*) and of segregating sites (*Ss*) of either of the Azorean species *F. francoi* or *F. petraea*, which were even more variable than *F. jubata* and *F. rivularis* (Table 1). A direct relationship was found among $\theta\pi$, *nh* and *Ss* within species. According to the average number of nucleotide differences (*d*; Table 2), haplotypes of *F. agustinii* were also the most divergent showing the highest values when compared to *F. jubata* and *F. francoi* haplotypes. Intermediate values were observed between *F. jubata* and Azorean haplotypes, whereas the lowest values were detected between *F. francoi*, *F. petraea* and *F. rivularis*. Shared mutation values highlighted the genetic proximity of the *F. francoi* and *F. petraea* Azorean haplotypes (*shm* = 3; Table 2), in contrast to other pairwise species' haplotype combinations that did not share mutations or only showed one (*F. agustinii* and *F. jubata*). Indels increased the number of haplotypes based on nucleotide polymorphism from 15 to 17 (Table S2, Supporting information). Specifically, haplotype n1 (without considering gaps) was subdivided into haplotypes h1, h3 and h9 (considering gaps). This increased the molecular diversity of *F. petraea*, and to a lesser extent that of *F. rivularis* (Table 1), although it did not suggest a different general pattern of molecular diversity from that observed without including indels.

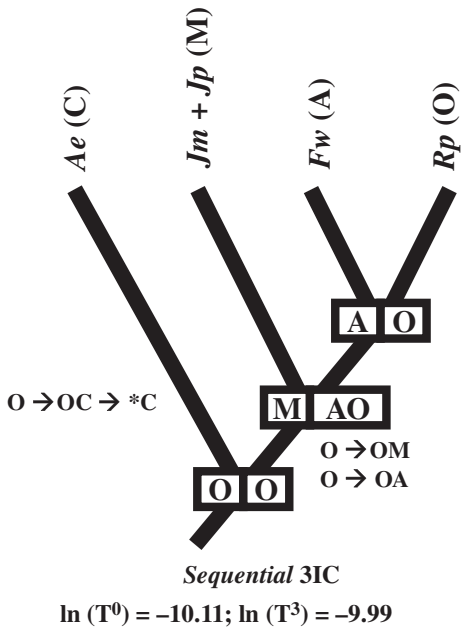


Fig. 5 Dispersal-extinction-cladogenesis (DEC) model with the highest global likelihood (−Ln). The *Sequential* species-history scenario was assumed as explained in the text. Codes for biogeographic areas are as follows: C (Canary Islands), M (Madeira), O (continental W Mediterranean) and A (Azores). Model parameters were adjusted assuming that the ancestral lineage of the Macaronesian red fescues was spread (fixed) in the biogeographic area O. The highest likelihood reconstruction of ancestral ranges for each node is represented by two contiguous boxes, each denoting ranges inherited by daughter lineages. Putative dispersal and extinction events (denoted by the symbol*) are given along branches. T^0 and T^3 indicate nodal ages following Fig. 3. All computations were performed using the program LAGRANGE version no. 2.0.1 (Ree & Smith 2008).

Table 6 Global likelihood (-lnL) of the 17 dispersal-extinction-cladogenesis (DEC) area-models imposed on the *Sequential* species-history tree. T^0 and T^3 represent divergence times as shown in the Fig. 4

	Sequential	
DEC model	T^0	T^3
3IC	10.11*	9.99*
OC, OAM	10.50	10.18
OM, OAC	10.59	10.41
OA, OMC	10.90	10.64
OA, OCM	11.20	11.22
OC, OMA	11.12	10.54
OM, OCA	12.00	11.84
OA(CM)	11.26	10.98
FIC	11.82	11.55
OAMC	11.61	11.34
OACM	11.73	11.66
OC(MA)	12.36	12.19
OM(CA)	12.52	11.83
OMAC	12.67	12.15
OCAM	12.21	11.84
OCMA	12.62	11.95
OMCA	13.34	12.77

O, C, A and M indicate the continental W Mediterranean, Canarian, Azorean and Madeiran archipelago biogeographic areas, respectively. 3IC is the three-independent-colonization model; FIC is the free-independent-colonization model. The constrained models were coded by a string of letters which is read in a left-to-right sense as successive dispersal steps. The leftmost letter (O) denoted the first step of the colonization route (fixed root). Letters within brackets indicate bi- or trifurcation dispersals from the immediate left-side letter. A comma separated two independent colonizations. Asterisk indicates the most likely model.

Evolutionary relationships of trnLF haplotypes

The evolutionary relationships of the diploid Macaronesian and Iberian red *Festuca* species and other close taxa of the fine-leaved Loliinae are shown in Fig. S1 (Supporting information). The maximum parsimony, maximum-likelihood and Bayesian inference analyses recovered a well-to-strongly-supported *Festuca* sect. *Aulaxyper* + *Vulpia* 2× clade. Nevertheless, the large sample size of diploid red *Festuca* taxa used in this study did not resolve the basal relationships within the *Festuca* sect. *Aulaxyper* crown, probably due to the retention of ancestral polymorphism and lack of resolution of classical phylogenetic methods as a consequence of recent speciation times.

Statistical parsimony was designed for scenarios where low genetic differences among sequences are expected (Templeton *et al.* 1992). This method has shown greater statistical power and more accurate organismal relationships than maximum parsimony when few characters are available, even at higher levels of divergence (Crandall 1994). In this sense, we relied upon the statistical parsimony network to characterize genealogical relationships of diploid red fescues haplotypes (Fig. 2). Overall, the network suggested three-two-step clades (Fig. 2a): a *Canarian* clade (2-2) that was exclusively formed by *F. agustinii* samples and the *Azorean–Iberian* (2-1) and the *Madeiran–Iberian* (2-3) clades showing haplotype connections between island and continental populations. The *Azorean–Iberian* clade was dominated by the large core of haplotype h1 from which seven terminal haplotypes (h2 to h8) radiated. This level of radiation was not evident in the other clades.

Phylogeographic signal

Nested clade phylogeographic analysis results are shown in Table S3 (Supporting information). In total, eight of the nine clades yielded significant associations

with geography. The null hypothesis of no geographic association was rejected at all higher-level categories, including the 2-step clades and the entire cladogram. Overall, the NCPA suggested long-distance colonizations (LDC) and/or past fragmentation plus extinction of intermediate populations (PF/ext) for Macaronesian red fescues. The prevalent pattern within the *Azorean–Iberian* clade was restricted gene flow by IBD. However, LDCs were also inferred within the 2-step *Azorean–Iberian* clade. Two inferences of allopatric fragmentation were performed for the *Madeiran–Iberian* (2-3) and the *Canarian* (2-2) clades, respectively. NCPA found a common pattern of IBD at all 1 step clades within the *Madeiran–Iberian* and the *Canarian* clades.

Statistical associations between the patristic distances of haplotypes and geographic locations of samples were established through the ParaFit host–parasite co-evolutionary analysis (Fig. 1, Table 3). Global F1 test detected a highly significant haplotype–location association (4.79, $P < 0.001$), suggesting a global stepping-stone pattern in the spatial distribution of the diploid red fescues. Specific association values and detailed inspection of specific haplotype–location combinations selected the regions that mostly contributed to this general pattern. A significant concentration of *Canarian* clade haplotypes in the Canary Islands and of *Azorean–Iberian* clade haplotypes in the Azores (Fig. 1, Table 3) was evidenced. In contrast, the *Madeiran–Iberian* clade haplotypes were broadly distributed in the Madeiran and Canarian archipelagos and in the Iberian Peninsula and did not show statistical association with geography.

Bayesian estimation of demographic parameters

The estimated absolute *trn*LF mutation rate μ of the *Festuca* sect. *Aulaxyper* taxa was 1.1684×10^{-9} substitution per site per year. This value was in agreement with synonymous substitution rates of six chloroplast genes for the maize/wheat comparison (1.1×10^{-9} to 1.6×10^{-9}; Wolfe *et al.* 1987).

IM HPP estimates of θ ($\mathbf{N_e}$), t (**t**) and m ($\mathbf{N_e}m$) are shown in Table 4. Molecular data contained sufficient information to generate posterior distributions for the θ parameter (data not shown) that could be circumscribed to the HPD bounds (Table 4). These bounds were very wide, hence precluding any statistical differentiation among HPP estimates, although some general trends could be extracted. First, species $\mathbf{N_e}$ values (average = 452.9) were larger than population $\mathbf{N_e}$ estimates (291.8), although this reduction was more pronounced for the *F. agustinii/Ae* and (*F. francoi* + *F. petraea*)/*Fw* comparisons. As these species (or pool of Azorean species) involved a large number of populations and sampled locations, we do not discard the possibility that some level of intraspecific genetic substructuring could have affected the estimation of $\mathbf{N_e}$ at the species level. Second, *F. agustinii* and *Ae* showed the largest θ estimate at their respective taxonomic level. *Ae* was the only population with a θ value larger than 1, representing a sufficiently large $\mathbf{N_e}$ for minimizing the impact of genetic drift (Crow & Kimura 1970).

Most estimates of m were not significantly different from zero as indicated by the LRT (Table 4). This test only detected one statistically positive migration rate (LRT = 1.18), starting from *F. jubata* (Madeira) to *F. agustinii* (Canary Islands). The HPP estimate was of 1.021×10^{-6} per gene per generation, with HPD bounds between 5.26×10^{-9} and 4.53×10^{-6} and a corresponding population migration rate ($\mathbf{N_e m}$) of 0.49 representing one migration event every two generations.

The posterior distributions of the number of generation since species/population divergence (**t**) are shown in Fig. S2a, b (Supporting information). For both analyses, HPD bounds were undefined (Table 4) making unreliable any statistical comparison of posterior distributions. In this sense, we limited the discussion of the results to the description of the general trends. First, species **t** HPP estimates circumscribe the more ancient Macaronesian red *Festuca* divergence to the last 609×10^3 generations or 1.22×10^6 years (*F. agustinii* vs. *F. petraea* + *F. francoi*). All comparisons gave relatively similar divergence times (range = 521×10^3 to 609×10^3 generations), except *F. rivularis* vs. *F. petraea* + *F. francoi* with a more recent splitting time of 309×10^3 generations or 618×10^3 years. This suggests that the *F. agustinii*, *F. jubata* and a putative *F. rivularis* + *Azorean* lineages could have simultaneously radiated approximately 558×10^3 generations ago or 1.12 Ma (averaged calculations excluding *F. rivularis* vs. *Azorean* species). On the other hand, population HPP estimates delimited the more ancient divergence time to the last 784×10^3 generations ago or 1.57 Ma (*Ae* vs. *Fw*). All comparisons involving *Ae* showed earlier splitting times (average = 732.7×10^3 generations ago or 1.47 Ma), suggesting that *F. agustinii* probably diverged first from the rest of the Macaronesian red fescues. When *Ae* was excluded from the analysis, more recent divergence times (average = 196×10^3 generations ago or 392×10^3 years ago) were obtained, with *F. jubata* (*Jm* + *Jp*) splitting at intermediate chronological times (average = 272×10^3 generations ago or 544×10^3 years ago) between the *Ae* average time and the very recent splitting event between *F. rivularis* (*Rp*) and the Azorean species (*Fw*; 44×10^3 generation ago or 88×10^3 years ago). Overall, IM analyses at both the population and the species levels indicated that the most ancient divergence time for the diploid red fescues could be circumscribed between 609×10^3 (1.22 Ma) and 784×10^3 (1.57 Ma) generations ago.

Hypotheses testing of species-history models

The hypotheses testing of species-history models was performed under the optimal scenarios that generated simulated distributions which clearly supported at least one species-history model over the others. This optimal scenario was chosen according to the three-criterion condition, which included (i) at least one nonrejected model and at least one rejected model; (ii) a lesser number of nonrejected models; and (iii) nonrejected model(s) where their observed probabilities were closest to 0.5. From IM HPP values, we constructed the initial PpN^0T^0 scenario, which was superseded by other scenarios according to the three-criterion condition, because none of the four species-tree models could be rejected according to the *p-distance* or the *S* distributions. Consequently, we moved at discrete steps from PpN^0T^0 (Fig. 4) progressively achieving the three-criterion condition as we simultaneously reduced $\mathbf{N_e}$ and increased T. The $PpN^{-1}T^2$ was chosen as the optimal scenario (Table 5) where the *p-distance* statistic supported the *Canarian* and *Sequential* models and rejected the *Dichotomous* and *Radiation* models. Additionally, the *S* statistic favoured the selection of the optimal $PpN^{-1}T^3$ combination, only rejecting the *Radiation* model, and suggested a poor fit of the *Dichotomous* model to the data (Table 5).

At the species level, we restricted the tests to four migration scenarios (Figs 3b and 4). First, two scenarios were computed from averaged values of $\mathbf{N_e}$ and **t** IM HPP estimates that included a migration rate of 1×10^{-6} (SpN^0T^0-m6) or 1×10^{-3} (SpN^0T^0-m3) events per gene per generation (Fig. 4). Second, based on previous findings at the Bayesian group level, which showed that the simultaneous increase in T values and reduction in $\mathbf{N_e}$ values generated *p-distance* and *S* distributions with better resolving power (Table 5), we restricted the exploration of the parameter space to only two new scenarios: $SpN^{-1}T^3$-m6 and $SpN^{-1}T^3$-m3 (Fig. 4). Similar to the results obtained at the population-level analyses, the starting combinations $SpN^{-0}T^0$-m6 and $SpN^{-0}T^0$-m3 produced distributions that were rejected by the three-criterion condition (Table 5). For *p-distance*, we achieved an optimal scenario when time was increased and the migration rate estimated by IM analyses was incorporated ($SpN^{-1}T^3$-m6), favouring again the *Sequential* and *Canarian* models (Fig. 4 and Table 5). Simulated *S* distributions showed less resolving power than the *p-distance* ones, rejecting only the *Radiation* model in T^3 scenarios (Table 5).

Overall, simulated *p-distance* and *S* distributions at both the population and species level supported the *Canarian* and *Sequential* species-history models when large divergence times and small effective population sizes were considered for tree parameter values ($PpN^{-1}T^2$, $PpN^{-1}T^3$, $SpN^{-1}T^3$-m6) (Fig. 4 and Table 5). When IM HPP estimates were used for simulations (N^0T^0), an inconclusive outcome arose as a consequence of lack of power of the test, although it does not preclude the support for the *Canarian* and *Sequential* models at low divergence times.

Biogeographic analysis

Global likelihood of the 17 DEC area-models imposed on the *Sequential* species-history tree model is shown in Table 6. For both T^0 and T^3, the most likely geographic distribution scenario corresponded to the 3IC model (Fig. 5), which suggested three independent colonization events from the continental area O to each of the Macaronesian archipelagos C, M and A. This model also suggested an extinction event (represented by an asterisk in Fig. 5) of the lineage that first dispersed to Canary Islands in O. The global maximum-likelihood dispersal rates of the 3IC model for T^0 and T^3 were, respectively, 0.87 and 0.47, representing approximately one dispersal event every 1×10^6 to 2×10^6 years. Ten and 13 solutions for T^0 and T^3 were, respectively, not statistically different from the most likely 3IC model, pointing to a failure of the DEC models to recover unique historical distributions of the recently diverged diploid Macaronesian and Iberian red fescues, probably due to the small number of species analysed and their recent divergences.

Discussion

Sampling effects on performance of IM estimates

The accuracy and precision of IM estimates, and hence the width of their HPD bounds, are expected to be affected by the number of loci and sequences used in the analysis (Strasburg & Riesenberg 2010). In our case, it could explain the wider or undefined HPD bounds, especially regarding the estimated splitting times and, to a lesser extent, the effective population sizes (Table 4; Fig. S2, Supporting information). Although we think that the use of a single locus might have been the main factor affecting our wide HPD intervals, we nevertheless adopted several grouping strategies to minimize violations of the IM model, that is, merging sequences from different biologic populations into Bayesian groups or species. In this way, we approached sample sizes similar to those used in previous studies (Hey *et al.* 2004; Hey 2005; Won & Hey 2005; Lee & Edwards 2008; Niemiller *et al.* 2008). Despite the merging of sequences from different island populations could have introduced some genetic substructure in those artificial groups, the IM model is relatively robust

to moderate violations by genetic structure (Strasburg & Riesenberg 2010). Consequently, working with larger groups with some level of substructure seems to be a better option than using small population sizes with almost no genetic variability. Apparently, there is not a specific threshold value of sampling size to generate good HPP estimates because in addition to the number of loci and number of sequences used, the genetic variability of the samples and the degree of haplotype sharing between the studied species should also be taken into account to obtain more precise IM estimates (Strasburg & Riesenberg 2010).

AFLP Bayesian groups vs. trnLF haplotype patterns

Some limitations that could preclude the resolution of the evolutionary processes that gave rise to the extant Macaronesian red fescues relate to the present restricted distribution of the close relative *F. rivularis* in SW Europe and the possible extinction of other close ancestral diploid lineages in this region and in NW Africa after the colonization of the oceanic islands. For example, excluding putative extinct continental population from analyses could have affected the possibility of circumscribing lineage divergent times to narrower time-windows or restricting the geographic range of colonization events to particular continental subareas. Extensive sampling of continental relatives is also of main interest for rigorously testing the monophyletic assemblage of populations within an archipelago, giving support to single-colonization biogeographic scenarios (Emerson 2002).

The chloroplast patterns first appeared to be apparently incongruent with the nuclear AFLP Bayesian groups [(sensu Díaz-Pérez *et al.* (2008)]. However, their respectively supported species-histories did not contradict each other; furthermore, the maternally inherited plastid data complemented the biparentally inherited AFLP models recovering two colonization waves of ancestors of *F. agustinii* in the Canary Islands that passed undetected in the AFLP data analysis. Incongruence among gene trees is an expected outcome of the coalescent process, especially for closely related species where short branch lengths generate a multitude of different gene topologies (Degnan & Salter 2005). The reconstruction of Bayesian groups from 189 nuclear AFLP markers reflected common ancestry, even if individual loci did not all reconstruct the same coalescent history (Díaz-Pérez *et al.* 2008). In this scenario, the nonrecombinant chloroplast chromosome is expected to show a single coalescent history, which might be partially congruent with the AFLP patterns detected in the diploid red *Festuca*. An apparent incongruence was reflected in the AFLP Bayesian groups that included haplotypes assigned to two 2-step plastid trnLF clades (Fig. 2a, Table S1, Supporting information), as shown for *Aw* and *Ae*. In this sense, the nonrecombinant nature of the cpDNA maternal lineage was useful to infer ancient independent colonization events in *F. agustinii*, whose traces were subsequently erased by sexual recombination at the nuclear AFLP loci. This was probably due to the admixture of the most recent colonizers with the previously established F. agustinii founders in the Gran Canaria island that, in spite of this, still leaved their plastid imprinting. It seems to be the case of the atypical clustering of the Gran Canaria h4 and h5 haplotypes within the *Azorean–Iberian* clade (Fig. 2a), obtained from individuals that were nonetheless clustered in the AFLP *Ae* Bayesian group (Table S1, Supporting information) without showing any signal of genetic differentiation from the other Tenerife populations. A different incongruence was observed in the *F. rivularis* Pyrenean population, with individuals showing haplotypes belonging to both the *Azorean–Iberian* (h2) and the *Madeiran–Iberian* (h9 and h12) clades. Consequently, we believe that the differences observed between the AFLP and *trn*LF patterns, or within the *trn*LF ones alone, could have been originated by incomplete lineage sorting mechanisms in one and the other genome, although more easily detected in the maternally inherited plastid genome, rather than artefacts produced by the AFLP Bayesian grouping criterion.

Additionally, the recent origin of the Azorean lineage, coupled with the low mutation rate of the *trn*LF region in *Festuca* sect. *Aulaxyper*, might explain why *F. petraea* and *F. francoi* did not show mutually exclusive haplotypes, despite they grow in different ecological niches and show an almost sister-clade monophyletic relationship according to multilocus AFLP markers (Díaz-Pérez *et al.* 2008). Interspecific hybridization might not be discarded, although AFLP data discouraged this scenario. It was also inferred that a very recent coastal halophytic *F. petraea* could have derived from relatively more ancestral mountain cliff-dweller *F. francoi* in the western island of Flores and that speciation could have occurred through ecological adaptation (Díaz-Pérez *et al.* 2008). This hypothesis is corroborated by the lack of differentiation of the *trn*LF haplotypes of the two species, which have been unable to accumulate specific mutations in such a short time span.

The present *trn*LF study has recovered a sister relationship between the continental *F. rivularis* and the Azorean *F. francoi* and *F. petraea* taxa (Fig. 2). This finding argues against the general hypothesis that continental diploid relatives should be older than (or as old as) the oceanic islands taxa or the eventual result of back-colonizations (Carine *et al.* 2004; Kim *et al.* 2008). Our study shows signatures that the supposedly 'relict' Iberian *F. rivularis* is of more recent origin than the

Canarian *F. agustinii* and the Madeiran *F. jubata*. Furthermore, it has been shown that the *F. rivularis* MRCA likely participated in the colonization of the Azorean archipelago in middle–late Pleistocene times (0.6–0.1 Ma). We could not rule out, however, the possibility that *F. rivularis* is a relict continental lineage from which oceanic lineages branched off over time. Nevertheless, it should be noted that this species showed the lowest level of genetic polymorphism among the diploid red fescues (see Ss, *nh* and $\theta\pi$ in Table 1), a fact that is incompatible with an ancient evolutionary lineages where the accumulation of a relative large level of variability is expected.

Phylogeographic signal

The NCPA suggested that the entire complex of haplotypic–geographic associations of Macaronesian red fescues is mainly explained by LDC (Table S3, Supporting information). This is in agreement with the general literature in that recently emerged biome-empty volcanic Macaronesian oceanic islands were filled through colonization events that were associated with long-distance dispersals (Whittaker & Fernández-Palacios 2007). Long-distance dispersals were detected within the Azorean–Iberian clade (2-1), subclade 1-1, involving haplotypes located in the Pyrenees, Azores and Gran Canaria (Fig. 1 and Table S3, Supporting information). A long-distance dispersal event is hypothesized to have resulted in the colonization of the central region of Azores, after which the western (1-3) and eastern (1-2 and 1-4) subarchipelagos were further colonized (Fig. 1). This pattern of colonization does not totally concur with AFLP inferences (Díaz-Pérez *et al.* 2008), where a directional rather than a centripetal (from central to western or eastern subarchipelagos) colonization pattern was inferred. The long-distance colonization of the Canary Islands is restricted to Gran Canaria and represents an independent lineage that probably colonized this archipelago in very recent times. To complete this scenario, it should be highlighted that there was a long-distance dispersion that was not detected by the NCPA, connecting the populations Frivu2 (Pyrenees) and Fjuba8 (Porto Santo, Madeira), placed at the 1-9 subclade within the *Madeiran–Iberian* 2-3 clade. Considering that Fjuba8 represented a small proportion of the total sample of clade 1-9, it is not difficult to assume that its contribution to the nested clade distance was not of main importance.

Two allopatric-fragmentation events were localized within the *Canarian* (2-2) and the *Madeiran–Iberian* (2-3) clades (Fig. 2a and Table S3, Supporting information). Each event involved at least two clades without overlapping distributions (Fig. 1), but without genetic discontinuity expressed by larger-than-average number of steps. It seems that recent fragmentation events took place in the Canarian and Madeiran archipelagos, but we may not rule out that this pattern could have been generated by the conjunction of lineages of independent origins by means of multiple colonization events. Multiple colonizations have probably occurred in Tenerife, an island with a complex geological history. This island was originally conformed by three isolated mountain systems, Anaga, Teno and Adeje, that originated 4–6 Ma and finally merged by volcanic activity approximately 1 Ma (Trusty *et al.* 2005 and references therein). The Fagus5 population (subclade 1-5; Teno) and Fagus3 (subclade 1-8; Anaga) (Fig. 1) could represent two isolated lineages that were brought together by posterior merging of both systems. The Anaga lineage could have experienced a posterior expansion to recent areas of Tenerife, as indicated by the geographic placement of subclade 1-8 populations Fagus4 and Fagus7. The other fragmentation event, detected in the Madeiran archipelago, possibly reflects recent oceanic barriers to gene flow between Pyrenean and Madeiran haplotypes.

Restricted gene flow by IBD was the main pattern found for all 1-step clades (Table S3, Supporting information). Some 1-step clades like 1-3 (Flores), 1-2 (Santa María) and 1-4 (Sao Miguel) in Azores involve populations located in the same island, implying limited dispersal abilities of red fescues on terrestrial landscapes (Fig. 1). In other clades with haplotypes located in at least two islands, NCPA suggested a restriction to gene flow within islands coupled with a stepping-stone colonization pattern (a variant of IBD), for example in the 1-6 (La Palma, Hierro and Tenerife), 1-7 (Madeira and La Palma), 1-8 (Tenerife and La Gomera) and the Azorean core of 1-1 (Terceira, Faial, Pico and Sao Jorge).

Some precautions should be observed when assuming NCPA inferences without taking into account alternative statistical methods supporting them. First, Panchal & Beaumont (2010) have recently shown that single-locus NCPA tends to give a high proportion of false inferences. Second, NCPA is not especially conceived for island populations located in different Macaronesian archipelagos that are isolated by large oceanic areas (Fig. 1), where the centre of the geographic distribution of some haplotypic clades was necessarily located at an intermediate highly improbable oceanic geographic location.

The ParaFit method

In this study, we have employed for the first time the ParaFit (Legendre 2001) host–parasite co-evolution model to detect phylogeographic signals in Macaronesian red fescues (Table 3, Fig. 1) as a complementary

statistical method to NCPA. ParaFit does not rely on nested clades to detect global and specific phylogeographic associations; instead, nonindependence of geographic and haplotypic distances was taken as evidence of a phylogeographic structure, as a consequence of IBD or a related variant like the SSM. The ParaFit method detected a highly significant global association for the entire complex of Macaronesian red fescues, suggesting IBD or SSM in addition to the long-distance colonization events inferred by NCPA. A significant concentration of regional haplotypes within the *Azorean–Iberian* clade is in agreement with the IBD outcome generated by NCPA, but also suggests that in addition to the allopatric-fragmentation events suggested by NCPA for the *Canarian* clade, there are also indications of IBD or a SSM. Contrary to NCPA, ParaFit was not able to detect any association by IBD within the geographically broadly distributed *Madeiran–Iberian* clade, given that some haplotypes were located simultaneously in different archipelagos (h10) or continental-island areas (h9) (Table 3 and Fig. 1).

The existence of a SSM scenario should not be interpreted as contradictory evidence to IM analyses, which basically did not detect any sign of conclusive migration events between Bayesian groups or species. A linear SSM pattern without further migration after arrival could have been originated by sequential founder events along a chain of geographically separated islands or archipelagos. According to this, they represent punctual historical events that IM would interpret as historical divergence and not as migrations as the splitting times clearly favoured the strict isolation model over the isolation-with-migration model.

Coalescence-based species-history hypotheses testing

For coalescent simulations, it was assumed that ancestral polymorphism, incomplete lineage sorting and migration after divergence were the main factors that shaped the evolution of the diploid red fescues, neglecting other sources such as hybridization and selection. Large oceanic distances, especially among archipelagos and continental areas, are strong barriers to the free interchange of genes, a fact that is in concordance with IM analyses, which only suggested an inconclusive localized gene flow from Madeira to the Canary Islands. Under this scenario, four species-history models were represented by species trees that conditioned different branching patterns of simulated gene trees (Fig. 3a). For diploid red fescues, the *Sequential* and *Canarian* models were consistently supported by empirical observations for different values of $\mathbf{N_e}$ and splitting time (T) values (Table 5). By contrast, rejection of the *Radiation* and *Dichotomous* models was achieved at lower $\mathbf{N_e}$ and T values proximal to the basal radiation of the *Festuca* sect. *Aulaxyper* clade (Inda *et al.* 2008). These results concurred with the expectations because a higher proportion of fixed exclusive mutations among lineages increase the power of the test statistics. The rejection of the *Dichotomous* model was only achieved by the *p-distance* statistics (Table 5), suggesting that this method is sensible enough to detect subtle genetic divergence under the relatively recent divergence times of the Macaronesian red fescues.

The *S* statistics showed a consistent failure to reject the *Dichotomous* model (Table 5). We do not discard that *S* could have not fully considered information about deeper population divergences (Knowles & Maddison 2002), but it is highly probable that in the reconstruction step of the gene trees through classical phylogenetic methods (Felsenstein 2004) the inferred topologies of *trn*LF trees could have reflected higher levels of statistical uncertainty. In fact, these methods have shown little statistical power to discriminate among alternative hypotheses when a scant level of polymorphism is present (Huelsenbeck *et al.* 2001) and to recover incorrect phylogenies when a small number of nucleotides or a large number of closely related sequences are used (Nei *et al.* 1998). By contrast, the *p-distance*-based statistic that we employed here does not rely on previous gene tree topology estimation and, at the same time, it takes into account species-tree branch length information, assuming a direct relationship between pairwise distances and divergence times of genes (Slatkin & Hudson 1991).

The history-species models proposed here (Fig. 3a) were oversimplified to capture the essence of past species divergence events. Hypotheses should be simple enough for them to be discriminated with available data, yet still capture the essence of the problem (Knowles 2004). Some criteria were employed to achieve this condition. First, both Azorean species were considered as one genetic unit to give relevance to evolution among different archipelagos. Second, we used averages of IM pairwise $\mathbf{N_e}$ and $\mathbf{t}$ estimates at different taxonomic levels (Fig. 3b), reducing in this way the number of alternatives that would need to have been tested (Fig. 4). Third, an IM population-level analysis was performed to minimize the putative effect of substructuring possibly found at the intraspecific level, approximating the theoretical assumptions of the IM model (Nielsen & Wakeley 2001; Hey & Nielsen 2004).

p-distance and *S* statistics coalescent distributions support the *Canarian* and *Sequential* species-history models (Fig. 3a and Table 5) that place *F. jubata* closely related to the Azorean + *F. rivularis* clade and establish a clear separation of *F. agustinii* from the rest of the group. These models have also been supported using Mantel

correlation average values between patristic distances obtained from the empirical statistical parsimony network and a random sample of patristic distances generated from coalescent-based networks (unpublished results). Even more, the *Sequential* model is in agreement with previous multilocus AFLP data (Díaz-Pérez *et al.* 2008) which placed *F. jubata* at intermediate placement in relation to *F. agustinii* and the Azorean species.

Dated historical and biogeographic scenarios of the diploid Macaronesian red fescues

Methods implemented to estimate the age of tree divergences from heterogeneous sequences, such as Bayesian relaxed-clock approaches, rely on fossil calibrations for minimum and maximum thresholds of basal nodes (Inoue *et al.* 2010; and references therein). The impact that priors on times, based on primary fossil calibrations, might inflict on the calculation of posterior time estimates has been stressed in recent studies (Inoue *et al.* 2010; Pyron 2010). The coalescence-based approach used in this study offers an alternative to estimate the ages of convergence of recently evolved groups that are derived from priors on mutation rates and effective population sizes. Using this approach, the HPP estimations of the times of divergence of the recently evolved Macaronesian red fescues indicate a much recent origin for all the oceanic island lineages and their continental relative *F. rivularis* (Table 4, Fig. S2a, b Supporting information) than the estimations derived from dated chronograms (Inda *et al.* 2008). Despite mutation rates of *Festuca* sect. *Aulaxyper trn*LF sequences have been derived from dated phylogenies (Inda *et al.* 2008) and that the estimated coalescent ages might be influenced by the short minimum generation times attributed to these perennial herbaceous grasses (2 years, Gaut *et al.* 1992; Catalán 2006), our results provide a feasible evolutionary scenario for rapid colonization events of these grass lineages in both oceanic islands and continental environments.

Evolution of diploid red fescues seems to be constrained to the last 1.57 Ma (population level) or 1.22 Ma (species level) as suggested by HPP estimates of divergence times obtained through IM analysis (Table 4). This represents a 49% to 63% reduction of the estimated value for the basal node of the *Festuca* sect. *Aulaxyper* radiation as previously dated by Inda *et al.* (2008) using Bayesian relaxed-clock methods (2.5 ± 0.9 Ma). HPP dates circumscribe the evolution of the diploid red fescues within the upper–middle Pleistocene (Futuyma 1998), which was characterized by repeated 100-Ky cycles of warm (interglacial) and cold (glacial) periods (Mudelsee & Shulz 1997), when the sea level dropped because much of the earth's water was frozen on land. These events could have promoted multiple colonizations of red *Festuca* lineages into the Macaronesian archipelagos within confined small-scale windows of opportunity (sensu Carine 2005; Kim *et al.* 2008). This fact seems to be supported by DEC models, where multiple independent colonization events were more likely than strictly stepping-stone colonization routes (Table 6 and Fig. 5). This pattern of colonization apparently was not influenced by the geological ages of the subaerial sections of the current archipelagos, which have been dated back between 14.5 and 2.0 Ma for Canary Islands, 14.0–5.3 Ma for Madeiran archipelago and 4.01–0.3 Ma for Azores (Díaz-Pérez *et al.* 2008), all of them older than the basal radiation of the *Festuca* sect. *Aulaxyper* ancestor that gave rise to diploid red fescues approximately 1.5 Ma. In this context, the 3IC model suggested by Díaz-Pérez *et al.* (2008) from AFLP data was the most likely supported biogeographic model, further suggesting that the correlations of AFLP-standardized residuals analyses can be a useful method for testing independent migration routes of the Macaronesian red fescues. It should be highlighted that we were not able to compute IM HPD bounds for divergence times, imposing a note of precaution to our conclusions about chronological events. The fact that Mesquite and SIMCOAL simulations suggested that a reasonable basal divergence time is proximal to 1.5×10^6 generations ago (or 3 Ma; Figs 3b and 4), indicate that we might have underestimated this parameter.

The 3IC model was enforced to contain three independent dispersal events from the continental area (O) to Canary Islands (C), Madeira (M) and Azores (A), respectively (Fig. 5). Interpreting its range inheritance scenarios following Ree *et al.* (2005) allowed us to infer that the *F. agustinii* lineage possibly was produced by an initial cladogenetic event in O. This lineage dispersed and colonized C followed by its extinction in O. Two independent colonizations originated from the continental sister lineage of *F. agustinii* without immediate speciation after arrival, generating a lineage of widespread distribution located in O, M and A. Subsequent allopatric speciation of M and A lineages is invoked to explain the origin of *F. jubata* and Azorean species, respectively.

On the other hand, given that dispersal-mediated allopatry (DMA; sensu Clark *et al.* 2008) has been invoked for several oceanic island plant species, it is probable that the divergence of Macaronesian red *Festuca* lineages occurred in association with the arrival of founders from mainland continental areas. At least, it seems probable for this to have occurred in Madeira, given a shift of geographic range at the base of the lineage leading to that archipelago. Nevertheless, DMA does not rule

out the possibility that further anagenetic evolution could have taken place after the arrivals as a consequence of founding effects and genetic drift (Whittaker & Fernández-Palacios 2007). According to Stuessy *et al.* (2006), the anagenesis model fits into scenarios where a genus is represented on an island or archipelago by a single endemic species and this situation matches that observed in the Canary Islands and Madeira.

Likelihood models as implemented in LAGRANGE version no. 2.0.1 (Ree & Smith 2008) cannot handle polytomic species trees. In this sense, we were not able to infer other likely biogeographic scenarios assuming the *Canarian* model species tree. Nevertheless, we explored approximate outcomes under this model artificially resolving the polytomy leading to *F. jubata*, *F. rivularis* and *F. francoi* + *F. petraea* (data not shown). In this case, we arbitrarily preserved the same branching pattern of the *Sequential* model, but fixing internal branch lengths closer but different from zero. T^0 parameters also supported the 3IC model, but T^3 parameters suggested a complex scenario involving other potential colonization routes. In this case, similar to the 3IC scenario, two independent colonizations were inferred from the continent leading to *F. agustinii* in Canary Islands and to the *F. francoi* + *F. petraea* complex in Azores, respectively. However, Madeira was colonized following stepping-stone events, comprising a dispersion from the continent to Azores and from there to Madeira where *F. jubata* speciated.

Species-history scenario of the Macaronesian red fescues

Two species-history models, which placed *F. agustinii* as the first branching lineage of the species tree, were supported by the *S* and *p-distance* statistics (Fig. 3a and Table 5). Its sister lineage involved a recent polytomy, following the *Canarian* model, or the sequential branching of lineages leading to *F. jubata* and finally to the sister clades *F. rivularis* and the Azorean species, following the *Sequential* model.

Additional genetic parameters seem to support *F. agustinii* as the first branching lineage. *Festuca agustinii* was the most variable species, having twice the number of haplotypes of the next variable species and showing the highest $\theta\pi$ and *Ss* values (Table 1), features compatible with a relatively ancient lineage that have accumulated more mutations since arrival than the rest of the species lineages. In addition, population substructure among islands and repeated colonizations from external sources might also explain high levels of genetic variation. The statistical parsimony network showed a *Canarian* clade that was clearly differentiated from the other clades (especially if indels were excluded) (Fig. 2a). Additionally, this clade showed an incipient nonstar subnetwork structuring (Fig. 2a, b) and a strong phylogeographic signal to the Canary Islands (Fig. 1, Table 3). Both elements are indicative of an *in situ* lineage divergence. This structured phylogeographic pattern of *F. agustinii* is incompatible with a scenario of a population that has recently expanded in size from a small number of founders (Slatkin & Hudson 1991; Avise 2001). The older origin of the Canarian species with respect to the more recent ones of other Macaronesian island congeners was also evidenced in the study of Kim *et al.* (2008), which showed the earlier chronological divergences of the Canarian lineages in five eudicot Macaronesian groups (woody *Sonchus* alliance, *Echium*, *Aeonium* alliance, *Sideritis* and *Crambe*).

According to the *Sequential* model of species history (Fig. 3a), graphical inspection of sampling location (Fig. 1) suggests two putative secondary colonizations of the Canary Islands from O + A (h4 and h5 haplotypes sampled in Gran Canaria) and M (h10 haplotype sampled in La Palma) biogeographic areas, respectively. Alternatively, those rare Canarian haplotypes could have originated from more ancient Canarian native haplotypes through convergent evolution. Both evolutionary scenarios suggest that the Canarian *F. agustinii* might have resulted from repeated colonization events such as those reported for other Canarian plants (Silvertown 2004; Carine 2005). Multiple colonizations to the Canary Islands have been explained by the relatively short geographic distance that separates them from northwestern Africa (approximately 70 Km), coupled with the direct influence of oceanic currents (Canary streams) and northeasterly trade winds, a likely scenario that could have fostered the dissemination of individuals from the continent (Fuertes-Aguilar *et al.* 2002; Carine *et al.* 2004). According to the window-of-opportunity hypothesis (Carine 2005), such colonizations probably occurred in repeated small periods of time with favourable conditions to colonizations. Therefore, an older event of c. 1.12 Ma (species level) or 1.47 Ma (population level) probably led to the main Canarian haplotypic lineage, whereas the rare variants might have been more recently acquired through different maternal introgressions (as discussed previously).

The *Sequential* model (Fig. 3a) suggests that the radiating structure of the *Azorean–Iberian* clade (Fig. 2a) could be explained from a relatively recent h1-like haplotype ancestor that was simultaneously inherited by *F. rivularis*, *F. francoi* and *F. petraea*. According to IM analyses, the 2-1 clade is of relative recent origin that probably originated in the last 100 000 (population-level) or 600 000 (species-level) years. Because the hypothetical reconstruction of the last 5 Ma of Macaro-

nesian palaeo-islands distributions does not suggest putative land bridges between Azores and the western coast of Europe (ca. 1300 Km) (Fernández-Palacios *et al.* 2011), adaptation to long-distance dispersal was a necessary condition for the colonization of this archipelago. According to this hypothesis, *F. francoi* probably reached Azores through bird zoochory (described as *F. jubata* in Schäfer 2003). Despite Azores is geographically separated from current migration routes along the cost of main continental massifs (Schäffer 2003), arrival of colonizers could have been favoured by palaeo-winds blowing in different direction than current westerly prevailing winds. The current restricted distribution of *F. rivularis* in the Iberian Peninsula and the Pyrenees and its close relationship to the Azorean taxa support the hypothesis of a recent single-colonization event of their MRCA to the Azores islands from SW Europe, followed by *in situ* speciation through ecological adaptation (cf. Díaz-Pérez *et al.* 2008). More recent mutations in some h1 lineages could be hypothesized to have occurred independently in these species originating the h2 haplotype in *F. rivularis* and the strongly phylogeographically associated h3, h6, h7 and h8 haplotypes in the Azorean taxa (Figs 1 and 2, Table 3).

The lack of other extant continental diploid red fescues precluded the identification of the potential ancestors of *F. jubata* and *F. agustinii*. However, the more ancestral and genetically differentiated divergence of the Canarian *F. agustinii* h-13/*n*-11 haplotype supports an independent ancient colonization event from a presumably continental ancestor (cf. Díaz-Pérez *et al.* 2008). In contrast, the more recent divergence of the *Madeiran–Iberian* 2-3 clade, its genetic similarity to the *Azorean–Iberian* 2-1 clade and its shared intermediate h9 haplotype with *F. rivularis* support a recent independent origin of *F. jubata* from a presumably SW European ancestor (cf. Díaz-Pérez *et al.* 2008).

The *Sequential* species tree and the molecular diversity indices of the *trn*LF region are compatible with previous results based on AFLP markers (Díaz-Pérez *et al.* 2008). Excluding *F. rivularis*, which was not included in that study, the unrooted AFLP Bayesian tree and the AFLP Bayesian structure analyses placed *F. agustinii* as the most divergent diploid red fescue. Assuming that the degree of population and close species differentiation based on AFLP markers are proportional to lineage ancestry (cf. Kropf *et al.* 2009; but see Crisp & Cook 2005), the AFLP-based Bayesian tree would have recovered a divergent scenario compatible with that of the *Sequential* model, in which lineages leading to *F. agustinii*, *F. jubata* and *F. francoi* + *F. petraea* diverged in chronological order.

Acknowledgements

We thank Isabel Sanmartin, Mark Carine and five anonymous reviewers for their valuable comments to an earlier version of the manuscript, and Juli Caujapé for providing us with *F. agustinii* samples from Gran Canaria. This work has been supported by two Spanish Ministry of Science and Innovation (CGL2006-00319/BOS and CGL2009-12955-C02-01) and one Portuguese Ministry of Science and Technology (POCTI/BME/39640/2001) research grant projects. A.J. Díaz-Pérez was supported by a Universidad Central de Venezuela CDCH PhD fellowship. *Part of this work was carried out by using the resources of the* BIFI Research Institute (University of Zaragoza, Spain) and by the *Computational Biology Service Unit from Cornell University which is partially funded by Microsoft Corporation.*

References

Avise J (2001) *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Bailey NW, Gwynne DT, Ritchie MG (2005) Are solitary and gregarious Mormon crickets (*Anabrus simplex*, Orthoptera, Tettigoniidae) genetically distinct? *Heredity*, **95**, 166–173.

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

Carine M, Schäffer H (2010) The Azores diversity enigma: why are there so few Azorean endemic flowering plants and why are they so widespread? *Journal of Biogeography*, **37**, 77–89.

Carine MA, Russell S, 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.

Carstens BC, Richards CL (2007) Integrating coalescence and ecological niche modelling in comparative phylogeography. *Evolution*, **61**, 1439–1454.

Carstens BC, Degenhardt JD, Stevenson L, Sullivan J (2005) Accounting for coalescent stochasticity in testing phylogeographical hypotheses: modelling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus*. *Molecular Ecology*, **14**, 255–265.

Catalán P (2006) Phylogeny and evolution of *Festuca* L. and related genera of subtribe Loliinae (Poeae, Poaceae). In: *Plant Genome: Biodiversity and Evolution* (eds Sharma A and Sharma A), pp. 255–303. Science Publishers, Enfield, NH.

Catalán P, Torrecilla P, López J, Müller J (2006) Molecular evolutionary rates shed new light on the relationships of *Festuca*, *Lolium*, *Vulpia* and related grasses (Loliinae, Pooideae, Poaceae). In: *Current Taxonomic Research on the British & European Flora* (eds Bailey J and Ellis R), pp. 45–70. BSBI, London.

Caujapé-Castells J, Tye A, Crawford DJ *et al.* (2010) Conservation of oceanic island floras: present and future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **12**, 107–129.

Clark J, Ree R, Alfaro M, King M, Wagner W, Roalson E (2008) A comparative study in ancestral range reconstruction

methods: retracing the uncertain histories of insular lineages. *Systematic Biology*, **57**, 693–707.
Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
Comes HP (2004) The Mediterranean region – a hotspot for plant biogeographic research. *New Phytologist*, **164**, 11–14.
Crandall K (1994) Intraspecific Cladogram estimation: accuracy at higher levels of divergence. *Systematic Biology.*, **43**, 222–235.
Crisp M, Cook L (2005) Do early branching lineages signify ancestral traits? *Trends in Ecology and Evolution*, **20**, 122–128.
Crow J, Kimura M (1970) *An Introduction to Population Genetics Theory*. Harper & Row, Publishers, New York.
Degnan J, Salter L (2005) Gene tree distributions under the coalescent process. *Evolution*, **59**, 24–37.
Díaz-Pérez A, Sequeira M, Santos-Guerra A, Catalán P (2008) Multiple colonizations, *in situ* speciation, and volcanic-associated stepping-stone dispersals shaped the phylogeography of the Macaronesian red fescues (Festuca L., Gramineae). *Systematic Biology*, **57**, 732–749.
Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, **11**, 951–966.
Emerson BC (2003) Genes, geology and biodiversity: faunal and floral diversity on the island of Gran Canaria. *Animal Biodiversity and Conservation*, **26**, 9–20.
Excoffier L, Novembre J, Schneider S (2000) SIMCOAL: a general coalescent program for the simulation of molecular data in interconnected populations with arbitrary demography. *Journal of Heredity*, **91**, 506–510.
Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.
Fan H, Kubatko LS (2011) Estimating species trees using approximate Bayesian computatio n. *Molecular Phylogenetics and Evolution*, **59**, 354–363.
Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.
Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates Inc., Sunderland, Massachusetts.
Fernández-Palacios JM, Nascimento L, Otto R *et al.* (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, Fuertes-Aguilar J, Kim SC, Santos-Guerra A, Crawford DJ, Jansen RK (2002) Phylogeny of the Macaronesian endemic *Crambe* section *Dendrocrambe* (Brassicaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *American Journal of Botany*, **89**, 1984–1990.
Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhicking and background selection. *Genetics*, **147**, 915–925.
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 Biology*, **27**, 74–83.
Futuyma D (1998) *Evolutionary Biology*. Sinauer, Sunderland, Massachusetts.
Gaut B, Muse S, Clark W, Clegg M (1992) Relative rates of nucleotide substitutions at the *rbc*L locus of monocotyledonous plants. *Journal of Molecular Evolution*, **35**, 292–303.
Hey J (2005) On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology*, **3**, e193.
Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.
Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.
Hey J, Won YJ, Asundar AS, Nielsen R, Markert J (2004) Using nuclear haplotypes with microsatellites to study gene flow between recently separated Cichlid species. *Molecular Ecology*, **13**, 909–919.
Hickerson MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.
Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.
Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, **294**, 2310–2314.
Inda LA, Segarra-Moragues JG, Müller J, Peterson PM, Catalán P (2008) Dated historical biogeography of the temperate Loliinae (Poaceae, Pooideae) grasses in the northern and southern hemispheres. *Molecular Phylogenetics and Evolution*, **46**, 932–957.
Inoue J, Donoghue PCJ, Yang Z (2010) The impact of the representation of fossil calibrations on Bayesian estimations of species divergence times. *Systematic Biology*, **59**, 74–89.
Kier G, Kreft H, Lee TM *et al.* (2009) A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Science USA*, **106**, 9322–9327.
Kim S, Crawford D, Francisco-Ortega J, Santos-Guerra 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 Science USA*, **93**, 7743–7748.
Kim SC, McGowen M, Lubinsky P, Barber J, Mort M, Santos-Guerra A (2008) Timing and tempo of early and successive adaptive radiations in Macaronesia. *PLoS One*, **3**, e2139.
Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.
Knowles LL (2001) Did the Pleistocene glaciation promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, **10**, 691–701.
Knowles LL (2004) The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, **17**, 1–10.
Knowles LL, Maddison W (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623–2635.

Knowles LL, Carstens BC, Keat M (2007) Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, **17**, 1–7.

Kreft H, Jetz W, Mutke J, Kier G, Barthlott W (2008) Global diversity of island floras from a macroecological perspective. *Ecology Letters*, **11**, 116–127.

Kropf M, Comes HP, Kadereit JW (2009) An AFLP clock for the absolute dating of shallow-time evolutionary history based on the intraspecific divergence of southwestern European alpine plant species. *Molecular Ecology*, **18**, 697–708.

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

Lee JY, Edwards SV (2008) Divergence across Australia's carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution*, **62**, 3117–3134.

Legendre P (2001) Test of host-parasite coevolution: program ParaFit user's guide. Département de sciences biologiques, Université de Montréal. 10 pp.

Legendre P, Desdevises Y, Bazin E (2002) A statistical test for host-parasite coevolution. *Systematic Biology*, **51**, 217–234.

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

Luo R, Hipp A, Larget B (2007) A Bayesian model of AFLP marker evolution and phylogenetic inference. *Statistical Applications in Genetics and Molecular Biology*, **6**, article 11. Available from http://www.bepress.com/sagmb/vol6/iss1/art11.

Maddison W, Knowles L (2006) Inferring phylogenies despite incomplete lineage sorting. *Systematic Biology*, **55**, 21–30.

Maddison W, Maddison D (2008) Mesquite: a modular system for evolutionary analysis. Version 2.5. Available from http://mesquiteproject.org.

Matsumura S, Yokoyama J, Fukuda T, Maki M (2009) Intraspecific differentiation of *Limonium wrightii* (Plumbaginaceae) on northwestern Pacific Islands: rate heterogeneity in nuclear rDNA and its distance-independent geographic structure. *Molecular Phylogenetics and Evolution*, **53**, 1032–1036.

Meimberg H, Abele T, Bräuchler C, McKay J, Pérez dePazP, Heubl G (2006) Molecular evidence for adaptive radiation of *Micromeria* Benth. (Lamiaceae) on the Canary Islands as inferred from chloroplast and nuclear DNA sequences and ISSR fingerprint data. *Molecular Phylogenetics and Evolution*, **41**, 566–578.

Moore M, Francisco-Ortega J, Santos-Guerra A, Jansen R (2002) Chloroplast DNA evidence for the roles of island colonization and extinction in *Tolpis* (Asteraceae: Lactuceae). *American Journal of Botany*, **89**, 518–526.

Mudelsee M, Shulz M (1997) The mid-Pleistocene climate transition: onset of 100 ka cycle lags ice volume build-up by 280 ka. *Earth and Planetary Science Letters*, **151**, 117–123.

Nei M, Kumar S, Takahashi K (1998) The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotide or amino acids used is small. *Proceedings of the National Academy of Science USA*, **95**, 12390–12397.

Nicholson G, Smith A, Jónsson F, Gústafsson O, Stefánsson K, Donnelly P (2002) Assessing population differentiation and isolation from single-nucleotide polymorphism data. *Journal of the Royal Statistical Society: Series B-Statistical Methodology*, **64**, 695–715.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Niemiller ML, Fitzpatrick BM, Miller B (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Molecular Ecology*, **17**, 2258–2275.

Panchal M, Beaumont M (2010) Evaluating nested clade phylogeographic analysis under models of restricted gene flow. *Systematic Biology*, **59**, 415–432.

Panero J, Francisco-Ortega J, Jansen R, Santos-Guerra A (1999) Molecular evidence for multiple origins of woodiness and a New World biogeographic connection of the Macaronesian Island endemic *Pericallis* (Asteraceae: Senecioneae). *Proceedings of the National Academy of Science USA*, **96**, 13886–13891.

Posada D, Crandall K (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Pritchard J, Wen W (2004) Documentation for *structure* software. Version 2. Chicago. Available from http://pritch.bsd.uchicago.edu.

Pritchard J, Stephens M, Donelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pyron RA (2010) A likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. *Systematic Biology*, **59**, 185–194.

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

Ree R, Moore B, Webb C, Donoghue M (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.

Rosenberg N (2003) The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution*, **57**, 1465–1477.

Schäfer H (2003) Chorology and Diversity of the Azorean flora. *Dissertationes Botanicae*, **374**, 1–130.

Sequeira MM, Díaz-Pérez A, Santos-Guerra A, Viruel J, Catalán P (2009) Karyological analysis of the five native Macaronesian *Festuca* (Gramineae) grasses supports a distinct diploid origin of two schizoendemic groups. *Anales del Jardín Botánico de Madrid*, **66**, 55–63.

Silvertown J (2004) The ghost of competition past in the phylogeny of island endemic plants. *Journal of Ecology*, **92**, 168–173.

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

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603–613.

Steele CA, Storfer A (2006) Coalescent-based hypothesis testing supports multiple Pleistocene refugia in the Pacific Northwest for the Pacific giant salamander (*Dicamptodon tenebrosus*). *Molecular Ecology*, **15**, 2477–2487.

Strasburg JL, Riesenberg LH (2010) How robust are ''Isolation with Migration'' analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, **27**, 297–310.

Stuessy T, Jakubowsky G, Salguero-Gómez R *et al.* (2006) An alternative model for plant evolution in islands. *Journal of Biogeography*, **33**, 1259–1265.

Swofford D (2002) *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods Ver. 4.0 beta 8*. Sinauer, Sunderland.

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

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437–460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

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

Templeton A (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Templeton A, Crandall K, Sing C (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Templeton A, Routman E, Phillips C (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.

Thompson J, Higgins D, Gibson T (1994) Clustal W – Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Torrecilla P, Lopez-Rodríguez JA, Catalán P (2004) Phylogenetic relationships of *Vulpia* and related genera (Poaeae, Poaceae) based on analysis of ITS and *trn*LF sequences. *Annals of the Missouri Botanical Garden*, **91**, 124–158.

Trusty J, Olmstead R, Santos-Guerra A, Sá-Fontinha S, Francisco-Ortega J (2005) Molecular phylogenetics of the Macaronesian-endemic genus *Bystropogon* (Lamiaceae): palaeo-islands, ecological shifts and interisland colonizations. *Molecular Ecology*, **14**, 1177–1189.

Whittaker R, Fernández-Palacios JM (2007) *Island Biogeography*. Oxford University Press, New York.

Wolfe K, Li WH, Sharp P (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Science USA*, **84**, 9054–9058.

Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, **22**, 297–307.

Woolley S, Posada D, Crandall KA (2008) A comparison of phylogenetic network methods using computer simulation. *PLoS One*, **3**, e1913.

A.D.-P. specializes on plant population genetics, evolution and breeding. M.S. investigation focuses on taxonomy, biogeography and ecology of the *M. flora* and historical aspects of Macaronesian botany. A.S.-G.'s main research concerns plant taxonomy and biodiversity in Macaronesia. P.C. investigates evolutionary mechanisms, speciation processes and biogeographical patterns of angiosperms.

Data accessibility

(a) GenBank Accessions: Loliinae DNA sequences included in the studies of Catalán *et al.* (2006) and Inda *et al.* (2008): AF478512, EF593000, EF592952, AF478516, AF478514, EF592995, EF592997, AY118098, AY118088, AF478522, AF533063, AF478510, AF478511, AF478508, AF487616, AF478526, AY118102, AF488773.

Our new generated *trn*LF GenBank Accession nos of *Festuca agustinii*, *Festuca francoi*, *Festuca jubata*, *Festuca petraea* and *Festuca rivularis* are shown in Table S1 (Supporting information).

(b) Final DNA sequence assembly; scripts, data files and supplemental results of IM and Biogeographic Analyses; and Coalescence-based *p-distance* and Slatkin and Maddison's *S* scripts were uploaded in the Dryad repository as http://dx.doi.org/10.5061/dryad.1k682sm1.

Supporting information

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

Fig. S1 Maximum Parsimony, Maximum Likelihood and Bayesian Inference (Felsenstein 2004) of *Festuca* sect. *Aulaxyper* species and closely related fine-leaved Loliinae taxa.

Fig. S2 Posterior distributions for species (a) and Bayesian Group (b) divergence times of diploid Macaronesian and Iberian red fescues.

Table S1 Geographic sampling location, GenBank Accession codes and haplotype absolute frequencies (*N*) of *trn*LF sequences of diploid red *Festuca* species populations.

Table S2 *trn*LF haplotypes of five diploid red *Festuca* species. *Ntrn* designation refers to nucleotide polymorphism.

Table S3 Nested Clade Phylogeographical Analysis (NCPA; Templeton *et al.* 1995; Templeton 2004) of *Festuca Ntrn*LF haplotypes.

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