Comparative phylogeography of a restricted and a widespread heather: genetic evidence of multiple independent introductions of *Erica mackayana* into Ireland from northern Spain

JAIME FAGÚNDEZ[1,2,*] and PILAR DÍAZ-TAPIA[1,3]

[1]*Universidade da Coruña, BIOCOST research group, Centro Interdisciplinar de Química e Bioloxía (CICA), Rúa As Carballeiras, 15071, A Coruña, Spain*
[2]*Universidade da Coruña, Departamento de Bioloxía, Facultade de Ciencias, 15071, A Coruña, Spain*
[3]*Instituto Español de Oceanografía (IEO-CSIC), Centro Oceanográfico de A Coruña, Paseo Marítimo Alcalde Francisco Vázquez, 10, 15001, A Coruña, Spain*

Species of flora and fauna occurring in the west of Ireland and south-west Europe, known as Lusitanian elements, constitute a puzzling case of isolated populations of uncertain origin. Here we studied the population genetic structure of the heather *Erica mackayana* in Ireland and northern Spain and compared it with its widespread close relative *Erica tetralix* using single nucleotide polymorphisms (SNPs). We reconstructed phylogenetic relationships using maximum likelihood (ML), inferred population genetic structure using cluster assignment and principal component analysis, and estimated population genetic diversity. The cluster analysis and ML phylogenetic tree showed a geographical pattern for *E. tetralix* supporting a post-glacial migration from Iberia to Ireland. In contrast, Irish populations of *E. mackayana* were supported in independent clades in the phylogenetic tree and shared clusters with Iberian populations in the structure analysis, and F_{ST} values were lower among Irish and Spanish populations than among Irish ones. This suggests that Irish populations of *E. mackayana* are the result of recent multiple independent introductions from its native area in northern Spain, probably assisted by humans. However, the origin of the largest Irish population at Roundstone Bog is unclear and should be further investigated. Post-glacial, long-distance dispersal is the most plausible explanation for Lusitanian species distribution in Ireland.

ADDITIONAL KEYWORDS: *Erica tetralix* – Hiberno-Iberian species – Lusitanian species – SNPs.

INTRODUCTION

The current distribution of extant plant species is the result of past events of migration, colonization and extinction (Taberlet *et al.*, 1998; Waltari *et al.*, 2007). Climatic changes, such as the Pleistocene glaciations in Europe, have forced geographical range contraction and local extinction of species in northern areas, followed by northward expansions from southern refugia after the retreat of the ice (Taberlet *et al.*, 1998; Hewitt, 2000). On islands, these past cold events have caused massive local extinctions and left bare emerged land to be colonized from adjacent populations (Bennike, 1999; Fernández-Palacios *et al.*, 2016).

The post-glacial migration history of many plants and animals across Eurasia is still unknown. An example is the disjunct distribution of the Lusitanian (also known as Hiberno-Iberian) elements, a recognized remarkable phenomenon of European biogeography. These floristic and faunistic elements occur in the west of Ireland and the south of Atlantic Europe, mainly in the north of the Iberian Peninsula, but nowhere else in Britain and Ireland (Perring, 1967; Webb, 1955; Moore, 1987; Preston & Hill, 1999). Why these species form isolated populations and, specifically, the

*Corresponding author. E-mail: jaime.fagundez@udc.es

origin of the extant Irish populations, has long been a subject of debate (Corbet, 1962; Moore, 1987). Three main hypotheses have been put forward that can be summarized as (1) the existence of a glacial refugium in Ireland, including land submerged today, or post-glacial colonization, either (2) through a terrestrial route, or (3) a long-distance dispersal event (Beatty & Provan, 2013).

The survival of viable populations in Ireland during the last glaciation has been historically advocated by many authors, for all or at least some of the Lusitanian species (Stapf, 1911; Praeger, 1932; Webb, 1955). However, during the Pleistocene ice ages the island was completely covered by an ice sheet, so that the potential habitats for *in situ* survival of terrestrial plants and animals were extremely restricted, particularly during the Last Glacial Maximum (LGM; 21 000–18 000 BP; Westley & Edwards, 2017; Clark *et al.*, 2018; Roberts *et al.*, 2020), when mean temperatures were 7.0 ± 1.0 ºC lower than the pre-industrial last millennium average (Osman *et al.*, 2021).

Alternatively, other authors have suggested that the origin of the Lusitanian elements in Ireland could be the result of a post-glacial migration (Reid, 1899, 1911; Corbet, 1962). This could have happened as a long-distance dispersal event or via land bridges. A terrestrial migration of these species to western Ireland would have been followed by a local extinction in Britain and eastern Ireland in more recent times. Land bridges between Britain and Ireland did not last beyond the retreat of the ice (Edwards & Brooks, 2008), and thus migration would have happened under extremely cold conditions. These land bridges allowed the colonization of Britain and Ireland by widespread northern species, such as *Quercus robur* L. (Kelleher *et al.*, 2004), but a similar process is less likely for southern species with limited tolerance to low temperatures.

An alternative migration route would have been from Iberia by long-distance dispersal, most probably mediated by humans, but also potentially by birds that could have transported seeds (Bennike, 1999; Popp *et al.*, 2011). The hypothesis of a recent long-distance dispersal event has been put forward by several authors, but long criticized and dismissed as inconceivable (Reid, 1899; Praeger, 1934; Webb, 1955). Other studies have proposed that modern human-mediated introductions, either inadvertent or deliberate, are responsible for Ireland's colonization by flora and fauna. For example, numerous exotic species have been introduced and naturalized in modern times [e.g. *Rhododendron ponticum* L. (Milne & Abbott, 2000)]. Others were introduced much earlier. For example, the low haplotypic diversity of Irish populations of the snail *Cepaea nemoralis* and the mitochondrial DNA lineage shared with Pyrenean populations support an ancient introduction by humans (Grindon & Davison, 2013). Divergence time could be as old as 8000 years for this snail, but this is consistent with dating of human colonization of Ireland after the LGM, at least 12 000 BP (Dowd & Carden, 2016).

The most recent studies performed on several Lusitanian taxa using molecular techniques have consistently suggested a scenario of post-glacial dispersal events to Ireland for a range of species (Grindon & Davison, 2013; Beatty *et al.*, 2015; Reich *et al.*, 2015; Santiso *et al.*, 2016a, b). Beatty and colleagues analysed chloroplast and nuclear molecular markers coupled with climatic niche modelling to study some of the most remarkable Lusitanian plants, including *Saxifraga spathularis* Brot., *Pinguicula grandiflora* Lam., *Daboecia cantabrica* (Huds.) K.Koch and *Euphorbia hyberna* L. (Beatty & Provan, 2013, 2014; Beatty *et al.*, 2015). Low levels of genetic diversity, the absence of private haplotypes in Irish populations and their affinity with continental populations did not support, with different levels of uncertainty, the hypothesis of the existence of a glacial refugium in Ireland.

Recent advances in genomics have facilitated the analysis of large single nucleotide polymorphism (SNPs) datasets for non-model organisms. Compared with previously used molecular markers, the application of these genomic tools has provided greater resolution for differentiating populations and for the analysis of the genetic structure in a wide variety of taxa (Andrews *et al.*, 2016; Parchman *et al.*, 2018). Likewise, these tools may help to clarify the phylogeography of the Lusitanian flora, as recommended for fauna (Carlsson *et al.*, 2014).

Erica mackayana Bab. (Ericaceae) is one of the most remarkable elements of the Lusitanian flora. It inhabits two disjunct areas: northern Spain and western Ireland. In Spain, it is frequent and even dominant in wet heathlands and bogs on the Cantabrian and northern Atlantic coasts and in mountain ranges of up to 1000 m a.s.l. (Fagúndez, 2006, 2016). In western Ireland, it grows on a few wet heaths and peat bogs of the coasts of Kerry, Mayo, Galway and Donegal (Sheehy Skeffington & Van Doorslaer, 2015; Sheehy Skeffington, 2017). Fossil remains of *E. mackayana* have been recovered in Ireland from the Gortian interglacial period (Jessen *et al.*, 1959) and the Late Quaternary period (Jessen, 1949). The species is sexually sterile in Ireland, only propagating itself vegetatively (Nelson, 2011).

Erica tetralix L., a close relative of *E. mackayana*, has a markedly different distribution, as it is one of the most widespread European heathers occurring in Western Europe from the southern Iberian Peninsula to Scandinavia, and it is frequent throughout Britain and Ireland (Nelson, 2011). These two species and *Erica*

andevalensis Cabezudo & J.Rivera, a narrow endemic from the south of the Iberian Peninsula, form a well-supported clade in the genus (Mugrabi de Kuppler *et al.*, 2015). In addition, the hybrid between *E. tetralix* and *E. mackayana* [*Erica × stuartii* (Macfarl.) Mast.] is frequent in the Irish populations and occasionally occurs in some overlapping populations in northern Spain (Fagúndez, 2006; Sheehy Skeffington, 2015, 2017). These species are diploids with a chromosome number of $2n = 24$, and polyploidy has not been reported in European species of *Erica* L. (Mugrabi de Kuppler *et al.*, 2015).

In this study we aimed to clarify the origin of the isolated Irish populations of *E. mackayana*. We applied a genome-wide approach using SNPs to analyse the phylogeny, genetic structure and genetic diversity of Iberian and Irish populations. We also included its close relative *E. tetralix* in our study, to compare patterns in species with contrasting present-day geographical distributions. The origin of the Irish populations of *E. mackayana* could be either a case of in situ survival during the last ice age, the result of independent multiple long-distance dispersal events from the northern Iberian Peninsula or the gradual fragmentation of a single large population settled from the same region after the LGM. The sexual sterility of this species and the high rates of hybridization with *E. tetralix* in Ireland could be a sign of a strong bottleneck and could constrain its ability to propagate, supporting the recent introduction hypothesis (Sheehy Skeffington, 2017). Moreover, molecular analyses by Kingston & Waldren (2006) and Pene Eftonga (2013) were inconclusive but congruent with a founder effect of the different Irish populations. Another hypothesis would be a coastal migration route established through France and southern England, including submerged land, followed by extinction in the connecting area.

Such a hypothesis has been proposed for explaining the distribution of other Lusitanian species [e.g. *D. cantabrica* (Beatty & Provan, 2013)] based on historical climatic niche reconstruction. Here we evaluate these hypotheses and the native status of *E. mackayana* in Ireland, which has recently been questioned in favour of modern human-mediated introductions, based on historical evidence of trading between both countries (Sheehy Skeffington, 2015, 2017).

MATERIAL AND METHODS

STUDIED MATERIAL

Sixty-five and 53 samples of *E. mackayana* and *E. tetralix*, respectively, were collected in the north of Spain and the west of Ireland between 2013 and 2017 (Table 1). We sampled all the known populations of *E. mackayana* in Ireland (Sheehy Skeffington, 2017) and six populations in the north of Spain. *Erica tetralix* was collected when it co-occurred with *E. mackayana*. Additionally, one sample of *Erica ciliaris* L., the sister species of the *E. tetralix*-*E. mackayana* clade (Mugrabi de Kuppler *et al.*, 2015), was collected to be used as an outgroup in the phylogenetic analysis. A flowering branch of each plant was dried in a bag with silica gel for DNA extraction, and the main part was mounted as a herbarium specimen to be stored in the herbarium of the University of Santiago de Compostela (SANT).

Genomic DNA was extracted from the silica-preserved samples (Table 1) using an adapted cetyltrimethylammonium bromide (CTAB) protocol (Doyle & Doyle, 1987). Genomic DNA was used to construct nextRAD genotyping-by-sequencing libraries (SNPsaurus LLC, Eugene, OR, USA) following Russello *et al.* (2015). Genomic DNA was first fragmented with Nextera reagent (Illumina

Table 1. Collection site information and number of samples of the two studied species included in the final analyses (excluding five specimens with high levels of missing data). Galway includes samples from Carna (numbers in parentheses). One sample of *E. ciliaris,* which served as an outgroup for the ML phylogenetic analysis, was collected from the Pimiango population. NP = not present.

Population	Region/ country	Latitude (°N)	Longitude (°W)	Elevation (m a.s.l.)	*E. mackayana* (*N*)	*E. tetralix* (*N*)
Pimiango	Asturias	43.39	4.54	122	2	NP
Peñas	Asturias	43.65	5.85	98	2	NP
Espina	Asturias	43.39	6.31	688	13	1
Bustantigo	Asturias	43.33	6.71	1011	7	11
Xistral	Galicia	43.43	7.48	753	8	NP
Loba	Galicia	43.28	7.94	632	10	13
Donegal	Ireland	55.03	8.16	57	3	10
Mayo	Ireland	54.11	9.52	85	3	4
Galway	Ireland	53.43	9.89	22	9(+1)	7(+2)
Kerry	Ireland	51.93	10.07	120	3	4

Inc, San Diego, CA, USA), which also ligates short adapter sequences to the ends of the fragments. The Nextera reaction was scaled for fragmenting 15 ng of genomic DNA. Fragmented DNA was then amplified for 27 cycles at 74 °C, with one of the primers matching the adapter and extending nine nucleotides into the genomic DNA with the selective sequence GTGTAGAGCC. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer will be efficiently amplified. The nextRAD libraries were sequenced on an Illumina HiSeq 4000 System with two lanes of 150-bp reads (at the University of Oregon).

The genotyping analysis used custom scripts (SNPsaurus LLC) that trimmed the reads using bbduk (BBMap tools v.38.79, http://sourceforge.net/projects/bbmap/) with the following parameters: ktrim = r, k = 17, hdist = 1, mink = 8, minlen = 100, ow = t, qtrim = r, trimq = 10. Next, a de novo reference was created by collecting 10 000 000 reads in total, evenly from the samples, and excluding reads that had fewer than six or more than 700 counts. The remaining loci were then aligned to each other using bbmap (BBMap tools) to identify allelic loci and collapse allelic haplotypes to a single representative. All reads were mapped to the reference using bbmap (BBMap tools) with the following parameters: minid = 0.95, ambig = toss and maxindel = 8. Genotype calling was done using callvariants (BBMap tools) with the following parameters: ploidy = 2, multisample = t, rarity = 0.05, minallelefraction = 0.05, usebias = f, ow = t, nopassdot = f, minedistmax = 5, minedist = 5, minavgmapq = 15, minreadmapq = 15, minstrandratio = 0.0, strandedcov = t. The mean number of reads was 4 193 195.69 (SD 2 477 783.97), and the mean of retained reads was 1 597 540.16 (SD 969 414.05) after filtering and mapping. The variant call format (vcf) was filtered to remove alleles with a population frequency of > 3%.

We obtained an aligned matrix of 2767 polymorphic loci. Five samples with more than 50% of missing data were excluded from further analyses, so that the final dataset comprised 52 and 61 samples of *E. tetralix* and *E. mackayana*, respectively.

PHYLOGENETIC ANALYSIS

The vcf file was exported as a fasta file using Tassel, and the heterozygous sites were encoded following the International Union of Pure and Applied Chemistry (IUPAC). Invariant sites were filtered using IQTREE v.2.1.3 so that the final alignment included 2051 base pairs. We built a ML phylogenetic tree using RAxML-NG v.8.2.12 (Kozlov *et al.*, 2019), using a GTGTR4 substitution model, a G among-site rate heterogeneity model and the ascertainment bias

correction ASC_LEWIS, with 1000 non-parametric bootstrap replicates.

ANALYSIS OF GENETIC DIVERSITY AND POPULATION GENETIC STRUCTURE

VCFTOOLS v. 0.1.13 (Danecek *et al.*, 2011) was used to prune out linked loci using 150 bp as the minimum distance between two sites, so that the final dataset comprised 791 and 747 variant sites for *E. mackayana* and *E. tetralix*, respectively. The overall level of missingness across all sites was 16.63% and 15.96% in the *E. mackayana* and *E. tetralix* data sets, respectively. Average expected (H_s) and observed (H_o) heterozygosity, and inbreeding coefficient (F_{IS}) were estimated using hierfstat v.0.04-6 (Goudet, 2005). Weir and Cockerham pairwise F_{ST} values among populations and their statistical significance were calculated using the R package SambaR v.1.06 with 1000 bootstrap replications (de Jong *et al.*, 2021).

The population genetic structure was separately analysed for the two studied species using two complimentary approaches. Firstly, a ML population structure assessment was inferred using ADMIXTURE v.1.3.0 with default parameters (Alexander & Lange, 2011) to estimate individual ancestries (Q matrix) for potential populations (K) ranging from K = 1 to 8 and 11 for *E. mackayana* and *E. tetralix*, respectively. Twenty independent analyses were run. The online web server of the program CLUMPAK (Kopelman *et al.*, 2015) was used to summarize the ADMIXTURE results into structure plots. Secondly, we analysed patterns of geographical differentiation using a principal component analysis (PCA) conducted with the R package adegenet v.2.1.3 (Jombart, 2008).

RESULTS

PHYLOGENETIC ANALYSES

In the ML phylogenetic tree, *E. mackayana* and *E. tetralix* were retrieved in two supported clades (Fig. 1). For the widespread species *E. tetralix* the tree shows a clear geographical pattern, as samples from Ireland were resolved in a neatly defined clade that received full support, whereas Spanish samples were placed in several earlier diverged lineages.

The Asturian populations (Bustantigo and Espina) were clustered in a clade nested in the basal Galician (Loba) population, and sister to the Irish clade. The topology of the Irish clade was structured according to the geographical distribution of populations following a latitudinal pattern, from Kerry in the south to Donegal in the north, with different levels of support for each population (Fig. 1).
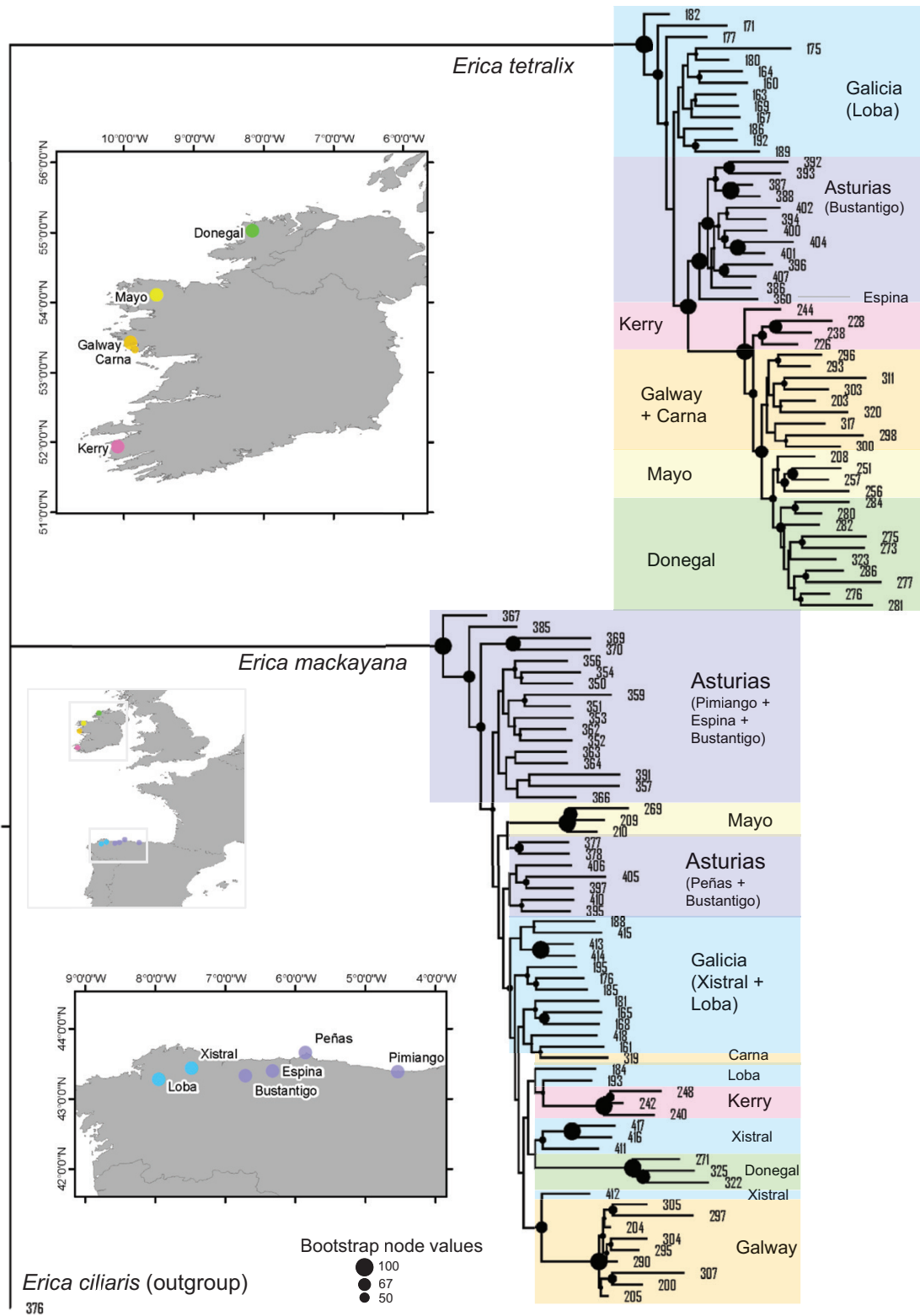
Figure 1. RAxML phylogenetic tree of the studied populations of *E. tetralix* and *E. mackayana*, with colours and labels for each population/region as in the maps. Codes are for samples, bootstrap node values are proportional to circle sizes (see legend).

Table 2. Estimates of genetic diversity calculated across sampling locations in which more than one sample was collected

Population	*E. tetralix*				*E. mackayana*			
	N	H_o	H_s	F_{IS}	N	H_o	H_s	F_{IS}
Donegal	10	0.205	0.169	-0.075	3	0.139	0.135	-0.045
Mayo	4	0.124	0.174	-0.09	3	0.222	0.153	-0.114
Carna	2	0.192	0.168	-0.055	1	-	-	-
Galway	7	0.204	0.171	-0.072	9	0.26	0.176	-0.167
Kerry	4	0.243	0.179	-0.125	3	0.232	0.159	-0.123
Loba	13	0.253	0.208	-0.089	10	0.29	0.235	-0.114
Xistral	-	-	-	-	8	0.282	0.239	-0.093
Bustantigo	12	0.265	0.217	-0.099	7	0.29	0.244	-0.093
Espina	1	-	-	-	13	0.276	0.227	-0.098
Peñas	-	-	-	-	2	0.189	0.188	-0.024
Pimiango	-	-	-	-	2	0.3	0.229	-0.149

Table 3. Pairwise F_{ST} estimates between *E. tetralix* populations, with 1000 bootstrap replicates. The colour gradient from green to red indicates increasing F_{ST} values. F_{ST} *P*-values were 0 for all pairwise estimates

	Galway	Kerry	Mayo	Donegal	Loba	Bustantigo
Kerry	0.07					
Mayo	0.06	0.10				
Donegal	0.09	0.14	0.06			
Loba	0.35	0.30	0.32	0.36		
Bustantigo	0.34	0.30	0.31	0.34	0.16	
Espina	0.37	0.35	0.36	0.38	0.12	0.05

The tree for *E. mackayana* was opposite to that of *E. tetralix* for the Iberian populations, with Galician populations (Xistral and Loba) nested in the early-diverging clade of Asturias (Pimiango, Peñas, Bustantigo and Espina). The four Irish populations, Roundstone Bog (Galway), Donegal, Kerry and Mayo, were placed in four strongly supported clades (node bootstrap value = 100), intermixed with the Iberian populations and not related to each other (Fig. 1). One sample collected in Carna, initially assigned to the Galway population as it was collected close to the Roundstone Bog population, was also placed independently in the tree.

GENETIC DIVERSITY AND POPULATION GENETIC STRUCTURE

Genetic diversity was in general higher in Spanish populations compared with Irish ones in the two studied species (Table 2). The inbreeding coefficient (F_{IS}) was negative for the two species in all sampling sites, indicating a small excess of heterozygotes.

The F_{ST} estimates showed relevant differences in patterns of genetic differentiation of populations between species. Differentiation among populations within each of the studied countries, Spain and Ireland, was lower (≤ 0.16) than between populations of these two countries (≥ 0.3) in *E. tetralix* (Table 3). Similarly low F_{ST} estimates were found among populations of *E. mackayana* from Spain (≤ 0.12, Table 4). However, F_{ST} estimates among Irish populations were similar or higher (≥ 0.36) when compared with Spanish populations (0.11–0.35). In particular, two Irish populations (Mayo and Kerry) had low F_{ST} estimates (0.11–0.31) with regard to Spanish populations. In contrast, some Irish populations showed the highest values (e.g. Kerry vs. Donegal, 0.43).

The clustering analysis revealed a highly structured genetic assemblage in populations for the two species (Fig. 2). According to the cross validations (Supporting Information, Fig. S1), the strongest support was found for two and three clusters in *E. mackayana* and *E. tetralix*, respectively, but cross validation values were highly similar in both species for K = 2–3 (ranging between 0.331–0.332 and 0.369–0.39 in *E. mackayana* and *E. tetralix*, respectively). However, the correspondence between clusters and the geographical origin of samples differed greatly between species.

In *E. tetralix*, the three clusters corresponded to the three main geographical areas [Ireland (blue), Galicia (green) and Asturias (red)] with limited admixture (Fig. 2A, C, E). The single plant analysed from Espina, in Asturias, showed an admixture of Galician and Asturias clusters. Irish populations had an increasing admixture of the Spanish clusters from north to south.

In *E. mackayana*, the three clusters identified did not correspond with the geographical origin of samples (Fig. 2B, D, F). One of the clusters (red) was dominant in plants from Galicia (Loba and Xistral populations) and the Irish populations from Kerry, Carna and Donegal. A second cluster (green) included plants from Asturias (Bustantigo, Espina, Peñas and Pimiango)

Table 4. Pairwise F_{ST} estimates between *E. mackayana* populations, with 1000 bootstrap replicates. The colour gradient from green to red indicates increasing F_{ST} values. One plant from Carna was not included. F_{ST} P-values were 0.01 for Bustantigo and Pimiango, 0 for all other pairwise estimates

	Donegal	Galway	Kerry	Mayo	Loba	Xistral	Bustantigo	Pimiango	Espina
Galway	0.42								
Kerry	0.43	0.36							
Mayo	0.4	0.38	0.37						
Loba	0.2	0.22	0.16	0.15					
Xistral	0.2	0.21	0.16	0.16	0.02				
Bustantigo	0.19	0.24	0.16	0.11	0.04	0.04			
Pimiango	0.29	0.3	0.28	0.27	0.06	0.05	0.02		
Espina	0.21	0.23	0.19	0.13	0.06	0.06	0.03	0.04	
Peñas	0.34	0.35	0.31	0.26	0.09	0.1	0.06	0.12	0.07

and the Irish population from Mayo. Some admixture of these two groups was detected in plants from these sites. The third cluster (blue) was represented by plants from Galway, and was marginally represented in the Galician populations of Xistral and Loba, the Asturian population of Espina and the Irish population of Donegal.

We further evaluated population structure using a PCA. Consistent with the structure clustering analysis, patterns again clearly differed between the two species. In *E. tetralix*, PC1 clearly differentiated Spanish and Irish populations, whereas PC2 and PC3 separated populations in each of these two regions (Fig. 3A). PC2 differentiated two groups of samples, corresponding to Galician (Loba) and Asturian samples. Likewise, PC3 moderately distinguished four groups of Irish samples, the placement of which in the PCA is congruent with their latitudinal origin. In contrast, in *E. mackayana* PC1 clearly separated the Galway (Roundstone Bog) population from all other Irish and Spanish samples (Fig. 3B). Moreover, PC2 and PC3 moderately distinguished north-western and northern Spanish populations, whereas the remaining Irish populations overlapped or were placed close to Spanish ones; the population from Kerry was the only exception.

DISCUSSION

Our study revealed conspicuous differences in population genetic structure between the widespread Atlantic *E. tetralix* and its restricted congener *E. mackayana*. We found a strong phylogeographic signal in populations of *E. tetralix*, congruent with a south-north migration and isolation-by-distance pattern. In contrast, the population structure of *E. mackayana* did not correspond with the geographical distribution of samples, and our analyses suggest that most Irish populations are the result of multiple recent dispersal events from different Spanish populations.

In *E. tetralix*, Irish and Iberian populations were clearly differentiated in all the analyses. The southernmost population in Ireland (Kerry) included a small percentage of the two Iberian clusters which were gradually lost northwards. Also, the genetic diversity was higher in Iberian populations. This pattern is consistent with the classical post-glacial migration from southern refugia to northern Europe described for other species (Hewitt, 2000; Kelleher *et al.*, 2004), including the most widespread heather *Calluna vulgaris* (L.) Hull. This species showed a genetic diversity decrease and a strong isolation-by-distance signal from south to north, from Iberia to northern Europe and at a regional scale in Scotland (Mahy *et al.*, 1999; Rendell & Ennos, 2002; Gil-López *et al.*, 2022). Although our study included only a limited number of populations of *E. tetralix*, considering its wide geographical range, the pattern found in the phylogenetic analysis is consistent with a south-north migration both at the broad scale (Iberia to Ireland) and the regional scale (Irish populations). However, additional populations covering the entire distribution of *E. tetralix* will be required to further understand the phylogeography of this widespread Atlantic species in a wider geographical context.

In contrast to *E. tetralix*, Spanish and Irish populations of *E. mackayana* were not discriminated at the country level; however, the Roundstone Bog (Galway) population was clearly distinct from the others. The phylogenetic tree showed a basal grade of Asturian populations with Galician ones nested in it with low support. The Irish populations were placed in unrelated positions among them. Surprisingly, the single specimen collected in Carna was genetically more similar to other populations albeit being only *c.* 15 km apart from the large Roundstone Bog population in Galway. Instead, Carna was placed independently from other Irish populations in the phylogenetic tree, and shared a cluster with Kerry, Donegal and populations from Galicia (Xistral, Loba) in the structure analysis.
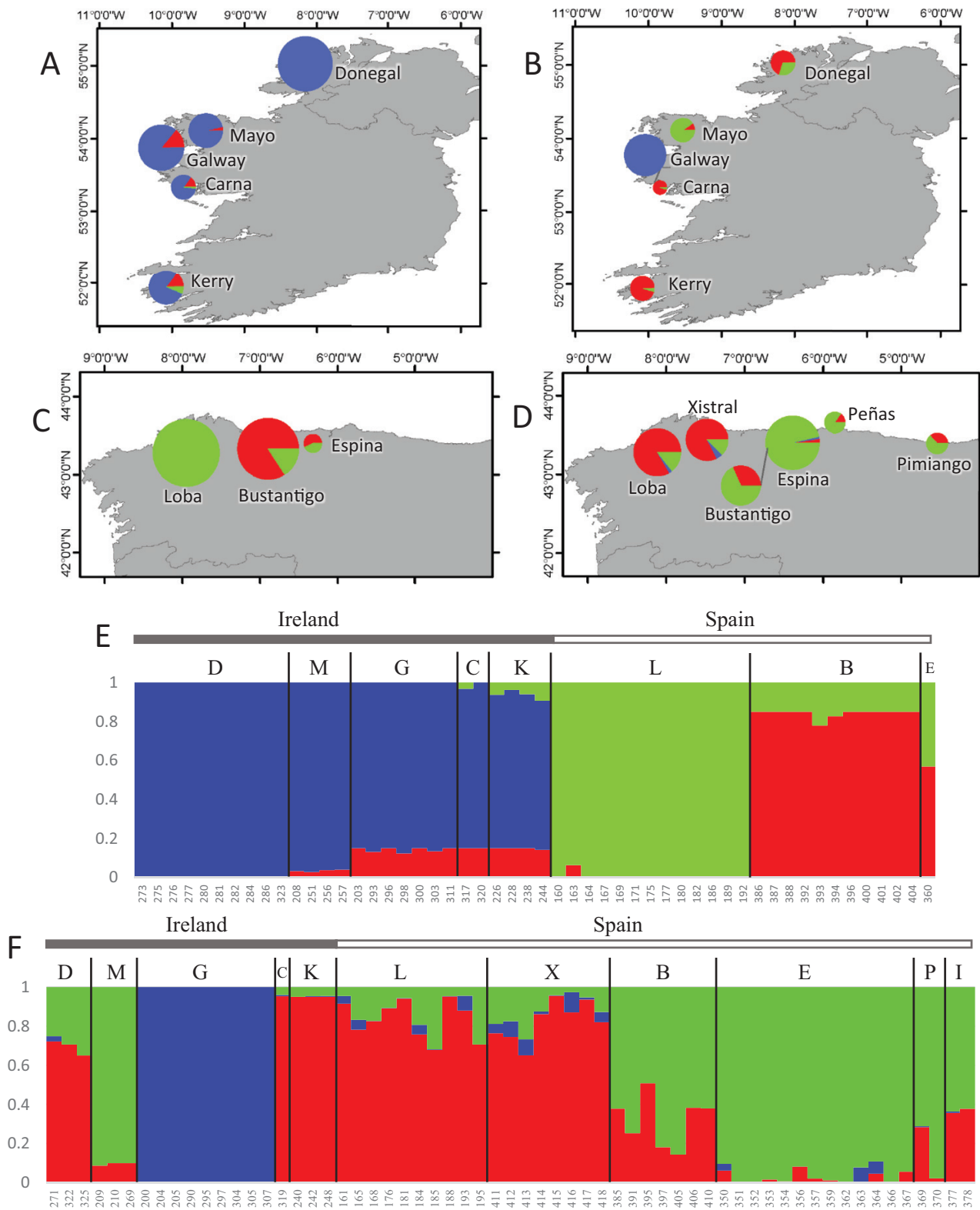
Figure 2. Population structure genetic matrix and geographical distribution of each cluster for populations of *E. tetralix* (A, C, E) and *E. mackayana* (B, D, F). K-values were selected based on cross-validation: *E. mackayana* and *E. tetralix* K = 3. Size of the circles in the maps are proportionate to number of samples per population. Y axes in E and F represent frequency (0-1). D = Donegal, M = Mayo, G = Galway, C = Carna, K = Kerry, L = Loba, X = Xistral, B = Bustantigo, E = Espina, P = Peñas, I = Pimiango.
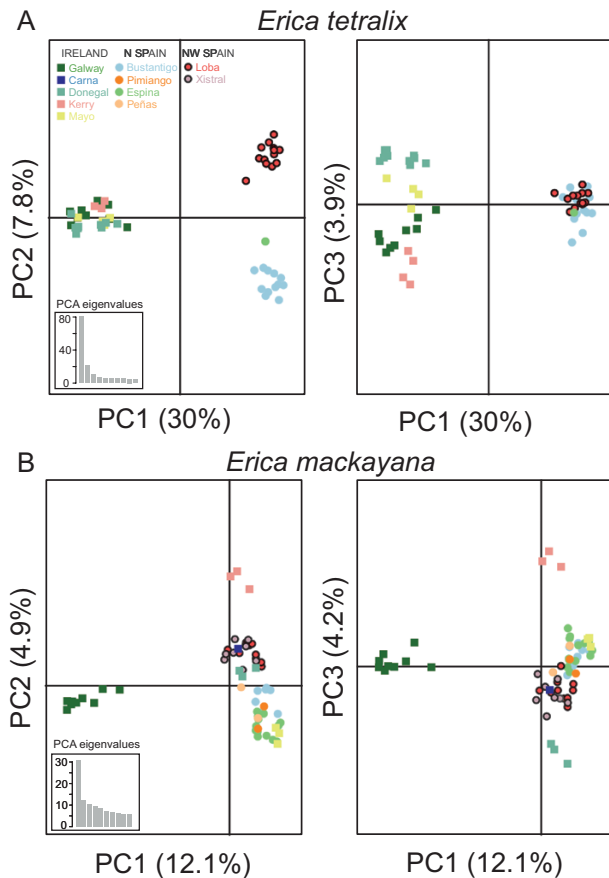
Figure 3. Principal component analysis (PCA) showing individual variation of the genomic data. Panels show PC1 plotted against PC2 and PC3 for *E. tetralix* (A) and *E. mackayana* (B). The amounts of variation explained by each PC are given as percentages. Eigenvalues for the first ten PCs obtained are shown in the insets.

Finally, the population of Mayo shared a cluster with the Asturian populations in the structure analysis, and it was placed among the Asturian samples in the phylogenetic tree, sharing an unsupported clade with plants from Peñas, in coastal central Asturias.

Our findings support the recent dispersal hypothesis for most Irish *E. mackayana* populations from multiple independent events. Trading of goods in the 19th and 20th centuries could have been a potential path of human-mediated introduction (Sheehy Skeffington, 2015; Sheehy Skeffington & Van Doorslaer, 2015). Smugglers at that time did not use the main roads, and they may have used heather to protect cautiously hidden goods in isolated places away from the main paths and at some distance from the coast, corresponding to the present-day location of most Irish *E. mackayana* populations (Sheehy Skeffington, 2017).

Other references to the use of heathers in trade would support this hypothesis. For example,

Reich *et al.* (2015) claimed that the slug *Geomalacus maculosus*, another Lusitanian species, may have been introduced unintentionally in materials used for packaging wine, which '…were probably protected by layers of heath or other vegetation which may have contained *G. maculosus* specimens and/or eggs'. Similarly, Foss & Doyle (1988) suggested that pilgrims in the 15th century may have introduced *Erica erigena* R.Ross in Ireland from Galicia, another heather species with a Lusitanian distribution. In another study, Santiso *et al.* (2016a, b) included an Irish population of *Arbutus unedo* L., one of the most iconic examples of the Lusitanian flora, in their wide phylogeographic study of the species using non-coding plastid DNA and amplified fragment length polymorphisms (AFLPs). The close linkage of Irish to northern Iberian populations is coherent with a long-distance colonization event, rather than a terrestrial colonization or *in situ* survival during the LGM as supported by previous authors (Sealy & Webb, 1950; Webb, 1955). An alternative hypothesis of introduction through mining activities has been given by Sheehy Skeffington & Scott (2021).

An alternative explanation for the foundation of Irish *E. mackayana* populations in the absence of human intervention could be a long-distance dispersal probably facilitated by migrant birds, as suggested for *C. vulgaris* (Mahy *et al.*, 1999). The seeds of *C. vulgaris* and many *Erica* spp., including *E. mackayana*, are small (< 1 mm), light (< 0.05 mg), have a prominent ornamentation and are produced in large numbers (Fagúndez *et al.*, 2010). These traits may facilitate long-distance travelling trapped among bird feathers, although they lack any specific structure for epizoochory (Sorensen, 1986).

Irish populations of *E. mackayana*, except Roundstone Bog (Galway), were genetically similar to Spanish plants from different regions suggesting that plants from Mayo could have recently dispersed from the coast of Asturias, whereas plants from Donegal, Carna and Kerry could have come from Galicia. This is in agreement with previous studies based on an AFLP analysis that revealed a mixed cluster of Irish and Spanish populations, interpreted as a potential effect of multiple founder events (Kingston & Waldren, 2006). However, the authors stated that samples could have been misidentified, and plants from Donegal could be of hybrid origin (Kingston & Waldren, 2006). Similarly, results from Pene Eftonga (2013) provided consistent results with the different origins of Irish populations, which also showed low genetic diversity compared to those from Spain.

The genetic differentiation among Irish populations, revealed by the highest F_{ST} values observed in the species, suggests that they are isolated. This result

is unsurprising considering that the species does not set viable seeds in Ireland (Webb, 1955; Nelson, 2011), contrary to Iberian populations (Fagúndez & Izco, 2010). In fact, Irish populations have been described as single clones, and sexual reproduction only takes place through crossing with *E. tetralix* (Nelson, 2005; Sheehy Skeffington & Van Doorslaer, 2015). The hybrid is frequent in Ireland but not in Spain where the contact area between the species is narrow, and even in co-occurring populations the hybrid is rare (Fagúndez, 2006). This markedly different pattern in hybridization rates between countries resembles other introduced plant species that frequently hybridize in non-native areas but rarely in their home range (Ellstrand & Schierenbeck, 2000; Prunera-Olivé *et al.*, 2019).

Our results showed a conspicuous differentiation of the population from Roundstone Bog (Galway) compared with the other Irish and Iberian populations. However, the phylogenetic tree showed this population nested in the main clade similar to the other Irish populations, and not as a sister clade, suggesting a similar origin to the others. The glacial refugium hypothesis is not supported by these results. There are fossil deposits of *E. mackayana* in Ireland, dating from both the Interglacial Gortian period (365 000 BP) and post-glacial times (Jessen, 1949; Jessen *et al.*, 1959). However, other non-native species registered in the fossil record are known to have been introduced recently in Ireland, such as *R. ponticum* (Jessen, 1948; Milne & Abbott, 2000). Several species, including *E. mackayana* and other heathers such as *Erica scoparia* L. and *Erica umbellata* L., are known to have occurred at northern latitudes during pre- and interglacial times, but did not recover their former range after the last glaciation (Jessen *et al.*, 1959, Wilson *et al.*, 1973). Sterility and hybridization are also major constraints for long-term population survival, even if vegetative and clonal growth is intense (Wolf *et al.*, 2001). The current Roundstone Bog (Galway) population could be another recent introduction from the Iberian Peninsula, but the source may have not been captured in our sampling, explaining the strong differences of this population shown in the structure analysis and the PCA. We covered the species range in the Iberian Peninsula, but the species inhabits several mountain ranges with some potential degree of isolation (Fagúndez, 2006, 2016).

CONCLUSION

Our results demonstrate, with a high level of confidence, that Irish populations of *E. mackayana* have independent origins and are not the result of fragmentation of a larger population. This confirms it is not a relict species that expanded its distribution from an Irish glacial refugium. However, there is some uncertainty about the origin of the largest Irish population at Roundstone Bog (Galway), which should be investigated further by dense sampling in northern Spain. The high resolution of the SNPs showed that Irish populations are probably the result of multiple independent colonization events, supporting the hypothesis of long-distance dispersal events for the Lusitanian flora and fauna, in line with recent works that found similar patterns for other elements with disjunct distributions (Grindon & Davison, 2013; Reich *et al.*, 2015; Beatty *et al.*, 2015). We suggest this is probably the biogeographic pattern for the majority or all of the Lusitanian species. Dating the arrival of these species and disentangling how humans may have been responsible of their introductions in historical times is another challenge for Iberian and Irish biogeography.

REFERENCES

Alexander DH, **Lange K. 2011**. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**: 1–6.

Andrews KR, **Good JM**, **Miller MR**, **Luikart G**, **Hohenlohe PA. 2016**. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**: 81–92.

Beatty GE, **Lennon JJ**, **O'Sullivan CJ**, **Provan J. 2015**. The not-so-Irish spurge: *Euphorbia hyberna* (Euphorbiaceae) and the Littletonian plant 'steeplechase'. *Biological Journal of the Linnean Society* **114**: 249–259.

Beatty GE, **Provan J. 2013**. Post-glacial dispersal, rather than *in situ* glacial survival, best explains the disjunct distribution of the Lusitanian plant species *Daboecia cantabrica* (Ericaceae). *Journal of Biogeography* **40**: 335–344.

Beatty GE, **Provan J**, **Provan J. 2014**. Phylogeographical analysis of two cold-tolerant plants with disjunct Lusitanian distributions does not support *in situ* survival during the last glaciation. *Journal of Biogeography* **41**: 2185–2193.

Bennike O. 1999. Colonisation of Greenland by plants and animals after the last ice age: a review. *Polar Record* **35**: 323–336.

Carlsson J, **Cross TF**, **McGinnity P**, **Prodöhl PA**, **McDevitt AD. 2014**. The use of genetics to infer the origins of the terrestrial and freshwater elements of the Irish fauna. *Irish Naturalists' Journal* **33**: 8–27.

Clark CD, **Ely JC**, **Greenwood SL**, **Hughes AL**, **Meehan R**, **Barr ID**, **Bateman MD**, **Bradwell T**, **Doole J**, **Evans DJ. 2018**. BRITICE Glacial Map, version 2: a map and GIS database of glacial landforms of the last British-Irish Ice Sheet. *Boreas* **47**: 11–e8.

Corbet G. 1962. The 'Lusitanian Element' in the British fauna. *Science Progress* **50**: 177–191.

Danecek P, **Auton A**, **Abecasis G**, **Albers CA**, **Banks E**, **DePristo MA**, **Handsaker RE**, **Lunter G**, **Marth GT**, **Sherry ST**, **McVean G**, **Durbin R**; **1000 Genomes Project Analysis Group. 2011**. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Dowd M, **Carden RF. 2016**. First evidence of a Late Upper Palaeolithic human presence in Ireland. *Quaternary Science Reviews* **139**: 158–163.

Doyle JJ, **Doyle JL. 1987**. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19**: 11–15.

Edwards R, **Brooks A. 2008**. The island of Ireland: drowning the myth of an Irish land-bridge? *The Irish Naturalists' Journal* **29**: 19–34.

Ellstrand NC, **Schierenbeck KA. 2000**. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences USA* **97**: 7043–7050.

Fagúndez J. 2006. Two wild hybrids of *Erica* L. (Ericaceae) from northwest Spain. *Botanica Complutensis* **30**: 131.

Fagúndez J. 2016. Grazing effects on plant diversity in the endemic *Erica mackayana* heathland community of northwest Spain. *Plant Ecology & Diversity* **9**: 207–217.

Fagúndez J, **Juan R**, **Fernández I**, **Pastor J**, **Izco J. 2010**. Systematic relevance of seed coat anatomy in the European heathers (Ericeae, Ericaceae). *Plant Systematics and Evolution* **284**: 65–76.

Fernández-Palacios JM, **Rijsdijk KF**, **Norder SJ**, **Otto R**, **de Nascimento L**, **Fernández-Lugo S**, **Tjørve E**, **Whittaker RJ. 2016**. Towards a glacial-sensitive model of island biogeography. *Global Ecology and Biogeography* **25**: 817–830.

Foss PJ, **Doyle GJ. 1988**. Why has *Erica erigena* (the Irish heather) such a markedly disjunct European distribution? *Plants Today* **1**: 161–168.

Gil-López MJ, **Segarra-Moragues JG**, **Casimiro-Soriguer R**, **Ojeda F. 2022**. From the Strait of Gibraltar to northern Europe: Pleistocene refugia and biogeographic history of heather (*Calluna vulgaris*, Ericaceae). *Botanical Journal of the Linnean Society* **198**: 41–56.

Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**: 184–186.

Grindon AJ, **Davison A. 2013**. Irish *Cepaea nemoralis* land snails have a cryptic Franco-Iberian origin that is most easily explained by the movements of Mesolithic humans. *PLoS One* **8**: e65792.

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

Jessen K. 1948. *Rhododendron ponticum* L. in the Irish interglacial flora. *The Irish Naturalists' Journal* **9**: 174–175.

Jessen K. 1949. Studies in Late Quaternary deposits and flora-history of Ireland. *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science* **52**: 85–290.

Jessen K, **Andersen ST**, **Farrington A. 1959**. The Interglacial Deposit near Gort, Co. Galway, Ireland. *Proceedings of the Royal Irish Academy, Section B: Biological, Geological, and Chemical Science* **60**: 1–77.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

de Jong MJ, **de Jong JF**, **Hoelzel AR**, **Janke A. 2021**. SambaR: an R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. *Molecular Ecology Resources* **21**: 1369–1379.

Kelleher CT, **Hodkinson TR**, **Kelly DL**, **Douglas GC. 2004**. Characterisation of chloroplast DNA haplotypes to reveal the provenance and genetic structure of oaks in Ireland. *Forest Ecology and Management* **189**: 123–131.

Kingston N, **Waldren S. 2006**. Biogeography of the Irish 'Lusitanian' heathers. In: *Botanical links in the Atlantic Arc. BSBI Conference Report No. 24*, 147–155.

Kopelman NM, **Mayzel J**, **Jakobsson M**, **Rosenberg NA**, **Mayrose I. 2015**. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15**: 1179–1191.

Kozlov AM, **Darriba D**, **Flouri T**, **Morel B**, **Stamatakis A. 2019**. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**: 4453–4455.

Mahy G, **Ennos RA**, **Jacquemart AL. 1999**. Allozyme variation and genetic structure of *Calluna vulgaris* (heather) populations in Scotland: the effect of postglacial recolonization. *Heredity* **82**: 654–660.

Milne RI, **Abbott RJ. 2000**. Origin and evolution of invasive naturalized material of *Rhododendron ponticum* L. in the British Isles. *Molecular Ecology* **9**: 541–556.

Moore PD. 1987. Snails and the Irish question. *Nature* **328**: 381–382.

Mugrabi de Kuppler AL, **Fagúndez J**, **Bellstedt DU**, **Oliver EGH**, **Léon J**, **Pirie MD. 2015**. Testing reticulate versus coalescent origins of *Erica lusitanica* using a species phylogeny of the northern heathers (Ericeae, Ericaceae). *Molecular Phylogenetics and Evolution* **88**: 121–131.

Nelson EC. 2005. *Erica mackaiana* Bab. and *Erica × stuartii* (MacFarl.) Mast. (Ericaceae): two heathers new to South Kerry (V.C. H1), Ireland. *Watsonia* **25**: 414–417.

Nelson EC. 2011. *Hardy heathers from the Northern Hemisphere*. Kew: Kew Publishing.

Osman MB, **Tierney JE**, **Zhu J**, **Tardif R**, **Hakim GJ**, **King J**, **Poulsen CJ. 2021**. Globally resolved surface temperatures since the Last Glacial Maximum. *Nature* **599**: 239–244.

Parchman TL, **Jahner JP**, **Uckele KA**, **Galland LM**, **Eckert AJ. 2018**. RADseq approaches and applications for forest tree genetics. *Tree Genetics & Genomes* **14**: 39.

Pene Eftonga RS. 2013. *Genetics, climate and phylogeography: understanding the Hiberno-Iberian distribution of Daboecia cantabrica (Huds). K. Koch and Erica mackaiana Bab. (Ericaceae)*. Ph. D. Thesis, Trinity College Dublin.

Perring FH. 1967. The Irish problem. In: Tüxen R, ed. *Pflanzensoziologie und Palynologie*. Dordrecht: Springer, 257–268.

Popp M, **Mirré V**, **Brochmann C. 2011**. A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences USA* **108**: 6520–6525.

Praeger RL. 1932. Recent views bearing on the problem of the Irish flora and fauna. An address delivered before the Academy, 30 November, 1932. *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science* **41**: 125–145.

Praeger RL. 1934. *The botanist in Ireland*. Dublin: Hodges, Figgis & Company.

Preston CD, **Hill MO. 1999**. The geographical relationships of the British and Irish flora: a comparison of pteridophytes, flowering plants, liverworts and mosses. *Journal of Biogeography* **26**: 629–642.

Prunera-Olivé J, **Galbany-Casals M**, **Cremades J**, **Fagúndez J. 2019**. A new hybrid between two alien *Helichrysum* species (Compositae, Gnaphalieae) from NW Spain. *Biological Invasions* **21**: 1481–1490.

Reich I, **Gormally M**, **Allcock AL**, **McDonnell R**, **Castillejo J**, **Iglesias J**, **Quinteiro J**, **Smith CJ. 2015**. Genetic study reveals close link between Irish and northern Spanish specimens of the protected Lusitanian slug *Geomalacus maculosus*. *Biological Journal of the Linnean Society* **116**: 156–168.

Reid C. 1899. *The origin of the British flora*. London: Dulau.

Reid C. 1911. The relation of the present plant population of the British Isles to the Glacial Period. *The Irish Naturalist* **20**: 201–209.

Rendell S, **Ennos RA. 2002**. Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology* **11**: 69–78.

Roberts DH, **Ó Cofaigh C**, **Ballantyne CK**, **Burke M**, **Chiverrell RC**, **Evans DJA**, **Clark CD**, **Duller GAT**, **Ely J**, **Fabel D**, **Small D**, **Smedley RK**, **Callard SL. 2020**. The deglaciation of the western sector of the Irish Ice Sheet from the inner continental shelf to its terrestrial margin. *Boreas* **49**: 438–460.

Russello MA, **Waterhouse MD**, **Etter PD**, **Johnson EA. 2015**. From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ* **3**: e1106.

Santiso X, **Lopez L**, **Retuerto R**, **Barreiro R. 2016a**. Population structure of a widespread species under balancing selection: the case of *Arbutus unedo* L. *Frontiers in Plant Science* **6**: 1264.

Santiso X, **Lopez L**, **Retuerto R**, **Barreiro R. 2016b**. Phylogeography of a widespread species: pre-glacial vicariance, refugia, occasional blocking straits and long-distance migrations. *AoB PLANTS* **8**: plw003.

Sealy JR, **Webb DA. 1950**. *Arbutus unedo* L. *Journal of Ecology* **38**: 223–236.

Sheehy Skeffington M. 2015. Ireland's Lusitanian heathers - an *Erica mackayana* perspective. *Ecological Questions* **21**: 13–15.

Sheehy Skeffington M. 2017. An alien immigrant? The story of Mackay's heath *Erica mackayana* in Ireland. *Moorea* **17**: 61–70.

Sheehy Skeffington M, **Scott N. 2021**. Is the strawberry tree, *Arbutus unedo* (Ericaceae), native to Ireland, or was it brought by the first copper miners? *British & Irish Botany* **3**: 385–418.

Sheehy Skeffington M, **Van Doorslaer L. 2015**. Distribution and habitats of *Erica mackayana* and *Erica × stuartii* (Ericaceae): new insights and ideas regarding their origins in Ireland. *New Journal of Botany* **5**: 164–177.

Sorensen AE. 1986. Seed dispersal by adhesion. *Annual Review of Ecology and Systematics* **17**: 443–463.

Stapf O. 1911. The southern element in the British flora. *Engler´s Botanische Jahrbücher (supp.)*: 509–525.

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

Waltari E, **Hijmans RJ**, **Peterson AT**, **Nyári AS**, **Perkins SL**, **Guralnick RP. 2007**. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One* **2**: e563.

Webb DA 1955. Biological flora of the British isles: *Erica Mackaiana* Bab. *Journal of Ecology* **43**: 319–330.

Westley K, **Edwards R. 2017**. Irish Sea and Atlantic Margin. In: Flemming NC, Harff J, Moura D, Burgess A, Bailey GN, eds. *Submerged landscapes of the European continental shelf: Quaternary paleoenvironments*. Hoboken, Chichester: John Wiley & Sons, 241–279.

Wilson DG, **Marchant R**, **Oldfield F. 1973**. Fossil seeds of *Erica* from the Cromer Forest bed series. *The New Phytologist* **72**: 1235–1237.

Wolf DE, **Takebayashi N**, **Rieseberg LH. 2001**. Predicting the risk of extinction through hybridization. *Conservation Biology* **15**: 1039–1053.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article on the publisher's website.

Figure S1. ADMIXTURE mean cross validation values across twenty independent runs for K = 1–8 and 1–11 in *E. tetralix* and *E. mackayana*, respectively.