

Collection/Special Issue: COST Action FP1202

“Strengthening conservation: a key issue for adaptation of marginal/peripheral populations of forest trees to climate change in Europe (MaP-FGR)”

Guest Editors: Fulvio Ducci, Kevin Donnelly

## Weak isolation by distance and geographic diversity gradients persist in Scottish relict pine forest

Patricia González-Díaz<sup>(1-2)</sup>,  
Stephen Cavers<sup>(2)</sup>,  
Glenn R Iason<sup>(3)</sup>,  
Allan Booth<sup>(3)</sup>,  
Joanne Russell<sup>(3)</sup>,  
Alistair S Jump<sup>(1-4)</sup>

Gene flow is one of the main factors shaping genetic diversity within and among tree populations, and occurs through pollen and seed dispersal. Recent findings of pollen-release asynchronies in distant populations of Scots pine (*Pinus sylvestris* L.) within Scotland suggest that gene dispersal among more distant populations might be less effective than previously thought. Limited gene dispersal is one of the major factors causing genetic structure for neutral markers, and pollen-release asynchrony could have driven isolation by distance (IBD) among Scottish populations. Previous studies of neutral markers found little differentiation among Scottish populations of Scots pine, however they did not consider IBD over the full Scottish range. We analysed data from 6 nuclear simple sequence repeats (SSR) and 5 chloroplast SSR loci in a total of 540 individuals of Scots pine from 18 populations across Scotland. Our aim was to assess contemporary levels and distribution of genetic variation and to test if the distribution of genetic diversity was consistent with IBD. We also analysed patterns of gene flow that could have contributed to the observed patterns of variation. Levels of genetic diversity were high, for both nuclear and chloroplast markers within populations, and there was no significant differentiation among populations. A weak signal of IBD was present. We found an increase in nuclear diversity towards the East along with greater gene flow in a West-East direction commensurate with the prevailing winds. Our findings suggest that this wind-driven gene flow is dominant over genetic drift and prevents differentiation among the Scottish populations. It may also counteract any pollen-release asynchronies among populations.

**Keywords:** *Pinus sylvestris*, Genetic Diversity, Gene Flow, Isolation by Distance, Prevailing Winds

□ (1) Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA (UK); (2) NERC Centre for Ecology and Hydrology Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB (UK); (3) The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH (UK); (4) CREA (Centre de Recerca Ecològica i Aplicacions Forestals), Campus UAB, Edifici C, E-08193, Belaterra, Barcelona (Spain)

@ Patricia González-Díaz  
([patricia.gonzalezdiaz@stir.ac.uk](mailto:patricia.gonzalezdiaz@stir.ac.uk))

Received: Apr 06, 2017 - Accepted: Apr 19, 2018

**Citation:** González-Díaz P, Cavers S, Iason GR, Booth A, Russell J, Jump AS (2018). Weak isolation by distance and geographic diversity gradients persist in Scottish relict pine forest. *iForest* 11: 449-458. - doi: [10.3832/ifor2454-011](https://doi.org/10.3832/ifor2454-011) [online 2018-07-02]

Communicated by: Fulvio Ducci

### Introduction

Genetic diversity provides the fundamental basis for the evolution of forest tree species and for their adaptation to change (FAO 2014). The importance of including genetic factors in sustainable forest management has been strongly supported in recent years (FAO 2014, Fady et al. 2016b) particularly in marginal populations (Fady et al. 2016a). Accordingly, several schemes have been proposed for monitoring genetic diversity at the European and international levels (Graudal et al. 2014), including the quantification of both adaptive and neutral genetic diversity in forest trees.

Neutral genetic diversity, which has little or no effect on the phenotype, is valuable for studying the effects of historical events such as population size changes, dispersal and vicariance, and of contemporary processes affecting gene flow, such as pollen and seed dispersal (Holderegger et al. 2006). Characterisation of the level and structure of neutral genetic diversity is considered an appropriate first step to designate conservation units. The amount of neutral genetic diversity and how it is parti-

tioned within and among populations typically results from the balance between gene flow and random genetic drift (Bursczyk et al. 2004). Gene flow can counteract genetic differentiation through genetic homogenization, whereas genetic drift (the random changes of allele frequencies over generations) is expected to lead to differentiation among populations (Slatkin 1985). Determining levels and structure of neutral genetic diversity of tree populations should be an essential step in the design of sustainable forest management plans, as well as to better understand the processes that are likely responsible for the maintenance of such diversity.

Gene flow occurs through pollen and seed dispersal, and in wind-pollinated trees is usually more extensive by pollen than by seeds, e.g., pollen flow in pines was up to approximately 60 times greater than seed flow (Ennos 1994). Although declining gene flow with distance is expected given that most pollen deposition, pollination events and seedling establishment occur near the parent plant (Deacon & Cavender-Bares 2015), genes and especially pollen

can travel long distances, resulting in large pollen-mediated gene flow. Pollen dispersal is, therefore, expected to shape the level and structure of genetic variation within and among tree populations at local and regional scales (Burczyk et al. 2004). For example, Scots pine pollen from southern and central Finland contributes to populations located hundreds of km further north (Lindgren et al. 1995, Varis et al. 2009).

Multiple ecological factors can restrict pollen-mediated gene flow among wind-pollinated tree populations, and hence lead to temporal or spatial increases in genetic drift. Physical barriers (e.g., mountains – Naydenov et al. 2011) and fragmentation (Provan et al. 2007) may restrict pollen flow. However, in some cases, increases of pollen flow counteract diversity loss resulting from fragmentation (Davies et al. 2013). Other significant barriers to pollen flow, that have received less attention, can be due to asynchrony in reproductive phenol-

ogy, e.g., timing of pollen or flower production (Aitken et al. 2008, Whittet et al. 2017), which may limit the randomness of mating. Indeed, there is evidence that plants mate with phenologically similar individuals more frequently than random (Ennos & Dodson 1987). Consequently, differences in timing of reproductive phenology in nearby tree populations are expected to favour mating between physically closer individuals by mating incompatibility among tree populations, thereby reinforcing isolation by distance (IBD). In addition, wind patterns might also influence the levels and direction of gene flow, such as the significant contribution to directional pollen flow in Scots pine populations from Scandinavia (Lindgren et al. 1995).

The fragmented natural Scots pine (*Pinus sylvestris* L.) forests of Scotland represent the westernmost extreme of the species' native range, separated by at least 500 km from natural stands in mainland Europe. Despite their geographical marginality, pre-

vious studies on the Scottish Scots pine populations found high levels of neutral genetic variation but little differentiation among populations (Provan et al. 1998, Sinclair et al. 1998, Wachowiak et al. 2011 – Tab. 1). The westernmost populations have previously been found to be somewhat distinct from the others (Forrest 1980, Kinloch et al. 1986, Sinclair et al. 1998 – Tab. 1), on the basis of allele frequency differences.

A recent study found mismatches in timing of pollen release between a Western and an Eastern population from Scotland. Pollen was released first in the west, between 9.8 to 15.8 days earlier than in the east (Whittet et al. 2017). Consistent asynchrony in pollen production among populations will limit gene flow; where the extent of asynchrony reflects geographic distance, a pattern of IBD will be established. Although several studies have addressed genetic variation in Scots pine from Scotland (Tab. 1), these studies have not investigated patterns of IBD in the full Scottish

**Tab. 1** - Studies assessing neutral genetic variation of Scots pine in Scotland using variable molecular markers. (RFLP): Restriction Fragment Length Polymorphism; (SSR): Simple Sequence Repeat or microsatellite; (SNP): Single Nucleotide Polymorphism; (No. pop): number of studied populations; (No. ind): number of genotyped individuals.

Marker	Location	No. pop.	No. ind.	No. markers	Gst*/Fst**/AMOVA#	Diversity	Reference	Main conclusions
Mono-terpene	Scotland	41	6705	11	Similarity matrix: 0-24 over 30	-	Forrest 1980	Variation between sites allowed divide the natural range into several areas of biochemical similarity, the most distinct being a north-western group of sites with Sheldal as its most distinctive site. A trend of gradually decreased of similarity from the north east from Scotland in a south westerly direction.
Mono-terpene	Europe	6	953	11	Similarity matrix: 3-24 over 30	-	Forrest 1982	Northern European populations were similar to each other but the three from middle and southern Europe showed large differences from them and from each other. Western region from Scotland showed similarity to middle Europe and south western populations to northern Europe.
	Scotland (from Forrest 1980)	41	953	11	Similarity matrix: 0-24 over 30	-		
Mono-terpene Isozymes	Scotland	40	5765	6	0.045*	0.272–0.378	Kinloch et al. 1986	High genetic diversity. Several populations in the Western region from Scotland were distinct from all others and each other. Scottish Scots pine forest originated from more than one refugium after the last glaciation.
		14	2177	9	0.028*	0.291–0.311		
RFLP (mtDNA)	Scotland	20	466	Coxl mitochondrial gene	0.37**	-	Sinclair et al. 1998	Two mitotypes were present: mitotype a is present at all sites, but that mitotype b is confined to three western populations
SSR (cpDNA)	Scotland	7	330	17	0.032 <sup>#</sup>	0.950–0.987	Provan et al. 1998	Higher levels of diversity for the Scottish populations than those for European population. A mutation in one loci occurred in the Western region of Scotland.
	Europe	8	185	17	Scot vs. Eur 0.0148 <sup>#</sup>	0.908–0.976		
SNP (nDNA)	Scotland	21	42	16	-0.017–0.023	0.754–0.819 (0.831 at 8 loci)	Wachowiak et al. 2011	High genetic diversity
	Europe	7	40	10	0.071–0.079	(0.795 at 8 loci)		
SSR (nDNA)	Scotland	21	1680	3	-	-	Salmela et al. 2011	High levels of outcrossing in Scottish populations
SNP (nDNA)	Scotland	12	120	-	0.009**	0.67	Wachowiak et al. 2013	High levels of nucleotide diversity within populations
SNP (mtDNA)	Scotland	-	-	-	-	0.81		
SSR (nDNA)	Scotland	2	647	12	0.004**	0.56–0.58	González-Díaz et al. 2017	High levels of genetic diversity and presence of moderate fine-scale spatial genetic structure

range. Forrest (1980) found a trend of gradually increasing genetic similarity from the southwest of the Scottish range toward the northeast. Other studies included estimates of  $F_{ST}$  (Kinloch et al. 1986, Sinclair et al. 1998, Wachowiak et al. 2011, 2013), which indicate the extent of genetic differentiation (genetic structure) among populations, regardless of the distance. Measures of IBD indicate the relationships between geographic and genetic distance, and can help us to identify where restricted gene dispersal may relate to geographic factors.

Based on the previous evidence, we hypothesised that, despite extensive pollen dispersal by wind, the presence of an East-West asynchrony in pollen production might result in IBD due to a higher probability that more synchronous populations will mate with each other. Alternatively, the absence of identifiable IBD would suggest that despite the presence of pollen asynchrony, effective gene flow at the regional scale, might have prevented differentiation. To test this hypothesis we characterised genetic diversity along the full East-West gradient of Scots pine within Scotland with two sets of neutral molecular markers, nuclear and chloroplast microsatellites (SSR). The use of both markers can allow greater understanding of the factors driving differentiation than based on either alone (Sjölund et al. 2017); chloroplast DNA is paternally inherited in conifers via dispersing pollen, and nuclear DNA is biparentally inherited via both pollen and seed dispersal. In addition, the coverage of the full native range of Scots pine in Scotland allows much greater resolution of geographical structuring of gene flow, including IBD, since most other SSR-based studies of this species incorporated relatively few Scottish populations (Provan et al. 1998, González-Díaz et al. 2017) or did not focus on these aspects of the species' biology (Salmela 2011). Specifically, we sought to answer the following questions: (1) what are the contemporary levels and structure of genetic variation of the Scots pine populations across Scotland? (2) Is there any evidence of IBD? And (3) what are the gene flow patterns among the studied populations?

## Material and methods

### Study species

Scots pine (*Pinus sylvestris* L.) in Scotland is a foundation species of the remnant Caladonian pine forests. It is typically a pioneer species that readily regenerates after major natural or human disturbances, if competition and grazing pressure are low. It grows well on most soils, nevertheless, due to low tolerance of shade and competition, it is often restricted to poor soils (Mason et al. 2004). Based on fossil data, this species reached its maximum extent in Scotland around 8,000 years ago, covering approximately 1.5 million ha (Birks 1989).

**Tab. 2** - Details of study sites. Population area was obtained from Mason et al. (2004).

Population name	Code	Seed zone	Pine area (ha)	Lat N (dec. deg.)	Long E (dec. deg.)	Altitude (m a.s.l.)
Cona Glen	COG	SW	189	56.78	5.33	148
Glen Loy	GLL	SW	74	56.91	5.13	170
Crannach	CRA	SW	70	56.5	4.77	296
Coille Coire Chuilc	CCC	SC	67	56.41	4.71	257
Meggernie	MEG	SC	277	56.58	4.35	306
Black Wood of Rannoch	BWR	SC	1011	56.67	4.32	275
Abernethy	ABE	EC	2452	57.24	3.66	341
Rothiemurchus	ROT	EC	1744	57.15	3.77	318
Allt Cul	ALT	NE	13	57.11	3.33	476
Glen Tanar	GLT	NE	1564	57.05	2.86	334
Glen Affric	GLA	NC	1532	57.27	4.92	256
Amat	AMA	NC	181	57.87	4.59	137
Loch Clair	LOC	NW	126	57.56	5.34	132
Shieldaig	SHI	NW	103	57.51	5.64	81
Beinn Eighe	BEE	NW	182	57.63	5.35	63
Glen Einig	GLE	N	27	57.95	4.76	55
Strath Oyckell	STO	N	14	57.95	4.64	103
Rhidorroch	RHI	N	103	57.93	4.97	182

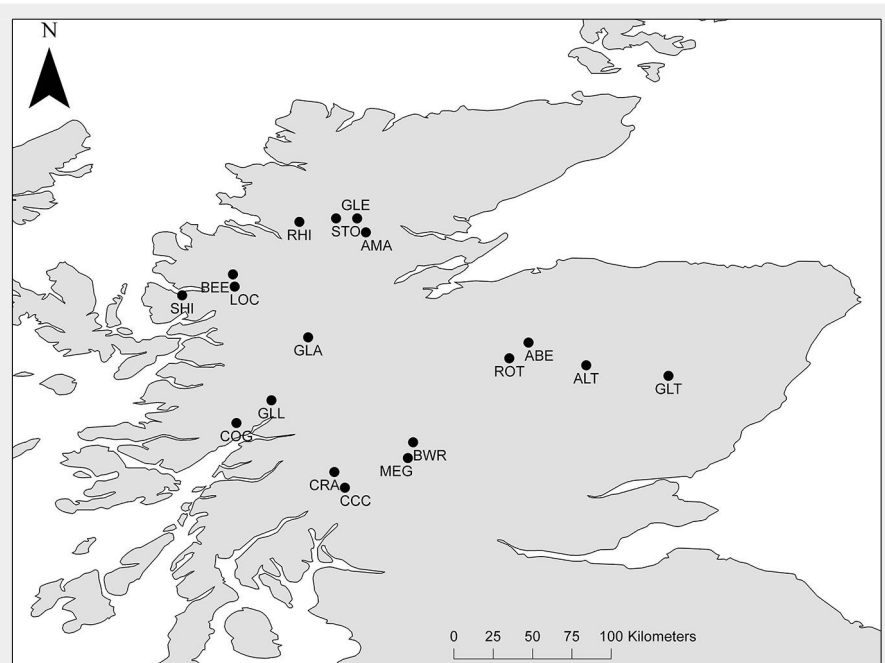
Nowadays, 84 fragments of the ancient native pine forest remain in this area, scattered over a total area of 17,882 ha (Mason et al. 2004).

### Study sites and microsatellite analysis

Eighteen populations were selected to cover the full native range of Scots pine within Scotland (Tab. 2, Fig. 1). We sampled a total of 30 randomly selected trees within each population. Sampled trees were old adults, thereby avoiding potential effects of gene flow from more recent Scots pine plantations that might otherwise obscure patterns of genetic diversity and divergence in the native stands.

Total genomic DNA was extracted from 50 mg silica gel-dried needles using a QIA-GEN DNeasy® plant extraction kit (QIAGEN

Ltd. Crawley, UK) following the manufacturer's protocol. All individuals from the eighteen Scottish populations were genotyped at six nuclear (nSSR) and five chloroplast (cSSR) microsatellite markers. We used six nSSR: PSAJ223770/SPAC11.14, PSAJ223766/SPAC11.8 (Soranzo et al. 1998), Ptx2146 (Auckland et al. 2002), SsrPt\_ctg4698, SsrPt\_ctg9249 (Chagné et al. 2004) and LOP3 (Liewlaksaneeyanawin et al. 2004); and five cSSR: PCP26106, PCP30277, PCP36567, PCP45071, PCP87314 (Provan et al. 1998). Reactions were carried out in a final volume of 10 µl with 1 µM fluorescently-labelled forward primer, 1 µM reverse primer, 200 µM each dNTPs, 0.5 units Taq® polymerase (Roche Applied Science), 1X PCR buffer (supplied with Taq) and 25 ng of template DNA. Annealing tem-



**Fig. 1** - Study populations for nSSR and cSSR analyses.



perature was 56 °C. Polymerase chain reactions (PCR) were performed in a Gene Amp PCR System 9700<sup>®</sup> thermo cycler (Applied Biosystems, Bleiswijk, Netherlands), with the following programme: 1 cycle at 95 °C for 4 min followed by 35 cycles at 95 °C for 45 s, 56 °C for 45 s, 72 °C for 45 s, and a final step at 72 °C for 5 min. Fragment analysis was performed by Genome Technology at James Hutton Institute, Dundee, UK, using a 3730 DNA Sequencer<sup>®</sup> (Applied Biosystems) with reference to a ROX 500 size standard. Fragments were scored using GeneMarker<sup>®</sup> v. 2.6.0. (SoftGenetics, State College, PA, USA). Null allele frequencies at nuclear loci for each locus and each population were checked by using the software Micro-Checker (Van Oosterhout et al. 2004). PSAJ223770/SPAC11.14 showed evidence of null alleles, however the frequency was below 0.2, which is the threshold over which null alleles can result in a significant underestimate of expected het-

erozygosity  $H_E$  (Chapuis & Estoup 2007, Belletti et al. 2012), therefore it was kept for further analysis.

#### Data analysis

##### Genetic diversity

Genetic diversity estimators within populations were estimated using FSTAT ver. 2.9.3.2 (Goudet 1995) and Arlequin ver. 3.5 (Excoffier & Lischer 2010). For nSSR, we estimated mean number of alleles per locus ( $n_A$ ), rarefied allelic richness ( $n_{A_r}$ ), number of private alleles ( $n_{Ap}$ ), observed heterozygosity ( $nH_o$ ), expected heterozygosity ( $nH_E$ ) and inbreeding coefficient ( $nF_{IS}$ ). Rarefaction of allelic richness controls for differences in sample size allowing comparison among sites and studies. For cSSR, we estimated mean number of alleles ( $cA$ ). For the remaining cSSR estimators, alleles were combined to compose a unique chloroplast haplotype for each individual. Indi-

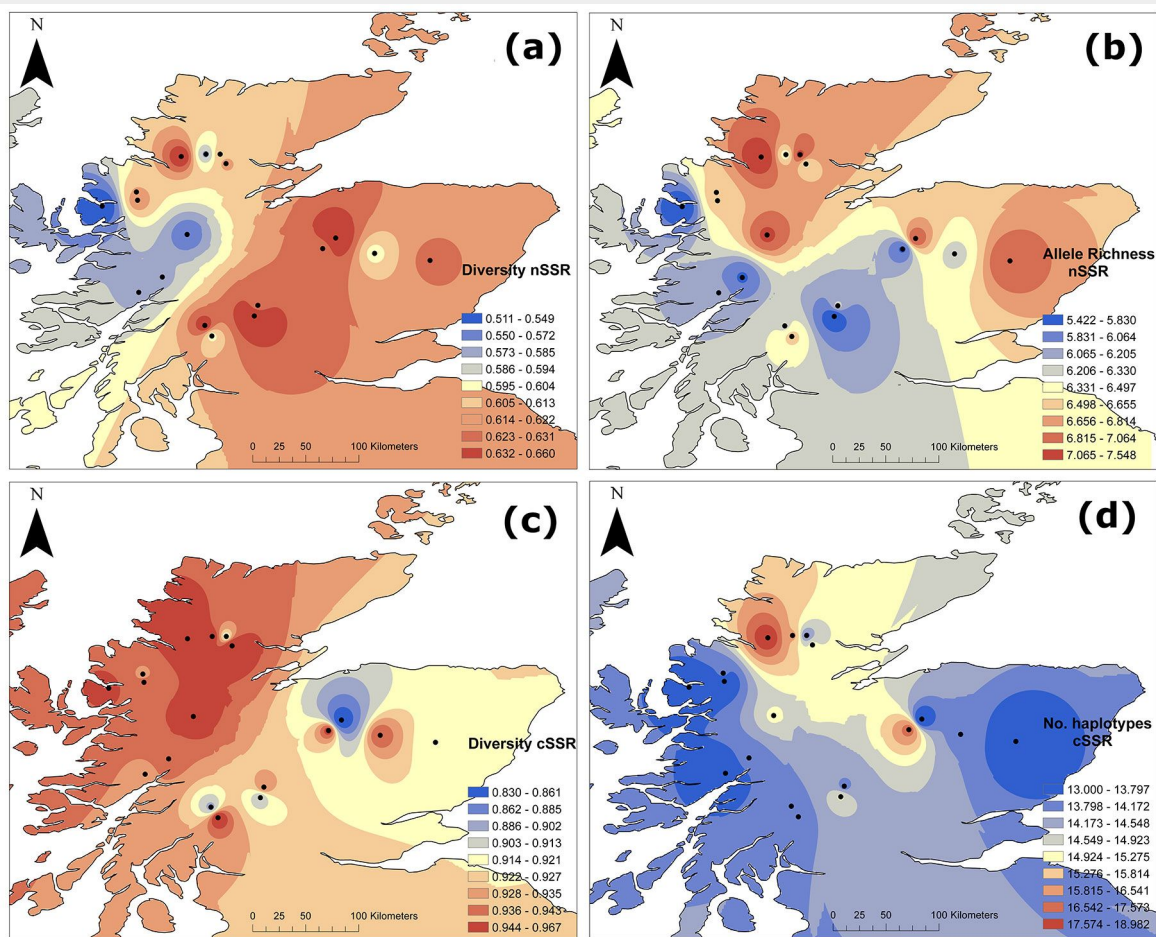
viduals with missing data were discarded from the inference of multilocus haplotypes. We estimated number of haplotypes ( $cH_N$ ), number of private haplotypes ( $cH_P$ ) and gene diversity ( $cH_E$ ), the latter based on haplotype frequencies.

##### Population differentiation and Bayesian clustering

To estimate population differentiation, we calculated  $nF_{ST}$  among populations using nSSR in Arlequin ver. 3.5 (Excoffier & Lischer 2010), and the differentiation index Jost's  $D$  implemented in the R package DEMETICS (Gerlach et al. 2010). In both cases, significance values were determined for a 5% nominal level after Bonferonni correction. Multilocus haplotypic and genetic estimators for both nSSR and cSSR ( $nH_E$ ,  $nA_r$ ,  $cH_E$  and  $cH_N$ ) were mapped using ArcMap<sup>®</sup> ver. 10 (ESRI, Redlands, CA, USA), using inverse distance weight methods available on the spatial analyst interpola-

**Tab. 3** - Genetic diversity estimators for nuclear (nSSR) and chloroplast (cSSR) markers. ( $nN$ ): no. of samples genotyped with nSSR; ( $nA$ ): no. of alleles; ( $nA_r$ ): rarefied allelic richness for 28 diploid individuals; ( $nAp$ ): no. of private alleles; ( $H_o$ ): observed heterozygosity; ( $H_E$ ): expected heterozygosity; ( $F_{IS}$ ): inbreeding coefficient; ( $cN$ ): no. of samples genotyped with cSSR; ( $cA$ ): no. of alleles; ( $cH_N$ ): no. of haplotypes; ( $cH_P$ ): no. of private haplotypes; ( $cH_E$ ): gene diversity corrected for sample size (Nei 1987).  $P$ -values for  $F_{IS}$  are obtained after 1,000 permutations of gene copies within individuals of each stand. Standard errors are reported in brackets. (\*):  $P < 0.05$ ; (\*\*):  $P < 0.01$ ; (\*\*\*):  $P < 0.001$ .

Population	nSSR							cSSR				
	$nN$	$nA$	$nA_r$	$nAp$	$nH_o$	$nH_E$	$nF_{IS}$	$cN$	$cA$	$cH_N$	$cH_P$	$cH_E$
COG	30	6.1670 (3.3120)	6.0992 (2.9791)	1	0.4944 (0.2352)	0.5779 (0.2402)	0.1460***	28	3.2 (1.095)	13	4	0.9339 (0.0227)
GLL	30	5.8330 (2.9270)	5.7860 (2.6523)	0	0.4944 (0.1500)	0.5801 (0.2743)	0.1500***	25	2.8 (0.837)	14	1	0.9400 (0.0280)
CRA	30	6.3330 (3.1410)	6.2860 (2.8557)	1	0.5151 (0.1960)	0.6389 (0.1990)	0.1960***	25	3.0 (1.225)	14	1	0.8967 (0.0431)
CCC	30	6.6670 (4.2270)	6.5777 (3.7589)	0	0.5722 (0.0460)	0.5995 (0.2201)	0.0460	23	2.8 (1.304)	14	1	0.9565 (0.0220)
MEG	30	5.5000 (2.8110)	5.4753 (2.5497)	0	0.5556 (0.1570)	0.6573 (0.1552)	0.1570***	28	3.4 (1.673)	15	3	0.9048 (0.0420)
BWR	30	6.3330 (2.9440)	6.2638 (2.6361)	0	0.5860 (0.0560)	0.6200 (0.1885)	0.0560	23	3.2 (1.304)	14	3	0.9368 (0.0331)
ABE	30	7.1670 (3.7640)	7.0532 (3.3582)	1	0.6111 (0.0590)	0.6490 (0.2224)	0.0590	27	2.8 (1.304)	13	0	0.8291 (0.0684)
ROT	30	5.8330 (2.3170)	5.7767 (2.0727)	1	0.5935 (0.0500)	0.6240 (0.2063)	0.0500	27	3.4 (1.140)	17	3	0.9487 (0.0257)
ALT	30	6.3330 (3.0770)	6.2532 (2.7636)	1	0.5937 (0.0120)	0.6006 (0.2542)	0.0120	25	2.8 (0.837)	14	0	0.9433 (0.0240)
GLT	30	7.1670 (2.9270)	7.0615 (2.5983)	0	0.5556 (0.1100)	0.6228 (0.2215)	0.1100*	25	3.4 (1.140)	13	1	0.9167 (0.0349)
GLA	30	7.3330 (4.2270)	7.1470 (3.7213)	0	0.4925 (0.1170)	0.5565 (0.2660)	0.1170**	22	3.4 (0.241)	15	2	0.9524 (0.0291)
AMA	30	6.5000 (3.6740)	6.4505 (3.3049)	0	0.5444 (0.1160)	0.6149 (0.2604)	0.1160**	21	3.2 (0.837)	15	1	0.9667 (0.0236)
LOC	30	6.6670 (3.4450)	6.5968 (3.0930)	0	0.5803 (0.0690)	0.6226 (0.2130)	0.0690	27	3.0 (1.000)	13	0	0.9373 (0.0220)
SHI	30	5.5000 (3.2710)	5.4198 (2.9151)	0	0.4944 (0.0330)	0.5110 (0.2737)	0.0330	29	3.0 (0.707)	13	2	0.9458 (0.0173)
BEE	30	6.6670 (3.7240)	6.6082 (3.3824)	1	0.5500 (0.1040)	0.6128 (0.2978)	0.1040*	29	3.2 (1.304)	14	0	0.9335 (0.0228)
GLE	30	6.5000 (2.8810)	6.3923 (2.5723)	2	0.4571 (0.2140)	0.5794 (0.2517)	0.2140***	28	2.8 (0.837)	16	2	0.9497 (0.0214)
STO	30	7.3330 (3.5020)	7.1848 (3.1003)	0	0.5222 (0.1450)	0.6095 (0.2079)	0.1450***	27	2.8 (1.304)	14	1	0.9145 (0.0334)
RHI	30	7.6670 (3.5020)	7.5522 (3.1469)	1	0.5714 (0.1130)	0.6430 (0.2847)	0.1130*	30	3.2 (1.304)	19	2	0.9586 (0.0209)
All populations	540	11.5000 (7.3420)	5.8400 (3.1362)	9	0.5434 (0.2068)	0.6179 (0.2327)	0.121***	469	4.6 (1.140)	64	27	-
Overall mean	30	6.5278 (3.3152)	6.4435 (2.9700)	0.5	0.5456 (0.1101)	0.6067 (0.2334)	0.1052	26.05	3.1 (1.077)	14.44	1.5	0.9315 (0.0298)



**Fig. 2** - Genetic diversity parameters for nSSR (above) and cSSR (below). (a) Genetic diversity for nSSR ( $nH_E$ ); (b) rarefied allele richness for nSSR ( $nA_R$ ); (c) gene diversity corrected for sample size for cSSR ( $cH_E$ ); (d) no. of haplotypes for cSSR ( $cH_N$ ).

tion tool. To map nuclear genetic differentiation, for each site we calculated the percentage of total sites that it was significantly differentiated from (e.g., percentage of differentiated sites,  $nDS$ , %), based on  $nF_{ST}$  values. To quantify the distribution of variation of nuclear genetic diversity and chloroplast haplotypes, we tested both marker sets in a hierarchical analysis of molecular variance (AMOVA) from the level of individual, population and cluster of populations (see directional relative migration rates section), performed in Arlequin ver. 3.5 (Excoffier & Lischer 2010).

We performed individual-based Bayesian assignment methods using data from nuclear loci in STRUCTURE ver. 2.3.4 (Pritchard et al. 2000). We used a model assuming correlated allele frequencies and admixed ancestry. We included the site location *a priori* (LOCPRIOR option) to improve the detection of weak population structure.  $K$  was set from 1 to 20, with 10 runs performed for each value of  $K$ . Runs consisted of 500,000 Markov Chain Monte Carlo (MCMC) iterations with a burn-in period of 100,000. We used STRUCTURE HARVESTER (Earl & Von Holdt 2011), an application that uses the Evanno method for assessing and visualizing likelihood values across multiple values of  $K$  and detecting the number of genetic groups that best fit the data.

### Isolation by distance and directional relative migration rates

For testing isolation by distance we used nuclear markers in SPAGeDi ver. 1.4 (Hardy & Vekemans 2002) with significance determined by permuting site locations among populations 10,000 times. Following Rousset (1997), we used the  $F_{ST} / (1 - F_{ST})$  ratio as a measure of genetic distance as it is expected to vary linearly with the natural log of the geographical distance.

To reduce the number of sites to an analytically tractable set for estimating migration patterns, we grouped sites with their nearest neighbours to give seven site clusters, corresponding to the biochemical zones described by Forrest (1980) and used as a seed zones (Tab. 2). Directional relative migration rates between site clusters were estimated using nuclear markers and the function DIVMIGRATE from the R-package DIVERSITY (Keenan et al. 2013) using Jost's  $D$  as a measure of genetic differentiation. To test whether relative migration is significantly higher in one direction than the other (e.g., asymmetric), 95% confidence intervals were calculated from 1,000 bootstrap iterations.

## Results

### Genetic diversity

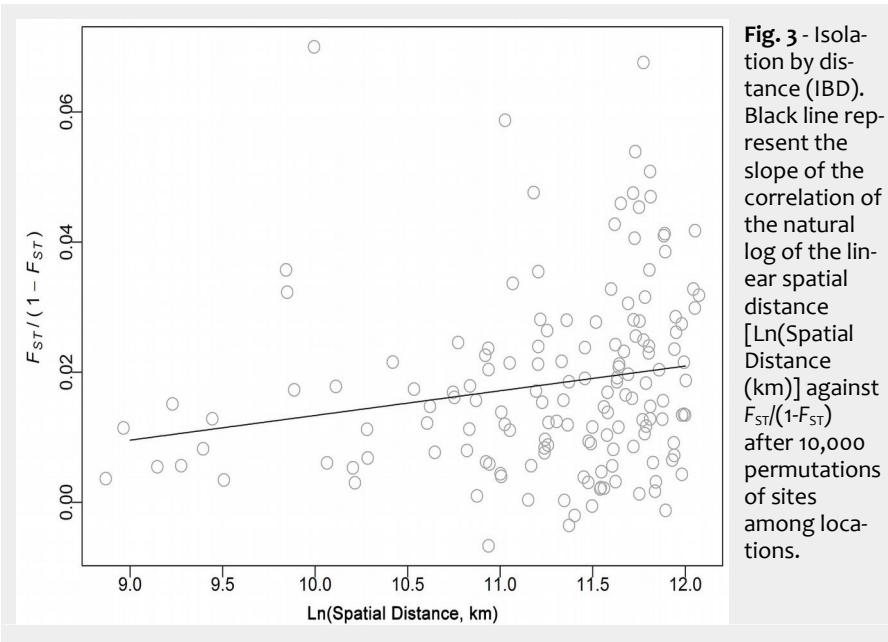
Among the six nuclear loci analysed, the number of alleles per locus ( $nA$ ) ranged

from 2 to 14, with a multilocus average of  $11.5 \pm 7.34$  for all populations combined. In total we obtained 9 private alleles. For rarefied allelic richness ( $nA_R$ ), multilocus estimates ranged from values of 5.42 to 7.55, based on a minimum number of 28 diploid individuals, with a multilocus average of 5.84 for all populations combined (Tab. 3 and Fig. 2b). The multilocus expected ( $nH_E$ ) and observed heterozygosity ( $nH_O$ ) was  $0.62 \pm 0.03$  and  $0.54 \pm 0.04$  respectively for all populations combined (Tab. 3, and Fig. 2a for  $nH_E$ ). A general trend of lower nuclear diversity ( $nH_E$ ) in the western sites and relatively higher diversity in the eastern sites was observed (Fig. 2a). This trend was positively correlated with longitude ( $F=6.703$ ,  $R^2=0.25$ ,  $p<0.05$ ), but not for  $nA_R$ . A general significant homozygote excess was found (overall  $F_{IS}=0.121$ ,  $p<0.05$  – Tab. 3).

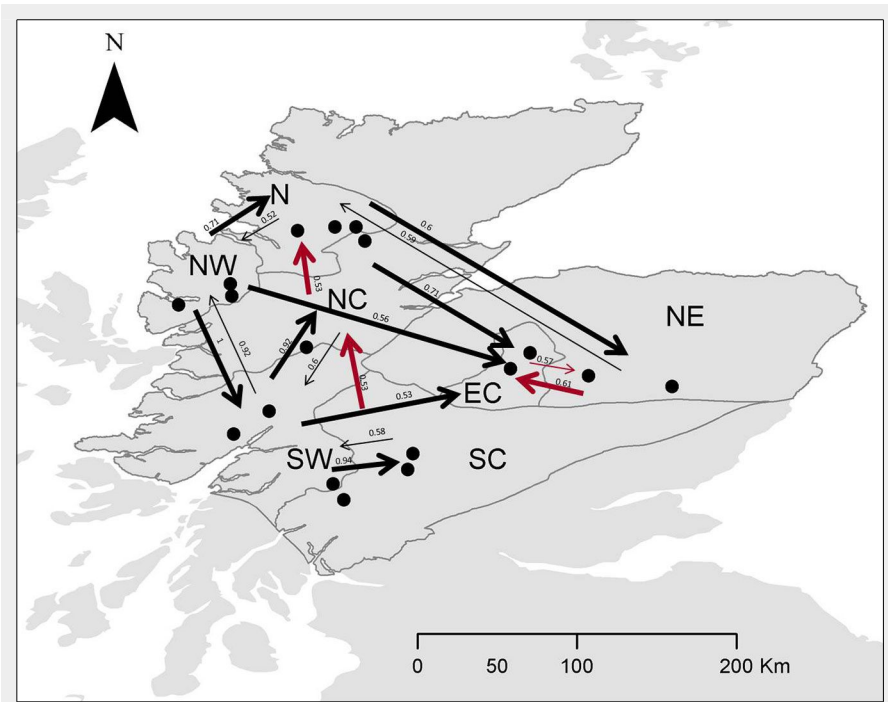
Among the five chloroplast loci analysed, the number of alleles per locus ( $cA$ ) ranged from 2 to 6, with a multilocus average of 3.08 for all populations combined. Gene diversity ( $cH_E$ ) ranged between 0.83 and 0.96 (Tab. 3 and Fig. 2c). The number of haplotypes ( $cH_N$ ) within populations ranged from 13 to 19 (Tab. 3 and Fig. 2d), with a total of 64 haplotypes recorded, of which 42% (27 haplotypes) were private, i.e., unique to a particular population (Tab. 3). Although most western sites showed high levels of chloroplast gene diversity ( $cH_E$ ), some of

**Tab. 4** - Hierarchical analysis of molecular variance (AMOVA) for nuclear (nSSR) and chloroplast (cSSR) markers at the individual, population and cluster of populations. The degrees of freedom (df), percentage of variation explained by each level (Variation, %), and the relevant P-values are indicated.

Source of variation	nSSR			cSSR		
	df	Variation (%)	p-value	df	Variation (%)	p-value
Among cluster of populations	6	0.17	0.25	6	0.25	0.11
Among populations	11	1.79	<0.001	11	1.89	<0.001
Within populations	1062	98.04	<0.001	453	97.86	<0.001



**Fig. 3** - Isolation by distance (IBD). Black line represent the slope of the correlation of the natural log of the linear spatial distance [Ln(Spatial Distance (km))] against  $F_{ST}/(1-F_{ST})$  after 10,000 permutations of sites among locations.



**Fig. 4** - Relative migration networks for nSSR with populations sorted in seed groups (see Tab. 2 for population details) using a threshold of 0.5. Arrows in black denote greater gene flow pattern in the W-E direction, whereas arrows in red denote greater gene flow in the E-W or S-N direction. Thicker arrows mean stronger gene flow in such direction.

the eastern sites showed also high levels of diversity and the spatial trend was not as clear as for nSSR (Fig. 2c). No significant correlation with longitude was found for chloroplast diversity.

*Population differentiation and Bayesian clustering*

When testing differences among populations, multilocus  $nF_{ST}$  values ranged from -0.005 (ROT and GLT) to 0.065 (SHI and MEG), being significant ( $p < 0.05$ ) in most cases (Tab. S1 in Supplementary material). Thus, for instance, of the most differentiated populations, SHI was significantly different from all populations and BWR from all but GLA (Tab. S1 and Fig. S1). RHI, BEE, CCC, MEG and ALT significantly differed from all but two or three populations (Tab. S1 and Fig. S1). Jost's  $D$  values were in agreement but greater than  $F_{ST}$  values and ranged between -0.012 (ROT and GLT) to 0.127 (SHI and MEG - Tab. S1 in Supplementary material). Interestingly, the populations with the greatest asynchrony in pollen release (Whittet et al. 2017), BEE and ALT, had some of the largest  $F_{ST}$  values ( $F_{ST} = 0.05$ ,  $p < 0.05$ ). The results of AMOVA showed that although most of the variation was found within populations, among-population variability is greater within groups than between them for both sets of molecular markers (Tab. 4).

Structure identified  $K=2$  as the more likely number of clusters (red and blue clusters hereafter), however, these clusters were not related to the East-West location of the individuals (Fig. S2 in Supplementary material). Most sites contained highly admixed individuals; however, the site SHI had more than 75% of individuals in the less common blue cluster. In addition, four sites had between 60 and 75% of individuals in the blue cluster STO, COG, BWR, ROT, ALT, followed by LOC, GLA and GLT with more than 50%. All the other sites had a majority of the red cluster.

*Isolation by distance and asymmetric migration*

Although weak, IBD was significant in the Scottish populations (slope = 0.0038,  $R^2 = 0.035$ ,  $P < 0.05$  - Fig. 3), indicating that when geographic distance increases, populations became more differentiated. Expressing geographic and/or genetic distance on a logarithmic scale did not improve the model fit (data not shown).

The relative migration network (Fig. S3a in Supplementary material) shows all relative migration rates between site clusters of Scottish populations of Scots pine. Those rates indicate that gene flow was, in most cases, greater in the West-East direction than in the East-West direction, although it was not significantly asymmetric. In Fig. 4 and Fig S3b (Supplementary material), directional relative migration rates below 0.50 were filtered out to emphasize the major gene flow networks.



## Discussion

Our study presents a detailed genetic survey of a subset of the remaining natural populations of Scots pine from Scotland using both bi-parentally inherited nuclear SSR and paternally inherited chloroplast SSR. Three main results were obtained: (i) high levels of genetic variation and low population differentiation; (ii) a weak pattern of isolation by distance; and (iii) an increase of nuclear diversity towards the East. While we detected some discrepancies (e.g., SHI), our findings suggest that the effects of gene flow dominate those of genetic drift and prevent differentiation in the Scottish populations. Our results suggest greater gene flow in the West-East direction, likely influenced by prevailing wind patterns.

### High levels of genetic variation and low population differentiation

In agreement with other molecular markers (Forrest 1980, Kinloch et al. 1986, Sinclair et al. 1998, Provan et al. 1998, Wachowiak et al. 2011), our results indicated relatively high levels of genetic variation across the Scottish native populations of Scots pine for both nuclear and chloroplast microsatellite markers. This is counter to expectations given their marginal distribution in global terms and the severe reduction in the extent of forest, and its fragmentation in recent centuries. Values of multilocus nuclear SSR diversity ( $nH_E = 0.51-0.66$ ) were comparable to those found in other Scottish and Eurasian populations ( $nH_E = 0.50-0.69$  – Naydenov et al. 2011, Bernhardsson et al. 2016, González-Díaz et al. 2017), although lower than in some mainland European populations ( $nH_E = 0.74-0.85$  – Scalfi et al. 2009, Nowakowska et al. 2014, García-Gil et al. 2015). In addition, the often high levels of chloroplast SSR diversity ( $CH_E = 0.83-0.97$ ) were similar to those reported for other Scottish, Iberian and Italian populations ( $CH_E = 0.92-0.99$  – Provan et al. 1998, Robledo-Arnuncio et al. 2004a, Scalfi et al. 2009) although somewhat higher than those reported from Finland ( $CH_E = 0.56$  – García-Gil et al. 2015). Some authors have hypothesised multiple origins for the Scottish Scots pine population, e.g., through contribution of cryptic glacial refugia (Forrest 1980, Kinloch et al. 1986, Sinclair et al. 1998, Provan et al. 1998). If true, this might help to explain the high levels of diversity we observed, reflecting a process of admixture following secondary contact.

We detected low levels of population differentiation, which represented less than 2% of the genetic variation among populations for both nuclear and chloroplast markers (Tab. 4). Outcrossing, long-lived trees with wind-mediated gene dispersal mechanisms usually harbour more diversity within populations than among them (Hamrick et al. 1992). High gene flow among populations counteracts differentiation due to drift and maintains levels of di-

versity (Slatkin 1985). Levels of differentiation among populations were very similar for both marker types (1.89% vs. 1.79% for chloroplast and nuclear markers respectively). Although chloroplast loci are expected to show greater differentiation than nuclear loci due to uniparental inheritance, smaller effective population size and higher susceptibility to genetic drift (Lendvay et al. 2014). However, the extensive pollen flow characteristic of conifers is likely to smooth such differences.

Despite evidence of extensive gene flow and consequently weak overall genetic structure, some significant differentiation among populations was apparent. In particular, the most western population SHI was most differentiated from other populations (greatest  $F_{ST}$ ). The distinctiveness of this population has been reported by other authors (Forrest 1980, Kinloch et al. 1986, Sinclair et al. 1998), although the basis for the difference remains uncertain. Some authors have observed unique, low frequency organelle haplotypes in western populations (Kinloch et al. 1986), suggesting contributions from a western refugium and possibly accounting for its distinctiveness. However, other local factors could explain the difference and resolution of the question awaits markers with sufficient power to conclusively characterise the postglacial colonisation patterns within Scotland.

### Weak isolation by distance

Our results showed a weak but significant pattern of IBD across Scotland. IBD occurs as a consequence of limited gene dispersal such that populations close to each other tend to be more genetically similar than populations farther apart (Wright 1943). Several factors are likely to limit seed and pollen flow among the Scottish populations of Scots pine. Firstly, the recently characterised asynchrony in pollen phenology between eastern and western populations within Scotland (Whittet et al. 2017) might limit gene transfer among populations even if pollen dispersal is physically possible across the distances concerned. Secondly, the drastic reduction in the effective population sizes and increased population isolation resulting from widespread deforestation in Scotland (Mason et al. 2004) may simply have reduced pollen availability to individual populations. Thirdly, the prevailing wind direction at the time of pollen release in Scotland is from the southwest (Dore et al. 2006). The resulting directional bias in pollen flow may enforce pollen limitation in populations to the west. Given the weak magnitude of IBD and since the described factors might probably contribute simultaneously to the observed pattern, we cannot unambiguously ascribe IBD patterns to a particular mechanism. On the other hand, if the weak magnitude of the IBD signal is driven by relatively recent pollen-release mismatches, then IBD might be expected to be exacerbated in future generations. Gene flow is

nevertheless clearly effective in minimising isolation, and IBD is limited.

Some of the Scottish populations showed significant levels of inbreeding ( $F_{IS} = 0.11-0.21$ ), indicating an excess of homozygotes. Other studies obtaining similar inbreeding values for Scots pine populations (0.07-0.22 – Scalfi et al. 2009, García-Gil et al. 2015) attributed the homozygote excess to the presence of null alleles (Scalfi et al. 2009). In our study, only one locus showed evidence of null alleles with a low frequency, therefore this seems unlikely to explain our  $F_{IS}$  estimates. Homozygote excess can be also the result of assortative mating, selection against heterozygotes, the Wahlund effect (subpopulation structure – Wahlund 1928, García-Gil et al. 2015), or drastic reduction in effective population sizes (Bagnoli & Buonamici 2009). However, as nearby populations obtained very different levels of inbreeding (e.g., 0.196 and 0.046 for CRA and CCC, respectively), assortative mating might not be the reason as it is unlikely to occur in adjacent populations because most trees flower at the same time. Furthermore, it seems unlikely to be selection against heterozygotes or a Wahlund effect given that nothing in our data or recent population history point to the existence of unrecognised population substructure. Therefore, the most likely explanation would be a drastic reduction in effective population sizes, as has been previously noted (Mason et al. 2004, González-Díaz et al. 2017).

### Geographic diversity gradients and predominant patterns of gene flow

Our results indicated increased nuclear genetic diversity towards the East. This West-East (W-E) trend does not fit theoretical expectations based on inferred patterns of post-glacial colonisation, which for Scots pine, and most other native species in Britain, has occurred from south to north (Birks 1989). Such a pattern of colonization is expected to leave greater levels of genetic diversity in the south than in the more recently colonised north (Hewitt 1999), and has been found in *Fagus sylvatica* in Britain (Sjölund et al. 2017), and in other tree species in Ireland (Kelleher et al. 2004) and mainland Europe (Petit et al. 2002). Equally, the W-E diversity trend does not follow the “central-marginal” hypothesis, which predicts reduced neutral genetic diversity and higher population differentiation towards distribution limits (Eckert et al. 2008). Where the S-N model predicts greater diversity in southern areas, the central-marginal model predicts greater diversity in the centre of the distribution, as populations located further from the centre (even in the south) are less connected. Levels of diversity from our populations, which represent the north-western distribution limit of Scots pine, were similar to those in mainland populations (Naydenov et al. 2011, Provan et al. 1998, Robledo-Arnuncio et al. 2004a), and no signatures of

migration were observed (i.e., a S-N or SE-NW trend). Therefore, it may be that the observed W-E diversity trend reflects more recent gene flow process. On the W-E axis, a predominant driver of gene flow in a species with wind-mediated gene dispersal (seed and pollen) is likely to be wind direction. In Scotland the predominant wind direction is west-south-westerly (Dore et al. 2006), and this would seem to be a likely explanation for the observed pattern.

While levels of diversity will be substantially dependent on past population sizes, asymmetric gene dispersal due to prevailing wind direction can also play an important role in shaping the current distribution of genetic diversity (Ennos 1997a). Gene flow patterns were not statistically asymmetric given that gene flow in Scots pine is extensive and can occur over substantial distances (Lindgren et al. 1995, Robledo-Arnuncio et al. 2004b). As there is no reason to suggest the existence of geographical or topographical barriers to gene dispersal, gene flow is likely to link nearby populations in any direction. However a strongly dominant wind direction (Dore et al. 2006) would likely favour consistent gene flow from western to eastern populations but not the converse, resulting over time in a gradient of diversity with lower levels in the populations to which gene flow is more restricted. This effect would be reinforced by the fact that western populations are at the very edge of the distribution, with only ocean beyond them, preventing gene input from upwind. The lack of West-East pattern in allele richness might be explained by the diverging trend in some populations, such as CRA, BWR, MEG and ROT, which showed low allele richness but high nuclear diversity. The explanation for this diverging trend might be due to selection or local bottlenecks (i.e., due to a substantial reduction in population size), which are characterized by large losses in allelic richness but only slight decrease of diversity if population size rebounds rapidly (Comps et al. 2001). Since no West-East pattern was observed in chloroplast haplotype diversity, seed dispersal may be more important than pollen in driving any such directionality in gene flow, an implication that tallies with the likely mobility of the two propagule types.

## Conclusions

Native pine forest in Scotland suffered a very substantial historic reduction in abundance. Despite this reduction and the resulting geographical isolation of populations, high levels of genetic variation and low levels of population differentiation still persist, suggesting that effective population size, together with extensive gene flow, has been high enough to limit the effects of genetic drift. This finding highlights the importance of maintaining large effective population sizes, especially in geographically marginal populations, to increase the probability of forest persis-

tence.

Despite potential barriers to gene flow such as population fragmentation or phenological asynchrony, gene flow among populations can still be sufficient to counteract their genetic isolation. However a weak signal of isolation by distance was detectable among the Scottish populations, suggesting that some spatial limitation of gene dispersal occurs, although gene flow is extensive. The detected gene flow patterns and geographic distribution of genetic variation were consistent with gene dispersal limitation due to prevailing wind patterns. From a practical point of view, taking into account such landscape impacts on genetic diversity is important when designing afforestation strategies or determining priorities in conservation and management plans.

Although western populations had relatively lower nuclear diversity, and there was greater differentiation and directional bias of gene flow towards the East, there was no evidence to suggest that any of the populations analysed here are genetically at risk. However, over recent decades, there has been extensive establishment of Scots pine plantations throughout the country and it would be interesting to understand the impact such plantations might have on the diversity and structure of subsequent generations of native Scots pine.

## Acknowledgements

We thank the Scottish Forestry Trust, University of Stirling and Centre for Ecology and Hydrology for co-funding this research. We are grateful to COST action FP102 MaP-FGR (<http://map-fgr.entecra.it/>) for supporting researcher mobility during this project.

## References

- Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* 1: 95-111. - doi: [10.1111/j.1752-4571.2007.00013.x](https://doi.org/10.1111/j.1752-4571.2007.00013.x)
- Auckland LD, Bui T, Zhou Y, Shepard M, Williams CG (2002). *Conifer microsatellite handbook*. Corporate Press, Raleigh, NC, USA, pp. 57. [online] URL: <http://www.cabdirect.org/cabdirect/abstract/20043034208>
- Bagnoli F, Buonamici A (2009). Is *Cupressus sempervirens* native in Italy? An answer from genetic and palaeobotanical data. *Molecular Ecology* 18: 2276-2286. - doi: [10.1111/j.1365-294X.2009.04182.x](https://doi.org/10.1111/j.1365-294X.2009.04182.x)
- Belletti P, Ferrazzini D, Piotti A, Monteleone I, Ducci F (2012). Genetic variation and divergence in Scots pine (*Pinus sylvestris* L.) within its natural range in Italy. *European Journal of Forest Research* 131: 1127-1138. - doi: [10.1007/s10342-011-0584-3](https://doi.org/10.1007/s10342-011-0584-3)
- Bernhardsson C, Floran V, Ganea SL, García-Gil MR (2016). Present genetic structure is congruent with the common origin of distant Scots pine populations in its Romanian distribution. *Forest Ecology and Management* 361: 131-143. -

doi: [10.1016/j.foreco.2015.10.047](https://doi.org/10.1016/j.foreco.2015.10.047)

- Birks HJB (1989). Holocene isochrone maps and patterns of tree-spreading in the British Isles. *Journal of Biogeography* 16: 503-540. - doi: [10.2307/2845208](https://doi.org/10.2307/2845208)
- Burczyk J, Difazio SP, Adams WT (2004). Gene flow in forest trees: How far do genes really travel? *Forest Genetics* 11: 179-192. [online] URL: <http://www.as.wvu.edu/~sdifazio/peer8.pdf>
- Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera MT, Vendramin GG, García V, Frigerio J-M, Echt C, Richardson T, Plomion C (2004). Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *Theoretical and Applied Genetics* 109: 1204-1214. - doi: [10.1007/s00122-004-1683-z](https://doi.org/10.1007/s00122-004-1683-z)
- Chapuis MP, Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621-631. - doi: [10.1093/molbev/msl191](https://doi.org/10.1093/molbev/msl191)
- Comps B, Gömöry D, Letouzey J, Thiébaud B, Petit RJ (2001). Diverging trends between heterozygosity and allelic richness during post-glacial colonization in the European Beech. *Genetics* 157: 389-397. [online] URL: <http://www.genetics.org/content/157/1/389.short>
- Davies SJ, Cavers S, Finegan B, White A, Breed MF, Lowe AJ (2013). Pollen flow in fragmented landscapes maintains genetic diversity following stand-replacing disturbance in a neotropical pioneer tree, *Vochysia ferruginea* Mart. *Heredity* 115: 125-129. - doi: [10.1038/hdy.2013.95](https://doi.org/10.1038/hdy.2013.95)
- Deacon NJ, Cavender-Bares J (2015). Limited pollen dispersal contributes to population genetic structure but not local adaptation in *Quercus oleoides* forests of Costa Rica. *PLoS ONE* 10 (9): e0138783. - doi: [10.1371/journal.pone.0138783](https://doi.org/10.1371/journal.pone.0138783)
- Dore AJ, Vieno M, Fournier N, Weston KJ, Sutton MA (2006). Development of a new wind-rose for the British Isles using radiosonde data, and application to an atmospheric transport model. *Quarterly Journal of the Royal Meteorological Society* 132: 2769-2784. - doi: [10.1256/qj.05.198](https://doi.org/10.1256/qj.05.198)
- Earl DA, Von Holdt BM (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361. - doi: [10.1007/s12686-011-9548-7](https://doi.org/10.1007/s12686-011-9548-7)
- Eckert CG, Samis KE, Lougheed SC (2008). Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* 17: 1170-1188. - doi: [10.1111/j.1365-294X.2007.03659.x](https://doi.org/10.1111/j.1365-294X.2007.03659.x)
- Ennos RA, Dodson RK (1987). Pollen success, functional gender and assortative mating in an experimental plant population. *Heredity* 58: 119-126. - doi: [10.1038/hdy.1987.16](https://doi.org/10.1038/hdy.1987.16)
- Ennos RA (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* 72: 250-259. - doi: [10.1038/hdy.1994.35](https://doi.org/10.1038/hdy.1994.35)
- Ennos AR (1997). Wind as an ecological factor. *Trends in Ecology and Evolution* 12: 108-111. - doi: [10.1016/S0169-5347\(96\)10066-5](https://doi.org/10.1016/S0169-5347(96)10066-5)
- Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567. - doi: [10.1111/j.1755-0998.2010.02847.x](https://doi.org/10.1111/j.1755-0998.2010.02847.x)
- Fady B, Aravanopoulos FA, Alizoti P, Mátyás C,



- Von Wühlisch G, Westergren M, Belletti P, Cvjetkovic B, Ducci F, Huber G, Kelleher CT, Khaldi A, Kharraat MBD, Kraigher H, Kramer K, Mühlethaler U, Peric S, Perry A, Rousi M, Sbay H, Stojnic S, Tijardovic M, Tsvetkov I, Varela MC, Vendramin GG, Zlatanov T (2016a). Evolution-based approach needed for the conservation and silviculture of peripheral forest tree populations. *Forest Ecology and Management* 375: 66-75. - doi: [10.1016/j.foreco.2016.05.015](https://doi.org/10.1016/j.foreco.2016.05.015)
- Fady B, Cottrell J, Ackzell L, Alía R, Muys B, Prada A, González-Martínez SC (2016b). Forests and global change: what can genetics contribute to the major forest management and policy challenges of the twenty-first Century? *Regional Environmental Change* 16: 927-939. - doi: [10.1007/s10113-015-0843-9](https://doi.org/10.1007/s10113-015-0843-9)
- FAO (2014). The state of the world's forest genetic resources. Commission on genetic resources for food and agriculture, Rome, Italy, pp. 276.
- Forrest GI (1980). Genotypic variation among native Scots pine populations in Scotland based on monoterpene analysis. *Forestry* 53: 101-128. - doi: [10.1093/forestry/53.2.101](https://doi.org/10.1093/forestry/53.2.101)
- Forrest GI (1982). Relationship of some European Scots pine populations to native Scottish woodlands based on monoterpene analysis. *Forestry* 55: 19-37. - doi: [10.1093/forestry/55.1.19](https://doi.org/10.1093/forestry/55.1.19)
- García-Gil MR, Floran V, Ostlund L, Mullin TJ, Gull BA (2015). Genetic diversity and inbreeding in natural and managed populations of Scots pine. *Tree Genetics and Genomes* 11 (2): 41. - doi: [10.1007/s11295-015-0850-5](https://doi.org/10.1007/s11295-015-0850-5)
- Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010). Calculations of population differentiation based on  $G_{ST}$  and  $D$ : forget  $G_{ST}$  but not all of statistics. *Molecular Ecology* 19: 3845-3852. - doi: [10.1111/j.1365-294X.2010.04784.x](https://doi.org/10.1111/j.1365-294X.2010.04784.x)
- González-Díaz P, Jump AS, Perry A, Wachowiak W, Lapshina E, Cavers S (2017). Ecology and management history drive spatial genetic structure in Scots pine. *Forest Ecology and Management* 400: 68-76. - doi: [10.1016/j.foreco.2017.05.035](https://doi.org/10.1016/j.foreco.2017.05.035)
- Goudet J (1995). Computer note. *Journal of Heredity* 86: 485-486. - doi: [10.1093/oxfordjournals.jhered.a111627](https://doi.org/10.1093/oxfordjournals.jhered.a111627)
- Graudal L, Aravanopoulos F, Bennadji Z, Changtragoon S, Fady B, Kjr ED, Loo J, Ramamonjisoa L, Vendramin GG (2014). Global to local genetic diversity indicators of evolutionary potential in tree species within and outside forests. *Forest Ecology and Management* 333: 35-51. - doi: [10.1016/j.foreco.2014.05.002](https://doi.org/10.1016/j.foreco.2014.05.002)
- Hamrick JL, Godt MJW, Sherman-Broyles SL (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95-124. - doi: [10.1007/BF00120641](https://doi.org/10.1007/BF00120641)
- Hardy OJ, Vekemans X (2002). Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618-620. - doi: [10.1046/j.1471-8286.2002.00305.x](https://doi.org/10.1046/j.1471-8286.2002.00305.x)
- Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87-112. - doi: [10.1111/j.1095-8312.1999.tb01160.x](https://doi.org/10.1111/j.1095-8312.1999.tb01160.x)
- Holderegger R, Kamm U, Gugerli F (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* 21: 797-807. - doi: [10.1007/s10980-005-5245-9](https://doi.org/10.1007/s10980-005-5245-9)
- Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013). DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 4: 782-788. - doi: [10.1111/2041-210X.12067](https://doi.org/10.1111/2041-210X.12067)
- 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. - doi: [10.1016/j.foreco.2003.07.032](https://doi.org/10.1016/j.foreco.2003.07.032)
- Kinloch B, Westfall RD, Forrest GI (1986). Caledonian Scots pine: origins and genetic structure. *New Phytologist* 104: 703-729. - doi: [10.1111/j.1469-8137.1986.tb00671.x](https://doi.org/10.1111/j.1469-8137.1986.tb00671.x)
- Lendvay B, Hohn M, Brodbeck S, Mindrescu M, Gugerli F (2014). Genetic structure in *Pinus cembra* from the Carpathian mountains inferred from nuclear and chloroplast microsatellites confirms post-Glacial range contraction and identifies introduced individuals. *Tree Genetics and Genomes* 10: 1419-1433. - doi: [10.1007/s11295-014-0770-9](https://doi.org/10.1007/s11295-014-0770-9)
- Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA, Ritland K (2004). Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* 109: 361-369. - doi: [10.1007/s00122-004-1635-7](https://doi.org/10.1007/s00122-004-1635-7)
- Lindgren D, Paule L, Xihuan S, Yazdani R, Segerström U, Wallin J-E, Lejdebö ML (1995). Can viable pollen carry Scots pine genes over long distances? *Grana* 34: 64-69. - doi: [10.1080/00173139509429035](https://doi.org/10.1080/00173139509429035)
- Mason WL, Hampson A, Edwards C (2004). Managing the pinewoods of Scotland. Forestry Commission, Edinburgh, UK, pp. 234. [online] URL: <http://www.cabdirect.org/cabdirect/abstract/20053055913>
- Naydenov KD, Naydenov MK, Tremblay F, Alexandrov A, Aubin-Fournier LD (2011). Patterns of genetic diversity that result from bottlenecks in Scots Pine and the implications for local genetic conservation and management practices in Bulgaria. *New Forests* 42: 179-193. - doi: [10.1007/s11056-010-9245-5](https://doi.org/10.1007/s11056-010-9245-5)
- Nei M (1987). *Molecular evolutionary genetics*. Columbia University Press, New York, pp. 512.
- Nowakowska JA, Zachara T, Konecka A (2014). Genetic variability of Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* L. Karst.) natural regeneration compared with their maternal stands. *Forest Research Papers* 75: 47-54. - doi: [10.2478/frp-2014-0005](https://doi.org/10.2478/frp-2014-0005)
- Petit J, Csaikl UM, Bordács S, Burg K, Coart E, Cottrell J, Van Dam B, Deans JD, Dumolin-Lapègue S, Fineschi S, Finkeldey R, Gillies A, Glaz I, Goicoechea P, Jensen J, König A, Lowe AJ, Madsen SF, Matyas G, Munro RC, Olalde M, Pomonge M-H, Popescu F, Slade D, Tabbener H, Turchini D, De Vries S, Ziegenhagen B, Kremer A (2002). Chloroplast DNA variation in European white oaks. *Phylogeography and patterns of diversity based on data from over 2600 populations*. *Forest Ecology and Management* 156: 5-26. - doi: [10.1016/S0378-1127\(01\)00645-4](https://doi.org/10.1016/S0378-1127(01)00645-4)
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959. [online] URL: <http://www.genetics.org/content/155/2/945>
- Provan J, Soranzo N, Wilson NJ, McNicol JW, Forrest GI, Cottrell JE, Powell W (1998). Gene pool variation in caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. *Proceedings of the The Royal Society, Biological sciences* 265: 1697-1705. - doi: [10.1098/rspb.1998.0491](https://doi.org/10.1098/rspb.1998.0491)
- Provan J, Beatty GE, Hunter AM, McDonald RA, McLaughlin E, Preston SJ, Wilson S (2007). Restricted gene flow in fragmented populations of a wind-pollinated tree. *Conservation Genetics* 9: 1521-1532. - doi: [10.1007/s10592-007-9484-y](https://doi.org/10.1007/s10592-007-9484-y)
- Robledo-Arnuncio JJ, Alía R, Gil L (2004a). High levels of genetic diversity in a long-term European glacial refugium of *Pinus sylvestris* L. *Forest Genetics* 11: 239-248. [online] URL: <http://agris.fao.org/agris-search/search.do?recordID=SK2006T00417>
- Robledo-Arnuncio JJ, Smouse PE, Gil L, Alía R (2004b). Pollen movement under alternative silvicultural practices in native populations of Scots pine (*Pinus sylvestris* L.) in central Spain. *Forest Ecology and Management* 197: 245-255. - doi: [10.1016/j.foreco.2004.05.016](https://doi.org/10.1016/j.foreco.2004.05.016)
- Rousset F (1997). Genetic differentiation and estimation of gene flow from F-Statistics under isolation by distance. *Genetics* 145: 1219-1228. [online] URL: <http://www.genetics.org/content/145/4/1219.short>
- Salmela MJ (2011). Adaptive genetic variation in Scots pine (*Pinus sylvestris* L.) in Scotland. PhD Thesis, The University of Edinburgh, Edinburgh, UK, pp. 159. [online] URL: <http://www.era.lib.ed.ac.uk/handle/1842/5662>
- Scalfi M, Piotti A, Rossi M, Piovani P (2009). Genetic variability of Italian southern Scots pine (*Pinus sylvestris* L.) populations: the rear edge of the range. *European Journal of Forest Research* 128: 377-386. - doi: [10.1007/s10342-009-0273-7](https://doi.org/10.1007/s10342-009-0273-7)
- Sinclair WT, Morman JD, Ennos RA (1998). Multiple origins for Scots pine (*Pinus sylvestris* L.) in Scotland: evidence from mitochondrial DNA variation. *Heredity* 80: 233-240. - doi: [10.1046/j.1365-2540.1998.00287.x](https://doi.org/10.1046/j.1365-2540.1998.00287.x)
- Sjölund MJ, González-Díaz P, Moreno-Villa JJ, Jump AS (2017). Understanding the legacy of widespread population translocations on the post-glacial genetic structure of the European beech, *Fagus sylvatica* L. *Journal of Biogeography* 44: 2475-2487. - doi: [10.1111/jbi.13053](https://doi.org/10.1111/jbi.13053)
- Slatkin M (1985). Gene flow in natural populations. *Annual Review of Ecology, Evolution, and Systematics* 16: 393-430. - doi: [10.1146/annurev.es.16.110185.002141](https://doi.org/10.1146/annurev.es.16.110185.002141)
- Soranzo N, Provan J, Powell W (1998). Characterization of microsatellite loci in *Pinus sylvestris* L. *Molecular Ecology* 7: 1260-1261.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538. - doi: [10.1111/j.1471-8286.2004.00684.x](https://doi.org/10.1111/j.1471-8286.2004.00684.x)
- Varis S, Pakkanen A, Galofré A, Pulkkinen P (2009). The extent of south-north pollen transfer in Finnish Scots pine. *Silva Fennica* 43: 717-726. - doi: [10.14214/sf.168](https://doi.org/10.14214/sf.168)
- Wachowiak W, Salmela MJ, Ennos RA, Iason G,

Cavers S (2011). High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland. *Heredity* 106: 775-787. - doi: [10.1038/hdy.2010.118](https://doi.org/10.1038/hdy.2010.118)

Wachowiak W, Iason GR, Cavers S (2013). Among population differentiation at nuclear genes in native Scots pine (*Pinus sylvestris* L.) in Scotland. *Flora* 208: 79-86. - doi: [10.1016/j.flora.2012.12.009](https://doi.org/10.1016/j.flora.2012.12.009)

Wahlund S (1928). The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity. *Hereditas* 11: 65-106. - doi: [10.1111/j.1601-5223.1928.tb02483.x](https://doi.org/10.1111/j.1601-5223.1928.tb02483.x)

Whittet R, Cavers S, Cottrell J, Rosique-Esplugas C, Ennos R (2017). Substantial variation in the timing of pollen production reduces reproductive synchrony between distant populations of *Pinus sylvestris* L. in Scotland. *Ecology and Evolution* 7: 5754-5765. - doi: [10.1002/ece3.3154](https://doi.org/10.1002/ece3.3154)

Wright S (1943). Isolation by distance. *Genetics* 28: 114-138.

### Supplementary Material

**Tab. S1** - Pairwise population differentiation ( $F_{ST}$ ) (below diagonal) and Jost's differentiation index,  $D$  (above diagonal).

**Fig. S1** - Percentage of differentiated sites within Scotland using nSSR, calculated as the percentage of total sites that were significantly differentiated based on  $F_{ST}$  values.

**Fig. S2** - The number of genetic clusters ( $K=2$ ) identified by STRUCTURE for nSSR.

**Fig. S3** - Relative migration networks for nSSR with populations sorted in seed groups.

**Link:** [Gonzalez-Diaz\\_2454@suppl001.pdf](mailto:Gonzalez-Diaz_2454@suppl001.pdf)