



Research Article

Phylogeography of Scots pine in Europe and Asia based on mtDNA polymorphisms

Witold Wachowiak^{1,2*} , Weronika B. Żukowska², Annika Perry³, Andrzej Lewandowski², Stephen Cavers³, and Bartosz Łabiszak¹ 

¹Institute of Environmental Biology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 6, Poznań 61-614, Poland

²Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, Kórnik 62-035, Poland

³UK Centre for Ecology and Hydrology Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

*Author for correspondence. E-mail: witwac@amu.edu.pl

Received 16 February 2022; Accepted 25 July 2022; Article first published 30 July 2022

Abstract We analyzed mitochondrial DNA polymorphisms to search for evidence of the genetic structure and patterns of admixture in 124 populations ($N = 1407$ trees) across the distribution of Scots pine in Europe and Asia. The markers revealed only a weak population structure in Central and Eastern Europe and suggested postglacial expansion to middle and northern latitudes from multiple sources. Major mitotype variants include the remnants of Scots pine at the north-western extreme of the distribution in the Scottish Highlands; two main variants (western and central European) that contributed to the contemporary populations in Norway and Sweden; the central-eastern European variant present in the Balkan region, Finland, and Russian Karelia; and a separate one common to most eastern European parts of Russia and western Siberia. We also observe signatures of a distinct refugium located in the northern parts of the Black Sea basin that contributed to the patterns of genetic variation observed in several populations in the Balkans, Ukraine, and western Russia. Some common haplotypes of putative ancient origin were shared among distant populations from Europe and Asia, including the most southern refugial stands that did not participate in postglacial recolonization of northern latitudes. The study indicates different genetic lineages of the species in Europe and provides a set of genetic markers for its finer-scale population history and divergence inference.

Key words: admixture, divergence, glacial refugia, mitochondrial DNA, *Pinus sylvestris*, population structure, postglacial recolonization.

1 Introduction

In Europe, tree populations have experienced recurring waves of extinction and migration driven by glacial cycles, and the consequences of these processes for contemporary genetic variation are still being identified (Hewitt, 2000; Petit et al., 2008). Although broad patterns of northern extinction and southern persistence are evident for many species, variations in the locations of refugial populations, and interactions between rates of migration, the distribution of suitable habitat, and topography have produced a wide range of outcomes for different species (Petit et al., 2003). For boreal species, in particular, which are potentially capable of having survived glacial maxima far to the north of recognized refugial areas, reconstruction of postglacial demographics in Europe has been difficult. However, to advance our understanding of the genetic basis of phenotypic divergence and adaptation within species, or

develop predictive models to forecast the likely impacts of environmental changes on our forests, the nature and dynamics of population structures in tree species must be more clearly resolved.

Scots pine (*Pinus sylvestris* L.) is a species of high ecological and economic importance that comprises >20% of the productive forest area in Europe. It has the broadest distribution of all pines, with natural populations covering large parts of Europe and Asia from western Scotland to eastern Siberia (San-Miguel-Ayanz et al., 2016). The ecological niche of Scots pine is also broad and spans several climatic zones and crosses more than 30 degrees of latitude from southern Spain (37°N) to the Arctic (70°N) and from sea level to >2500 m above sea level. Despite the vast geographic and ecological breadth, and numerous recognized ecotypes reflecting its phenotypic and ecophysiological variations (Giertych, 1991; Oleksyn et al., 1992; Shutyaev & Giertych,

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

2000; Zadworny et al., 2016), Scots pine remains taxonomically homogeneous. Clear clinal patterns of adaptation to local environmental conditions have been established since postglacial migration, as evident in growth traits, phenology, abiotic stress resistance, or responses to light spectra (Hurme et al., 1997; Savolainen et al., 2007; Salmela et al., 2011; Ranade & García-Gil, 2013; Wachowiak et al., 2017; Ramírez-Valiente et al., 2021). Patterns of population differentiation are generally clinal across the range and vary with latitude (Hurme et al., 1997; Kujala & Savolainen, 2012), although local adaptation at a fine scale is evident where landscapes are environmentally heterogeneous across short spatial scales (Salmela et al., 2011; Donnelly et al., 2016; Perry et al., 2016).

Although the patterns of Scots pine morphological variations are known, the underlying demographic history of populations contributing to the establishment of the present-day distribution is not fully resolved. Pollen and macrofossil data indicate that the widespread paradigm of retreat to southern refugia during the last glacial period does not provide the full picture for Scots pine and that low-density persistence at high northern latitudes, cryptic refugia, and early postglacial dispersal are all likely to have played a role (Birks, 1989; Willis & van Andel, 2004; Cheddadi et al., 2006; Parducci et al., 2012; Tzedakis et al., 2013). Generally, low between-population genetic differentiation has been found across the distribution using isozymes, microsatellites, and nuclear sequence variation (Prus-Glowacki & Stephan, 1994; Cheddadi et al., 2006; Wachowiak et al., 2014), probably resulting from widespread species persistence during the Last Glacial Maximum (LGM) and its high dispersal ability by seeds and pollen (Robledo-Arnuncio et al., 2004). Genetic approaches using sequence polymorphism in nuclear genes and coalescent simulation approaches indicate an ancient bottleneck of Scots pine populations in Europe; however, this event likely predates the last glacial period (Pyhäjärvi et al., 2007).

Population structure assessments using maternally inherited seed-dispersed mitochondrial (mtDNA) markers have been largely limited to variations in a few polymorphic markers, mostly the *nad1* and *nad7* regions (Sinclair et al., 1998; Sinclair et al., 1999; Soranzo et al., 2000; Cheddadi et al., 2006; Naydenov et al., 2007; Pyhäjärvi et al., 2007). However, due to the low spatial resolution of these markers, population structure has only been resolved at a broad scale (Soranzo et al., 2000; Naydenov et al., 2007; Pyhäjärvi et al., 2008; Wójcikiewicz et al., 2016a; Dering et al., 2017; Semerikov et al., 2018; Zimmer & Sønstebo, 2018). Southern refugial populations were certainly present and are consistently evident in marker data sets as locations of distinctive mitotypes in Iberia, Italy, and Turkey (Cheddadi et al., 2006; Dering et al., 2017; Donnelly et al., 2017; Semerikov et al., 2020). North of these locations, little genetic structure is evident, as much of western Europe was found to be dominated by a single mitotype, while a distinctive group of populations was evident from elevated frequencies of a different mitotype in eastern Sweden, Finland, and western Russia (Sinclair et al., 1999; Soranzo et al., 2000; Cheddadi et al., 2006; Naydenov et al., 2007; Pyhäjärvi et al., 2008; Buchovska et al., 2013; Dering et al., 2017; Semerikov et al., 2018). Additional work has been performed with more

variable pollen mediated, paternally inherited chloroplast DNA markers, and with nuclear DNA markers. However, these are generally less suitable for phylogeographic inference in pines as compared to seed-dispersed, maternally inherited mtDNA markers, due to their high dispersal ability and homoplasmy in the case of cpSSRs, which tends to obscure the structure at wide spatial scales. These markers have shown little structure and primarily demonstrate that gene flow among Scots pine populations is extensive (Wachowiak et al., 2009; Belletti et al., 2012; Kujala & Savolainen, 2012; Dering et al., 2017; Tóth et al., 2017; Tyrmi et al., 2020).

Novel genomic resources have recently been developed for Scots pine and its close relatives with the use of next-generation sequencing approaches (Wachowiak et al., 2015; Donnelly et al., 2017). They include information on the polymorphisms in large numbers of fragments of resequenced mitochondrial DNA (Donnelly et al., 2017), which were used for the development of a set of novel mitochondrial markers (Wachowiak et al., 2018; Łabiszak et al., 2019; Zaborowska et al., 2020). In this study, we aimed to use polymorphisms in seed-mediated and maternally transmitted mitochondrial DNA to investigate genetic variations within and among Scots pine populations across its range. We applied the markers to search for evidence of genetic structure of the population in Scots pine and the possible admixture of populations of different origins that might have contributed to the establishment of different ecotypes of the species in Europe and Asia.

2 Material and Methods

2.1 Sampling, DNA extraction, and PCR amplification

Plant material was collected from 124 autochthonous populations of Scots pine in Europe and Asia: 1407 trees were analyzed, comprising 5–30 different individuals from each location (Table 1; Fig. 1). Genomic DNA was extracted from needles or a single haploid megagametophyte per mother tree using a DNeasy Plant Mini Kit (Qiagen) and a Genomic Mini AX Plant kit (A&A Biotechnology) following the standard manufacturer protocol, and its quality was evaluated using a BioPhotometer (Eppendorf). A set of 10 polymorphic mtDNA markers (*PR5*, *PR7*, *PR15*, *PR19*, *PR20*, *PR21*, *PR24*, *PR25*, *PR29*, *PR31*) and one insertion/deletion locus in *nad1* intron B/C were genotyped (Table S1) according to the methods described in Łabiszak et al. (2019) and Soranzo et al. (2000), respectively. Full details of PCR amplification and genotyping of the markers are provided in the Supporting Information.

2.2 Nucleotide and haplotype variations

Multilocus genotypes were recorded for all 11 polymorphic markers. The level of missing data was assessed; however, no sample exceeded the cut-off level set at 10% and thus all individuals were retained. For haplotype analysis, in 198 individuals, single missing positions in genotypes were imputed according to the most similar haplotypes and/or the variants dominating in the same population: the creation of *de novo* haplotypes was avoided. For subsequent analyses, the indel in the *nad1* locus was coded as a single mutation event. The spatial distribution of mtDNA haplotypes

Table 1 Location of the investigated populations

Acr.	Location	N	Lat.	Long.	Acr.	Location	N	Lat.	Long.
FL.1	Finland, Joutsa	24	61.74	26.14	PL.20	Poland, Miłomłyn	10	53.75	19.83
FL.2	Finland, nr Temmes	24	64.69	25.71	PL.21	Poland, Tarnawa	10	49.10	22.49
FL.3	Finland, nr Rovaniemi	24	66.57	26.21	PL.22	Poland, Liski reserve	10	50.93	23.87
FL.4	Finland, nr Kielajoki	24	69.65	29.07	PL.23	Poland, Jarocin	10	51.97	17.48
FL.5	Finland, Punkaharju	10	61.76	29.39	ES.1	Spain, Cuenca Una	12	40.44	1.98
FL.6	Finland, Kolari	10	67.18	24.05	ES.2	Spain, Sierra de Gudar,	12	40.36	0.72
FL.7	Finland, Usjoentie	10	69.55	27.23	ES.3	Spain, Cuenca	12	40.44	1.90
FL.8	Finland, Ruka	10	66.18	29.15	ES.4	Spain, Sierra Nevada	12	37.13	3.43
FL.9	Finland, Viitostie	10	65.44	29.03	ES.5	Spain, Trevenque	10	37.10	-3.55
FL.10	Finland, Oulu	10	65.07	25.46	ES.6	Spain, Valsain	10	40.87	-4.04
FL.11	Finland, Savonlinna	10	61.84	28.96	SCO.1	Scotland, Abernethy	10	57.21	-3.61
FL.12	Finland, road to Rovaniemi	10	66.46	25.59	SCO.2	Scotland, Allt Cul	10	57.04	-3.35
FL.13	Finland, Rovaniemi ACHA	10	66.60	26.12	SCO.3	Scotland, Amat	10	57.87	-4.60
FL.14	Finland, Ivalontie	10	67.71	27.02	SCO.4	Scotland, Ballochuie	10	56.99	-3.30
FL.15	Finland, Ivalontie	10	68.15	27.10	SCO.5	Scotland, Beinn Eighe	10	57.63	-5.35
FL.16	Finland, Siida	10	68.91	27.07	SCO.6	Scotland, Black	5	56.67	-4.32
FL.17	Finland, Usjoentie	10	69.88	27.00	SCO.7	Scotland, Coille Coire Chuilc	10	56.41	-4.71
FL.18	Finland, Kamaen	10	69.34	27.22	SCO.8	Scotland, Cona Glen	10	56.79	-5.33
FL.19	Finland, Noryskontie	10	67.74	29.02	SCO.9	Scotland, Crannach	10	56.58	-4.68
FL.20	Finland, Salla Reindeer Park	10	66.74	28.82	SCO.10	Scotland, Glen Affric	5	57.27	-4.92
FL.21	Finland, Kuopio-Helsinki	10	61.45	26.69	SCO.11	Scotland, Glen Cannich	10	57.33	-4.96
FL.22	Finland, Naantali	10	60.47	22.01	SCO.12	Scotland, Glen Derry	10	57.03	-3.58
SE.1	Sweden, Väster Mjöingenn	6	62.75	13.58	SCO.13	Scotland, Glen Einig	10	57.95	-4.76
SE.2	Sweden, Tjärnbergsheden	9	64.62	20.80	SCO.14	Scotland, Glen Loy	7	56.91	-5.13
SE.3	Sweden, Backsjon	12	63.95	20.38	SCO.15	Scotland, Glen Tanar	5	57.05	-2.86
NO.1	Norway, Alta	12	69.99	20.38	SCO.16	Scotland, Loch Clair	10	57.56	-5.36
NO.2	Norway, Ostamarka	12	59.87	11.05	SCO.17	Scotland, Meggernie	10	56.58	-4.35
NO.3	Norway, Smoldalen	12	61.45	12.42	SCO.18	Scotland, Rhidorroch	10	57.89	-4.98
EE.1	Estonia, Vardi	10	58.98	24.48	SCO.19	Scotland, Rothiemurcys	5	57.15	-3.77
EE.2	Estonia, Pärnu rannaniidu	10	58.36	24.56	SCO.20	Scotland, Shieldaig	13	57.51	-5.64
LV	Latvia, Dunezers Lake	10	57.15	24.35	SCO.21	Scotland, Strath Oykel	10	57.98	-4.61
LT.1	Lithuania, Ciapkeliai	12	54.05	24.42	HU	Sopron	12	47.67	16.52
LT.2	Lithuania, Marcinkonys	12	54.02	24.55	RO	Dedulesti	5	45.01	24.53
FR.1	France, Arlanc-Mayres	12	45.40	3.69	BG	Pirin, Bansko Razlog	12	41.88	23.36
FR.2	France, Chabrioux	12	45.57	3.93	RS.1	Serbia, Zlatibor Mt.	12	43.68	19.67
DE.1	Germany, Godendorf	12	53.27	13.13	RS.2	Serbia, Velika Plec	12	44.18	20.03
DE.2	Germany, Joachimsthal	12	52.97	13.67	GR	Greece, Fteri	12	40.19	22.13
DE.3	Germany, Enzenauer Kopf	12	47.73	11.46	TR.1	Turkey, Savsat Ardahan	12	41.23	42.43
AT	Austria, Pernitz	10	47.91	16.00	TR.2	Turkey, Catacik	12	39.96	31.11
CH	Switzerland, Follateres	12	46.12	7.07	TR.3	Turkey, Tosya Ilgazi	12	41.12	34.06
IT.1	Italy, Casina	10	44.54	10.52	TR.4	Turkey, Artvin	12	41.15	41.76
IT.2	Italy, Casina	12	44.54	10.54	UA.1	Ukraine, Yalta	12	44.55	34.20
IT.3	Italy, Cella di Palmia	12	44.63	10.17	UA.2	Ukraine, Sokil	11	51.17	23.55
PL.1	Poland, Węgliniec reserve	14	51.28	15.23	UA.3	Ukraine, Yavorivskiy National Park	12	51.00	23.30
PL.2	Poland, Chojnik	14	50.83	15.63	UA.4	Ukraine, Luck	12	50.75	25.33
PL.3	Poland, Zieleniec reserve	15	50.35	16.41	RU.1	Russia, Bobrovka	12	53.17	83.88
PL.4	Poland, Szczeliniec	30	50.43	16.23	RU.2	Russia, Khanty-Mansi	12	62.35	70.68
PL.5	Poland, Skalniak	11	50.47	16.30	RU.3	Russia, Western Siberia	12	62.76	62.31
PL.6	Poland, Głowa Króla	16	50.48	16.38	RU.4	Russia, Goritsy	12	59.86	38.27
PL.7	Poland, Zwierzyniec	10	50.60	22.96	RU.5	Russia, Volga Kama Nature Reserve	12	55.18	49.17
PL.8	Poland, Gubin	10	51.95	14.72	RU.6	Russia, Kadnikow	8	59.56	40.47
PL.9	Poland, Barlinek	10	52.98	15.23	RU.7	Russia, Pavlovsk	20	53.29	82.94
PL.10	Poland, Pusta Wielka	10	49.40	20.82	RU.8	Russia, Karmanka	22	51.87	87.10
PL.11	Poland, Koryciska Wielkie	11	49.27	19.83	RU.9	Russia, Artybasz	11	51.79	87.25
PL.12	Poland, Pieniński National Park	12	49.42	20.36	RU.10	Russia, Minusinsk	16	53.57	91.80
PL.13	Poland, Woziwoda	10	53.67	17.92	RU.11	Russia, Tanzybey	12	53.16	92.88

Continued

Table 1 Continued

Acr.	Location	N	Lat.	Long.	Acr.	Location	N	Lat.	Long.
PL.14	Poland, Syców	10	51.30	17.72	RU.12	Russia, Stolby Nature Sanctuary	6	55.95	92.77
PL.15	Poland, Pisz	10	53.60	21.77	RU.13	Russia, Murmansk	10	68.97	33.15
PL.16	Poland, Bolewice	10	52.38	16.12	RU.14	Russia, Alu	10	67.18	32.42
PL.17	Poland, Wyszaków	12	52.67	21.48	RU.15	Russia, Yakkima	10	61.52	30.14
PL.18	Poland, Tabórz	14	53.75	20.03	RU.16	Russia, Apatyty	8	67.58	33.43
PL.19	Poland, Hajnówka	10	52.74	23.58	RU.17	Russia, Zoja	8	67.41	32.65

Acr., acronym for the population; lat., latitude (N); long., longitude (E); N, number of samples analyzed.

(mitotypes) was assessed between populations. The phylogenetic relationship among the obtained haplotypes was inspected using a median-joining network constructed with POPART v.1.7 (Bandelt et al., 1999). The levels of diversity and differentiation were calculated for each population and each of the 12 geographical regions. The regions were defined based on the geographic location of the populations and/or putative refugial or distinctive stands. These regions are (1) Spain; (2) Italy; (3) Scotland; (4) Western Europe (France, Germany, Austria, Switzerland); (5) Poland; (6) Ukraine; (7) the Balkans (Hungary, Serbia, Romania, Bulgaria, Greece); (8) the Baltic region (Lithuania, Latvia, Estonia); (9) Turkey; (10) Russia; (11) Scandinavia (Norway, Sweden); and (12) Finland (including neighboring populations RU.13-17; see Table 1). Basic statistics including the number of haplotypes detected in each population (A), number of private haplotypes (P), effective number of haplotypes (N_e), haplotypic richness (R_H), and haplotype diversity (H_d) were calculated at the population level and averaged across regions using HaplotypeAnalysis 1.05 software (Eliades & Eliades, 2009). Additionally, to determine whether there was a relationship between genetic diversity and latitude, we performed a linear regression analysis between four different measures of genetic diversity: number of haplotypes (A), effective number of haplotypes (N_e), haplotypic richness (R_H), and haplotype diversity (H_d). We used the `stat_smooth` function using the method "lm" in `ggplot2` to calculate and visualize the regression lines (Wickham, 2016). An inverse relationship between genetic diversity and latitude is commonly found for plant and animal species, especially in the Northern Hemisphere, and is an indicator of postglacial recolonization from southern refugia at the Pleistocene/Holocene boundary. We looked at this relationship for the whole data set and for each region. All calculations were performed in R v.4.0.3 (R Core Team, 2022) and visualized using `ggplot2`.

2.3 Population differentiation

We used cladistic analysis based on the unweighted pair group method with the arithmetic mean (UPGMA) to inspect the genetic relationships between populations to find whether there is a tendency of populations to group by their geographic locations. The genetic distance among populations, based on the mean number of differences across all mtDNA sites, was calculated in MEGA 7 (Kumar et al., 2016) and used in Principal Coordinate Analysis (PCoA) in GENALEX v.6.503 software (Peakall & Smouse, 2012) both at the population level and on the regional scale. Additionally, we used two different approaches to detect genetically distinct

clusters of populations. First, we used Bayesian population clustering based on linked data implemented in BAPS (Corander & Martinen, 2006; Corander & Tang, 2007) to delineate the main genetic clusters in the data set, with the maximum number of potential clusters set to 100. The population and individual admixture were subsequently tested on the basis of these population clustering results and visualized using the `pophelper` package in R (Francis, 2017). Then, to formally check for the presence of phylogeographic structure in our data set, two measures of population differentiation including G_{ST} (Nei, 1973) and N_{ST} (Lynch & Crease, 1990) were calculated and compared using a permutation test with 1000 replicates in `PermutCpSSR v.2.0` (Pons & Petit, 1996; Burban et al., 1999). The tests take into account haplotype frequency alone (G_{ST}) and haplotype similarity (N_{ST}), and the presence of phylogeographic structure is confirmed when the N_{ST} value is higher than G_{ST} . The test was applied to the entire data set and within each region to verify the presence of large- and fine-scale population structure. Finally, to verify the hypothesis of isolation by distance (IBD) in *P. sylvestris* resulting from postglacial recolonization of a few refugial areas, we used the Mantel test separately in each region and for all populations. Correlations between genetic and geographic distances among populations were performed using 999 permutations of Euclidean distances in R using the `mantel.randtest` function in the `adegenet` package (Jombart, 2008).

3 Results

3.1 Haplotype variation

All 11 markers were polymorphic, providing 73 novel haplotypes in 1407 trees across the *P. sylvestris* populations (Table S2). The four most frequent haplotypes (H8, H14, H4, H2) were present in almost 60% of the trees and in the majority of regions (with the exception of H14, the remaining three haplotypes were found in seven to nine regions). We also recorded an excess of minor frequency haplotypes—39 of 73 haplotypes were present only in one or two individuals (27 and 17, respectively). The majority of singleton mitotypes were found in populations from Poland and Finland. As these low-frequency haplotypes are not phylogenetically informative, we used only 29 major haplotypes present in at least three individuals to show their geographic distribution (Fig. 1) and to reconstruct the haplotype network (Fig. 2).

Overall, the haplotype network revealed two main groups of haplotypes. These include haplotype H8 and several other less frequent haplotypes, mostly from Finland, in one group

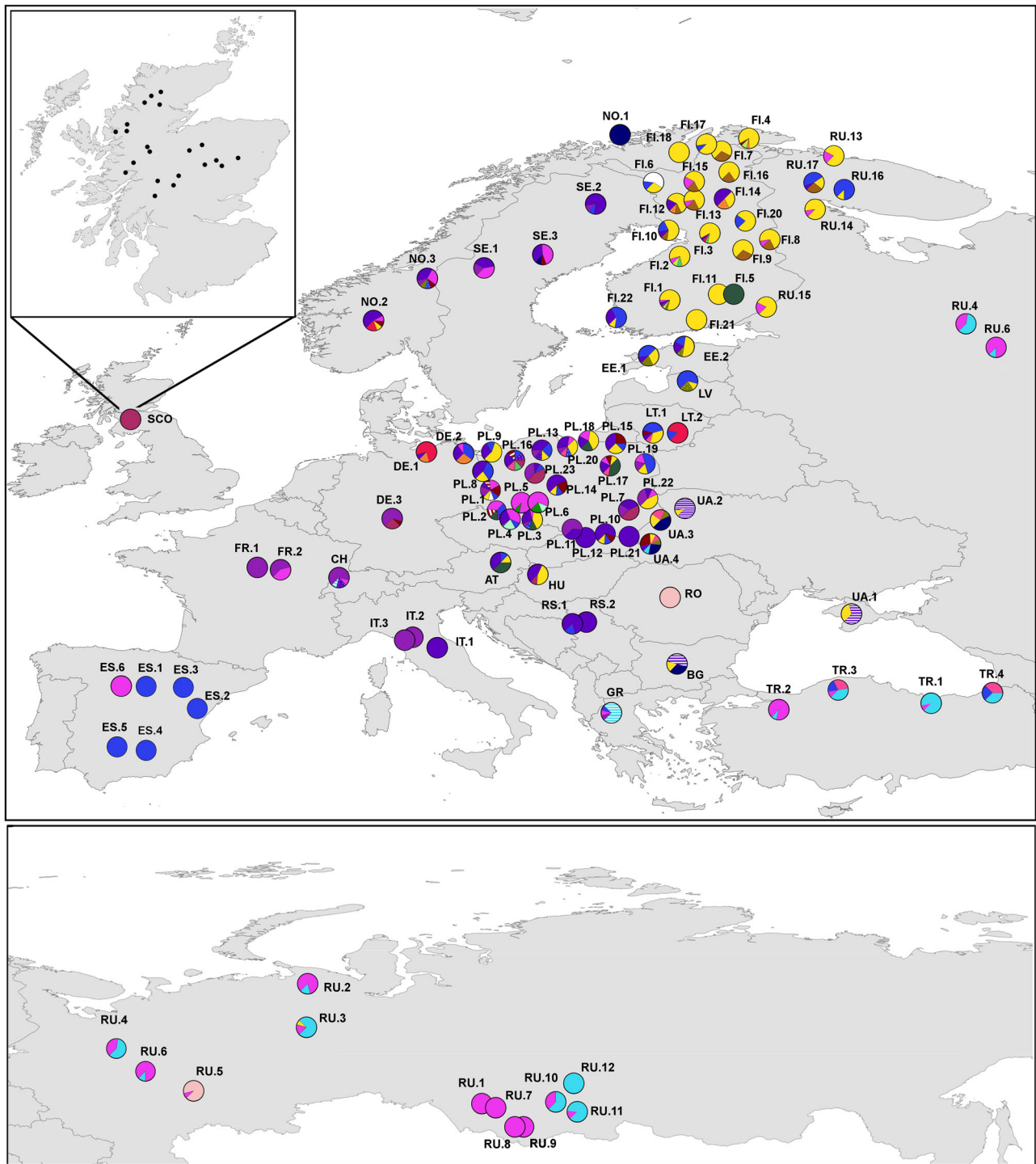


Fig. 1. The geographic distribution of 29 major mitochondrial DNA haplotypes in the populations analyzed. The colors correspond to those used in the haplotype network in Fig. 2. The figure was created with Natural Earth vector data (public domain) and modified.

and haplotypes H1, H2, H4, H6, and H14 from the other regions in the other group. However, the distribution of haplotypes from those two groups did not constitute a clear phylogeographic pattern as haplotypes were shared by individuals from different regions across the sampled range (Fig. 1). Furthermore, only 6 of the 29 major haplotypes were

region specific and shared between different populations within regions. However, in light of putative migration routes, the co-occurrence of some haplotypes in different regions might also be informative. Among the most insightful examples are (1) the distribution of haplotype H14, which is fixed in all populations from Scotland and present only in a

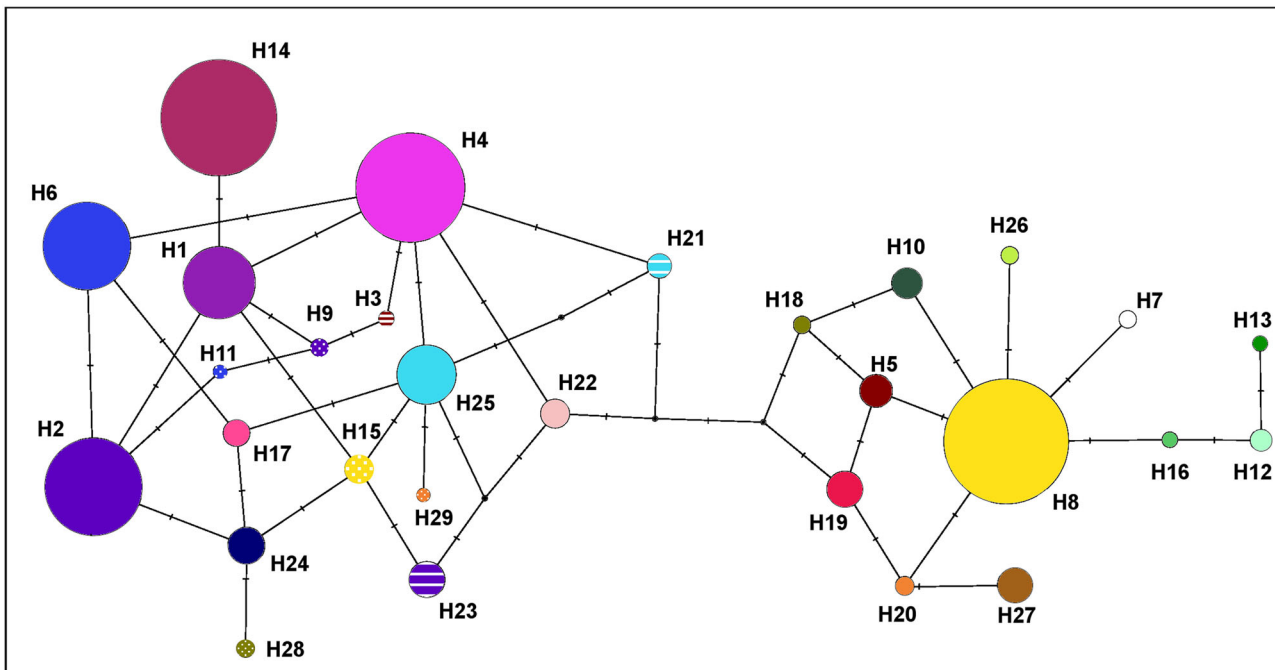


Fig. 2. The median-joining network of haplotypes detected at 11 mitochondrial DNA regions analyzed for 124 Scots pine populations. Areas of the circles are proportional to haplotype frequencies, and hatch marks represent numbers of nucleotide differences between them.

Table 2 Haplotype diversity at 11 mtDNA loci in 12 *Pinus sylvestris* averaged for regional groups

Regional group	N	N_p	A	P	N_e	R_h	H_d
The Balkans	65	6	1	1	4.884	8.264	0.808
The Baltic Region	54	5	10	2	6.451	7.022	0.861
Finland	322	27	29	13	3.850	6.761	0.743
Italy	34	3	2	0	1.710	1.000	0.428
Poland	279	23	40	22	7.676	9.853	0.873
Russia	155	12	6	1	2.015	2.322	0.507
Scandinavia	63	6	12	1	4.309	6.680	0.780
Scotland	190	21	1	0	1.000	0.000	0.000
Spain	68	6	3	0	1.380	1.342	0.280
Turkey	48	4	6	1	3.716	4.242	0.746
Ukraine	47	4	13	4	5.697	8.631	0.842
Western Europe	82	7	15	2	4.537	8.031	0.789

A, number of haplotypes; H_d , haplotype diversity; N, sample size; N_e , effective number of haplotypes; N_p , number of populations; P, number of private haplotypes; and R_h , haplotype richness. For details of the composition of regional groups, see Table S2.

few individuals from continental Europe (populations from Poland and Germany); (2) the disjunct occurrence of haplotype H22, which is unique to individuals from Romania and from a population in Russia; and (3) the sharing of haplotypes H4 and H25 between multiple populations from Russia and Turkey.

The levels of diversity and differentiation were calculated for each population and averaged for each of the 12 regions (Tables 2, S3). The mean haplotype diversity for the entire

collection was $H_d = 0.390$, although there was a lot of variation between populations and regions. The lowest average haplotype diversity was detected in Scotland, Italy, and Spain ($H_d = 0.280$ and 0.428 , respectively), while the highest average haplotype diversity was detected in Poland, the Baltic region, and Ukraine (0.873 , 0.861 , and 0.842 , respectively; Table 2). About 33% of the populations were fixed for one haplotype (all Scottish, five Spanish and Russian, three Italian and Finnish, two Polish, and single populations from France and Serbia; Fig. 1; Table S2). Both the highest number of haplotypes in the regions and the highest number of haplotypes detected in a single population were in Poland (haplotype number on a regional scale $A = 40$ and $A = 11$ in population PL.1), and almost 44% of all populations studied had haplotype diversity greater than 0.500 (Fig. 1; Table S2).

Contrary to the expectation under the “southern richness northern purity” model of genetic diversity following postglacial migration, we found low to moderate levels of mtDNA variation in southern refugial areas (Spain, Italy, the Balkans, Turkey), moderate levels of variation in the northernmost regions of Fennoscandia (Norway, Sweden, Finland, and Russian Karelia), and the highest mtDNA diversity in middle-latitude regions (western Europe, Poland, Ukraine, the Baltic region). Therefore, the linear regression analysis did not indicate any significant relationships between four different measures of genetic diversity and latitude for the entire data set (even after accounting for the excess population with $H_d = 0$) and the weak but statistically significant inverse relationship in the case of Northern Europe groups (Fig. S1).

3.2 Population differentiation

Analysis of genetic relationships between populations using UPGMA revealed a similar division between two groups as in

close proximity to the relatively temperate conditions of the Atlantic Ocean (Sinclair et al., 1999; Parducci et al., 2012). Our results are in line with those suggestions and further indicate that scattered populations that have survived in central and southern Europe, including the Danube River basin, could explain the high genetic diversity observed in that area in our study. This conclusion is supported by the lack of a formal large-scale phylogeographic structure, the lack of the expected southern richness northern purity pattern, the relatively low variability in southern areas usually considered as refugial sites, and the greatest variability in the central European part of the Scots pine distribution where individuals from various waves of migration from many scattered refugial populations have putatively met and mixed with each other. Consequently, populations from the Alps, Poland, western Ukraine, and the Baltic countries have large haplotype diversity compared to other regions. Furthermore, the Polish populations had the largest number of haplotypes, the highest haplotype diversity, and the largest number of unique haplotypes. These patterns of distribution reflect different waves of migration and crossing of populations of different origins in the area. However, not all of the Polish populations were sampled from protected or inaccessible mountain regions, and some lowland locations representing local ecotypes were collected from commercial forests. Therefore, the influence of human-mediated transfer of plant material cannot be completely ruled as a cause of variability of some of the analyzed populations. Studies on Scots pine and other forest tree species document evidence for historic transfer of reproductive material in Europe (Myking et al., 2016; Jansen et al., 2017).

4.2 Different population history of Fennoscandia

The data indicate a few major genetic lineages that have contributed most to the recolonization process at middle and northern latitudes in Europe. Our data support earlier observations of heterogeneous patterns of genetic variation in different species of forest trees in Fennoscandia (Parducci et al., 2012; Salojärvi et al., 2017; Wang et al., 2020). The combination of mitotypes was clearly different between Norwegian and Swedish stands as compared to Finland and Russian Karelia. The latter part of Scots pine distribution was inhabited during colonization by a variable group of initial populations and the data fit the IBD stepping stone model. Consequently, we observed a weak but statistically significant correlation between geographic and genetic distance and a decrease in all measures of variation with latitude for Finland. The combination of haplotypes (especially H8, H6, H5, and H10) suggests that the source and direction of colonization in Finland and Russian Karelia were from the south, but not from the east. For instance, frequently present haplotypes in Finland (H8 and H10) are present in Austria and Hungary. Interestingly, pines from postglacial peat bogs in the Stołowe Mountains in Poland contained haplotype H8 (Łabiszak et al., 2019), which is common in Poland, and it is the most frequent haplotype in the Finnish populations. Although it seems clear that the Baltic countries, Finland, and the area of Russian Karelia were not recolonized from the east, our sampling from the western part of Russia is not dense enough to exclude the possibility of a putative source of recolonization of that area from a

refugium to the south-east of Moscow (Buchovska et al., 2013) or southern refugium in the Carpathian Mountains or the Balkan region. The Carpathians have been suggested as a refugial area for another boreal forest conifer *Picea abies* (Gugerli et al., 2001); however, this region did not contribute to recolonization of the discontinuous northern part of the Norway spruce range (Dering et al., 2012). In contrast, populations from Norway and Sweden have a different composition of haplotypes than those from Finland. These patterns suggest that northward colonization was likely to have reached western Scandinavia from western and central Europe, which have a high frequency of haplotypes present (such as H4 and H2, respectively). Patterns of division between Western and Central versus Eastern European populations and between Swedish and Norwegian versus Finnish populations were also observed in the distribution of mitotypes detected at *nad1* and *nad7* introns in earlier studies (Naydenov et al., 2007; Pyhäjärvi et al., 2008; Dering et al., 2017).

4.3 A glacial refugium in the northwest?

Additional insight into the population structure might be gained by examining the current distribution of fixed haplotypes, especially those found only in particular regions and/or populations. Overall, such haplotypes in our data set were restricted mainly to marginal parts of the species distribution. The most genetically homogeneous but distinct region as compared to the others was Scotland. All 190 samples from 21 stands representative of the remnant Scottish Scots pine populations from the most north-western parts of the species distribution were fixed for only one haplotype (H14). This variant shows a marginal frequency on the continent that is present in just a few individuals in one Polish and one German population, and it is only one mutation away from H1, which is very common in Germany and France. A relatively recent and rapid wave of migration over the land bridge connecting the British Isles with mainland Europe, either by a small number of colonizers or by a very homogeneous group, could explain such a pattern. The British Isles were largely ice-covered during the LGM and the contemporary population is certainly a product of postglacial establishment. However, in other molecular data sets, there is no sign of the severe founder effect that would be expected by a rapid wave of migration from a small source population (Wachowiak et al., 2011; Sejrup et al., 2016; Semerikov et al., 2020). Some heterogeneity of the western populations in Scotland compared to other locations was shown based on mtDNA RFLP variation (Sinclair et al., 1998) and, although no differentiation between Scottish and mainland European populations was observed at the cpSSR loci, a unique, low-frequency allele was identified in some western sites (Provan et al., 1998). Furthermore, variation at nuclear gene loci showed that Scottish populations have equally high or higher genetic diversity than populations from continental Europe (Wachowiak et al., 2011). Therefore, our results do not exclude the existence of a northwestern refugium during the LGM, which may have contributed to Scots pine recolonization of Scotland and partially western Europe. However, it is equally possible that the homogeneity of the Scottish mtDNA data reflects the founder effect of colonization from a central European source, while wider

genetic diversity is accounted for by extensive subsequent pollen flow into the founder populations.

4.4 Genetic variation in southern refugia

Previous studies at different biochemical and molecular markers provide several lines of evidence indicating low or nonexistent contributions of more or less isolated southern glacial refugia located in the Iberian Peninsula, Apennine Peninsula, and Asia Minor to northward colonization of Europe (Soranzo et al., 2000; Cheddadi et al., 2006; Labra et al., 2006; Pyhäjärvi et al., 2008; Prus-Głowacki et al., 2012; Jasińska et al., 2014; Pavia et al., 2014). Our data indicate distinct haplotypes between the Iberian and Apennine Peninsulas and show genetic variation between populations in each of these areas. Patterns of between population divergence among the patchily distributed Iberian populations were reported for chloroplast SSR loci (Robledo-Arnuncio et al., 2005). We found different mitotypes between the Apennine populations that showed low genetic diversity in earlier studies (Scalfi et al., 2009; Belletti et al., 2012). Significant differentiation between populations from the Apennines compared to populations from the southern slopes of the Italian Alps likely reflects their isolation and low population size, as well as limited gene flow between the regions. Together, this also suggests that Scots pine populations from that area have not contributed to the postglacial recolonization of Central and Northern Europe (Cheddadi et al., 2006; Scalfi et al., 2009). Our data show that the Iberian and Apennine Peninsulas share mitotypes present in other distinct parts of the species range. However, the presence of haplotype H4 in Iberia and populations from Asia Minor to Western Siberia indicates its ancestral origin and probable widespread occurrence before the last glacial period. Interestingly, this haplotype (together with H2) is present at high frequency in populations from the Stołowe and the Karkonosze Mountains in southern Poland, which are presumed to have been established during the early Holocene (Boratyrńska et al., 2021). The existence of different haplotypes in the Apennine Peninsulas present at high frequency in the eastern European distribution and in the Balkans suggests a connection between those regions that may predate the last glacial period, as found in other forest tree species (Zaborowska et al., 2020).

4.5 The Balkans and the Black Sea basin

Our data also show some less frequent haplotypes present in highly localized distributions. Such endemic mitotypes may represent new variants or remnants of isolated stands that survived locally, but did not contribute to the major recolonization events. Recent mtDNA and cpDNA analyses suggested that the eastern Black Sea region, including Crimea, Caucasus, and Asia Minor, has a distinct origin and might have been isolated from the populations of the main range for a long time (over hundreds of thousands of years, Semerikov et al., 2020; Dering et al., 2021). Our data show that populations across a large area of the Black Sea basin, including Crimea, western Ukraine, Romania, Bulgaria, and western Russia, share haplotypes that were not observed elsewhere in our sample. The results suggest their distinct genetic character but relatively limited dispersal during the postglacial expansion period. Additionally, populations from

Asia Minor had mitotypes observed in Western Siberia and most eastern Scots pine populations that were distinct from major haplotypes observed in the Balkans and other parts of the Black Sea basin. Greater connectivity between the Asia Minor populations and eastern parts of the distribution was suggested in earlier mtDNA studies (Pyhäjärvi et al., 2008) and isolation between the Balkan and Asia Minor populations of Scots pine was observed at SSR loci (Wójcikiewicz et al., 2016b). The Balkan populations were possibly influenced by migrants from the regions of eastern Italian Alps and this connection was observed for other forest tree species (Zaborowska et al., 2020).

4.6 Different population history of European and Asian pines

It is not clear so far to what extent the populations from the main species range in Europe versus eastern European and Asian populations contributed mutually to the recolonization process. Haplotypes in the eastern European/Russian part of the distribution (Rus-4 and 6) are similar to those found in populations from the far eastern territory of Russia and are distinct from the most frequent haplotypes in Russian Karelia and Finland. Clearly, pines have not migrated from Fennoscandia eastward to the Siberian steppes, and most European Scots pine have a different origin and postglacial history than eastern European/Russian and Siberian populations. Our data indicate distribution of the same mitotypes across large areas from eastern European parts of Russia to southwestern regions of the Altai Mountains (meridians 38–93, respectively). Populations in that area show a much lower haplotype diversity and a relatively uniform haplotype structure compared to most European locations, suggesting a much simpler recolonization history and migration from limited sources. Populations from the most eastern distribution occur in regions indicated as refugial stands for recolonization of the West Siberian Plain by other conifer species of the boreal forest, including the genus *Larix* (Semerikov et al., 2013), and an eastern refugium in the region of southern Urals was previously suggested (Semerikov et al., 2018). Results of other studies including patterns of polymorphisms at cpSSR loci across eastern Fennoscandia and the north-eastern European part of the Russian distribution support the conclusion that those regions have their own independent history (Semerikov et al., 2014).

5 Conclusions

The presented study is a step forward toward understanding of the complex population history of Scots pine across its European and Asian range. We reveal a high genetic variation of populations in central and eastern Europe, including the Alps, the Hungarian plain, and the Balkans, which most likely contained many patchily distributed populations that survived the LGM and expanded to the middle and northern latitudes in Europe. Not all survivors succeeded equally in the post-glacial recolonization process. Major mitotype variants identified include the remnants of Scots pine in most western distributions of the species that colonized the Scottish Highlands, two main western and central European variants that contributed to the present populations in

Norway and Sweden, the central-eastern European lineage that colonized Baltic countries, Finland, and most western areas of Russia, and separate one common to most eastern European parts of Russia and western Siberia. We also observed signs of a distinct refugium located in the northern parts of the Black Sea basin that contributed to the genetic variation observed in several populations in the Balkans, Ukraine, and Western Russia. Shared haplotypes between individuals from geographically distant and more or less isolated refugial stands in the Iberian Peninsula, Asia Minor, Asia, and the major continuous European distribution of the species are most likely of ancient origin with a distribution that predates the Quaternary glaciations. The developed markers applied in other populations of Scots pine distribution not covered in the presented study might be helpful in providing a more complete picture of the species postglacial history.

Acknowledgements

We thank Dr. Marta Kempf and students Anna Mikołajczyk, Maciej Weigt, Daria Jadwisieńczyk, and Simona Jeziorowska from the Interdepartmental Students Scientific Group of the University of Agriculture in Kraków, Poland, members of the GenTree consortium (EU Horizon 2020 Grant Agreement No. 676876), and NERC (Grant No. NE/H003959/1) for providing plant material from several populations included in the study. The research was financially supported by the Polish National Science Centre (UMO-2017/27/B/NZ9/00159). Weronika B. Żukowska acknowledges financial support from NCN Preludium project (UMO-2016/21/N/NZ9/01499).

Conflict of Interest

We have no conflict of interest to declare.

References

- Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48.
- 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.
- 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.
- Birks HJB. 1989. Holocene isochrone maps and patterns of tree-spreading in the British isles. *Journal of Biogeography* 16: 503–540.
- Boratyńska K, Gołąb Z, Łabiszak B, Niemczyk W, Sobierajska KI, Ufnalski K, Wachowiak W, Boratyński A. 2021. Are there any traces of *Pinus uliginosa* in the Stołowe mountains outside the Wielkie Torfowisko Batorowskie and Błędne Skały? *Acta Societatis Botanicorum Poloniae* 90: 904.
- Buchowska J, Danusevicius D, Baniulis D, Stanys V, Siksnianiene JB, Kavaliauskas D. 2013. The location of the northern glacial refugium of Scots pine based on mitochondrial DNA markers. *Baltic Forestry* 19: 2–12.
- Burban C, Petit RJ, Carcreff E, Jactel H. 1999. Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology* 8: 1593–1602.
- Cheddadi R, Vendramin G, Litt T, François L, Kageyama M, Lorentz S, Laurent J, Beaulieu J, Sadori L, Jost A, Lunt D. 2006. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography* 15: 271–282.
- Corander J, Marttinen P. 2006. Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* 15: 2833–2843.
- Corander J, Tang J. 2007. Bayesian analysis of population structure based on linked molecular information. *Mathematical Biosciences* 205: 19–31.
- Dering M, Baranowska M, Beridze B, Chybicki IJ, Danelia I, Iszkuło G, Kvartskhava G, Kosiński P, Rączka G, Thomas PA, Tomaszewski D, Walas Ł, Sękiewicz K. 2021. The evolutionary heritage and ecological uniqueness of Scots pine in the Caucasus ecoregion is at risk of climate changes. *Scientific Reports* 11: 22845.
- Dering M, Kosiński P, Wyka TP, Pers-Kamczyc E, Boratyński A, Boratyńska K, Reich PB, Romo A, Zadworny M, Żytowski R, Oleksyn J. 2017. Tertiary remnants and Holocene colonizers: Genetic structure and phylogeography of Scots pine reveal higher genetic diversity in young boreal than in relict Mediterranean populations and a dual colonization of Fennoscandia. *Diversity and Distributions* 23: 540–555.
- Dering M, Misiorny A, Lewandowski A, Korczyk A. 2012. Genetic and historical studies on the origin of Norway spruce in Białowieża Primeval Forest in Poland. *European Journal of Forest Research* 131: 381–387.
- Donnelly K, Cavers S, Cottrell JE, Ennos RA. 2016. Genetic variation for needle traits in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes* 12: 40.
- Donnelly K, Cottrell J, Ennos RA, Vendramin GG, A'Hara S, King S, Perry A, Wachowiak W, Cavers S. 2017. Reconstructing the plant mitochondrial genome for marker discovery: A case study using *Pinus*. *Molecular Ecology Resources* 17: 943–954.
- Eliades N, Eliades D. 2009. *Haplotype analysis: Software for analysis of haplotypes data*. Distributed by the authors: Forest genetics and forest tree breeding. Germany: Georg-August University Göttingen.
- Feurdean A, Tantau I, Farcas S. 2011. Holocene variability in the range distribution and abundance of *Pinus*, *Picea abies*, and *Quercus* in Romania; implications for their current status. *Quaternary Science Reviews* 30: 3060–3075.
- Francis RM. 2017. Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* 17: 27–32.
- Giertych M. 1991. Chapter 6—Provenance variation in growth and phenology. In: Giertych M ed. *Developments in plant genetics and breeding*. Amsterdam, The Netherlands and Akadémiai Kiado, Budapest, Hungary: Elsevier Science Publishers, 87–101.
- Gugerli F, Sperisen C, Büchler U, Magni F, Geburek T, Jeandroz S, Senn J. 2001. Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious founder effect during postglacial re-colonization of the western Alps. *Molecular Ecology* 10: 1255–1263.
- Haesaerts P, Borziac I, Chekha VP, Chirica V, Drozdov NI, Koulakovska L, Orlova LA, van der Plicht J, Damblon F. 2010. Charcoal and wood remains for radiocarbon dating Upper Pleistocene loess

- sequences in Eastern Europe and Central Siberia. *Palaeogeography, Palaeoclimatology, Palaeoecology* 291: 106–127.
- Hewitt GM. 2000. The genetic legacy of the quaternary ice ages. *Nature* 405: 907–913.
- Hurme P, Repo T, Savolainen O, Paakkonen T. 1997. Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris*). *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 27: 716–723.
- Jansen S, Konrad H, Geburek T. 2017. The extent of historic translocation of Norway spruce forest reproductive material in Europe. *Annals of Forest Science* 74: 56.
- Jaramillo-Correa JP, Bousquet J. 2005. Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics* 171: 1951–1962.
- Jasińska AK, Boratyńska K, Dering M, Sobierajska KI, Ok T, Romo A, Boratyński A. 2014. Distance between south-European and south-west Asiatic refugial areas involved morphological differentiation: *Pinus sylvestris* case study. *Plant Systematics and Evolution* 300: 1487–1502.
- Jombart T. 2008. Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Kujala S, Savolainen O. 2012. Sequence variation patterns along a latitudinal cline in Scots pine (*Pinus sylvestris*): Signs of clinal adaptation? *Tree Genetics & Genomes* 8: 1451–1467.
- Kumar S, Stecher G, Tamura K. 2016. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33: 1870–1874.
- Łabiszak B, Zaborowska J, Wachowiak W. 2019. Patterns of mtDNA variation reveal complex evolutionary history of relict and endangered peat bog pine (*Pinus uliginosa*). *AoB Plants* 11: plz015.
- Labra M, Grassi F, Sgorbati S, Ferrari C. 2006. Distribution of genetic variability in southern populations of Scots pine (*Pinus sylvestris* L.) from the Alps to the Apennines. *Flora—Morphology, Distribution, Functional Ecology of Plants* 201: 468–476.
- Lynch M, Crease TJ. 1990. The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution* 7: 377–394.
- Myking T, Rusanen M, Steffenrem A, Kjær ED, Jansson G. 2016. Historic transfer of forest reproductive material in the nordic region: Drivers, scale and implications. *Forestry: An International Journal of Forest Research* 89: 325–337.
- Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J. 2007. Glacial vicariance in Eurasia: Mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology* 7: 233.
- Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* 70: 3321–3323.
- Oleksyn J, Tjoelker MG, Reich PB. 1992. Growth and biomass partitioning of populations of European *Pinus sylvestris* L. Under simulated 50° and 60° n daylengths: Evidence for photoperiodic ecotypes. *New Phytologist* 120: 561–574.
- Parducci L, Jørgensen T, Tollefsrud MM, Elverland E, Alm T, Fontana SL, Bennett KD, Haile J, Matetovici I, Suyama Y, Edwards ME, Andersen K, Rasmussen M, Boessenkool S, Coissac E, Brochmann K, Taberlet P, Houmark-Nielsen M, Larsen NK, Orlando L, Gilbert MTP, Kjær KH, Alsos IG, Willerslev E. 2012. Glacial survival of boreal trees in northern Scandinavia. *Science* 335: 1083–1086.
- Pavia I, Mengl M, Gaspar MJ, Carvalho A, Heinze B, Lima-Brito J. 2014. Preliminary evidence of two potentially native populations of *Pinus sylvestris* L. in Portugal based on nuclear and chloroplast SSR markers. *Austrian Journal of Forest Science* 131: 1–22.
- Peakall R, Smouse PE. 2012. Genalex 6.5: Genetic analysis in excel. Population genetic software for teaching and research—An update. *Bioinformatics* 28: 2537–2539.
- Perry A, Brown AV, Cavers S, Cottrell JE, Ennos RA. 2016. Has Scots pine (*Pinus sylvestris*) co-evolved with *Dothistroma septosporum* in Scotland? Evidence for spatial heterogeneity in the susceptibility of native provenances. *Evolutionary Applications* 9: 982–993.
- Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Muller-Starck G, Demesure-Musch B, Palme A, Martin JP, Rendell S, Vendramin GG. 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* 300: 1563–1565.
- Petit RJ, Hu FS, Dick CW. 2008. Forests of the past: A window to future changes. *Science* 320: 1450–1452.
- Polezhaeva MA, Lascoux M, Semerikov VL. 2010. Cytoplasmic DNA variation and biogeography of *Larix Mill.* in northeast Asia. *Molecular Ecology* 19: 1239–1252.
- Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144: 1237–1245.
- Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* 23: 564–571.
- Provan J, Soranzo N, Wilson NJ, McNicol JW, Forrest GI, Cottrell J, Powell W. 1998. Gene-pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. *Proceedings of the Royal Society of London Series B-Biological Sciences* 265: 1697–1705.
- Prus-Głowacki W, Stephan BR. 1994. Genetic-variation of *Pinus sylvestris* from Spain in relation to other European populations. *Silvae Genetica* 43: 7–14.
- Prus-Głowacki W, Urbaniak L, Bujas E, Curtu AL. 2012. Genetic variation of isolated and peripheral populations of *Pinus sylvestris* (L.) from glacial refugia. *Flora* 207: 150–158.
- Pyhäjärvi T, Garcia-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O. 2007. Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics* 177: 1713–1724.
- Pyhäjärvi T, Salmela MJ, Savolainen O. 2008. Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genetics & Genomes* 4: 247–254.
- R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramírez-Valiente JA, Solé-Medina A, Pyhäjärvi T, Savolainen O, Heer K, Opgenoorth L, Danusevicius D, Robledo-Arnuncio JJ. 2021. Adaptive responses to temperature and precipitation variation at the early-life stages of *Pinus sylvestris*. *New Phytologist* 232: 1632–1647.
- Ranade SS, García-Gil MR. 2013. Ecotypic variation in response to light spectra in Scots pine (*Pinus sylvestris* L.). *Tree Physiology* 33: 195–201.
- Robledo-Arnuncio JJ, Collada C, Alía R, Gil L. 2005. Genetic structure of montane isolates of *Pinus sylvestris* L. in a mediterranean refugial area. *Journal of Biogeography* 32: 595–605.
- Robledo-Arnuncio JJ, Smouse PE, Gil L, Alía R. 2004. 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.

- Salmela MJ, Cavers S, Cottrell JE, Iason GR, Ennos RA. 2011. Seasonal patterns of photochemical capacity and spring phenology reveal genetic differentiation among native Scots pine (*Pinus sylvestris* L.) populations in Scotland. *Forest Ecology and Management* 262: 1020–1029.
- Salojärvi J, Smolander O-P, Nieminen K, Rajaraman S, Safronov O, Safdari P, Lamminmäki A, Immanen J, Lan T, Tanskanen J, Rastas P, Amiryousefi A, Jayaprakash B, Kammonen JI, Hagqvist R, Eswaran G, Ahonen VH, Serra JA, Asiegbu FO, de Dios Barajas-Lopez J, Blande D, Blokhina O, Blomster T, Broholm S, Brosché M, Cui F, Dardick C, Ehonen SE, Elomaa P, Escamez S, Fagerstedt KV, Fujii H, Gauthier A, Gollan PJ, Halimaa P, Heino PI, Himanen K, Hollender C, Kangasjärvi S, Kauppinen L, Kelleher CT, Kontunen-Soppela S, Koskinen JP, Kovalchuk A, Kärenlampi SO, Kärkönen AK, Lim K-J, Leppälä J, Macpherson L, Mikola J, Mouhu K, Mähönen AP, Niinemets Ü, Oksanen E, Overmyer K, Palva ET, Pazouki L, Pennanen V, Puhakainen T, Poczai P, Possen BJHM, Punkkinen M, Rahikainen MM, Rousi M, Ruonala R, van der Schoot C, Shapiguzov A, Sierla M, Sipilä TP, Sutela S, Teeri TH, Tervahauta AI, Vaattovaara A, Vahala J, Vetchinnikova L, Welling A, Wrzaczek M, Xu E, Paulin LG, Schulman AH, Lascoux M, Albert VA, Auvinen P, Helariutta Y, Kangasjärvi J. 2017. Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch. *Nature Genetics* 49: 904–912.
- San-Miguel-Ayanz J, de Rigo D, Caudullo G, Durrant TH, Mauri A. 2016. *European atlas of forest tree species*. Luxembourg: European Commission.
- Savolainen O, Pyhäjärvi T, Knürr T. 2007. Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics* 38: 595–619.
- 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.
- Sejrup HP, Clark CD, Hjelstuen BO. 2016. Rapid ice sheet retreat triggered by ice stream debulking: Evidence from the North Sea. *Geology* 44: 355–358.
- Semerikov NV, Petrova IV, Sannikov SN, Semerikova SA, Tashev AN, Lascoux M, Semerikov VL. 2020. Cytoplasmic DNA variation does not support a recent contribution of *Pinus sylvestris* L. from the Caucasus to the main range. *Tree Genetics & Genomes* 16: 59.
- Semerikov VL, Semerikova SA, Dymshakova OS, Zatsepina KG, Tarakanov VV, Tikhonova IV, Ekart AK, Vidyakin AI, Jamiyansuren S, Rogovtsev RV, Kalchenko LI. 2014. Microsatellite loci polymorphism of chloroplast DNA of Scots pine (*Pinus sylvestris* L.) in Asia and eastern Europe. *Russian Journal of Genetics* 50: 577–585.
- Semerikov VL, Semerikova SA, Polezhaeva MA, Kosintsev PA, Lascoux M. 2013. Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): A range-wide analysis of cytoplasmic markers. *Molecular Ecology* 22: 4958–4971.
- Semerikov VL, Semerikova SA, Putintseva YA, Tarakanov VV, Tikhonova IV, Vidyakin AI, Oreshkova NV, Krutovsky KV. 2018. Colonization history of Scots pine in Eastern Europe and North Asia based on mitochondrial DNA variation. *Tree Genetics & Genomes* 14: 8.
- Shutyaev AM, Giertych M. 2000. Genetic subdivisions of the range of Scots pine (*Pinus sylvestris* L.) based on a transcontinental provenance experiment. *Silvae Genetica* 49: 137–151.
- Sinclair WT, Morman JD, Ennos RA. 1998. Multiple origins for Scots pine (*Pinus sylvestris* L.) in Scotland: Evidence from mitochondrial DNA variation. *Heredity (Edinb)* 80: 233–240.
- Sinclair WT, Morman JD, Ennos RA. 1999. The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: Evidence from mitochondrial DNA variation. *Molecular Ecology* 8: 83–88.
- Soranzo N, Alia R, Provan J, Powell W. 2000. Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. *Molecular Ecology* 9: 1205–1211.
- Tóth EG, Köbölkuti ZA, Pedryc A, Höhn M. 2017. Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. *Journal of Forestry Research* 28: 637–651.
- Tyrmi JS, Vuosku J, Acosta JJ, Li Z, Sterck L, Cervera MT, Savolainen O, Pyhäjärvi T. 2020. Genomics of clinal local adaptation in *Pinus sylvestris* under continuous environmental and spatial genetic setting. *G3 Genes|Genomes|Genetics* 10: 2683–2696.
- Tzedakis PC, Emerson BC, Hewitt GM. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution* 28: 696–704.
- Wachowiak W, Balk P, Savolainen O. 2009. Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes* 5: 117–132.
- Wachowiak W, Perry A, Donnelly K, Cavers S. 2017. Early phenology and growth trait variation in closely related European pine species. *Ecology and Evolution* 8: 655–666.
- 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 (Edinb)* 106: 775–787.
- Wachowiak W, Trivedi U, Perry A, Cavers S. 2015. Comparative transcriptomics of a complex of four European pine species. *BMC Genomics* 16: 234.
- Wachowiak W, Wójkiewicz B, Cavers S, Lewandowski A. 2014. High genetic similarity between Polish and north European Scots pine (*Pinus sylvestris* L.) populations at nuclear gene loci. *Tree Genetics & Genomes* 10: 1015–1025.
- Wachowiak W, Zaborowska J, Łabiszak B, Perry A, Zucca G, González-Martínez S, Cavers S. 2018. Molecular signatures of divergence and selection in closely related pine taxa. *Tree Genetics & Genomes* 14: 83.
- Wang B, Wang X-R. 2014. Mitochondrial DNA capture and divergence in *Pinus* provide new insights into the evolution of the genus. *Molecular Phylogenetics and Evolution* 80: 20–30.
- Wang X, Bernhardsson C, Ingvarsson PK. 2020. Demography and natural selection have shaped genetic variation in the widely distributed conifer Norway spruce (*Picea abies*). *Genome Biology and Evolution* 12: 3803–3817.
- Wickham H. 2016. *ggplot2: Elegant graphics for data analysis*. Cham: Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Willis KJ, van Andel TH. 2004. Trees or no trees? The environments of central and eastern Europe during the last glaciation. *Quaternary Science Reviews* 23: 2369–2387.
- Wójkiewicz B, Cavers S, Wachowiak W. 2016a. Current approaches and perspectives in population genetics of Scots pine (*Pinus sylvestris* L.). *Forest Science* 62: 343–354.
- Wójkiewicz B, Litkowiec M, Wachowiak W. 2016b. Contrasting patterns of genetic variation in core and peripheral populations of highly outcrossing and wind pollinated forest tree species. *AoB Plants* 8: plw054.
- Zaborowska J, Łabiszak B, Wachowiak W. 2020. Population history of European mountain pines *Pinus mugo* and *Pinus uncinata*

revealed by mitochondrial DNA markers. *Journal of Systematics and Evolution* 58: 474–486.

Zadworny M, McCormack ML, Mucha J, Reich PB, Oleksyn J. 2016. Scots pine fine roots adjust along a 2000-km latitudinal climatic gradient. *New Phytologist* 212: 389–399.

Zimmer K, Sønstebo JH. 2018. A preliminary study on the genetic structure of northern European *Pinus sylvestris* L. by means of neutral nuclear microsatellite markers. *Scandinavian Journal of Forest Research* 33: 6–13.

Supplementary Material

The following supplementary material is available online for this article at <http://onlinelibrary.wiley.com/doi/10.1111/jse.12907/supinfo>:

Table S1. Mitochondrial markers analysed and frequency of the variants in the probe. In 198 individuals from the total of 1407, at most two unidentified alleles were recorded and missing sites in those individuals were imputed from the most closely related complete haplotype.

Table S2. Major haplotypes, their frequency and presence in the studied regions.

Table S3. Haplotype diversity at 11 mtDNA loci in 124 *Pinus sylvestris* populations. Acr. – Population acronym, N – sample size, A – number of haplotypes, P – number of private haplotypes, N_e – effective number of haplotypes, R_h – haplotype richness, H_d – haplotype diversity, and regional groups. See Table 1 for details of populations location.

Fig. S1. Linear regression analysis of four different measures of genetic diversity (A–D) and latitude of population origin for (1) the whole dataset (including population with $H_d = 0$) and (2) Northern Europe (including Finland, Sweden, Norway and Russian Karelia (populations R.13–R.17)). Genetic diversity measures used: H_d – haplotype diversity, R_h – haplotype richness, N_e – effective number of haplotypes, A – number of haplotypes.

Fig. S2. UPGMA phylogenetic tree of 124 *P. sylvestris* populations. Colors indicate defined regional groups of pines. All populations from Scotland are represented together as SCO.

Fig. S3. PCoA based on mean number of nucleotide differences averaged across populations within twelve geographical regions defined including Spain; Italy; Scotland; Western Europe (France, Germany, Austria, Switzerland); Poland; Ukraine; Balkans (Hungary, Serbia, Romania, Bulgaria Greece); Baltic region (Lithuania, Latvia, Estonia); Turkey; Russia; Scandinavia (Norway, Sweden); Finland.

Fig. S4. Output of clustering analysis of Scots pine populations performed using all polymorphic sites for $K = 29$ in BAPS software.

Fig. S5. The results of the Mantel test for (1) all populations analyzed and (2) regions were the statistically significant relationships between genetic variation and geographical distances were observed.