



RESEARCH PAPER

Open Access



Genetic variation of Scots pine (*Pinus sylvestris* L.) in Eurasia: impact of postglacial recolonization and human-mediated gene transfer

Weronika Barbara Żukowska^{1*} , Błażej Wójkiewicz¹ , Andrzej Lewandowski¹ , Richárd László²  and Witold Wachowiak^{1,3} 

Abstract

Key message The range-wide level of genetic variation of Scots pine (*Pinus sylvestris* L.) is geographically structured. High admixture and low genetic structure of populations in Central Europe and Fennoscandia suggest past recolonization from multiple sources and the influence of human-mediated gene transfer. Gene pools of marginal and isolated stands require active conservation. Some areas of Scots pine distribution need further genetic studies.

Context Scots pine (*Pinus sylvestris* L.) seems to be a species of low conservation priority because it has a very wide Eurasian distribution and plays a leading role in many forest tree breeding programs. Nevertheless, considering its economic value, long breeding history, range fragmentation, and increased mortality, which is also projected in the future, it requires a more detailed description of its genetic resources.

Aims Our goal was to compare patterns of genetic variation found in biparentally inherited nuclear DNA with previous research carried out with mitochondrial and chloroplast DNA due to their different modes of transmission.

Methods We analyzed the genetic variation and relationships of 60 populations across the distribution of Scots pine in Eurasia (1262 individuals) using a set of nuclear DNA markers.

Results We confirmed the high genetic variation and low genetic differentiation of Scots pine spanning large geographical areas. Nevertheless, there was a clear division between European and Asian gene pools. The genetic variation of Asian populations was lower than in Europe. Spain, Turkey, and the Apennines constituted separate gene pools, the latter showing the lowest values of all genetic variation parameters. The analyses showed that most populations experienced genetic bottlenecks in the distant past. Ongoing admixture was found in Fennoscandia.

Conclusions Our results suggest a much simpler recolonization history of the Asian than European part of the Scots pine distribution, with migration from limited sources and possible founder effects. Eastern European stands seem to have descended from the Urals refugium. It appears that Central Europe and Fennoscandia share at least one glacial refugium in the Balkans and migrants from higher latitudes, as well as from south-eastern regions. The low genetic structure between Central Europe and Fennoscandia, along with their high genetic admixture, may result

Handling editor: Véronique Jorge.

*Correspondence:

Weronika Barbara Żukowska
wzukowska@man.poznan.pl

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

at least partially from past human activities related to the transfer of germplasm in the nineteenth and early twentieth centuries. In light of ongoing climate changes and projected range shifts of Scots pine, conservation strategies are especially needed for marginal and isolated stands of this species. Genetic research should also be complemented in parts of the species distribution that have thus far been poorly studied.

Keywords *Pinus sylvestris*, Admixture, Gene flow, Homogenization, Conservation, Nuclear markers

1 Introduction

The unique characteristics of forest trees—their long-life spans, relatively late reproductive maturity, economically important traits controlled by many genes, and the fact that they are generally undomesticated with large effective population sizes—make genetic studies hard to carry out and interpret. However, due to the high ecological and economic importance of forests and the social services they provide, forest tree genetics has rapidly evolved throughout the last few decades. The genetic diversity of trees serves two important purposes: (i) ensuring the stability and sustainability of forest ecosystems (Schaberg et al. 2008) by maintaining pre-existing genetic variation subjected to natural selection (Savolainen et al. 2007); and (ii) being the source of variability used by humans during breeding practices and improvement programs to increase forest productivity and adaptability to rapidly changing environments (Hedrick 2001).

The level and patterns of genetic variation observed today in forest tree species are mainly the consequence of their migration and range shifts after the Last Glacial Maximum (LGM), which took place approximately 25,000–18,000 years ago (Petit et al. 2003). A lot of genetic studies have been carried out to verify and accompany findings based on palynological and macrofossil data with an aim to trace the location of glacial refugia and postglacial recolonization routes. Most pine species constitute a unique study system in this matter because their mitochondrial and chloroplast genomes are transmitted in different ways (maternally via seeds and paternally via pollen). Therefore, the level and patterns of genetic variation revealed by mitochondrial and chloroplast markers are different (Petit et al. 2005). Mitochondrial DNA (mtDNA) still did not allow a clear inference of recolonization routes for many pine species, though a study by Donnelly et al. (2017) provides a set of new markers for finer-scale population genetic studies in pines. Their usefulness has already been shown for phylogeographic inferences in Scots pine by Wachowiak et al. (2022). Chloroplast DNA (cpDNA) is far less useful for such purposes due to its hypervariability, long-distance migration of pollen, and possible homoplasy (Provan et al. 2001). The nuclear genome, on the other hand, shows greater variability due to recombination and higher mutation rates (Wolfe et al. 1987).

Here, we focused on Scots pine (*Pinus sylvestris* L.), one of the most important forest-forming tree species in the Northern Hemisphere, which has high ecological and economic value, as well as breeding significance. Scots pine is a pioneer species that grows on poor and sandy soils (Farjon 2005). It is a key component of forest ecosystems in temperate and boreal zones, both in lowlands and mountain areas, growing either in pure stands or in mixed forests with other conifers and broad-leaf species (Richardson 1998). The vast Eurasian distribution of Scots pine spans a ca. 8000-km-long area from substantial populations in Scotland and the Iberian Peninsula in the west (6°W) to very far eastern areas of Siberia (150°E). The species crosses several climatic zones, starting from southern Spain and Turkey (40°N) up to the Arctic Circle (70°N) (Richardson 1998; Farjon 2005). It forms continuous populations across Central Europe, Fennoscandia, and Siberia, but the distribution along the western and southern range boundaries is fragmented. The gradual decline of populations and the expansion to higher latitudes are expected along the southern limit of Scots pine, mostly due to drought and biotic stresses (Tóth et al. 2017b).

The genetic variation and postglacial recolonization history of Scots pine have been extensively studied with molecular markers by numerous research groups, both in Europe and Asia (e.g., Buchovska et al. 2013; Cheddadi et al. 2006; Dering et al. 2017; Naydenov et al. 2007; Pyhäjärvi et al. 2008; Sannikov et al. 2014; Semerikov et al. 2014, 2018; Sinclair et al. 1999; Wachowiak et al. 2011, 2022; Wójkiewicz and Wachowiak 2016). The results show a high level of genetic variation of Scots pine, which is accompanied by remarkably low genetic structure across most of its distribution, especially based on cpDNA marker data. This situation seems to result from the high dispersal ability of Scots pine via pollen, but also from the postglacial migration history of this species (Semerikov et al. 2014, 2018; Wachowiak et al. 2022b). Still, there are no studies spanning large areas of Scots pine distribution based on nuclear markers, making it impossible to compare the level and patterns of genetic variation assessed with markers having different modes of transmission. Phylogeographic research indicates that Scots pine in Europe most likely survived the LGM in southern refugia—in the Iberian, Italian, and

Balkan Peninsulas, as well as in Turkey. Most of Western Europe is dominated by just a few mitotypes, one of which is shared with the Iberian Scots pine populations (Cheddadi et al. 2006; Pyhäjärvi et al. 2008; Dering et al. 2017; Wachowiak et al. 2022), whereas Central and Northern Europe seem to have been recolonized from multiple sources, including cryptic refugia located at higher latitudes, but mainly by populations originating in the Balkan Peninsula (Pyhäjärvi et al. 2008; Wójcikiewicz and Wachowiak 2016; Dering et al. 2017; Wachowiak et al. 2022). For the Asian part of the Scots pine distribution, genetic variation seems to be lower than that in the European part of the range (Semerikov et al. 2014, 2018; Wachowiak et al. 2022). The hypothetical Pleistocene refugia include numerous locations along the border of Russia with Kazakhstan and Mongolia and as far east as the Amur region (Sannikov and Petrova 2012; Sannikov et al. 2014, 2020).

In this study, we investigated the genetic diversity of 60 Scots pine populations in Eurasia. We analyzed the level and geographic patterns of genetic variation using nuclear microsatellite markers. Our goal was to verify if nuclear markers show similar genetic variation patterns

as mitochondrial and chloroplast markers used in previous research considering their different modes of transmission. We also aimed to assess the role of admixture of populations from distinct recolonization lineages in shaping the level and distribution of genetic variation in Scots pine, especially in Central Europe and Fennoscandia. Finally, we discuss which parts of Scots pine distribution require special attention in terms of conservation of their gene pools.

2 Materials and methods

2.1 Sampling and DNA extraction

We sampled 60 populations of Scots pine in Eurasia, yielding a total of 1262 trees (Appendix Table 4; Fig. 1). Our dataset comprised locations that have different postglacial histories, as shown in previous studies. Needles were sampled from randomly selected, standing trees. Fresh needles collected from all individuals were stored at -20°C until extraction. DNA was subsequently extracted from 50 to 70 mg of tissue ground in liquid nitrogen according to the CTAB protocol (Dumolin et al. 1995). The protocol was modified to remove RNA from the samples by adding the final incubation step with $5\ \mu\text{l}$

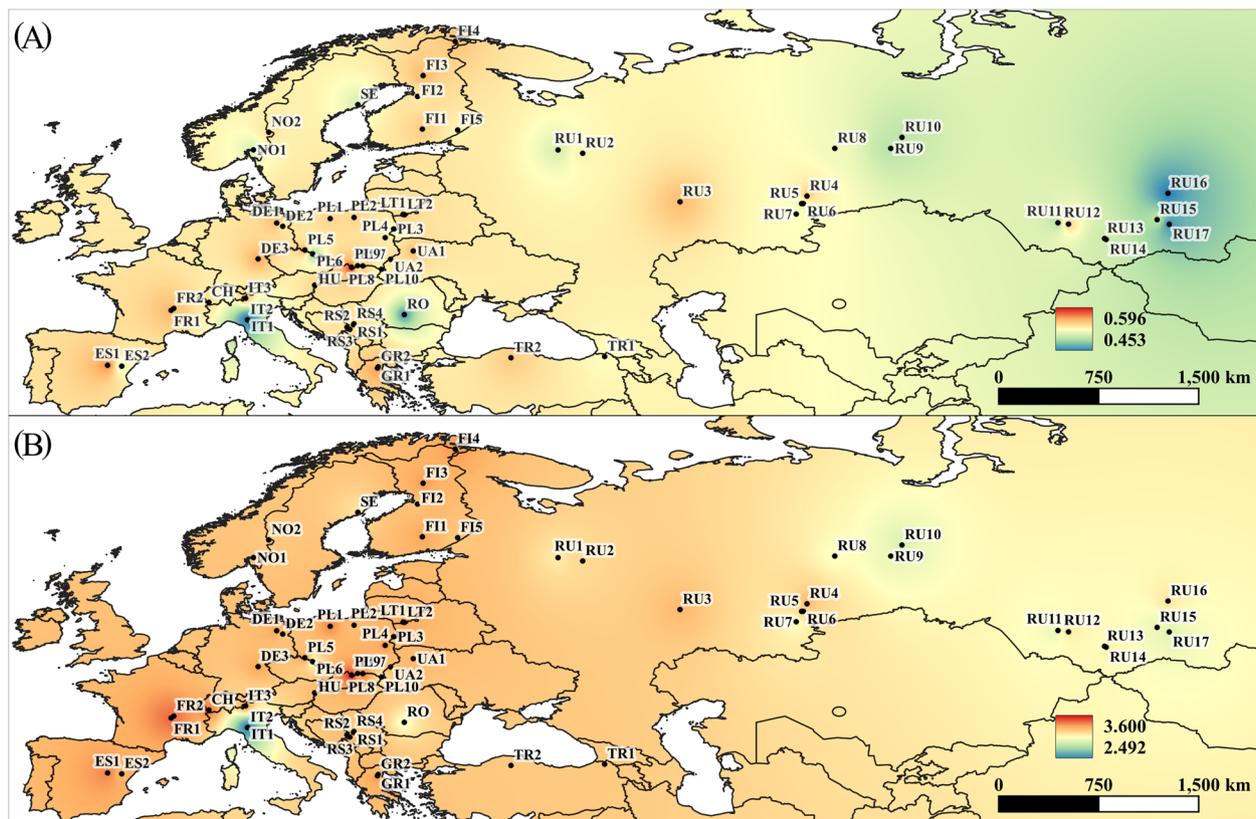


Fig. 1 Genetic diversity maps of the analyzed Scots pine populations: IDW interpolation of **A** expected heterozygosity values (H_e) and **B** allelic richness values (A_r). Acronyms as in Appendix Table 4

of RNase A (10 mg/ml) at 60 °C for 30 min. The DNA concentration was measured with a BioPhotometer plus spectrophotometer (Eppendorf AG, Germany) and adjusted to 15–20 ng/μl.

2.2 Molecular analysis

All samples were genotyped with 19 nuclear microsatellite loci (Appendix Table 5). The selected markers were amplified in three multiplex PCRs using the Qiagen Multiplex PCR Kit (QIAGEN, Germany). A specific fluorescent dye was attached to the forward primer in each primer pair. The concentration of each primer was adjusted to yield even fluorescent signals. The multiplexes comprised the following markers: (I) *psyl2*, *psyl16*, *psyl18*, *psyl25*, *psyl36*, *psyl42*, *psyl44*, and *psyl57*; (II) *spac11.4*, *spag7.14*, *ptTX2146*, *ptTX3025*, *ptTX4001*, and *ptTX4011*; (III) *spac11.8*, *spac12.5*, *pTctg4363*, *ptTX8446*, and *psyl17*. Each multiplex PCR was performed in a total volume of 10 μl using the protocols described by Żukowska et al. (2017). The fluorescently labeled PCR products, along with the GeneScan™ 500 LIZ™ internal size standard (Thermo Fisher Scientific, USA), were separated on the Applied Biosystems® 3130xl Genetic Analyzer (Thermo Fisher Scientific, USA). The allele sizes were determined using GeneMapper™ software ver. 4.0 (Thermo Fisher Scientific, USA). The raw data were assessed manually and converted into discrete allele sizes using the automated binning software Tandem (Matschiner and Salzburger 2009). All size variants were also checked and adjusted manually.

2.3 Genetic variation and differentiation

To detect loci that deviated from Hardy–Weinberg equilibrium (HWE) in each population separately, we used the score test (U test) with a Markov chain (MC) algorithm to estimate the p values (Guo and Thompson 1992), implemented in the online version of GENEPOP ver. 4.7.5 (Rousset 2008). Both the heterozygote deficiency and heterozygote excess hypotheses were tested. In addition, we also employed the multisample version of the score test to check each locus across all populations. The same software was used to test linkage disequilibrium (LD) between each pair of loci in each population by using the log-likelihood ratio statistic (G -test) with the MC algorithm of Raymond and Rousset (1995). As we suspected the presence of null alleles in loci with significant deficiencies of heterozygotes, we checked the frequency of null alleles with a maximum likelihood method using the expectation–maximization (EM) algorithm of Dempster et al. (1977) available in FreeNA (Chapuis and Estoup 2007).

For each locus, population, and geographic region, we calculated basic genetic parameters, including the mean

number of alleles (A), mean effective number of alleles (A_E), number of private alleles (A_P), and observed (H_O) and expected heterozygosity (H_E) using GenAlEx ver. 6.5 (Peakall and Smouse 2006, 2012). FSTAT ver. 2.9.4. (Goudet 2003) was used to compute the inbreeding coefficients (F_{IS}) and mean rarefied allelic richness (A_R). Subsequently, the values of H_E and A_R calculated for all populations were interpolated and presented on maps by using an inverse distance weighting (IDW) interpolation available in QGIS ver. 3.22 (QGIS Development Team 2021). In addition, we performed a linear regression analysis between the values of A , A_E , A_R , H_O , H_E , and the latitude of all populations using JMP® Pro ver. 16.1.0 (SAS Institute Inc. 2021). Considering the presence of null alleles, we also corrected the F_{IS} values (F_{ISnull}) with a Bayesian approach implemented in the INEst software ver. 2.2 (Chybicki and Burczyk 2009). The following parameters were set: 200,000 MC Monte Carlo cycles with every 20th cycle updated and a burn-in of 20,000. The deviance information criterion (DIC) was used to compare the full model (“nfb”—with null alleles, $F_{IS} > 0$, and genotyping failures) with the random mating model (“nb”—with null alleles and genotyping failures; $F_{IS} = 0$), thereby testing the significance of inbreeding.

2.4 Population structure

We assessed the genetic differentiation among the studied Scots pine populations grouped into geographic regions by computing pairwise fixation indexes (F_{ST}) and pairwise Slatkin’s analogues of F_{ST} (R_{ST}) using the SPAGeDi ver. 1.5 computer package (Hardy and Vekemans 2002). A total of 10,000 permutations were performed to calculate the p values. Subsequently, principal coordinate analysis (PCoA), implemented in GenAlEx, was carried out to infer the genetic relationships among the geographic regions. Additionally, we used a Bayesian clustering approach implemented in BAPS ver. 6.0 (Corander et al. 2008) and STRUCTURE ver. 2.3.4 (Pritchard et al. 2000; Falush et al. 2003). Both programs perform well when the clusters are not well differentiated ($F_{ST} = 0.02–0.03$), but they use different methods to search for the most likely number of clusters (Latch et al. 2006). For BAPS, we performed nonspatial genetic mixture clustering at the population level to define genetically divergent groups of populations. To this end, we initially tested all possible numbers of groups ($K = 1–60$) with 10 independent runs and repeated the analysis with 50 replicates for $K = 1–12$. Finally, we carried out admixture clustering using the obtained result of mixture clustering with parameters recommended by the authors. With STRUCTURE, we performed 500,000 iterations following a burn-in period of 50,000. We chose the admixture ancestry model with correlated allele frequencies. This configuration is

considered best when the population structure is subtle (Falush et al. 2003). Twenty independent runs were set for $K=1-12$. The output from STRUCTURE was visualized in the STRUCTURE HARVESTER web application (Earl and vonHoldt 2012). The optimum value of K was determined according to the Evanno method (Evanno et al. 2005). We aligned multiple runs for the optimum K value using the LargeKGreedy algorithm implemented in StructureSelector web-based software (Li and Liu 2018). StructureSelector was also used to generate STRUCTURE bar plots. Finally, to check the differentiation among populations, geographic regions, and the groups defined by the PCoA and Bayesian clustering, we ran the analysis of molecular variance (AMOVA) in Arlequin ver. 3.5.2.2 (Excoffier and Lischer 2010). A total of 10,000 permutations were performed to calculate the p values.

2.5 Phylogeographic structure, isolation by distance, and demographic history

We tested for a phylogeographic pattern among all populations, geographic regions, and populations within each geographic region using the permutation test by Hardy et al. (2003). Pairwise R_{ST} and permuted R_{ST} (pR_{ST}) values were calculated and compared in SPAGeDi. P values were computed with 10,000 permutations. We subsequently checked whether differentiation among the Scots pine populations, as well as populations grouped into geographic regions, resulted from their isolation by distance (IBD) by performing a Mantel test implemented in GenAlEx. A total of 10,000 permutations were performed to calculate the p values. In the next step, we checked the possibility of bottlenecks, first by using the test developed by Cornuet and Luikart (1996) and implemented in INEst. This test is based on the excess of heterozygosity that can be observed as a consequence of a recent bottleneck compared to a demographically stable population (with constant size). As microsatellite markers do not mutate under a strict single-step mutation model (SMM), we chose a two-phase model (TPM) that incorporates both single-step and multistep mutations (Di Rienzo et al. 1994). The test was run with the parameters recommended by Peery et al. (2012). The number of coalescent simulations was set to 10,000. P values were determined with the Wilcoxon signed-rank test based on 1,000,000 permutations. We also used INEst to perform the test for the deficiency in M-Ratio (MR) (Garza and Williamson 2001), which is analogous to the test for the excess of heterozygosity but indicates bottlenecks that were more severe and took place in the distant past (Williamson-Natesan 2005). The simulation of a demographically stable population (MR^{eq}) was calculated as the mean of 10,000 coalescence replicates. The values of MR and MR^{eq} were compared for each population and the

significance of $MR < MR^{eq}$ was checked using the Wilcoxon signed-rank test based on 1,000,000 permutations.

We used an approximate Bayesian computation (ABC) framework implemented in DIYABC ver. 2.1.0 (Cornuet et al. 2014) to carry out demographic history modeling. We tested six demographic scenarios (Appendix Fig. 4). Populations were combined into four groups according to the results revealed by Bayesian clustering and PCoA: Spain/France (populations ES1-2, FR1-2); Fennoscandia/Central Europe/Balkans (NO1-2, SE, FI1-5, LT1-2, PL1-10, DE1-3, UA1-2, HU, RO, RS1-4, GR1-2); Eastern Europe (RU1-3); Urals/Asia (RU4-17). The old refugial Apennine and Turkish populations were excluded as outliers. Five scenarios were based on various versions of divergence events, whereas the sixth scenario assumed that populations from Eastern Europe result from admixture between European and Asian stands. Parameters used for computations are given in Appendix Table 6. Prior parameters for generation time, effective population sizes, the timing of events, and mutation rates were defined based on previous genetic and palynological studies. We performed 6×10^5 coalescent simulations for all scenarios. Simulations were repeated for the two best scenarios. For details on the best scenario choice methodology, please refer to Dering et al. (2021).

3 Results

3.1 Genetic polymorphism at the locus level

The final dataset consisted of 16 polymorphic loci for which the obtained electropherograms were of good quality, unambiguous, and reproducible (Żukowska et al. 2023b). Three markers were excluded due to inconclusive results (psyl16), insufficient quality (spag7.14), and no amplification in a large number of samples (spac11.8). The test for the presence of LD was significant in some populations for a few pairs of loci. Furthermore, some markers showed deviations from HWE. The mean frequency of null alleles averaged over all populations was $Null=0.018$ (range, 0.001–0.088). Hence, the cut-off of $Null=0.19$ (Chapuis et al. 2008), above which the value of H_E is significantly underestimated, was not exceeded in any case. Inbreeding coefficients corrected for the presence of null alleles (F_{ISnull}) were low for almost all loci, except psyl25, for which $F_{ISnull}=0.285$ (mean, 0.051). The results of the analysis using the INEst program showed that inbreeding was not a significant part of the model for ptTX4011, ptTX3025, and ptTX8446. In general, the markers from the “psyl” series were less polymorphic than the rest. A total of 242 alleles (A_{tot}) were detected, resulting in a mean $A_{tot}=15.13$ per locus (range, 3–45). Among them, 38 were private. The mean number of alleles calculated across all populations was $A=5.675$ (range, 1.267–16.200). Due to the presence

of many alleles with low frequencies, the mean effective number of alleles across all populations was lower ($A_E=3.298$; range, 1.016–10.812). The mean rarefied allelic richness values were slightly higher for most loci (mean $A_R=3.208$; range, 1.059–6.313). The mean value of observed heterozygosity almost equalled the mean value of expected heterozygosity ($H_O=0.526$ and $H_E=0.528$). However, these values differed considerably for individual markers (Appendix Table 5).

3.2 Genetic variation and differentiation

The values of genetic variation parameters were generally high and geographically structured. The western parts of the Scots pine distribution (Scandinavia, Finland, Central Europe, Western Europe, the Balkans, and Spain) had the highest genetic variation, the eastern European parts and Turkey (Eastern Europe, Urals, and Turkey) had intermediate values, whereas the remaining Asian part had the lowest genetic variation (Western Siberia I, Western Siberia II, and Eastern Siberia) (Table 1). This western-eastern pattern was visible on the maps showing the IDW of H_E and A_R (Fig. 1). The Apennines were one exception, with the lowest values of all calculated parameters ($A=4.125$, mean 8.212; $A_R=3.902$, mean 6.428; $A_E=2.346$, mean 3.595; $H_O=0.426$, mean 0.518; $H_E=0.462$, mean 0.532). The greatest variation was found among the values of A (mean, 8.212; range, 4.125–12.063), but it was due to the differences in the number of individuals from particular geographic regions, as the variation in the values of A_R was much lower (mean, 6.428; range, 3.902–7.475) (Table 1). We found more private alleles (A_p) in putative glacial refugial regions than in other parts of the Scots

pine distribution (including Spain, the Italian Alps, Serbia, and southern part of the Urals) (Appendix Table 7). A linear regression analysis found no correlation between the values of genetic variation parameters and latitude (results not shown).

The genetic variation parameters were similar for individual populations within geographic regions, with few exceptions. Slightly lower values of the calculated genetic parameters were found for the Swedish population, one mountain and one peat-bog stand from Poland (PL6 and PL10, respectively), the Romanian population (RO, with only five individuals analyzed), and three locations in Asia (RU10, RU11, and RU13). Notably, the PL8 population from the Tatra Mountains had the highest values of $A_R=3.600$ (mean, 3.208), $H_O=0.607$ (mean, 0.526) and $H_E=0.596$ (mean, 0.528). The F_{ISnull} values were very low in almost all stands (range, 0.004–0.061; mean, 0.029), except PL6 ($F_{ISnull}=0.102$) and PL10 ($F_{ISnull}=0.443$). Comparisons between the full model and the random mating model showed that inbreeding had a significant effect on the genetic variability of 26 out of 60 Scots pine populations (Appendix Table 7).

The level of interpopulation differentiation was low but significant (Table 2). The global value of the fixation index was $F_{ST}=0.0276$ ($p<0.001$). The global value of the Slatkin’s analogue of F_{ST} was higher ($R_{ST}=0.0332$; $p<0.001$). Comparisons among geographic regions showed the distinctiveness of the Spanish and Apennine populations, with moderate pairwise values of R_{ST} . The only exception was the greater similarity between Spain and Western Europe. Surprisingly, the R_{ST} values between Eastern Siberia and Scandinavia, Finland, and Central Europe

Table 1 Values of genetic variation parameters in the studied Scots pine populations grouped into geographic regions

Region	Acronym	Populations	N	A	A_R	A_E	A_p	H_O	H_E
Scandinavia	Sca	NO1, NO2, SE	60	8.688	7.091	3.692	4	0.522	0.530
Finland	Fin	FI1-FI5	115	9.688	7.174	4.048	2	0.543	0.555
Central Europe	CEu	LT1, LT2, PL1-PL10, DE1-DE3, UA1, UA2	280	12.063	7.175	4.018	8	0.537	0.563
Balkans	Bal	HU, RO, RS1-RS4, GR1, GR2	156	11.063	7.163	4.029	11	0.546	0.563
Western Europe	WEu	FR1, FR2, CH, IT3	103	9.563	7.475	4.115	6	0.546	0.563
Spain	Spa	ES1, ES2	47	8.063	7.192	3.899	4	0.537	0.550
Apennines	Ape	IT1, IT2	40	4.125	3.902	2.346	1	0.426	0.462
Turkey	Tur	TR1, TR2	48	7.000	6.369	3.687	1	0.518	0.522
Eastern Europe	EEu	RU1-RU3	56	7.688	6.536	3.816	0	0.545	0.540
Urals	Ura	RU4-RU7	96	8.188	6.287	3.483	3	0.531	0.541
Western Siberia I	WSil	RU8-RU10	91	7.125	5.720	3.109	1	0.496	0.513
Western Siberia II	WSIil	RU11-RU14	140	7.750	5.723	3.378	1	0.496	0.519
Eastern Siberia	ESi	RU15-RU17	30	5.750	5.750	3.114	1	0.498	0.497
		Mean	97.08	8.212	6.428	3.595	3.31	0.518	0.532

N Number of individuals, A Mean number of alleles, A_R Mean rarefied allelic richness, A_E Mean effective number of alleles, A_p Number of private alleles, H_O Mean observed heterozygosity, H_E Mean expected heterozygosity

Table 2 Pairwise F_{ST} (below diagonal) and R_{ST} (above diagonal) values among the studied Scots pine populations grouped into geographic regions. Values that are not significant ($p > 0.05$) are written in italics. Acronyms as in Table 1

	Sca	Fin	CEu	Bal	WEu	Spa	Ape	Tur	EEu	Ura	WSiI	WSiII	ESi
Sca		0.0098	<i>0.0027</i>	0.0369	0.0723	0.1578	0.1699	0.0387	0.0157	0.0183	0.0285	0.0375	<i>0.0100</i>
Fin	0.0032		<i>0.0013</i>	0.0086	0.0375	0.0943	0.0951	0.0313	<i>0.0026</i>	0.0078	0.0076	0.0272	<i>0.0033</i>
CEu	<i>0.0017</i>	0.0028		0.0133	0.0405	0.0983	0.1081	0.0171	0.0086	0.0143	0.0094	0.0378	<i>0.0062</i>
Bal	0.0075	0.0062	0.0026		0.0191	0.0688	0.0549	0.0248	0.0191	0.0259	0.0141	0.0489	0.0178
WEu	0.0136	0.0116	0.0115	0.0092		0.0279	0.0679	0.0230	0.0647	0.0557	0.0363	0.0829	0.0508
Spa	0.0343	0.0284	0.0307	0.0138	0.0302		0.0733	0.0892	0.1211	0.1042	0.0740	0.1274	0.1131
Ape	0.0594	0.0529	0.0573	0.0665	0.0565	0.0898		0.1427	0.1097	0.1188	0.1046	0.1318	0.1224
Tur	0.0142	0.0112	0.0130	0.0193	0.0148	0.0278	0.0688		0.0579	0.0503	0.0351	0.0888	0.0506
EEu	0.0106	0.0074	0.0068	0.0223	0.0108	0.0350	0.0791	0.0242		<i>0.0070</i>	0.0104	0.0223	<i>0.0049</i>
Ura	0.0211	0.0172	0.0175	0.0266	0.0234	0.0408	0.0898	0.0341	0.0050		<i>0.0055</i>	<i>0.0036</i>	<i>-0.0056</i>
WSiI	0.0190	0.0165	0.0156	0.0265	0.0207	0.0420	0.0817	0.0362	0.0062	0.0050		0.0221	<i>0.0017</i>
WSiII	0.0316	0.0256	0.0266	0.0381	0.0280	0.0586	0.1031	0.0509	0.0136	0.0138	0.0122		<i>-0.0015</i>
ESi	0.0200	0.0217	0.0221	0.0308	0.0254	0.0480	0.0927	0.0412	0.0172	0.0163	0.0072	0.0071	

were not significant ($p > 0.05$). Scots pine from the Balkans was genetically closer to Scots pine from Finland and Central Europe than to the remaining geographic regions. In most cases, the values of F_{ST} and R_{ST} between adjacent geographic regions were low or nonsignificant (Table 2).

3.3 Population and phylogeographic structure, demographic history

The results of Bayesian clustering with BAPS and STRUCTURE were consistent. The most likely genetic structure identified with STRUCTURE consisted of two genetic groups (Fig. 2B): one comprising the Russian populations and the second including all other locations. The second most likely number of K was K=5

(Fig. 2C), with the following genetic clusters: (1) Spain and France; (2) the Apennines; (3) Turkey; (4) the Urals and Asian part of Russia; and (5) the admixed group of the remaining populations. The analysis carried out in BAPS showed similar results but only three genetic clusters: (1) the Apennines; (2) the Urals and Asian part of Russia + one easternmost European population; and (3) the remaining European populations (Fig. 2D). The results of the Bayesian clustering with K=5 were similar to the results of the PCoA (Fig. 3). The AMOVA outcomes confirmed that grouping into five genetic clusters best reflected the identified patterns of genetic differentiation, although the calculated values were low ($F_{ST} = 0.0228$, $R_{ST} = 0.0371$; $p < 0.001$) (Table 3).

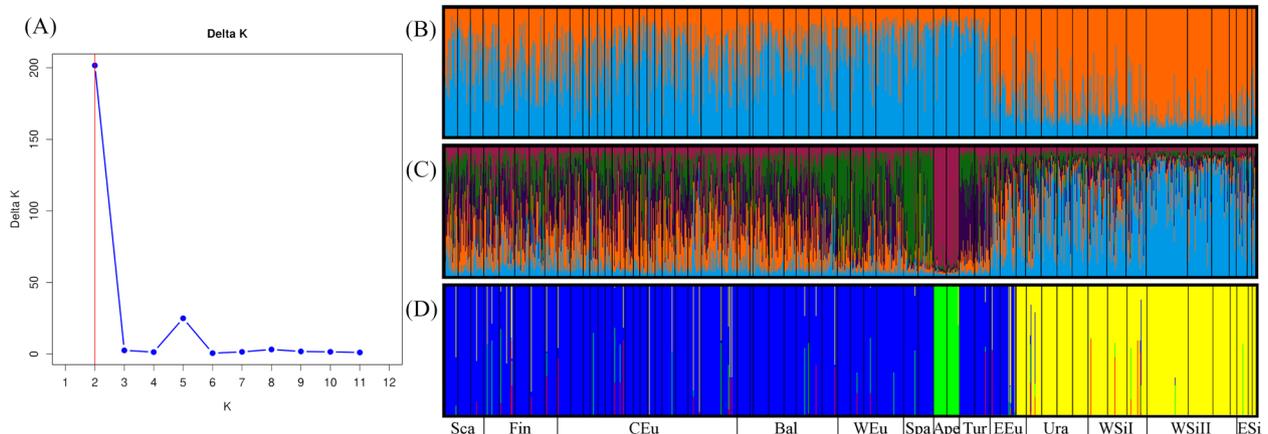


Fig. 2 Bayesian clustering of the Scots pine populations analyzed in the study. **A** Distribution of ΔK over $K = 1-12$ generated by STRUCTURE HARVESTER. **B** Bar plot showing STRUCTURE results for $K = 2$. **C** Bar plot showing STRUCTURE results for $K = 5$. **D** Bar plot showing BAPS results. The populations are separated by black vertical lines. The order of the populations corresponds to Appendix Table 4. Acronyms as in Table 1

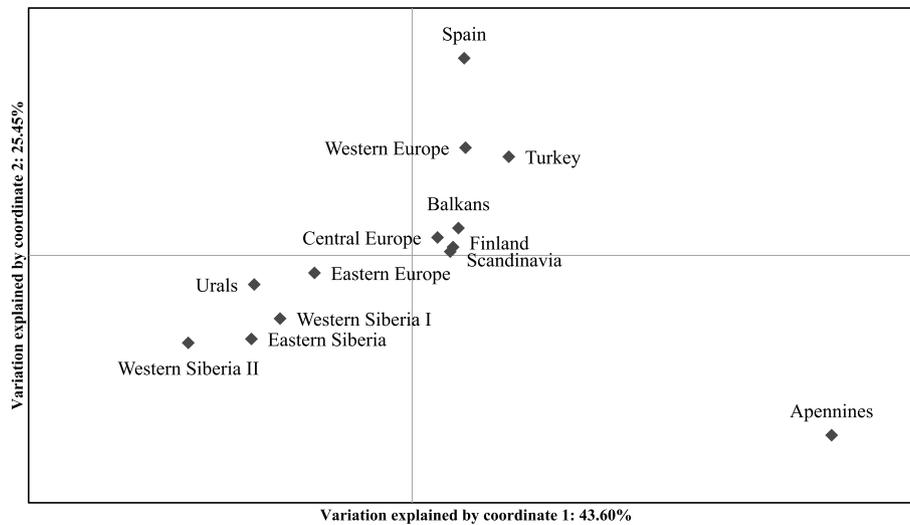


Fig. 3 PCoA among the studied Scots pine populations grouped into geographic regions

Table 3 AMOVA partitioning of genetic variation: (a) among geographic regions; (b) among groups defined by STRUCTURE (K=5); (c) among groups defined by BAPS (K=3). All values are significant at $p < 0.001$ (except * $p = 0.003$)

Source of variation	d.f	F_{ST}		R_{ST}		
		% variation	F statistics	% variation	F statistics	
(a)	Among groups	12	1.86	0.0186	3.42	0.0343
	Among populations within groups	47	1.07	0.0109	0.80	0.0083*
	Within populations	2464	97.07	0.0293	95.78	0.0423
(b)	Among groups	4	2.28	0.0228	3.71	0.0371
	Among populations within groups	55	1.38	0.0141	1.65	0.0171
	Within populations	2464	96.34	0.0366	94.64	0.0536
(c)	Among groups	2	2.27	0.0277	2.93	0.0293
	Among populations within groups	57	1.67	0.0171	2.49	0.0256
	Within populations	2464	96.06	0.0394	94.58	0.0542

The comparison of the values of R_{ST} vs. pR_{ST} confirmed the existence of a phylogeographic structure among geographic regions ($p = 0.0432$) and populations ($p = 0.0525$). However, this is primarily the effect of marginal stands because at the level of geographic regions, the phylogeographic structure was significant only among populations within Western Europe ($p = 0.0588$), Western Siberia I ($p < 0.001$), and Western Siberia II ($p = 0.0405$). The Mantel test, although significant ($p = 0.015$), revealed almost no correlation between genetic and geographic distances for all populations ($R^2 = 0.018$) (Appendix Fig. 5). The test for the excess of heterozygosity did not reveal any recent bottlenecks in any of the analyzed populations. Conversely, the calculated MRs were significantly below the mean MRs^{eq} for 41 out of 62 populations tested (Appendix Table 7). The only geographic region where none of the tests showed past bottlenecks was Spain.

The initial results of the ABC analysis showed that scenarios 3 and 4 are more likely than the others, although the estimated probability values were not high (21.07% and 20.93%, respectively). A re-analysis of these two best scenarios identified scenario 3 as the most likely (probability value 55.73%). The PCA for simulated data overlapped with the PCA results for the observed data (Appendix Fig. 6). According to the best scenario 3, populations from Eastern Europe diverged from the Urals/Asia group. The estimated divergence time was 1580 generations, i.e. (assuming a generation time of ~ 25 years) ~ 39,500 years BP. Divergence time back to the common ancestor was calculated to be ~ 4740 generations/118,500 years BP. The effective population size was the lowest for the Urals/Asia group (5180) and highest for the Fennoscandia/Central Europe/Balkans group (8790) (Appendix Table 8).

4 Discussion

Studies of genetic variation are particularly important in the era of climate change. Knowledge of neutral genetic variation, population history, and genetic relationships among different geographical regions is the first step of further research on the genetic basis of adaptive variation. To the best of our knowledge, this is the first comprehensive nuclear microsatellite-based study in the Eurasian range of Scots pine to address the question of admixture and homogenization of gene pools among populations of presumably distinct origins. Our results point to the differences between European and Asian Scots pine gene pools. We found that overall, Scots pine has high genetic variation, but this variation is significantly lower in its easternmost stands as well as in the Apennines and a few isolated areas. It also seems that the postglacial history of this species in Europe is very different from that in Asia, including recolonization from a higher number of glacial refugia, significant admixture, and the influence of human activities. Nuclear microsatellites seem to be useful in showing areas of secondary contact between once separated populations or the ones that have different origins.

4.1 The level and geographic pattern of genetic variation

Our results meet the expectations for a wind-pollinated tree species with a large census population size (Pyhäjärvi et al. 2020) and are in line with previous microsatellite and isoenzyme studies that showed substantial genetic variation but low among-population differentiation of Scots pine (Goncharenko et al. 1993; Prus-Głowacki et al. 1993; Provan et al. 1998; Bernhardsson et al. 2016; Wójkiewicz and Wachowiak 2016; Dering et al. 2017). We could distinguish the European and Asian gene pools, which differed from each other. These differences were evident both in the values of genetic variation parameters and in the Bayesian clustering, as well as in the PCoA. At the same time, the populations from Eastern Europe had a transitional character with intermediate genetic variation parameter values. This pattern definitely differs from the once widely accepted pattern of “southern richness to northern purity” (Hewitt 2000), which shows that Scots pine has a more complicated postglacial recolonization history.

Modes of transmission and mutation and migration rates differ for various marker types (mitochondrial, chloroplast, and nuclear) in pines. Thus, it is desirable to carry out phylogeographic studies with combined types of markers, especially since mtDNA, the only genome inherited in pines solely by seeds, evolves slowly and has a low level of intraspecific variation (Tóth et al. 2017b). The results of our study confirm the widely

recognized locations of southern glacial refugia for Scots pine around the Mediterranean basin. In our study, we observed clear distinctiveness of the Apennine Peninsula. Likewise, Spanish populations also formed a separate genetic cluster. The pairwise F_{ST}/R_{ST} values between these two regions and the remaining areas were the highest. Among them, we detected less differentiation between Spain and Western Europe. Furthermore, Spanish and French stands grouped in the Bayesian clustering, showing a limited contribution of Iberian Scots pine populations to the recolonization of Western Europe. Demographic history modeling with DIYABC indicated that French and Spanish populations split from other parts of Europe already at the very beginning of the Last Glacial period (~118,500 years BP). Nevertheless, the best scenario (scenario 3) was only slightly more likely than scenario 4 which assumed that this divergence took place ~39,500 BP. Regardless of the time of divergence, the uniqueness of the gene pool of Iberian Scots pine, which we still observe for various types of genetic markers, confirms that the Pyrenees are a strong geographical barrier to gene flow (see also Sect. 4.3). According to our results, Turkish populations were only partly similar to Central and Northern European stands but were different from the eastern part of the Scots pine range, as evidenced by the PCoA, Bayesian clustering with STRU CTURE, and pairwise F_{ST}/R_{ST} . These results confirmed earlier findings of no or only little contribution of populations from the Iberian and Apennine Peninsulas, as well as from Turkey, to the northwards recolonization of Europe (Prus-Głowacki et al. 2003; Cheddadi et al. 2006; Pyhäjärvi et al. 2008; Wójkiewicz and Wachowiak 2016; Dering et al. 2017).

The origin of Scots pine in Central and Northern Europe is still a subject of much debate. The outcomes of this study support earlier reports indicating that Scots pine could migrate to Central Europe and far north from glacial refugia located in the Balkan Peninsula and in the Alps (e.g., Wachowiak et al. 2014; Wójkiewicz and Wachowiak 2016). In our research, populations from the Balkan Peninsula and Central and Northern Europe formed one genetic cluster with high admixture. Growing evidence of cryptic refugia at mid-latitudes makes it also possible that Scots pine in Central and Northern Europe descends from more refugial areas. The fact that we detected a high degree of admixture in Central and Northern Europe supports this hypothesis. Palynological and macrofossil data, as well as earlier studies with genetic markers, indicate the possible existence of northern, cryptic refugia of Scots pine in Central Europe in the Hungarian Plains, the Carpathians, and the Circum-Alpine area (e.g. Binney et al. 2009; Feurdean et al. 2011; Tóth et al. 2019), but also in northern latitudes in

Scandinavia (Stewart et al. 2010). Unfortunately, our sampling was too limited to complete this picture, as we had only two populations from the Alps (CH and IT3), and just a few individuals from the Romanian Carpathians (RO). We also lack samples from the British Isles, where the origin of Scots pine is still a subject of debate (McGeever and Mitchell 2016; Sassoon et al. 2021).

Our study did not reveal the dual recolonization of Fennoscandia, which was shown in earlier works (Naydenov et al. 2007; Pyhäjärvi et al. 2008; Wachowiak et al. 2022). This result is particularly interesting, indicating that pollen-mediated gene transfer in areas of secondary contact may be efficient enough to blur a phylogeographic signal detected by mitochondrial markers transmitted by seeds. Therefore, nuclear microsatellites might be useful in showing areas of secondary contact between once separated populations or the ones that have different origins. This finding should be considered with caution as it may depend on the number and set of markers/populations used (Zimmer and Sønstebo 2018). The existence of the northeastern refugium somewhere in western Russia, hypothesized in several studies (Naydenov et al. 2007; Pyhäjärvi et al. 2008; Buchovska et al. 2013; Wachowiak et al. 2022), cannot be unequivocally confirmed or rejected based on our findings. Our demographic history modeling implied that it is more likely that Eastern European populations have descended not from the European refugium but rather from the Urals/Asian stands. It should be emphasized, however, that due to the high level of admixture, drawing more definitive conclusions would require the use of complex modeling and an increase in both the sampling geographical range as well as the number of trees from individual populations. A more detailed analysis of our results showed that populations from the Eastern European Plain (RU1-3) could in fact be divided into two groups (western: RU1-2 and eastern: RU3), the former being genetically more similar to Central and Northern Europe. It seems that the southern Urals refugium had a greater contribution to the genetic variation of RU3 than RU1-2, which is consistent with the findings of Vasilyeva et al. (2021). As the IBD was non-significant in this area, the observed pattern does not result from a geographic distance. These outcomes are in line with previous mtDNA studies that suggested a split between western and eastern populations of Scots pine approximately along the 38th meridian (Semerikov et al. 2018). Nevertheless, an unambiguous reconstruction of the demographic history of this region definitely requires exhaustive sampling, which takes into account also areas located in southwest Asia.

Our results point to the lower genetic variation of Scots pine in Asia and its rather uniform genetic structure. This is most likely the effect of a relatively simple recent

demographic history of Scots pine in Asia, accompanied by limited admixture during expansion after the last glacial period. These results agree with previous findings based on mtDNA and cpDNA markers (Semerikov et al. 2014, 2018; Wachowiak et al. 2022). It should be mentioned that major parts of Russia, except the westernmost areas, remained ice-free during the LGM (Svendsen et al. 2004). For this reason, the number of glacial refugia was probably lower than in Europe. Moreover, gene flow could have had more time to homogenize Asian gene pools as compared to the European Scots pine. Our research indicates that this part of Scots pine distribution was recolonized by migrants from the southern Urals by a series of founder effects. Nevertheless, this conclusion needs further support with more thorough sampling from the Asian part of the Scots pine's range. Apart from that, it seems that populations from the Urals have at least partially contributed not only to the contemporary variation in Eastern Europe but also, although to a lesser extent, to genetic variation in Central Europe and Finland. We cannot exclude the possibility that the area at the borders of Russia, Kazakhstan, and Mongolia served as another source of migrants. Nevertheless, our dataset comprised only several populations from Siberia and no samples from the far eastern parts. Unfortunately, palaeobotanical and genetic data on the Russian part of the Scots pine distribution are not sufficient to draw definite inferences (Sannikov et al. 2020). Therefore, more comprehensive studies are required to resolve the postglacial history of this region.

4.2 Genetic admixture and human-mediated gene transfer

When discussing admixture patterns, it is difficult to omit the recent history of Scots pine, characterized by intensive breeding that may have significantly affected original postglacial patterns and levels of genetic variation. As forests were overexploited during the medieval and early modern periods (Hosius et al. 2006), the reduction in forest cover eventually forced European countries to develop large-scale reforestation programs, leading to the increase in a forest area during the nineteenth and twentieth centuries (Mather 2001). In many cases, the remaining forests could not meet the high demand for seeds; thus, large amounts of germplasm, especially seeds, were transferred across Europe (Tulstrup 1959). It is usually impossible to trace back the historical transfer routes of forest reproductive material (FRM). Nevertheless, historical data about the time and geographical scales of FRM transfer are enough to conclude that they have undeniably affected the initial gene pools of Scots pine and other forest tree species. The trade and transfer of seeds of Scots pine, Norway spruce, European larch, and oaks started in the eighteenth century and continued

throughout the nineteenth and early twentieth centuries (Tulstrup 1959; Jansen et al. 2017). Large-scale imports of non-native FRM to Fennoscandia began in the nineteenth century. During later afforestation efforts, non-autochthonous FRM of beech, oak, and Scots pine was widely introduced to Denmark, FRM of Norway spruce was transported to Sweden and Norway, and FRM of Scots pine was imported to Sweden (Myking et al. 2016). Most FRM was transferred from Germany, with little contribution of FRM from France, Belgium, and Hungary (Jansen et al. 2017). In contrast, Finland has limited records of FRM introductions (Myking et al. 2016). Reforestation efforts were not always successful. Hence, some countries tried to limit the use of foreign seeds and FRM from unknown sources. These failures eventually led to the establishment of the first international provenance trials and subsequent breeding programs. Today, the trade of FRM is regulated by European law and national regulations, but it is not obligatory to keep records about seed sources. Unfortunately, foresters and forest owners still tend to treat FRM as a cost that should be as low as possible. Furthermore, most Central and Northern European forests are intensely managed. Taking all these into account, almost no primary natural forests are left in Europe (Lorenz et al. 2005). Thus, virtually all genetic resources of Northern and Central European tree species have been shaped by a combination of natural processes, such as postglacial recolonization and local adaptation, and human impacts, including seed transfer, fragmentation, and silviculture (Ratnam et al. 2014; Żukowska et al. 2023a). Substantial admixture found in our study, especially in Central Europe and Fennoscandia, supports the view that the time and geographical scales of FRM transfer have undeniably affected the native gene pools of Scots pine.

4.3 Conservation implications

Genetic conservation of Scots pine seems to be of low priority because of its very wide distribution spanning several climatic zones. However, Scots pine in Europe has already experienced retreats from several sites in western and southern parts of continental Europe, as well as from the British Isles. These movements occurred due to a combination of climate change and anthropogenic disturbance, favoring the spread of broad-leaf species (Tóth et al. 2017b). Our research shows clear separation and a lower level of genetic variation of Scots pine from the Apennines than of populations occupying other geographic regions. It remains to be verified whether this is a matter of lower effective population sizes, range fragmentation, geographic isolation, or a combination of all these factors. Similar results were previously obtained by Belletti et al. (2012). The authors used nine nuclear

microsatellites to analyze 20 populations of Scots pine from the Italian Alps and a single stand from the Apennines. The only population from the Apennines had the lowest values of genetic diversity parameters and constituted a separate genetic cluster. In another study, Scalfi et al. (2009) examined one Alpine and three Apennine Scots pine stands using three nuclear and two chloroplast microsatellites. The authors noticed the genetic distinctiveness of Apennine Scots pine and its possible origin from a separate glacial refugium. Nevertheless, they concluded that the studied Apennine populations still retain a high level of genetic diversity. As our study comprised two of the three nuclear microsatellites used by Scalfi et al. (2009), it emphasizes how the choice and number of markers can lead to distinct conclusions.

Pairwise comparisons of R_{ST} values between geographic regions in our study showed that, apart from the Apennines, also Spanish populations have a unique gene pool. R_{ST} indexes are supposed to be higher between populations that have been isolated from each other for a longer time (Slatkin 1995). Therefore, our study is another that confirms the long-lasting isolation and small contribution of Iberian populations to the recolonization of other parts of Europe due to the Pyrenees acting rather as a barrier than a bridge for gene exchange (Cheddadi et al. 2006; Pyhäjärvi et al. 2008; Dering et al. 2017; Prus-Głowacki et al. 1993; Wachowiak et al. 2022). Scots pine in the Iberian Peninsula inhabits disjunct mountainous areas with different climatic and ecological conditions. Individual populations are separated by large distances (Prus-Głowacki et al. 2003). Our study comprised only two stands from Spain but they are over 100 km apart with distinct microhabitats. Of these, population ES2 from Sierra de Gúdar is particularly interesting, as it is located at the eastern end of the Iberian System at an altitude of 2000 m a.s.l., which is usually covered with snow in winter, but it is also near the Mediterranean coast which makes the local climate milder than far inland. Both ES1 and ES2 had similar values of parameters connected with allelic diversity, but they slightly differed in heterozygosity values ($H_O=0.555$ vs. 0.519; $H_E=0.567$ vs. 0.516, respectively). It is also worth pointing out that none of the tests showed past bottlenecks, either in ES1 or ES2.

Geographically isolated populations are usually small and fragmented, making them more prone to genetic drift, inbreeding, and environmental changes. The study of the evolutionary potential of marginal populations is also important due to projected range shifts (Dyderski et al. 2018), which may lead to the loss of valuable gene pools. For Scots pine, many such populations occupy mountain ranges in southern peripheral areas. Studies from the Carpathian Mountains (Tóth

et al. 2017a; Şofletea et al. 2020) showed that Scots pine growing in these territories has higher genetic variation than in non-mountainous areas, but the gene pools are not homogenous. They are also far less likely to be affected by the past human-mediated transfer of FRM. Recognition and conservation of the genetic pools in hard-to-reach and marginal areas should now be the priority. Whether the conservation policies should take the form of ex situ actions or rather of genetic rescue by promoting gene exchange, should be evaluated separately for each region. Unfortunately, our research comprised just a few isolated stands of Scots pine. We suppose that if we had more samples from the Carpathians and the Alps, the genetic variability of the European Scots pine range compared to its Asian part would be even higher. This would also allow more accurate inferences about human impact on the level and distribution of Scots pine’s genetic variation. In the context of comparing genetic variation between Europe and Asia, it would also be extremely important to thoroughly examine the Urals region.

5 Conclusion

This study provides important insights into the patterns of genetic variation of Scots pine across its Eurasian distribution. We revealed a high level of genetic admixture in Central and Northern Europe, which seems to be partially the result of postglacial recolonization from multiple glacial refugia but also human-mediated gene transfer. It appears that the Asian part of the Scots pine distribution has a different postglacial demographic history. The lower and homogenous genetic variation of Asian stands suggests their establishment by founder effects, possibly from the southern Urals refugium. This refugium had probably also served as the source of migrants that recolonized Eastern Europe, but this conclusion requires further research. Populations from the Apennines require special attention reflecting a general need for recognition and conservation of fragmented and marginal stands. Certainly, more comprehensive studies are necessary to unambiguously identify midlatitude hotspots of Scots pine genetic variation. In light of ongoing climate change, it is important to carry out genetic research in isolated, mountainous areas and those that thus far have been poorly covered: (1) Eastern Europe (Russia, Belarus, and Ukraine); (2) the southwestern Balkans (Bosnia and Herzegovina, Montenegro, Albania, and North Macedonia); (3) mid and high latitudes in Asia; and (4) the Russian Far East.

Appendix

Table 4 List of the Scots pine populations analyzed in the study

Acr	Location	N	Lat	Long	Acr	Location	N	Lat	Long
NO1	Norway, Ostamarka	20	59.87	11.05	RS4	Serbia, Velika Plec	20	44.18	20.03
NO2	Norway, Smoldalen	20	61.46	12.42	GR1	Greece, Fteri	20	40.19	22.13
SE	Sweden, Backsjön	20	63.95	20.38	GR2	Greece, Pieria Ori	24	40.31	22.20
FI1	Finland, Joutsa	24	61.74	26.14	FR1	France, Arlanc-Mayres	20	45.40	3.69
FI2	Finland, nr Temmes	24	64.69	25.71	FR2	France, Chabrioux	20	45.57	3.93
FI3	Finland, nr Rovaniemi	23	66.57	26.21	CH	Switzerland, Follatères	20	46.12	7.07
FI4	Finland, nr Kielajoki	24	69.65	29.07	IT3	Italy, Alps near Bornio	43	46.49	10.35
FI5	Finland, Savonlinna	20	61.66	29.30	ES1	Spain, Cuenca Una	23	40.44	-1.98
LT1	Lithuania, Ciapkeliai	20	54.05	24.42	ES2	Spain, Sierra de Gúdar	24	40.36	-0.72
LT2	Lithuania, Marcinkonys	20	54.02	24.55	IT1	Italy, Casina	20	44.54	10.52
PL1	Poland, Woziwoda	10	53.67	17.91	IT2	Italy, Casina	20	44.54	10.54
PL2	Poland, Tabórz	13	53.77	20.04	TR1	Turkey, Savsat Ardahan	24	41.23	42.43
PL3	Poland, Hajnówka	11	52.74	23.58	TR2	Turkey, Tosya Ilgazi	24	41.12	34.06
PL4	Poland, Liski Reserve	11	51.95	22.82	RU1	Russia, Goritsy	16	59.86	38.27
PL5	Poland, Chojnik	20	50.83	15.64	RU2	Russia, Kadnikov	25	59.56	40.47
PL6	Poland, Szczeliniec	13	50.48	16.34	RU3	Russia, Volga-Kama Nature Reserve	15	55.18	49.17
PL7	Poland, Pusta Wielka	10	49.40	20.82	RU4	Russia, Kysztym	24	55.70	60.51
PL8	Poland, Koryciska Wielkie	12	49.27	19.83	RU5	Russia, Lake Turgoyak	24	55.02	60.03
PL9	Poland, Pieniński National Park	12	49.42	20.36	RU6	Russia, Miasskiy Trakt	24	55.02	60.19
PL10	Poland, Tarnawa	11	49.10	22.49	RU7	Russia, Karagayskiy Bor	24	54.08	59.55
DE1	Germany, Godendorf	20	53.27	13.13	RU8	Russia, Kondinskie	30	60.00	63.00
DE2	Germany, Joachimsthal	20	52.97	13.67	RU9	Russia, Mukhrino	30	60.00	68.00
DE3	Germany, Rotmaital	21	50.03	11.47	RU10	Russia, Khanty-Mansiysk	31	61.00	69.00
UA1	Ukraine, Luck	32	50.75	25.33	RU11	Russia, Pavlovsk	64	53.29	82.94
UA2	Ukraine, Yavorivskiy National Park	24	50.00	23.30	RU12	Russia, Bobrovka	38	53.17	83.88
HU	Hungary, Sopron	20	47.67	16.52	RU13	Russia, Karmanka	27	51.87	87.10
RO	Romania, Dedulesti	5	45.01	24.53	RU14	Russia, Artybasz	11	51.79	87.25
RS1	Serbia, Divčibare Mts	24	44.10	19.99	RU15	Russia, Minusinsk	17	53.57	91.80
RS2	Serbia, Zaovine-Paljevine road	24	43.87	19.41	RU16	Russia, Krasnoyarsk Pillars	6	55.95	92.77
RS3	Serbia, Zlatibor Mt	19	43.68	19.67	RU17	Russia, Tanzybey	7	53.16	92.88

N Number of individuals

Table 5 Values of genetic variation parameters calculated for each microsatellite marker

Locus	A_{tot}	A	A_R	A_E	A_P	H_O	H_E	Null	F_{ISnull}
psyl42	6	4.067	3.152	2.891	0	0.673	0.644	0.010	0.023
psyl25	3	1.267	1.059	1.016	0	0.010	0.015	0.008	0.285
psyl2	7	2.467	1.861	1.442	0	0.290	0.293	0.028	0.038
psyl18	7	2.350	1.382	1.124	0	0.101	0.098	0.001	0.023
psyl57	9	4.600	2.867	2.126	2	0.527	0.509	0.013	0.018
psyl36	5	3.033	1.748	1.280	0	0.205	0.203	0.010	0.027
psyl44	5	2.667	1.497	1.165	0	0.135	0.129	0.005	0.023
ptTX4011	8	4.783	3.238	2.732	1	0.489	0.617	0.088	0.088 ^{ns}
spac11.4	24	9.517	4.983	5.815	6	0.822	0.818	0.015	0.043
ptTX3025	11	4.567	2.865	2.320	0	0.535	0.551	0.027	0.051 ^{ns}
ptTX2146	29	7.083	3.977	3.814	6	0.739	0.723	0.016	0.035
ptTX4001	21	7.550	4.046	3.745	3	0.737	0.718	0.016	0.028
spac12.5	45	16.200	6.313	10.812	7	0.925	0.898	0.004	0.019
psyl17	19	6.367	4.120	4.150	4	0.763	0.753	0.010	0.035
pTctg4363	25	6.217	3.617	3.260	6	0.705	0.681	0.011	0.022
ptTX8446	18	8.067	4.608	5.078	3	0.766	0.790	0.027	0.062 ^{ns}
Mean	15.13	5.675	3.208	3.298	2.38	0.526	0.528	0.018	0.051
Total	242								

A_{tot} Total number of alleles, A Mean number of alleles, A_R Mean rarefied allelic richness, A_E Mean effective number of alleles, A_P Number of private alleles, H_O Mean observed heterozygosity, H_E Mean expected heterozygosity, $Null$ Frequency of null alleles, F_{ISnull} Inbreeding coefficient corrected for the presence of null alleles, ^{ns} Inbreeding is not a significant part of the model

Table 6 Parameters used for ABC demographic history modeling

Parameters	Distribution	Min–Max	Mean	Shape
Genetic				
Set 1 (psyl2, psyl17, psyl18, psyl25, psyl36, psyl42, psyl44, psyl57, spac11.4, pTctg4363, ptTX4001, ptTX4011, ptTX8446)				
Mean mutation rate	Log uniform	1.10 ⁻⁷ -1.10 ⁻⁴		
Individual mutation rate	Gamma	1.10 ⁻⁹ -1.10 ⁻³	Mean mutation rate	2
Mean coefficient P	Uniform	0.1–9		
Individual locus coefficient P	Gamma	0.1–10	Mean coefficient P	2
Set 2 (spac12.5, ptTX2146, ptTX3025)				
Mean mutation rate	Log uniform	1.10 ⁻⁴ -1.10 ⁻²		
Individual mutation rate	Gamma	1.10 ⁻⁵ -1.10 ⁻²	Mean mutation rate	2
Mean coefficient P	Uniform	0.1–5		
Individual locus coefficient P	Gamma	0.1–10	Mean coefficient P	2
Historical				
Generation time [years]		25		
Effective population size		10–150,000		
t1		10–5,000		
t2		500–10,000		
ra		0.001–0.999		
Conditions	t2 < t1			

Table 7 Values of genetic variation parameters in the studied Scots pine populations

Pop	N	A	A _R	A _E	A _P	H _O	H _E	F _{ISnull}	MR	MR ^{eq}
NO1	20	6.563	3.252	3.594	0	0.530	0.510	0.007 ^{ns}	0.700	0.791
NO2	20	6.000	3.228	3.471	0	0.555	0.536	0.018	0.684	0.802
SE	20	5.938	3.138	3.154	4	0.482	0.510	0.031 ^{ns}	0.627*	0.802
FI1	24	6.875	3.335	3.634	0	0.579	0.546	0.009	0.716	0.793
FI2	24	6.563	3.256	3.535	0	0.533	0.534	0.013 ^{ns}	0.696*	0.810
FI3	23	6.375	3.347	3.726	1	0.518	0.554	0.047	0.614*	0.800
FI4	24	6.375	3.354	3.576	0	0.548	0.550	0.021 ^{ns}	0.721	0.801
FI5	20	6.125	3.233	3.622	1	0.538	0.530	0.009	0.680*	0.789
LT1	20	6.500	3.337	3.619	3	0.553	0.542	0.018 ^{ns}	0.670*	0.795
LT2	20	6.188	3.332	3.714	0	0.538	0.544	0.031 ^{ns}	0.692	0.793
PL1	10	5.000	3.384	3.248	0	0.563	0.543	0.021	0.661	0.750
PL2	13	5.250	3.269	3.216	1	0.529	0.543	0.009	0.634*	0.770
PL3	11	4.813	3.237	3.245	0	0.528	0.518	0.030 ^{ns}	0.699	0.791
PL4	11	5.250	3.366	3.344	1	0.517	0.534	0.016 ^{ns}	0.632*	0.767
PL5	20	5.938	3.349	3.509	2	0.539	0.557	0.019 ^{ns}	0.637*	0.801
PL6	13	4.438	3.020	2.889	0	0.440	0.483	0.102	0.682*	0.811
PL7	10	5.188	3.413	3.351	0	0.538	0.548	0.019 ^{ns}	0.636	0.757
PL8	12	5.438	3.600	3.602	0	0.607	0.596	0.043 ^{ns}	0.606*	0.765
PL9	12	5.000	3.223	2.982	0	0.510	0.516	0.035	0.605*	0.790
PL10	11	4.500	3.105	2.814	0	0.466	0.524	0.443 ^{ns}	0.608*	0.789
DE1	20	6.313	3.349	3.312	0	0.576	0.555	0.010	0.631*	0.788
DE2	20	5.750	3.219	3.359	0	0.534	0.532	0.012 ^{ns}	0.747	0.820
DE3	21	6.188	3.331	3.601	0	0.531	0.558	0.015	0.669*	0.811
UA1	32	7.125	3.282	3.507	2	0.553	0.551	0.014	0.635*	0.817
UA2	24	6.625	3.272	3.480	1	0.547	0.538	0.015	0.632*	0.816
HU	20	6.000	3.274	3.238	1	0.559	0.542	0.013 ^{ns}	0.647*	0.798
RO	5	3.250	3.022	2.572	0	0.534	0.466	0.023	0.635	0.727
RS1	24	5.938	3.190	3.277	3	0.555	0.541	0.010	0.695*	0.827
RS2	24	6.063	3.270	3.679	1	0.538	0.534	0.006	0.601*	0.802
RS3	19	6.188	3.308	3.613	1	0.549	0.540	0.019	0.677	0.793
RS4	20	6.125	3.208	3.156	4	0.556	0.534	0.004 ^{ns}	0.638*	0.794
GR1	20	6.313	3.463	3.643	2	0.536	0.580	0.061 ^{ns}	0.672*	0.791
GR2	24	5.188	3.171	3.476	0	0.533	0.532	0.024 ^{ns}	0.674*	0.835
FR1	20	6.313	3.460	3.683	1	0.586	0.559	0.007	0.696	0.793
FR2	20	6.500	3.409	3.593	0	0.540	0.545	0.010 ^{ns}	0.708	0.801
CH	20	6.313	3.414	3.854	0	0.523	0.538	0.014 ^{ns}	0.673*	0.799
IT3	43	7.438	3.339	3.535	5	0.541	0.550	0.013 ^{ns}	0.727	0.826
ES1	23	6.313	3.396	3.675	2	0.555	0.567	0.020	0.796	0.806
ES2	24	7.000	3.294	3.566	2	0.519	0.516	0.010	0.709	0.803
IT1	20	3.375	2.492	2.268	2	0.426	0.453	0.042	0.651*	0.868
IT2	20	3.688	2.554	2.311	1	0.426	0.456	0.019 ^{ns}	0.628*	0.847
TR1	24	5.563	3.231	3.360	2	0.513	0.526	0.027 ^{ns}	0.677*	0.810
TR2	24	6.063	3.233	3.185	0	0.523	0.546	0.021 ^{ns}	0.748	0.802
RU1	16	5.063	3.119	3.344	0	0.504	0.509	0.055 ^{ns}	0.691	0.792

Pop	<i>N</i>	<i>A</i>	<i>A_R</i>	<i>A_E</i>	<i>A_P</i>	<i>H_O</i>	<i>H_E</i>	<i>F_{ISnull}</i>	<i>MR</i>	<i>MR^{eq}</i>
RU2	25	6.000	3.177	3.637	0	0.549	0.530	0.005 ^{ns}	0.669*	0.811
RU3	15	5.250	3.274	3.322	0	0.579	0.552	0.011	0.662*	0.808
RU4	24	6.313	3.313	3.390	0	0.535	0.550	0.037 ^{ns}	0.680*	0.809
RU5	24	6.188	3.154	3.353	5	0.500	0.513	0.038 ^{ns}	0.536*	0.808
RU6	24	5.750	3.134	3.235	2	0.523	0.523	0.021	0.649*	0.821
RU7	24	5.063	3.004	2.793	1	0.564	0.525	0.005 ^{ns}	0.618*	0.820
RU8	30	5.625	3.039	3.056	0	0.528	0.519	0.016 ^{ns}	0.601*	0.833
RU9	30	5.438	2.989	3.031	1	0.491	0.497	0.013	0.635*	0.841
RU10	31	5.438	2.941	2.897	0	0.471	0.498	0.024 ^{ns}	0.630*	0.850
RU11	64	6.438	2.973	3.119	1	0.480	0.496	0.011 ^{ns}	0.608*	0.838
RU12	38	6.250	3.210	3.479	0	0.535	0.543	0.034 ^{ns}	0.605*	0.835
RU13	27	5.063	2.987	3.030	0	0.475	0.506	0.054 ^{ns}	0.628*	0.830
RU14	11	4.500	3.156	3.228	0	0.511	0.504	0.029 ^{ns}	0.592*	0.795
RU15	17	4.813	2.993	2.870	0	0.504	0.495	0.030	0.599*	0.822
RU16	6	3.750	3.123	2.659	1	0.479	0.454	0.019	0.606*	0.736
RU17	7	3.625	2.956	2.658	0	0.500	0.462	0.010	0.666	0.776
Mean	21.03	5.675	3.208	3.298	0.90	0.526	0.528	0.029		

N Number of individuals, *A* Mean number of alleles, *A_R* Mean rarefied allelic richness, *A_E* Mean effective number of alleles, *A_P* Number of private alleles, *H_O* Mean observed heterozygosity, *H_E* Mean expected heterozygosity, *F_{ISnull}* Inbreeding coefficient corrected for the presence of null alleles, ^{ns} Inbreeding is not a significant part of the model, *MR* Observed M-Ratio, *MR^{eq}* M-Ratio under mutation-drift equilibrium, * significantly lower than *MR^{eq}* ($p < 0.05$)

Table 8 Parameters estimated for the best demographic scenario (scenario 3) in the ABC demographic history modeling

Parameter	Q25	Median	Q75
N1	5690	6760	7760
N2	8150	8790	9310
N3	4670	5870	7150
N4	4170	5180	6310
t1	716	1580	2850
t2	3000	4740	7000
μmic1	7.77E-05	8.75E-05	9.47E-05
pmic1	0.825	0.883	0.9
snimic1	5.45E-07	3.40E-06	8.27E-06
μmic2	6.41E-04	1.02E-03	1.73E-03
pmic2	0.211	0.314	0.416
snimic2	9.28E-08	5.48E-07	2.63E-06

N1-4—effective population size for Pop1-4, respectively; t—time in generations

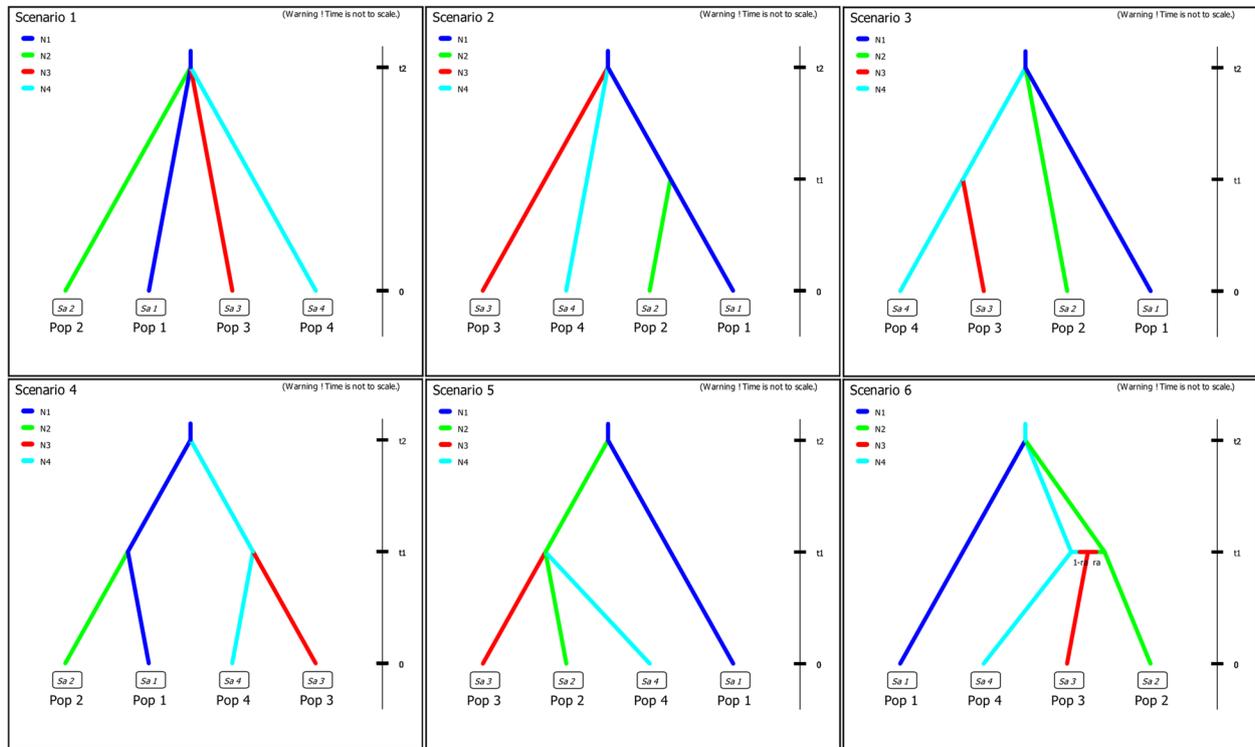


Fig. 4 Demographic scenarios tested with the ABC demographic history modeling. Pop1—Spain/France; Pop2 –Fennoscandia/Central Europe/Balkans; Pop3—Eastern Europe; Pop4—Urals/Asia

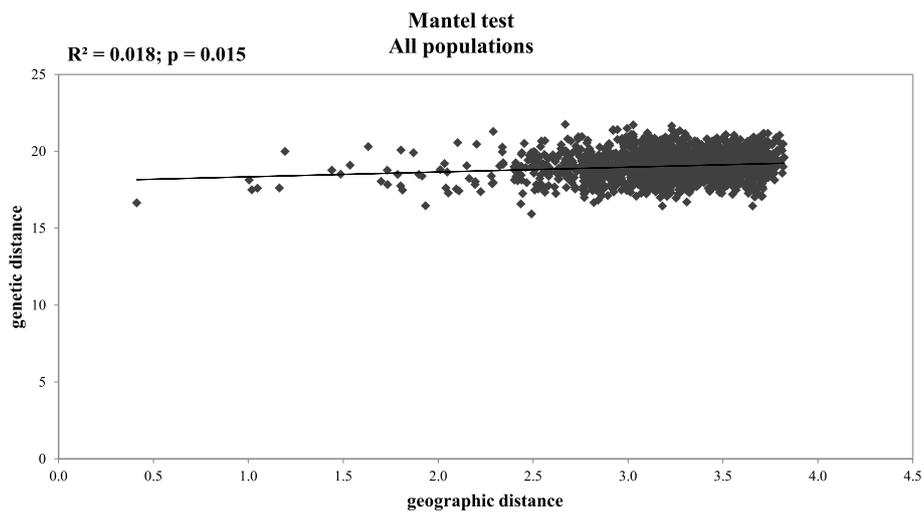


Fig. 5 Mantel test results for all Scots pine populations analyzed in this study

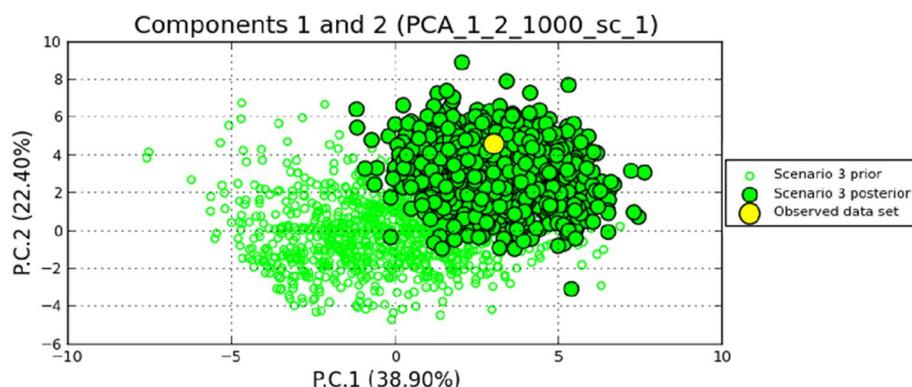


Fig. 6 Principal component analysis (PCA) carried out in DIYABC for the best demographic scenario (scenario 3)

Acknowledgements

We thank members of the GreenTree 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. We also express our great gratitude to Dr Łukasz Walas from the Institute of Dendrology, Polish Academy of Sciences, for his help with demographic history modeling.

Code availability

Not applicable.

Consent to participate

Not applicable.

Authors' contributions

Conceptualization: WBŻ, WW; methodology: WBŻ, WW; formal analysis and investigation: WBŻ, BW; writing—original draft preparation: WBŻ; writing—review and editing: WBŻ, BW, AL, RL, WW; funding acquisition: WBŻ, WW; resources: AL, RL, WW. The authors read and approved the final manuscript.

Funding

The research was financially supported by the Polish National Science Centre (UMO-2016/21/N/NZ9/01499) and the statutory activity of the Institute of Dendrology, Polish Academy of Sciences. WW acknowledges financial support from the Polish National Science Centre (UMO-2017/27/B/NZ9/00159).

Availability of data and materials

The dataset analyzed during the current study is available in the Zenodo repository, <https://zenodo.org/record/8405255>.

Declarations

Ethics approval and consent to participate

The authors declare that they obtained the approval of the Polish Ministry of Climate and Environment for conducting the study in protected areas in Poland.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, 62-035 Kórnik, Poland. ²Institute of Wildlife Biology and Wildlife Management, Faculty of Forestry, University of Sopron, Bajcsy-Zsilinszky U. 4, 9400 Sopron, Hungary. ³Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University in Poznań, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland.

Received: 25 February 2023 Accepted: 5 October 2023

Published online: 23 October 2023

References

Belletti P, Ferrazzini D, Piotti A, Moteleone I, Ducci F (2012) Genetic variation and divergence in Scots pine (*Pinus sylvestris* L.) within its natural range in Italy. *Eur J Forest Res* 131:1127–1138. <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. *For Ecol Manage* 361:131–143. <https://doi.org/10.1016/j.foreco.2015.10.047>

Binney HA, Willis KJ, Edwards ME, Bhagwat SA, Anderson PM, Andreev AA, Blaauw M, Damblon F, Haesaerts P, Kienast F, Kremenetski KV, Krivonogov SK, Lozhkin AV, MacDonald GM, Novenko EY, Oksanen P, Sapelko TV, Väiranta M, Vazhenina L (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quatern Sci Rev* 28:2445–2464. <https://doi.org/10.1016/j.quascirev.2009.04.016>

Buchovska J, Danusevičius D, Baniulis D, Stanyš V, Šikšnianienė J, KAvaliauskas D (2013) The location of the northern glacial refugium of Scots Pine based on mitochondrial DNA Markers. *Balt for* 19:2–12

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24:621–631. <https://doi.org/10.1093/molbev/msl191>

Chapuis M-P, Lecoq M, Michalakakis Y, Loiseau A, Sword GA, Piry S, Estoup A (2008) Do outbreaks affect genetic population structure? A worldwide survey in *Locusta migratoria*, a pest plagued by microsatellite null alleles. *Mol Ecol* 17:3640–3653. <https://doi.org/10.1111/j.1365-294X.2008.03869.x>

Cheddadi R, Vendramin GG, Litt T, François L, Kageyama M, Lorentz S, Laurent J-M, de Beaulieu J-L, Sadori L, Jost A, Lunt D (2006) Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Glob Ecol Biogeogr* 15:271–282. <https://doi.org/10.1111/j.1466-822X.2006.00226.x>

Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100:106–113. <https://doi.org/10.1093/jhered/esn088>

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. <https://doi.org/10.1186/1471-2105-9-539>

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001–2014. <https://doi.org/10.1093/genetics/144.4.2001>

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30:1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>

- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J Roy Stat Soc Ser B (Methodol)* 39:1–38
- Dering M, Kosiński P, Wyka TP, Pers-Kamczyc E, Boratyński A, Boratyńska K, Reich PB, Romo A, Zadworny M, Żytkowiak 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. *Divers Distrib* 23:540–555. <https://doi.org/10.1111/ddi.12546>
- Dering M, Baranowska M, Beridze B, Chybicki IJ, Danelia I, Izsakulo G, Kwartskhava 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. *Sci Rep* 11:22845. <https://doi.org/10.1038/s41598-021-02098-1>
- Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proc Natl Acad Sci U S A* 91:3166–3170. <https://doi.org/10.1073/pnas.91.8.3166>
- 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*. *Mol Ecol Resour* 17:943–954. <https://doi.org/10.1111/1755-0998.12646>
- Dumolin S, Demesure B, Petit RJ (1995) Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor Appl Genet* 91:1253–1256. <https://doi.org/10.1007/BF00220937>
- Dyderski MK, Paź S, Frelich LE, Jagodziński AM (2018) How much does climate change threaten European forest tree species distributions? *Glob Change Biol* 24:1150–1163. <https://doi.org/10.1111/gcb.13925>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587. <https://doi.org/10.1093/genetics/164.4.1567>
- Farjon A (2005) Pines: drawings and descriptions of the genus *Pinus*, 2nd edn. Brill, Leiden
- Feurdean A, Tanțău I, Fărcaș S (2011) Holocene variability in the range distribution and abundance of *Pinus*, *Picea abies*, and *Quercus* in Romania; implications for their current status. *Quatern Sci Rev* 30:3060–3075. <https://doi.org/10.1016/j.quascirev.2011.07.005>
- Garza J, Williamson E (2001) Detecting the reduction in population size using data from microsatellite DNA. *Mol Ecol* 10:305–318. <https://doi.org/10.1046/j.1365-294x.2001.01190.x>
- Goncharenko GG, Padutov VE, Silin AE (1993) Allozyme variation in natural populations of Eurasian pines. II. Genetic variation, diversity, differentiation and gene flow in *Pinus sibirica* Du Tour in some lowland and mountain populations. *Silvae Genetica* 42:246–253
- Goudet J (2003) Fstat (ver. 2.9.4), a program to estimate and test population genetics parameters. Available from <https://www2.unil.ch/popgen/softwares/fstat.htm>. Accessed 23 Feb 2023
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361. <https://doi.org/10.2307/2532296>
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2:618–620. <https://doi.org/10.1046/j.1471-8286.2002.00305.x>
- Hardy OJ, Charbonnel N, Fréville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* 163:1467–1482
- Hedrick PW (2001) Conservation genetics: Where are we now? *Trends Ecol Evol* 16:629–636. [https://doi.org/10.1016/S0169-5347\(01\)02282-0](https://doi.org/10.1016/S0169-5347(01)02282-0)
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913. <https://doi.org/10.1038/35016000>
- Hosius B, Leinemann L, Konnert M, Bergmann F (2006) Genetic aspects of forestry in the Central Europe. *Eur J Forest Res* 125:407–417. <https://doi.org/10.1007/s10342-006-0136-4>
- Jansen S, Konrad H, Geburek T (2017) The extent of historic translocation of Norway spruce forest reproductive material in Europe. *Ann for Sci* 74:1–17. <https://doi.org/10.1007/s13595-017-0644-z>
- Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv Genet* 7:295–302. <https://doi.org/10.1007/s10592-005-9098-1>
- Li Y-L, Liu J-X (2018) StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* 18:176–177. <https://doi.org/10.1111/1755-0998.12719>
- Lorenz M, Fischer R, Mues V (2005) Forest resources in Europe and their condition. In: Geburek T, Turok J (eds) Conserving and management of forest genetic resources in Europe. Arbora Publishers, Slovakia, pp 111–126
- Mather AS (2001) The transition from deforestation to reforestation in Europe. In: Angelsen A, Kaimowitz D (eds) Agricultural technologies and tropical deforestation. CAB in association with Centre of International Forestry Research, United Kingdom, p 35–52
- Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics* 25:1982–1983. <https://doi.org/10.1093/bioinformatics/btp303>
- McGeever AH, Mitchell FJG (2016) Re-defining the natural range of Scots Pine (*Pinus sylvestris* L.): a newly discovered microrefugia in western Ireland. *J Biogeogr* 43:2199–2208. <https://doi.org/10.1111/jbi.12761>
- 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* 89:325–337. <https://doi.org/10.1093/forestry/cpw020>
- 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 Evol Biol* 7:233. <https://doi.org/10.1186/1471-2148-7-233>
- Peakall R, Smouse PE (2006) GenAEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Peakall R, Smouse PE (2012) GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Béer E, Robinson S, Vásquez-Carrillo C, Pauli JN, Palsbøll PJ (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Mol Ecol* 21:3403–3418. <https://doi.org/10.1111/j.1365-294X.2012.05635.x>
- Petit RJ, Aguinalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300:1563–1565. <https://doi.org/10.1126/science.1083264>
- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) INVITED REVIEW: comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol* 14:689–701. <https://doi.org/10.1111/j.1365-294X.2004.02410.x>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- 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. *Proc Biol Sci* 265:1697–1705. <https://doi.org/10.1098/rspb.1998.0491>
- Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends Ecol Evol* 16:142–147. [https://doi.org/10.1016/S0169-5347\(00\)02097-8](https://doi.org/10.1016/S0169-5347(00)02097-8)
- Prus-Glowacki W, Urbaniak L, Zubrowska-Gil M (1993) Allozyme differentiation in some European populations of Scots pine (*Pinus sylvestris* L.). *Genetica Polonica* 34:159–176
- Prus-Glowacki W, Stephan BR, Bujas E, Alia R, Marciniak A (2003) Genetic differentiation of autochthonous populations of *Pinus sylvestris* (Pinaceae) from the Iberian Peninsula. *Plant Syst Evol* 239:55–66. <https://doi.org/10.1007/s00606-002-0256-3>
- Pyhäjärvi T, Salmela MJ, Savolainen O (2008) Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genet Genomes* 4:247–254. <https://doi.org/10.1007/s11295-007-0105-1>
- Pyhäjärvi T, Kujala ST, Savolainen O (2020) 275 years of forestry meets genomics in *Pinus sylvestris*. *Evol Appl* 13:11–30. <https://doi.org/10.1111/eva.12809>

QGIS Development Team (2021) QGIS geographic information system. Open Source Geospatial Foundation Project. Available from: <http://qgis.osgeo.org>. Accessed 23 Feb 2023

Ratnam W, Rajora OP, Finkeldey R, Aravanopoulos F, Bouvet J-M, Vaillancourt RE, Kanashiro M, Fady B, Tomita M, Vinson C (2014) Genetic effects of forest management practices: global synthesis and perspectives. *For Ecol Manag* 333:52–65. <https://doi.org/10.1016/j.foreco.2014.06.008>

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* 49:1280–1283. <https://doi.org/10.1111/j.1558-5646.1995.tb04456.x>

Richardson DM (1998) Ecology and biogeography of *Pinus*. Cambridge University Press, Cambridge

Rousset F (2008) Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Mol Ecol Resour* 8:103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>

Sannikov SN, Petrova IV (2012) Phylogenogeography and genotaxonomy of *Pinus sylvestris* L. populations. *Russ J Ecol* 43:273–280. <https://doi.org/10.1134/S1067413612040145>

Sannikov SN, Petrova IV, Egorov EV, Sannikova NS (2014) A system of Pleistocene refugia for *Pinus sylvestris* L. in the southern marginal part of the species range. *Russ J Ecol* 45:167–173. <https://doi.org/10.1134/S1067413614030102>

Sannikov SN, Petrova IV, Egorov EV, Sannikova NS (2020) Searching for and revealing the system of Pleistocene refugia for the species *Pinus sylvestris* L. *Russ J Ecol* 51:215–223. <https://doi.org/10.1134/S1067413620030133>

SAS Institute Inc. (2021) JMP[®], Version 16.1.0. SAS Institute Inc., USA, 1989–2021

Sassoon D, Fletcher WJ, Hotchkiss A, Owen F, Feng L (2021) Scots pine (*Pinus sylvestris*) dynamics in the Welsh Marches during the mid to late-Holocene. *Holocene* 31:1033–1046. <https://doi.org/10.1177/0959683621994654>

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annu Rev Ecol Syst* 38:595–619. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095646>

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. *Eur J Forest Res* 128:377–386. <https://doi.org/10.1007/s10342-009-0273-7>

Schaberg PG, DeHayes DH, Hawley GJ, Nijensohn SE (2008) Anthropogenic alterations of genetic diversity within tree populations: implications for forest ecosystem resilience. *For Ecol Manage* 256:855–862. <https://doi.org/10.1016/j.foreco.2008.06.038>

Semerikova VL, Semerikova SA, Dymshakova OS, Zatssepina 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. *Russ J Genet* 50:577–585. <https://doi.org/10.1134/S1022795414040127>

Semerikova 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 Genet Genomes* 14:8. <https://doi.org/10.1007/s11295-017-1222-0>

Sinclair WT, Morman JD, Ennos RA (1999) The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. *Mol Ecol* 8:83–88. <https://doi.org/10.1046/j.1365-294X.1999.00527.x>

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462. <https://doi.org/10.1093/genetics/139.1.457>

Șofletea N, Mihai G, Ciocîrlan E, Curtu AL (2020) Genetic diversity and spatial genetic structure in isolated Scots Pine (*Pinus sylvestris* L.) populations native to Eastern and Southern Carpathians. *Forests* 11:1047. <https://doi.org/10.3390/f11101047>

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proc Biol Sci* 277:661–671. <https://doi.org/10.1098/rspb.2009.1272>

Svendsen J, Alexanderson H, Astakhov V, Demidov I, Dowdeswell JA, Funder S, Gataullin V, Henriksen M, Hjort C, Houmark-Nielsen M, Hubberten HW, Ingólfsson Ó, Jakobsson M, Kjær KH, Larsen E, Lokrantz H, Lunkka JP, Lyså A, Mangerud J, Matoriouchkov, Murray AA, Möller P, Niessen F, Nikolskaya O, Polyak L, Saarnisto M, Siegest C, Siegest MJ, Spielhagen RF, Stein R (2004) Late Quaternary ice sheet history of Northern Eurasia. *Quat Sci Rev* 23: 1229–1271. <https://doi.org/10.1016/j.quascirev.2003.12.008>

Tóth EG, Vendramin GG, Bagnoli F, Cseke K, Höhn M (2017a) High genetic diversity and distinct origin of recently fragmented Scots pine (*Pinus sylvestris* L.) populations along the Carpathians and the Pannonian Basin. *Tree Genet Genomes* 13:47. <https://doi.org/10.1007/s11295-017-1137-9>

Tóth EG, Köbölkuti ZA, Pedryc A, Höhn M (2017b) Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. *J for Res* 28:637–651. <https://doi.org/10.1007/s11676-017-0393-8>

Tóth EG, Bede-Fazekas Á, Vendramin GG, Bagnoli F, Höhn M (2019) Mid-Pleistocene and Holocene demographic fluctuation of Scots pine (*Pinus sylvestris* L.) in the Carpathian Mountains and the Pannonian Basin: Signs of historical expansions and contractions. *Quatern Int* 504:202–213. <https://doi.org/10.1016/j.quaint.2017.11.024>

Tulstrup N (1959) International trade in forest tree seed. *Unasylva* 13:196–201

Vasilyeva Y, Chertov N, Nechaeva Y, Sboeva Y, Pystogova N, Boronnikova S, Kalendar R (2021) Genetic structure, differentiation and originality of *Pinus sylvestris* L. populations in the east of the East European Plain. *Forests* 12:999. <https://doi.org/10.3390/f12080999>

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. <https://doi.org/10.1038/hdy.2010.118>

Wachowiak W, Wójcikiewicz 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 Genet Genomes* 10:1015–1025. <https://doi.org/10.1007/s11295-014-0739-8>

Wachowiak W, Żukowska WB, Perry A, Lewandowski A, Cavers S, Łabiszak B (2022) Phylogeography of Scots pine in Europe and Asia based on mtDNA polymorphisms. *J Syst Evol*. <https://doi.org/10.1111/jse.12907>

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conserv Genet* 6:551–562. <https://doi.org/10.1007/s10592-005-9009-5>

Wójcikiewicz B, Wachowiak W (2016) Substructuring of Scots pine in Europe based on polymorphism at chloroplast microsatellite loci. *Flora – Morphology, Distribution, Functional Ecology of Plants* 220: 142–149. <https://doi.org/10.1016/j.flora.2016.03.005>

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A* 84:9054–9058. <https://doi.org/10.1073/pnas.84.24.9054>

Zimmer K, Sørnstebø JH (2018) A preliminary study on the genetic structure of Northern European *Pinus sylvestris* L. by means of neutral nuclear microsatellite markers. *Scand J for Res* 33:6–13. <https://doi.org/10.1080/02827581.2017.1337919>

Żukowska WB, Wójcikiewicz B, Litkowiec M, Wachowiak W (2017) Cross-amplification and multiplexing of cpSSRs and nSSRs in two closely related pine species (*Pinus sylvestris* L. and *P. mugo* Turra). *Dendrobiology* 77:59–64. <https://doi.org/10.12657/denbio.077.005>

Żukowska WB, Lewandowski A, Wójcikiewicz B, Litkowiec M, Rożkowski R, Urbaniak L, Kowalczyk J (2023a) Changes in the gene pool composition of Scots pine depending on the mode of regeneration. *Dendrobiology* 89:46–55. <https://doi.org/10.12657/denbio.089.005>

Żukowska WB, Wójcikiewicz B, Lewandowski A, Łászló R, Wachowiak W (2023b) Genetic variation of Scots pine in Eurasia: impact of postglacial recolonisation and human-mediated gene transfer. *Zenodo Repository* V2:2023. <https://doi.org/10.5281/zenodo.8405255>. Accessed 04 Oct 2023

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

