

## Article

# Novel post-glacial haplotype evolution in Birch – a case for conserving local adaptation

Samuel Belton <sup>1</sup>, Philippe Cubry <sup>1,2</sup>, Erica Fox <sup>1</sup> and Colin T. Kelleher <sup>1,\*</sup>

<sup>1</sup> DBN Plant Molecular Laboratory, National Botanic Gardens of Ireland, Glasnevin, Dublin 9, Ireland

<sup>2</sup> Current address : DIADE, Univ Montpellier, CIRAD, IRD, Montpellier, France

\* Correspondence: [colin.kelleher@opw.ie](mailto:colin.kelleher@opw.ie)

**Abstract:** Despite constituting the western-most edge of the population distributions for several native European flora, Ireland has largely been left out of key Europe-wide phylogeographic studies. This is true for birch (*Betula pubescens* Ehrh. and *Betula pendula* Roth.), for which the genetic diversity has yet to be mapped for Ireland. Here, we used eight cpDNA (two RFLP and six SSR) markers to map the genetic diversity of *B. pubescens*, *B. pendula* and putative hybrid individuals sampled from 19 populations spread across most of the island of Ireland. Within Ireland, 11 distinct haplotypes were detected, the most common of which (H1) was also detected in England, Scotland, France and Norway. A moderate level of population structuring ( $G_{ST} = 0.282$ ) was found across Ireland and the genetic diversity of its northern populations was twice that of its southern populations. This indicates that, unlike other native Irish trees, such as oak and alder, post-glacial recolonisation by birch did not begin in the south (*i.e.*, from Iberia). Rather, and in agreement with palynological data, birch most likely migrated in from eastern populations in Britain. Finally, we highlight Irish populations with comparatively unique genetic structure which may be included as part of European genetic conservation networks.

**Keywords:** Birch, *Betula pubescens*, *Betula pendula*, haplotype, Ireland, phylogeography

## 1. Introduction

Two species of birch are native to Ireland, *Betula pubescens* Ehrh. and *Betula pendula* Roth, whereas in Britain, the more cold-tolerant *Betula nana* L. can also be found in Scotland and some upland parts of England. Birch tends to be a pioneer species, either in forest gaps or forest edges or in wetlands and areas of acidic soils [1]. The distinction between *B. pubescens* and *B. pendula* is not clear-cut, but *B. pubescens* is the predominant species in Ireland, with *B. pendula* occurring less frequently [2]. Hybrids are also evident, but a detailed study has yet to be undertaken [3]. Indeed, the occurrence of shared haplotypes in all three species indicates a species complex of hybrids and introgressed individuals rather than distinct taxa [4].

Palynological evidence shows that birch was present in Ireland in localized populations from c. 12,000 years before present (BP) and had completely colonized the island by 9,500 BP [5,6]. For more temperate tree species, recolonisation of northern Europe following the last glacial maximum (LGM) generally involved individuals moving north from refugia populations in the south. The consequence of this is a “southern richness and northern purity” model of genetic diversity, as new colonisation typically involves only a few individuals [7]. For oak, this model holds true [8], and in Ireland oak haplotype diversity is even lower than in Britain [9]. However, for cold tolerant species such as birch, which could survive nearer to the edge of the glacial fronts during the LGM, this model tends not to fit.

First of all, palaeoecological evidence on European birch points to a scenario in which it persisted in refugia at mid latitudes [10,11], unlike oak, which persisted mainly in Iberia, Italy and in the Balkans [8]. Secondly, initial phylogeographic analyses of birch haplotypes

in Europe show a general northwest-southeast divide, with haplotype diversity being particularly high in eastern Europe and Russia [12,13], an observation which is typically taken to be indicative of the presence of refugial populations [14]. A more recent genetic analysis of *Betula* spp. at the nuclear level confirmed that the main refugia during the LGM for *B. pubescens* and *B. pendula* were most likely in Russia and western Siberia [15], a finding which is supported in the fossil record [16].

In Ireland, the direction of pollen influx shows a westward migration from Britain[5], making Ireland the western most point of a recolonisation progression originating as far east as Russia. If this is the case, a founder effect may be observable in the cpDNA diversity, which may mean that Ireland contains only a subset of haplotypes which are observable elsewhere in Europe. Such a scenario would not be surprising given that Ireland has a limited flora and therefore a limited gene pool.

In this paper we aim to provide genetic data to inform conservation initiatives and pre-breeding efforts for *Betula* species in Ireland [17]. In the current study, the following questions were posed: i) What is the origin of Irish birch populations? ii) What level of genetic diversity is there in Irish populations? iii) Is there genetic structure in the populations or between the species?

2. Materials and Methods

2.1 Sampling

Leaves were sampled from putatively native Irish populations (6-14 per location), as identified in the 2008 National Survey of Native Woodlands [2]. These were primarily located on state-owned land and were managed by either local authorities, the National Parks and Wildlife Services or the commercial forestry business, Coillte. Samples were also taken from national breeding programmes (2-20 per collection). These were maintained as Coillte nursery collections, the provenances of which were known (Table 1). Finally, a small number of samples ( $\leq 2$  per location) from wild populations in France, Spain and Norway were analysed as references but were not included in statistical analyses due to their small sample sizes. Individual trees were designated as either *B. pubescens* or *B. pendula* according to morphological characteristics. If an individual had an intermediate morphology it was designated as a ‘putative hybrid’. For natural populations in Ireland, mature trees were sampled which were separated by approximately 15 metres apart. Leaf material (typically three-four young leaves per individual) were immediately placed in silica gel following removal.

**Table 1.** Details of *Betula* spp. sampling locations. Bars in the species column correspond to the frequency of *B. pubescens* (green), *B. pendula* (red) and putative hybrid (blue) individuals in each sampling population.

Site	Location	Provenance	Lat., Long.	Samples/ haplotypes	Gene Diversity $\pm$ S.D.	Species
Alberes, Pyrenees Orientales	France	Wild	42.48, 2.95	1/ NA	0.0000 $\pm$ 0.0000	<div><div></div></div>
Annamarron	Monaghan, Ireland	Wild	53.93, -6.66	11/ 2	0.1640 $\pm$ 0.0334	<div><div></div><div></div></div>
Glenbeg (Coillte)	Scotland	Breeding	56.69, -5.94	13/ 5	0.1982 $\pm$ 0.0210	<div><div></div><div></div></div>
Brownstown Wood	Kilkenny, Ireland	Wild	52.42, -7.03	11/ 2	0.0826 $\pm$ 0.0313	<div><div></div></div>
Carnpark	West-meath, Ireland	Wild	53.42, -7.82	10/ 3	0.1450 $\pm$ 0.0373	<div><div></div></div>
Cork (Coillte)	Cork, Ireland	Breeding	51.94, -8.73	11/ 3	0.2025 $\pm$ 0.0229	<div><div></div></div>

Scariff (Coillte)	Clare, Ireland	Breeding	52.71, -2.76	19/ 4	0.0727 ± 0.0248	
Scottish (Coillte)	Scotland	Breeding	52.96, -8.58	2/ 2	0.0625 ± 0.0313	
Shropshire/ Shrewsbury (Coillte)	England	Breeding	56.49, -4.2	20/ 5	0.1836 ± 0.0222	
Crappagh	Monaghan, Ireland	Wild	54.14, -7.1	12/ 3	0.1985 ± 0.0217	
Cronybyrne	Wicklow, Ireland	Wild	52.97, -6.25	11/ 3	0.1010 ± 0.0333	
Deputy's Pass	Wicklow, Ireland	Wild	52.95, -6.16	10/ 2	0.0675 ± 0.0303	
Derrygoul	Clare, Ireland	Wild	52.91, -8.85	5/ 1	0.0000 ± 0.0000	
Derrysheridan	Meath, Ireland	Wild	53.78, -7.33	14/ 2	0.0546 ± 0.0252	
Killarney	Kerry, Ireland	Wild	51.98, -9.57	7/ 1	0.0000 ± 0.0000	
Lac des Camboux, Lozere,	France	Wild	44.49, 3.58	2/ 1	0.1250 ± 0.0431	
Lough Gill, Slishwood	Sligo, Ireland	Wild	54.24, -8.4	14/ 3	0.2777 ± 0.0331	
Lough Slevin	Westmeath, Ireland	Wild	53.56, -7.32	8/ 4	0.1832 ± 0.0437	
Moods	Kildare, Ireland	Wild	53.27, -6.79	8/ 4	0.2305 ± 0.0426	
Norway	Norway	Wild	60.47, 8.47	2/ 1	0.0625 ± 0.0312	
Ongenstown wood	Meath, Ireland	Wild	53.64, -6.82	8/ 4	0.2653 ± 0.0377	
Rostrevor forest	Down, Ireland	Wild	54.11, -6.18	5/ 2	0.3556 ± 0.0452	
Scragh Bog	Westmeath, Ireland	Wild	53.58, -7.36	10/ 2	0.0225 ± 0.0172	
Spain	Spain	Wild	40.46, 3.75	1/ 1	0.0000 ± 0.0000	
Stormanstown Bog	Louth, Ireland	Wild	53.88, -6.62	12/ 5	0.3212 ± 0.0289	
Sylaun	Galway, Ireland	Wild	53.53, -8.93	13/ 5	0.3113 ± 0.0320	

## 2.2. DNA extraction

For each sample, approximately 200 mg of dried leaf tissue was disrupted for 2 min using a bead mill (30 Hz) and a single 3 mm tungsten carbide bead. Extraction of DNA was performed using a DNeasy Plant mini kit (QIAGEN, cat. no. 69204) according to manufacturer's instructions. DNA was quantified using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, cat. no. ND-2000) and DNA quality was determined by agarose (1.5%) gel electrophoresis and staining with SYBR<sup>TM</sup> Safe DNA Gel Stain (Invitrogen<sup>TM</sup>).

### 2.3 Chloroplast DNA sequencing and polymorphism discovery

To allow for more rapid haplotype identification, the detection of sequence polymorphisms defining PCR-RFLPs was performed using high resolution melting (HRM) for medium-throughput genotyping [18], as done by Cubry *et al.* [19]. To identify candidate polymorphic cpDNA regions around which to design HRM primers, a preliminary PCR-RFLP screen was performed on a discovery set of samples from seven geographically distant Irish sites. Regions targeted were *trn*C-D (CD), *psaA-trn*S (AS) and *trn*T-F (TF) (as per Palme *et al.* [4] and Maliouchenko *et al.* [12]). PCRs were performed in 10 × NH<sub>4</sub> reaction buffer (BIOLINE), 5 units of BIOTAQ DNA polymerase (BIOLINE), 0.3 mM dNTP mix, 3 mM MgCl<sub>2</sub>, 0.2 μM primer mix and ~30 ng gDNA. Targets were amplified using “subcycling” PCR conditions for targets with relatively low GC content according to Guido *et al.* [20]. This included an initial incubation at 95°C for 5 min, 30 cycles of 98°C for 20 sec followed by 4 subcycles (i.e., 30 × 4) of 60°C and 65°C for 15 sec each. This was ended by a final extension for 5 min at 65°C. Amplicons were digested directly using *Taq*I and *Hinf*I (separately), except for AS, which was only digested with *Taq*I [4]. All amplicons were then analysed on an 8% TBE non-denaturing polyacrylamide gel (Novex™, Thermo Fisher Scientific), which was stained as before. Samples which captured the different RFLPs were sent for sequencing at Macrogen Europe (Macrogen Corporation, Amsterdam, The Netherlands). To ensure good quality contigs for each region, multiple internal primers were used (Table 2). Sequences were trimmed and mapped to *in silico* PCR amplicons from the *Betula pubescens* chloroplast reference sequence (NC\_039996) using Geneious Prime® 2021.1.1. *In silico* digests were then performed before aligning variable fragments to identify sequence polymorphisms around which to design suitable HRM primers (Table 2).

**Table 2.** List of primers used for sequencing and HRM analysis.

Primer	Region	Location <sup>a</sup>	Sequence (5'-3')	Reference
trnC-f	<i>trn</i> C-D	1-20 bp	CCAGTTCAAATCTGGGTGTC	[21]
trnC_int_F	<i>trn</i> C-D	708-734 bp	TCCAGGGGTGTATCTACGTATTTTGCT	This work
CD_int_birch_seq1	<i>trn</i> C-D	1,617-1590 bp	CTTACAATTCGAATTCCTAGAATTCTCG	This work
psbMF_Shaw	<i>trn</i> C-D	2,068-2,097bp	AGCAATAAATGCGAGAA-TATTACTTCCAT	[22]
Ag_trnC-D_indel_R	<i>trn</i> C-D	2,237-2,215 bp	TCATGATATTGCTCCGATTCGAT	[19]
CD_int_birch_seq2	<i>trn</i> C-D	3,009-2,985 bp	CTATACGTTTACAGGAGGCTATACA	This work
trnD-M	<i>trn</i> C-D	3,408-3,389 bp	GGGATTGTAGTTCAATTGGT	[21]
psaA-f	<i>psaA-trn</i> S	1-22 bp	ACTTCTGGTTCCGGCGAACGAA	[21]
Birch_AS_indel1_b-F	<i>psaA-trn</i> S	892-872 bp	TGGTTGAAGATCACAAGGCGT	This work
Birch_AS_SNP3_R	<i>psaA-trn</i> S	1,076-1,095 bp	CGGCTCAGCAGTCAATTCTT	This work
Birch_AS_SNP3_F	<i>psaA-trn</i> S	1,275-1,252 bp	GCTTATTCTTCTAAAGGTGGGAA	This work
Birch_AS_SNP2_F	<i>psaA-trn</i> S	1,845-1,826 bp	AGGGCACTAGAACGAAACCC	This work
Birch_AS_SNP1_F	<i>psaA-trn</i> S	2,292-2,272 bp	TCCTGGAAATTAAGGGGTGCT	This work
AS_int_birch_seq_1	<i>psaA-trn</i> S	2,840-2,816 bp	CCCAGATCTCGGATAAATGGAAATT	This work
Tab_a	<i>trn</i> T-F	1-20 bp	CATTACAAATGCGATGCTCT	[23]
TF11_Rv	<i>trn</i> T-F	633-610 bp	GTGTAATTTGAGATACTCGAACGG	This work
Tab_b	<i>trn</i> T-F	968-949 bp	TCTACCGATTTCGCCATATC	[23]
trnL(UAA)h	<i>trn</i> T-F	1,155-1,134 bp	CCATTGAGTCTCTGCACCTATC	[24]
Tab_d	<i>trn</i> T-F	1,399-1,380 bp	GGGGATAGAGGGACTTGAAC	[23]
Tab_f	<i>trn</i> T-F	1,855-1,836 bp	ATTTGAACTGGTGACACGAG	[23]
Birch_AS_indel1_F <sup>b</sup>	<i>psaA-trn</i> S	885-866 bp	AGATCACAAGGCGTTTCGAA	This work
Birch_AS_indel1_R <sup>b</sup>	<i>psaA-trn</i> S	693-712 bp	TGGGGACAACAAACAAACT	This work

Birch_CD_indel1_b-F <sup>b</sup>	trnC-D	2,575-2,595 bp	AAGGAGAGTCCGGGTATAAAA	This work
Birch_CD_indel1_b-R <sup>b</sup>	trnC-D	2,746-2,725 bp	TCCAAAGAACAAAGAAATGGGA	This work

<sup>a</sup> Location from 5' end of marker region  
<sup>b</sup> Primers used in HRM analysis

2.4 HRM experiments

HRM analysis was performed on the whole sample set using a QIAGEN Rotor-Gene Q 2-plex HRM platform (QIAGEN GmbH, Germany). Each PCR was performed in a final volume of 15 µl comprising 2 × Type-it HRM mix (QIAGEN, cat. no. 206546), 0.8 µM primer mix and ~5 ng gDNA. PCR conditions were as follows: 95°C for 5 mins, 40 cycles of 95°C for 10 sec, 55°C for 30 sec and 72°C for 10 sec. For HRM, fluorescence was continually monitored at a ramp rate of 0.1°C for 2 sec between 65°C and 80°C. Haplotypes were assigned manually using the Rotor-Gene Q – Pure Detection software (v2.3.5). Grouping consistency was verified by comparison with PCR-RFLP gels and sequence alignments from the discovery sample set.

2.5 Chloroplast microsatellites

Six chloroplast microsatellite markers were used in this study; ccmp2, ccmp4, ccmp5, ccmp6, ccmp7 and ccmp10, as used by Maliouchenko *et al.* [12]. These were amplified using the universal primers designed by Weising and Gardner [25]. PCRs were performed according to Maliouchenko *et al.* [12]. PCR products were analysed on an ABI 310 Genetic Analyser (Applied Biosciences, Lincoln, USA) using a 1:50 dilution of each PCR product. Allele sizes were called using a ROX 500S size standard in GeneMarker v2.4.0 software (SoftGenetics LLC, State College, PA16803, USA).

2.6 Data analysis

Raw HRM and microsatellite data were combined prior to data analysis, as done by others [12,26]. Species were analysed together unless specified otherwise. Haplotype calling and frequency estimates per sampling site were calculated using a custom R script. These were mapped using QGIS (v3.18). General data handling, visualisation and statistical analyses were performed using the R package adegenet (v2.1.3) [27,28]. The packages ade4 (v.1.7-16) [29], hierfstat (v.0.5-7) [30], mmod (v.1.3.3) [31] and poppr (v.2.9.1) [32,33] were also used to estimate population differentiation and diversity statistics as well as to bootstrap samples for tests of statistical significance. Before statistical analysis, sites which were sampled less than five times were removed; this resulted in the removal of the Norwegian, Spanish, French and “Scottish (Coillte)” samples. In addition, individuals with more than two null alleles (alleles that did not PCR-amplify) were removed prior to analyses.

Typological differences between sampling sites were investigated by submitting an allele contingency table to a factorial correspondence analysis using adegenet. Results were plotted using plot3D (v.1.3). Global population differentiation statistics were estimated using mmod, with estimates of significance being provided by a 95% confidence interval (CI) computed on 1,000 bootstrap permutations of the dataset. For pairwise comparisons, Hedrick’s  $G'_{ST}$  was calculated, again using mmod.  $G'_{ST}$  is a standardised version of Nei’s  $G_{ST}$ , which is itself an estimate of the fixation index,  $F_{ST}$  for multiallelic markers [34].  $G'_{ST}$  considers the maximum theoretical  $G_{ST}$  based on observed heterozygosity for a given marker, thereby dealing with biases towards low estimates for highly variable loci [35]. For easier interpretation of the pairwise estimates, each 95% CI was converted to a  $p$  value according to the method outlined by Altman and Bland [36] for deriving a  $p$  value when a CI is given for an estimate of difference in effect.

Analysis of molecular variance (AMOVA) tests were performed using the adegenet, poppr and ade4 packages for R. Specifically, the data in adegenet ("genind") format was passed through the poppr wrapper of the ade4 AMOVA function. AMOVA significance was calculated on 1,000 permutations of the data. To test for isolation by distance (IBD), samples containing null alleles were removed before using the ade4 Mantel test function to test for significant correlations between Slatkin's linearised pairwise  $F_{ST}$  [37] and geographic distances – as suggested by Rousset [38] – following 1,000 permutations of the data.

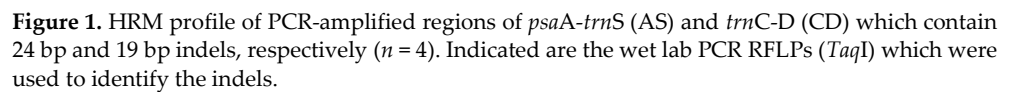
To estimate the extent to which all haplotypes were captured by our sampling, 10,000 permutations of our haplotype frequency distribution were used to calculate haplotype accumulation curves using the iterative extrapolation simulation algorithm, HACSIm, available through the R package HACSIm (v.1.0.5) [39]. HACSIm calculates the number of samples needed to recover 95% of all haplotypes.

### 3. Results

#### 3.1 Chloroplast DNA variation

A total of 240 birch individuals were sampled across 26 sites, 19 of which were in Ireland (Table 1). Morphologically, 201 of these were identified as *Betula pubescens*, 31 as *Betula pendula* and eight as putative hybrids. Of the 26 sampling populations, 21 were wild and occurred in putative native woodland. The remainder were sourced from Coillte national breeding programmes, although their provenances were known. Haplotypes were identified based on variation at two PCR-RFLP (HRM) (Figure 1) and six microsatellite loci, as used in previous phylogeographic works on European birch [4,12,13]. The former were selected from a preliminary PCR-RFLP screen of a discovery set of Irish samples across three variable loci; *trnC-D*, *psaA-trnS* and *trnT-F* (Table 2). Sequencing these samples revealed that the RFLP variation for all three lies primarily in indels of 19, 24 and 10 bp in length, respectively. For faster throughput haplotyping, primers flanking each region were designed and successfully tested in HRM experiments (Table 2). The exception, however, was the *trnT-F* indel, for which we were unable to design suitable primers due to the AT-richness of the flanking sequences. Even without screening at this region, 16 distinct haplotypes could nonetheless be identified, six of which occurred five times or more in the whole sample set (Table 3). The third (H3) and fourth (H4) most abundant haplotypes only occurred in Ireland, roughly spanning from Cronybyrne (Co. Wicklow) in the east to Lough Gill, Slishwood (Co. Sligo) in the northwest (Figure 2). The fifth most abundant haplotype (H5) was only identified in the English and Scottish sample set. A minimum spanning network (MSN) based on Euclidean distances between haplotypes indicates that the haplotypes are relatively closely related and that the most abundant, H1 is also the most geographically widespread (Figure 3). Based on the distribution pattern, H1 is most likely to be equivalent to Haplotype A from Palmé *et al.* [13], which is the dominant haplotype in northern European populations.

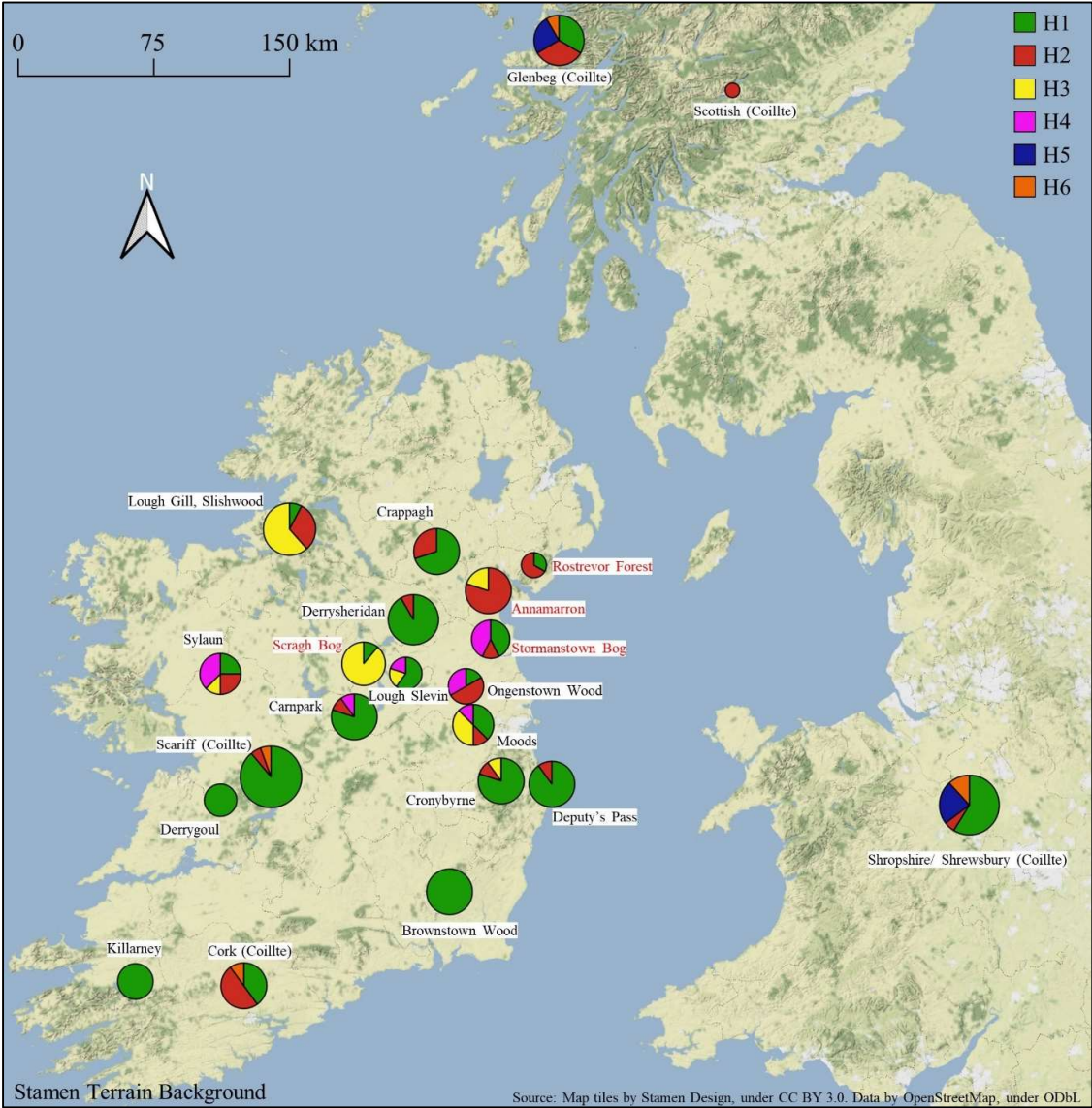




	<i>trnC</i> - D	<i>psaA</i> - <i>trnS</i>	<i>ccmp</i> 5	<i>ccmp</i> 10	<i>ccmp</i> 6	<i>ccmp</i> 2	<i>ccmp</i> 4	<i>ccmp</i> 7	Found in:	Frequency <sup>a</sup>			
										All	Pb	Pn	Un
H1	1	1	105	118	100	205	117	147	Ireland, England, Scotland, France, Norway	115	94	20	1
H2	2	2	106	118	100	205	117	147	Ireland, England, Scotland	40	35	3	2
H3	1	1	105	118	100	205	118	147	Ireland	24	24	0	0
H4	2	2	108	118	100	205	117	154	Ireland	11	11	0	0
H5	2	1	106	118	100	205	117	147	England, Scotland	7	7	0	0
H6	2	1	105	118	100	205	118	147	Ireland, England, Scotland	5	2	3	0
H7	1	1	105	118	100	205	118	148	Ireland	4	4	0	0
H8	2	2	106	118	100	205	118	147	Ireland	2	2	0	0
H9	1	1	105	118	100	205	119	147	Ireland	2	2	0	0
H10	1	1	104	118	100	205	117	147	Ireland	1	0	1	0
H11	2	2	106	118	100	205	119	147	Scotland	1	0	1	0
H12	2	2	106	118	100	205	116	147	England	1	0	1	0
H13	2	2	105	118	100	205	117	154	Spain	1	0	0	1
H14	1	1	105	118	100	205	119	148	Ireland	1	1	0	0
H15	1	1	105	118	98	205	118	147	Ireland	1	1	0	0
H16	1	2	105	118	100	205	117	147	Scotland	1	1	0	0

<sup>a</sup>Of the 240 individuals analysed, 23 could not be haplotyped due to a lack of PCR amplification at one or more loci.

Statistical analysis of population differentiation and genetic diversity was only performed on populations containing five or more sampled individuals. Additionally, individuals with more than two null alleles were removed prior to analysis; in effect, this meant that only populations from Ireland, England and Scotland were analysed ( $n = 228$ ). Unlike total diversity, intra-population diversity was low and not statistically significant, whereas a global  $G'_{ST}$  estimate of 0.336 indicated a moderate level of population structure (Table 4). This estimate was negligibly different when only the Irish populations were considered ( $G_{ST} = 0.282$ ,  $G'_{ST} = 0.353$ ). In agreement with this, results of a nested AMOVA revealed that 19.16% of variation was from differences between sampling populations (Table 5).



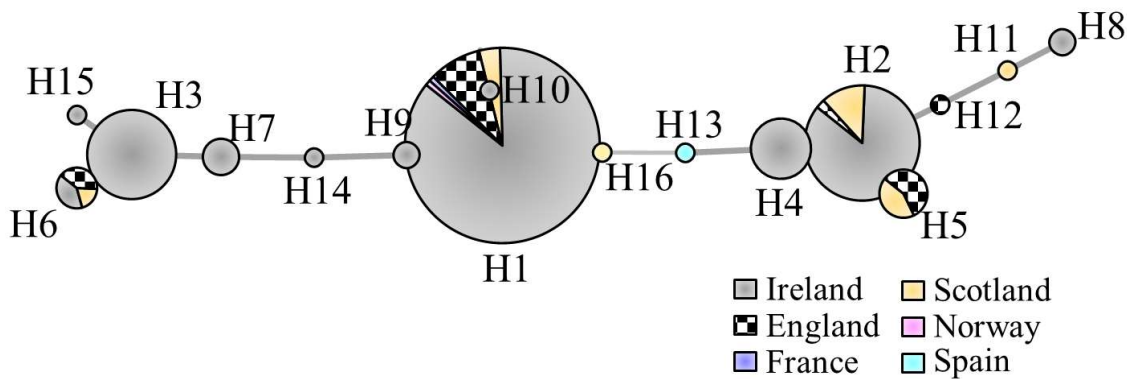
**Figure 1.** Geographic distribution of the six main haplotypes (those which occurred five or more times) identified in this study. Pie chart size corresponds to the number of each haplotype which could be identified in a particular population. Names highlighted in red are sites which appeared more genetically differentiated relative to each other and to the other populations (Figure 3 and Figure 4).

Effectively all population structure was attributable to differentiation between *B. pubescens* populations (Table 4). This is likely because there were considerably fewer *B. pendula* individuals, which meant that only two populations of *B. pendula* could be compared



after removing those which possessed less than five individuals. In agreement with this, the nested AMOVA results showed that only 1.58% of variation could be explained at the species level, which was not statistically significant (Table 5).

To investigate which sites, if any, possessed unique allelic variation, population typology was investigated using a factorial correspondence analysis (FCA) (Figure 4). Most of the allelic variation (77.05%) was explained by the first three FCA axes. While most populations were not clearly distinct from one another, there are exceptions, such as Ros-trevor Forest (Co. Down), Annamarron (Co. Monaghan), Stormanstown Bog (Co. Louth) and Scragh Bog (Co. Westmeath), all of which appeared distinct from both each other and from the other populations. These sites also stood out in a pairwise  $G'_{ST}$  comparison of



**Figure 2.** MSN based on Euclidean distances between haplotypes. Node size corresponds to the relative abundance of each haplo-type.

**Table 3.** Population differentiation statistics estimated across all haplotypes in the Irish, English and Scottish sampling populations as well as separately across *B. pubescens*, *B. pendula* and putative hybrids. Statistics include estimates of intra-population diversity ( $h_s$ ), total diversity ( $h_T$ ), diversity which apporitions between populations ( $G_{ST}$ ) and  $G_{ST}$  adjusted for the theoretical maximum based on mean heterozygosity ( $G'_{ST}$ ). Bootstrapped 95% confidence intervals ( $n = 1,000$  permutations) are provided in brackets. Bold val-ues indicate statistical significance ( $p \leq 0.05$ ).

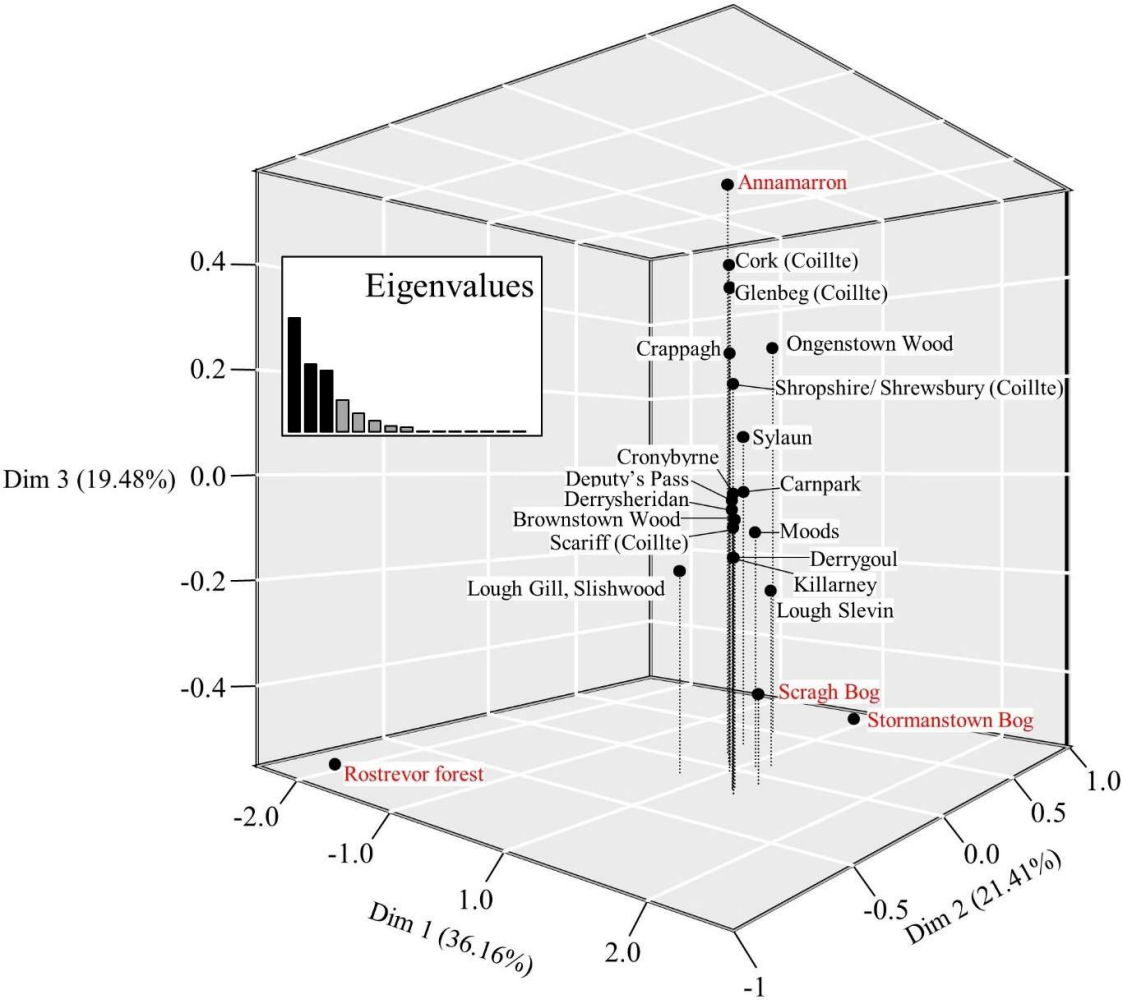
	$h_s$	$h_T$	$G_{ST}$	$G'_{ST}$
All species (n = 236/228 <sup>a</sup> )	0.173	0.236	<b>0.268 [0.214, 0.324]</b>	<b>0.336 [0.276, 0.397]</b>
<i>B. pubescens</i> (n = 193/192 <sup>a</sup> )	0.160	0.217	<b>0.261 [0.205, 0.320]</b>	<b>0.323 [0.259, 0.387]</b>
<i>B. pendula</i> (n = 29/20 <sup>a</sup> )	0.136	0.144	0.056 [-0.042, 0.198]	0.122 [-0.071, 0.356]
Putative hy- brid (n = 6/0 <sup>a</sup> )	NA	NA	NA	NA

<sup>a</sup>Sample size before/after removing populations comprising less than five individuals

**Table 4. (Top)** Results of a nested AMOVA between species nested within all populations (sites occurring in Ireland and England). **(Bottom)** Nested AMOVA between Irish populations nested within either a north or south (N-S) location, or an east or west (E-W) location. *P* values were computed across 1,000 permutations of the data.

All sites	d.f.	SSD	Variance	Variance (%)	<i>p</i> value
Between sites	20	98.56	0.33	19.16	0.015
Species within sites	8	11.94	0.03	1.58	0.301
Within all	183	251.07	1.37	79.25	<0.001
Total <sup>a</sup>	211	361.57	1.73	100.00	

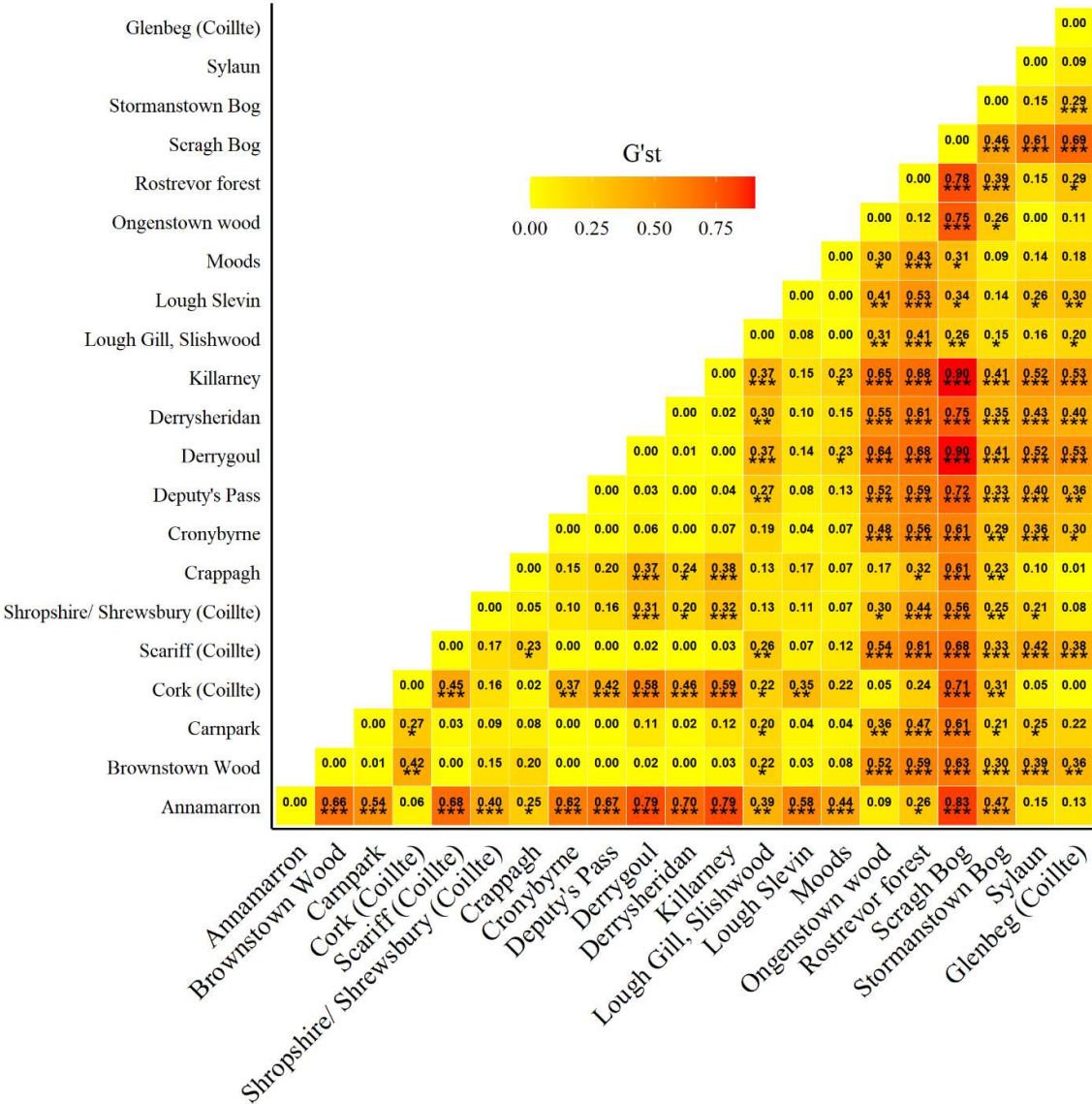
Irish only	d.f.	SSD	Variance	Variance (%)	<i>p</i> value					
	N-S	E-W	N-S	E-W	N-S	E-W	N-S	E-W	N-S	E-W
Between	1	1	15.82	0.65	0.13	0.00	7.03	0.00	0.034	1.000
Sites within	17	17	74.50	89.66	0.33	0.42	18.18	23.92	<0.001	<0.001
Within all	162	162	216.73	216.73	1.34	1.34	74.79	76.08	<0.001	<0.001
Total <sup>a</sup>	180	180	307.04	307.04	1.79	1.76	100.00	100.00		



**Figure 3.** Results of an FCA performed on an allele contingency table at the population level. Plotted are the first three eigenvalues (subset, black), which account for 77.05% of the allelic variation between sample populations. Highlighted in red are populations which are indicated to be particularly differentiated relative to each other and to the other populations.

populations (Figure 5), with the highest  $G'_{ST}$  values being for Scragh Bog, a site which had the highest frequency of the Irish-specific H3 haplotype (Figure 2).

The apparently lower levels of haplotype diversity among the more southern sampling populations in Ireland prompted us to test whether there was any statistical backing for either a north-south or an east-west effect. For this, populations south of Moods (a site nearest to the mid latitude point in Ireland) were deemed to be southern populations, whereas populations west of (and including) Carnpark (near the longitudinal midpoint) were deemed to be western populations. When populations were nested within either a northern or a southern location, a significant level (7.03%) of the variation could be explained, whereas no variation could be explained by an east-west division (Table 5). Indeed, genetic diversity was more than twice as high for northern ( $0.2806 \pm 0.0133$ ) compared to southern ( $0.1121 \pm 0.0154$ ) populations. To test for evidence that this geographic variation could be caused by IBD, a Mantel test was performed to test for a correlation between genetic and geographic distances. Neither before ( $r^2 = 0.069$ ,  $p$  value = 0.347) nor after ( $r^2 = 0.004$ ,  $p$  value = 0.539) removing the English and Scottish populations could a significant effect of IBD be observed. This latter result suggests that the north-south difference in Ireland is not attributable to IBD.



**Figure 4.** Pairwise  $G'_{ST}$  comparisons between sampling populations. Values which are significantly different from 0 are marked with asterisks (\* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ ), as determined from 1,000 random permutations of the data.

3.2 Sufficiency of haplotype capture in a European context

Using the HACSim algorithm developed by Phillips *et al.* [39] for estimating the sufficiency of haplotype sample sizes, 10,000 permutations of the frequency distribution of all 16 haplotypes were used to extrapolate a haplotype accumulation curve (Figure A1, Appendix A). From this, it was inferred that a total of 452 (95% CI: 449.74 - 454.26) individuals would need to have been sampled in order to have sufficiently captured 95% of the actually occurring haplotypes. Rather, it was estimated that 84% were captured instead, which suggests that up to three additional rare haplotypes may not have been identified at  $n = 217$  (*i.e.*, individuals with no null alleles).

According to Maliouchenko *et al.* [12], who used the same markers employed here (although including *trnT-F*) to identify 66 haplotypes in *B. pubescens* and *B. pendula* sampled across Western Europe (excluding Ireland) and Russia, at least 50 haplotypes ought to be identifiable across the regions represented in our sampling data. When we entered

the adjusted the number of expected haplotypes into the HACSim algorithm and assigned sampling probability frequencies of zero to the haplotypes which we did not identify (*i.e.*,  $50 - 16 = 34$ ), only 27% of all possible haplotypes were identified. Realistically however, if only the samples from Ireland and Britain are considered, then at  $n = 214$  we can assume that our sampling reflects 67% of the potential haplotypes in these regions. This estimation is based on an assumption by Maliouchenko *et al.* [12] that there are 20 observable haplotypes in Britain, that they can also be found on the island of Ireland, and that they should include those which we identify here as Irish-specific haplotypes (H3 and H4). This estimate still suggests that most haplotypes for Ireland were captured in our study.

#### 4. Discussion

Molecular studies are key elements in characterising populations for conservation and for monitoring sustainable forest management [40]. The cpDNA markers employed here have been widely used in previous studies to effectively map out the phylogeographies of several important European tree species, including birch [8,12,19,26,41]. However, populations from Ireland have rarely, if at all, been included in most of these works. With added studies, greater resolution of Europe-wide phylogeographic patterns can be achieved and the data can be used to select populations for use in conservation networks [42]. In the case of Ireland, there is an added urgency that the genetic diversity of its native tree populations be mapped given that it has one of the lowest levels of native forestry cover in Europe, which stands currently at only approximately 2% [43].

Native birch has been designated as “high priority” for conservation in Ireland [44]. In this study, 11 Irish haplotypes could be identified (Table 3). Based on abundance, the two most frequent (H1 and H2) may correspond to haplotypes “1” and “26”, respectively, which were identified in Britain by Maliouchenko *et al.* [12]. Overall, of the most abundant haplotypes ( $\geq 5$ ; Figure 2), two (H3 and H4) were not identified outside of Ireland. However, as the size and number of the non-Irish sampling populations was considerably smaller, it cannot be ruled out that these haplotypes were simply missed. Nonetheless, it is clear that the genetic diversity of Irish birch is relatively high compared to what has been reported for other native Irish trees such as oak (*Quercus petraea* and *Quercus robur*), alder (*Alnus glutinosa*) and ash (*Fraxinus excelsior*), for which fewer haplotypes appear to be present [9,19,41]. For Irish oak, however, diversity has been estimated based on variation at only two (*trnD*-T and *trnT*-F) cpDNA regions [9], meaning that greater diversity may be revealed with the addition of more markers.

We are confident that our estimate of the genetic diversity of Irish birch is, if anything, slightly underestimated, as suggested by the extrapolation of haplotype accumulation curves. With all sampling populations included, it was estimated that 84% of all haplotypes were captured. Intuitively, this seems like a major overestimation given that regions other than Ireland (England, Scotland, France, Norway and Spain) were represented in our data. Undoubtedly, this is attributable to the very small number of samples from these regions ( $n = 33, 2, 2, 2$  and  $1$ , respectively). When we factored in the number of haplotypes observed in these regions by Maliouchenko *et al.* [12], then haplotype capture dropped to only 27%. Inputting only the number of haplotypes which Maliouchenko *et al.* [12] observed in Britain, then this figure increased to 67% when considering only the British and Irish populations. However, it is worth mentioning that based on our markers alone, 15 haplotypes were identified across Ireland and Britain. The figure of 20 from Maliouchenko *et al.* [12] however, did not come directly from extensive sampling across Britain ( $n = 36$ ), but rather from haplotypes identified in other regions which were presumed to be present because they were shown to be closely related to their actually observed British haplotypes in a MSN. This might suggest that our figure of 15 is a more accurate approximation. If so, then our sampling at  $n = 214$  for Ireland and Britain ought to have captured 85% of all haplotypes. Focusing only on the Irish samples ( $n = 181$ ) and the associated 11 haplotypes which we detected, then 87% of all possible Irish haplotypes were identified. Therefore, at the depth of the eight cpDNA markers used here, we are



reasonably confident that a good representation of the genetic diversity in Ireland in the context of Europe-wide diversity has been revealed.

Owing to its lighter, wind-dispersed seeds, the genetic structuring and differentiation in birch compared to oak was expected to be lower [26]. Indeed, even with the English and Scottish populations included, cpDNA differentiation between sampling populations was lower for birch ( $G_{ST} = 0.268$ ) than for Irish oak ( $G_{ST} = 0.730$ ) [9]. This is more in line with other wind-dispersed species such as alder ( $G_{ST} = 0.283$ ) [19] and goat willow (*Salix caprea*;  $G_{ST} = 0.38$ ) [45]. At the species level, significant population structuring was only detected for *B. pubescens* (Table 4). We attributed this to an insufficient number of *B. pendula* populations from which diversity estimates could be calculated, which itself reflected the fact that this species is significantly less common than *B. pubescens* in Ireland [2].

Both species are well known to display high levels of hybridisation and haplotype sharing, so much so that interspecific cpDNA variation tends to be lower within the same forest compared to intraspecific variation between different forests [4,12]. Here, of the most common haplotypes, H1, H2 and H6 could be detected in both species. The occurrence of H3 and H4 in *B. pubescens* only is likely explained by their increased frequency in more northern populations where *B. pendula*, by contrast, becomes increasingly less frequent. As is the case for oak at the European level [8], the rarer haplotypes tended to be restricted to a single species.

It was not expected that the selected cpDNA markers would differentiate between species, as cpDNA variation in birch had already been extensively demonstrated to show no clear species delimitation [12,46,47]. For the most common shared haplotypes, this has been explained by incomplete lineage sorting, whereas for more rarer haplotypes it has been argued that interspecific backcrossing and sympatric introgression are responsible [46,47]. This is distinct from convergent evolution, which is not thought to play a role given the slow mutation rate of cpDNA and the asymmetric sharing of chloroplast alleles between the species [46]. Conversely, using nuclear DNA markers, strong species delimitation has been observed [15,46,48]. This has been explained by a model which states that higher gene flow (*i.e.*, through pollen) within species will lead to better differentiation between species [49,50]. For species delimitation in Ireland, the well-validated nSSR markers for birch originally developed by Kulju *et al.* [51] ought to be tested.

The results of this work help to answer questions relating to the origins and phylogeographic patterns of birch in Ireland in the context of Europe and post-glacial recolonisation following the LGM. Studies to date have shown a mixture of origins for tree populations in Ireland. Oak and ash populations have been shown to have originated in the Iberian Peninsula [9,41], whereas palynological and genetic data for alder indicate a two-pronged re-colonisation from the Iberian Peninsula and the Carpathians [19]. For birch in Ireland, a significant level of variation could be explained by a north-south divide. However, diversity was significantly greater in the more northern populations. This is congruent with earlier Europe-wide works which have demonstrated that birch did not recolonise from the south, as if this were the case, we would instead expect declining diversity at higher latitudes. Therefore, it is probably more realistic to account for the northern haplotype richness in Ireland as being a result of migration from Britain, in which case Ireland may be part of a western leading edge for birch in Europe. Potentially, the absence of the UK-specific haplotype (H5) in Ireland could mark a declining westward genetic diversity as might be expected as part of a founder effect. Within Ireland however, this effect was not observed as there was no significant east-west difference.

Another possible explanation for the higher northern diversity could be remnant *B. nana* haplotypes from ancient introgression events. *B. nana* is absent in Ireland today, but the macrofossil record reveals that it was present early following the LGM [6]. Moreover, pollen records show that *B. nana* and tree birch (presumably *B. pubescens*) likely co-occurred during this period [52]. In Scotland, extant *B. nana* have been shown to share more nuclear alleles with *B. pubescens* than with *B. pendula* [48]. This, in conjunction with fossil evidence, led Wang *et al.* [48] to conclude that as *B. nana* moved northwards post-LGM

with climate warming, “a footprint of introgressed genes in the genome of [advancing] *B. pubescens*” was left behind [48]. Indeed, triploid hybrids are readily observed where the species continue to co-occur ([53]. Gene flow from *B. nana* into *B. pubescens* (but not the other way around) increases with latitude. Interestingly, this has led to a scenario being suggested in which pollen swamping of *B. nana* by *B. pubescens* creates hybrids which then backcross with the latter, resulting in haplotype capture from *B. nana* [46]. This agrees with findings from Currat *et al.* [49], in which it was demonstrated that introgression is almost always unidirectional from the local into the invading species. Therefore, the novel haplotype variation in the more northern Irish populations may be a genetic legacy of now-extinct Irish *B. nana*. Investigating whether these haplotypes occur in Scottish populations of *B. nana* could be useful in testing the validity of this hypothesis.

A main aim of this work was to select conservation units for birch in Ireland. Indeed, selection at the population level (rather than species) may be more sensible and practical for conserving FGR given that the haplotypes were distributed geographically and not interspecifically. Towards this end, we suggest that populations sampled in the more northern and north-eastern areas (Figure 2), for example Scragh Bog, be prioritised as both FCA and pairwise  $G'_{ST}$  analysis suggest these to be genetically the most differentiated (Figure 4-5). We recommend that these sites be prioritised for conservation as they may be possible sites of local adaptation and potentially contain unique allele combinations.

## 5. Conclusions

By mapping the genetic diversity of birch in Ireland, this work fills a gap in the phylogeographic structure of birch in Europe. Contrary to expectations based on other native Irish trees, haplotype richness in Irish birch is comparatively high. Building on previous work by Maliouchenko *et al.* [12], which estimated that up to 20 haplotypes may be observable in Britain, we empirically show that at least 11 of these can be found in Ireland. In addition, the strikingly lower genetic diversity of southern populations supports the hypothesis that post-glacial recolonisation did not involve migration from the Iberian Peninsula. Instead, and in agreement with pollen data, an eastern migration route from Britain is most likely. Moreover, based on findings from more recent works, we formulate a hypothesis which suggests that the greater northern diversity is, in part, attributable to historic sympatric introgression events between *B. pubescens* and now-extinct Irish *B. nana*. Finally, we suggest populations which may be particularly worthy of selection as part of European conservation networks for birch in Ireland.

**Author Contributions:** Conceptualisation, S.B., P.C. and C.K.; Methodology, S.B., P.C. and C.K.; Software, S.B. and E.F.; Validation, C.K.; Formal Analysis, S.B.; Investigation, P.C.; Resource, C.K.; Data Curation, S.B.; Writing – Original Draft Preparation, S.B. and C.K.; Writing – Review and Editing, P.C. and C.K.; Visualisation, S.B. and E.F.; Supervision, C.K.; Project Administration, C.K.; Funding Acquisition, C.K.

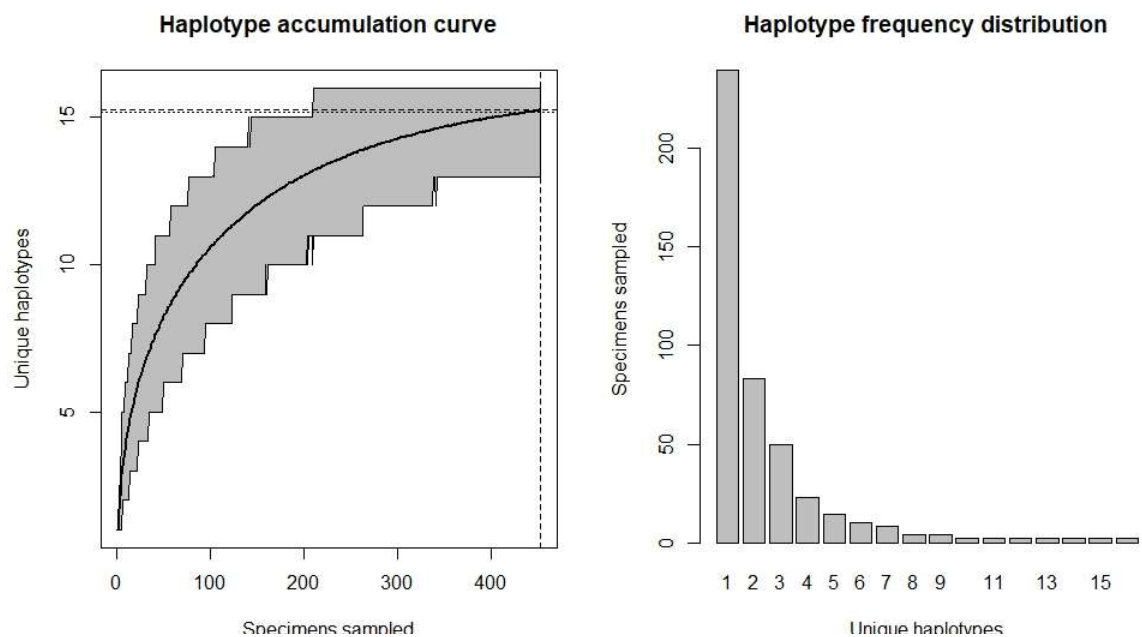
**Funding:** Funding for the FORGEN and GeneNet projects was used for this work. This funding was part of the CoFoRD Forestry Research programme of the Department of Agriculture, Food and The Marine (DAFM).

**Data Availability Statement:** GenBank accession numbers for the different haplotype sequences are as follows: (AS1) MZ293192, MZ293194, (AS2) MZ293193, MZ293195, (CD1) MZ293196, MZ293197, MZ293199, (CD2) MZ293198, (TF1) MZ293201, MZ293202, MZ293203 and (TF2) MZ293200.

**Acknowledgments:** We thank all landowners for access to sampling sites.

**Conflicts of Interest:** The authors claim no conflicts of interests between the work reported here and the funding provided for it.

## Appendix A



**Figure A1.** Output of a HACSIm simulation of haplotype sampling based on the 16 haplotypes identified in this study.

## References

1. Atkinson, M.D. *Betula Pendula* Roth (*B. Verrucosa* Ehrh.) and *B. Pubescens* Ehrh. *J. Ecol.* **1992**, *80*, 837–870.
2. Perrin, P.; Martin, J.; Barron, S.; O'Neill, F.; McNutt, K.; Delaney, A. *National Survey of Native Woodlands 2003–2008*; Dublin, Ireland, 2008; Vol. 1, p. 177.
3. Praeger, R.L. Hybrids in the Irish Flora: A Tentative List. *Proc. R. Ir. Acad. [B]* **1951**, *54*, 1–14.
4. Palme, A.E.; Su, Q.; Palsson, S.; Lascoux, M. Extensive Sharing of Chloroplast Haplotypes among European Birches Indicates Hybridization among *Betula Pendula*, *B. Pubescens* and *B. Nana*. *Mol Ecol* **2004**, *13*, 167–78.
5. Birks, H.J.B. Holocene Isochrone Maps and Patterns of Tree-Spreading in the British Isles. *J. Biogeogr.* **1989**, *16*, 503–540.
6. Mitchell, F.J.G. WHERE DID IRELAND'S TREES COME FROM? *Biol. Environ. Proc. R. Ir. Acad.* **2006**, *106B*, 251–259.
7. Hewitt, G.M. Genetic Consequences of Climatic Oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* **2004**, *359*, 183–95; discussion 195.
8. Petit, R.J.; Csakl, U.M.; Bordács, S.; Burg, K.; Coart, E.; Cottrell, J.; van Dam, B.; Deans, J.D.; Dumolin-Lapègue, S.; Fineschi, S.; et al. Chloroplast DNA Variation in European White Oaks: Phylogeography and Patterns of Diversity Based on Data from over 2600 Populations. *For. Ecol. Manag.* **2002**, *156*, 5–26.
9. Kelleher, C.T.; Hodgkinson, T.R.; Kelly, D.L.; Douglas, G.C. Characterisation of Chloroplast DNA Haplotypes to Reveal the Provenance and Genetic Structure of Oaks in Ireland. *For. Ecol. Manag.* **2004**, *189*, 123–131.
10. Bennett, K.D.; Tzedakis, P.C.; Willis, K.J. Quaternary Refugia of North European Trees. *J. Biogeogr.* **1991**, *18*, 103–115.
11. Lascoux, M.; Palmé, A.E.; Cheddadi, R.; Latta, R.G. Impact of Ice Ages on the Genetic Structure of Trees and Shrubs. *Philos Trans R Soc Lond B Biol Sci* **2004**, *359*, 197–207.

12. Maliouchenko, O.; Palmé, A.E.; Buonamici, A.; Vendramin, G.G.; Lascoux, M. Comparative Phylogeography and Population Structure of European *Betula* Species, with Particular Focus on *B. Pendula* and *B. Pubescens*. *J. Biogeogr.* **2007**, *34*, 1601–1610.
13. Palmé, A.E.; Su, Q.; Rautenberg, A.; Manni, F.; Lascoux, M. Postglacial Recolonization and CpDNA Variation of Silver Birch, *Betula Pendula*. *Mol Ecol* **2003**, *12*, 201–12.
14. Hewitt, G.M. Some Genetic Consequences of Ice Ages, and Their Role in Divergence and Speciation. *Biol. J. Linn. Soc.* **1996**, *58*, 247–276.
15. Tsuda, Y.; Semerikov, V.; Sebastiani, F.; Vendramin, G.G.; Lascoux, M. Multispecies Genetic Structure and Hybridization in the *Betula* Genus across Eurasia. *Mol Ecol* **2017**, *26*, 589–605.
16. Binney, H.A.; Willis, K.J.; Edwards, M.E.; Bhagwat, S.A.; Anderson, P.M.; Andreev, A.A.; Blaauw, M.; Damblon, F.; Haesaerts, P.; Kienast, F.; et al. The Distribution of Late-Quaternary Woody Taxa in Northern Eurasia: Evidence from a New Macrofossil Database. *Quat. Sci. Rev.* **2009**, *28*, 2445–2464.
17. Kelleher, C.T. A National Forest Tree Gene Conservation Strategy and Action Plan for Ireland. *Ir. For. J.* **2020**, *77*, 7–32.
18. Dang, X.D.; Kelleher, C.T.; Howard-Williams, E.; Meade, C.V. Rapid Identification of Chloroplast Haplotypes Using High Resolution Melting Analysis. *Mol Ecol Resour* **2012**, *12*, 894–908.
19. Cubry, P.; Gallagher, E.; O'Connor, E.; Kelleher, C.T. Phylogeography and Population Genetics of Black Alder (*Alnus Glutinosa* (L.) Gaertn.) in Ireland: Putting It in a European Context. *Tree Genet. Genomes* **2015**, *11*, 99.
20. Guido, N.; Starostina, E.; Leake, D.; Saaem, I. Improved PCR Amplification of Broad Spectrum GC DNA Templates. *PLoS One* **2016**, *11*, e0156478.
21. Demesure, B.; Sodzi, N.; Petit, R.J. A Set of Universal Primers for Amplification of Polymorphic Non-Coding Regions of Mitochondrial and Chloroplast DNA in Plants. *Mol Ecol* **1995**, *4*, 129–31.
22. Shaw, J.; Lickey, E.B.; Beck, J.T.; Farmer, S.B.; Liu, W.; Miller, J.; Siripun, K.C.; Winder, C.T.; Schilling, E.E.; Small, R.L. The Tortoise and the Hare II: Relative Utility of 21 Noncoding Chloroplast DNA Sequences for Phylogenetic Analysis. *Am J Bot* **2005**, *92*, 142–66.
23. Taberlet, P.; Gielly, L.; Pautou, G.; Bouvet, J. Universal Primers for Amplification of Three Non-Coding Regions of Chloroplast DNA. *Plant Mol Biol* **1991**, *17*, 1105–9.
24. Taberlet, P.; Coissac, E.; Pompanon, F.; Gielly, L.; Miquel, C.; Valentini, A.; Vermet, T.; Corthier, G.; Brochmann, C.; Willerslev, E. Power and Limitations of the Chloroplast TrnL (UAA) Intron for Plant DNA Barcoding. *Nucleic Acids Res* **2007**, *35*, e14, doi:10.1093/nar/gkl938.
25. Weising, K.; Gardner, R.C. A Set of Conserved PCR Primers for the Analysis of Simple Sequence Repeat Polymorphisms in Chloroplast Genomes of Dicotyledonous Angiosperms. *Genome* **1999**, *42*, 9–19.
26. Petit, R.; Aguinalalde, I.; de Beaulieu, J.L.; Bittkau, C.; Brewer, S.; Cheddadi, R.; Ennos, R.; Fineschi, S.; Grivet, D.; Lascoux, M.; et al. Glacial Refugia: Hotspots but Not Melting Pots of Genetic Diversity. *Science* **2003**, *300*, 1563–5.
27. Jombart, T. Adegnet: A R Package for the Multivariate Analysis of Genetic Markers. *Bioinformatics* **2008**, *24*, 1403–1405.
28. Jombart, T.; Ahmed, I. Adegnet 1.3-1: New Tools for the Analysis of Genome-Wide SNP Data. *Bioinformatics* **2011**, *27*, 3070–3071.
29. Dray, S.; Dufour, A.-B. The Ade4 Package: Implementing the Duality Diagram for Ecologists. *2007* **2007**, *22*, 20.
30. Goudet, J. Hierfstat, a Package for r to Compute and Test Hierarchical F-Statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186.
31. Winter, D.J. Mmod: An R Library for the Calculation of Population Differentiation Statistics. *Mol. Ecol. Resour.* **2012**, *12*, 1158–1160.

32. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction. *PeerJ* **2014**, *2*, e281.
33. Kamvar, Z.N.; Brooks, J.C.; Grünwald, N.J. Novel R Tools for Analysis of Genome-Wide Population Genetic Data with Emphasis on Clonality. *Front Genet* **2015**, *6*, 208.
34. Nei, M. Analysis of Gene Diversity in Subdivided Populations. *Proc Natl Acad Sci U S A* **1973**, *70*, 3321–3.
35. Hedrick, P.W. A Standardized Genetic Differentiation Measure. *Evolution* **2005**, *59*, 1633–1638.
36. Altman, D.G.; Bland, J.M. How to Obtain the P Value from a Confidence Interval. *BMJ* **2011**, *343*, d2304.
37. Slatkin, M. Inbreeding Coefficients and Coalescence Times. *Genet Res* **1991**, *58*, 167–75.
38. Rousset, F. Genetic Differentiation and Estimation of Gene Flow from  $F_{ST}$ -Statistics Under Isolation by Distance. *Genetics* **1997**, *145*, 1219.
39. Phillips, J.D.; French, S.H.; Hanner, R.H.; Gillis, D.J. HACSIm: An R Package to Estimate Intraspecific Sample Sizes for Genetic Diversity Assessment Using Haplotype Accumulation Curves. *PeerJ Comput Sci* **2020**, *6*, e243.
40. Fussi, B.; Westergren, M.; Aravanopoulos, F.; Baier, R.; Kavaliauskas, D.; Finzgar, D.; Alizoti, P.; Bozic, G.; Avramidou, E.; Konnert, M.; et al. Forest Genetic Monitoring: An Overview of Concepts and Definitions. *Env. Monit Assess* **2016**, *188*, 493.
41. Heuertz, M.; Fineschi, S.; Anzidei, M.; Pastorelli, R.; Salvini, D.; Paule, L.; Frascaria-Lacoste, N.; Hardy, O.J.; Vekemans, X.; Vendramin, G.G. Chloroplast DNA Variation and Postglacial Recolonization of Common Ash (*Fraxinus Excelsior* L.) in Europe. *Mol Ecol* **2004**, *13*, 3437–52.
42. Lefèvre, F.; Koskela, J.; Hubert, J.; Kraigher, H.; Longauer, R.; Olrik, D.C.; Schüller, S.; Bozzano, M.; Alizoti, P.; Bakys, R.; et al. Dynamic Conservation of Forest Genetic Resources in 33 European Countries. *Conserv Biol* **2013**, *27*, 373–84.
43. DAFM *Forest Statistics Ireland 2020*; Ireland, 2020; p. 81.
44. Cahalane, G.; Doody, P.; Douglas, G.; Fennessy, J.; O'Reilly, C.; Pfeifer, A. Sustaining and Developing Ireland's Forest Genetic Resources. An Outline Strategy. **2007**, 30.
45. Perdereau, A.C.; Kelleher, C.T.; Douglas, G.C.; Hodkinson, T.R. High Levels of Gene Flow and Genetic Diversity in Irish Populations of *Salix Caprea* L. Inferred from Chloroplast and Nuclear SSR Markers. *BMC Plant Biol.* **2014**, *14*, 202.
46. Eidesen, P.B.; Alsos, I.G.; Brochmann, C. Comparative Analyses of Plastid and AFLP Data Suggest Different Colonization History and Asymmetric Hybridization between *Betula Pubescens* and *B. Nana*. *Mol Ecol* **2015**, *24*, 3993–4009.
47. Thomson, A.M.; Dick, C.W.; Dayanandan, S. A Similar Phylogeographical Structure among Sympatric North American Birches (*Betula*) Is Better Explained by Introgression than by Shared Biogeographical History. *J. Biogeogr.* **2015**, *42*, 339–350.
48. Wang, N.; Borrell, J.S.; Bodles, W.J.; Kuttapitiya, A.; Nichols, R.A.; Buggs, R.J. Molecular Footprints of the Holocene Retreat of Dwarf Birch in Britain. *Mol Ecol* **2014**, *23*, 2771–82.
49. Currat, M.; Ruedi, M.; Petit, R.J.; Excoffier, L. The Hidden Side of Invasions: Massive Introgression by Local Genes. *Evolution* **2008**, *62*, 1908–20.
50. Petit, R.J.; Excoffier, L. Gene Flow and Species Delimitation. *Trends Ecol Evol* **2009**, *24*, 386–93, doi:10.1016/j.tree.2009.02.011.
51. Kulju, K.K.M.; Pekkinen, M.; Varvio, S. Twenty-Three Microsatellite Primer Pairs for *Betula Pendula* (Betulaceae). *Mol. Ecol. Notes* **2004**, *4*, 471–473.



- 
52. Molloy, K.; O'Connell, M. Post-Glaciation Plant Colonisation of Ireland: Fresh Insights from An Loch Mór, Inis Oírr, Western Ireland. *Ir. Nat. J.* **2014**, *33*, 66–88.
  53. Þórsson, Æ.T.; Pálsson, S.; Lascoux, M.; Ananthawat-Jónsson, K. Introgression and Phylogeography of *Betula Nana* (Diploid), *B. Pubescens* (Tetraploid) and Their Triploid Hybrids in Iceland Inferred from CpDNA Haplotype Variation. *J. Biogeogr.* **2010**, *37*, 2098–2110.