BALTIC FORESTRY

http://www.balticforestry.mi.lt
ISSN 1392-1355
eISSN 2029-9230

Baltic Forestry 2022 28(2): 166–178
Category: Research article
https://doi.org/10.46490/BF652

DNA markers reveal genetic association between the sea-side Lithuanian and Bavarian Scots pine populations

DARIUS KAVALIAUSKAS [1]*, DARIUS DANUSEVIČIUS [1], BARBARA FUSSI [2] AND REMIGIJUS ŽALKAUSKAS [1]

[1] *Institute of Forest Biology and Silviculture, Vytautas Magnus University Agriculture Academy, Studentų str. 11, Akademija LT-53361, Kaunas district, Lithuania*

[2] *Bavarian Office for Forest Genetics, Am Forstamtsplatz 1, Teisendorf D-83317, Germany*

** Corresponding author: darius.kavaliauskas@vdu.lt; phone: +370 69873833*

Kavaliauskas, D., Danusevičius, D., Fussi, B. and Žalkauskas, R. 2022. DNA markers reveal genetic association between the sea-side Lithuanian and Bavarian Scots pine populations. *Baltic Forestry* 28(2): 166–178. https://doi.org/10.46490/BF652.

Received 8 June 2022 **Revised** 19 September 2022 **Accepted** 30 September 2022

Abstract

Due to complex history of European forests, natural populations may not necessarily represent autochthonous genepools for forest trees. Eastern Prussian forests were famous for using non-local sources for afforestation. We studied efficiency of a set of nuclear microsatellite markers (nSSR) for genetic association and diversity studies of 194 adult trees from 5 populations genotyped at 11 nSSR loci. The Bayesian and UPGMA clustering revealed two genetically distinct groups: (a) the Baltic group, and (b) the Bavarian one and an over 200-years-old sea-side Lithuanian population of Juodkrante from the sea-side Curonian spit of Neringa. We interpret this result as putative introduction of Bavarian Scots pine back in the 18th century, when reforestation efforts were made to sustain moving sands in the dunes of Neringa. The genomic SSRs were more variable than the EST SSRs. However, the association between the variability of the nuclear microsatellite loci and their efficiency in detecting population differentiation was not strong.

Keywords: nSSR, *Pinus sylvestris* (L.), population structure, provenance transfer, FRM transfer, molecular markers

Introduction

In wind-pollinated wide-spread conifers such as Scots pine exhibiting weak among population differentiation (Kavaliauskas et al. 2022), occurrence of distinct morphotypes in the absence of obvious environmental gradients may indicate human interferences such as seed transfer and artificial afforestation. Examples of such distinct morphotypes of Scots pine in the Baltic region could be the "Riga pine" morphotype in western Latvia notorious for its stem height and straightness (Shutyaev and Giertych 1997, Krakau et al. 2013) or a specific bark morphotype of Scots pine at the seaside of Lithuania distinguishing by flat shallow bark flakes, remaining a turtle coat (Ramanauskas 1994, Danusevicius 2001). Such locally outstanding populations may originate from introductions by humans in the past (Krakau et al. 2013). Absence of records indicating such human interventions does not necessarily guaranty the natural origin or development of a forest stand (Krakau et al. 2013, Kembrytė et al. 2021, Kembrytė et al. 2022). In this study we aimed to assess the efficiency of a set of nuclear SSR markers and to investigate putative genetic association between the sea-side Lithuanian and the Bavarian and Polish Scots pine populations. This case is interesting from the historical perspective as a reflection of the human efforts to restore the forests on the unique landscape of the sea-side spit Neringa (Curonian spit) in western Lithuania. Another aspect is the age of over 200 years of the studied stand at the sea-side Lithuania, dating back to the period when the artificial regeneration of Scots pine was rare in this region (Krakau et al. 2013). At the same occasion, we discuss the efficiency of a set of nuclear SSR loci for population genetic studies by comparing the polymorphism and population differentiation indices.

The forest degradation preceding the numerous military conflicts of the 18th century and the long-lasting efforts for stabilizing the moving sands in the Lithuanian sea-side spit Neringa provide a possibility for occasional artificial establishment of Scots pine forests in Neringa. Historically, Neringa was a part of eastern Prussia but since 1757, as an aftermath of the wars, for a decade fell under the

Russian rule (Strakauskaite 2004). During this period massive forest exploitation took place at the sea-side spit in Neringa followed by desertification and disastrous effects of moving sands driven by the sea winds. After regaining the land, the Prussian administration forced a large-scale afforestation initiative, mainly with *Pinus mugo* (Turra) (Danusevicius et al. 2012, 2013). We are not aware of any historical records professionally describing the origin of the material used for this afforestation back in the 18th century. Our recent study on the cpSSR polymorphism of the *Pinus mugo* populations in Neringa revealed high genetic diversity, low risks for inbreeding and indicates that the seed were imported from a restricted geographic area such as Denmark rather than collected from early plantations in Neringa (Danusevicius et al. 2012). These large-scale introductions of *Pinus mugo* (Turra) in Neringa leave a possibility for artificial afforestation with Scots pine as early as in the 18th century. At that time, Scots pine was usually established by direct seeding, which results in a close to natural spatial structure of trees within such seeded stands.

Furthermore, there are numerous examples of historical transfer of forest reproductive material (FRM) within Europe (e.g. Koskela et al. 2014, Jansen and Geburek 2016, Myking et al. 2016, Jansen et al. 2017, Geburek and Myking 2018, Jansen et al. 2019). According to Koskela et al. (2014 and references therein) in the 18th century, seeds of the main forest tree species such as *Pinus sylvestris*, *Picea abies*, *Larix decidua* (Jansen and Geburek 2016) and *Quercus* spp. were widely traded across Europe and transfer of seed contributed significantly to reforestation with *Pinus sylvestris* at least in 13 European countries (Koskela et al. 2014). Myking et al. (2016) presented a study on FRM transfer into the Nordic region, where FRM of four main tree species (Scots pine, Norway spruce, European beech and oaks) was relocated and traded mainly in Sweden, Norway and Denmark from various sources in Central and Eastern Europe. Jansen et al. (2017) tracked the trade and transfer of Norway spruce FRM within Europe and showed historical routes of FRM transfer and expansion of Norway spruce distribution area. Furthermore, Jansen et al. (2019) studied recent rates of FRM trade in Europe based on national datasets and EU Council directive 1999/105/EC and showed that ca. 30 million plants are traded in Europe every year. Thus, FRM transfer is an ongoing process which can have positive effects to countries' economies and help to cope with climate change. However, there are a number of risks related to FRM transfer in the complex chain of FRM production and handling. These risks mainly include origin tracking errors of any nature. Therefore, DNA based FRM tracking technologies must be developed for safeguarding the origin identification within various seed transfer schemes (e.g., Konnert and Behm 2006, Finkeldey et al. 2010, Konnert et al. 2015, Kavaliauskas et al. 2021).

Microsatellites (SSR – simple sequence repeats) in recent decades have been the most commonly used highly polymorphic co-dominant markers for assessment of population structure and differentiation (Li et al. 2002, Oliveira et al. 2006, Ellis and Burke 2007, Putman and Carbone 2014, etc.). The genomic function of microsatellites is diverse from selectively neutral to functional as chromatin organisation, recombination and regulation of gene expression (Kashi and Soller 1999, Li et al. 2002, Li et al. 2004). These diverse functions lead to variable evolutionary assumptions each with specific statistical approaches when analyzing the SSR data (Selkoe and Toonen 2006, Putman and Carbone 2014, etc.). Thus, selection of microsatellite loci may have a significant effect on the efficiency to detect population structure and differentiation (Epperson 2004). However, the choice of loci becomes a complex issue when considering the evolutionary assumptions and a variety of the statistical approaches (e.g. Luikart and England 1999, Pearse and Crandall 2004, Jost 2008, etc.).

It is generally accepted in population genetic studies that more loci provide more reliable estimates of population genetic parameters (e.g. Hedrick 1999). However, the statistical significance may not be evolutionarily meaningful and including loci, for instance, with high homoplasy rates may even obscure the precision of the genetic estimates (Blankenship et al. 2002, Epperson 2005). Using few polymorphic loci may provide similar efficiency as more but less polymorphic loci (Kalinowski 2005). High number of polymorphic loci may increase the technical genotype scoring errors (Buchnan et al. 2005) and the bias due to homoplasy resulting in underestimation of population differentiation and overestimation of geneflow (Gaggiotti et al. 1999, Epperson 2005). On the other hand, an unjustified elimination of loci may reduce the statistical power and the representation of the genome studied (Kalinowski 2002, Epperson 2004).

Particularly interesting is the issue of connection between the variability of SSR loci and their power to detect population differentiation (Balloux and Lugon-Moulin 2002, Kalinowski 2005). Though it may obviously seem a positive association, it may not be because of high incidence of homoplasy associated mutations in highly variable loci, which may cause a negative association between variability and differentiation (Olsen et al. 2004, O'Reilly et al. 2004). In case of Scots pine, it is also unclear is there a connection between the loci variability and ability to differentiate populations, as for instance, using highly polymorphic loci can increase the error of allele frequency estimates, especially under small sample sizes (Ruzzante 1998, Gomez-Unchida and Banks 2005, Kalinowski 2005). Because species, loci, and even populations have own evolutionary pathways, specific responses to evolutionary forces such as mutation rates shaping population parameters, the choice of loci and their properties may be case specific (Li et al. 2002, Oliveira et al. 2006, Selkoe and Toonen 2006). The problem, therefore, is how efficient are the specific SSR loci used in other species in revealing population differentiation in Scots pine and how variable

are the estimates of differentiation calculated with different evolutionary assumptions and statistical algorithms (Danusevicius et al. 2016).

Two main SSR mutation rate models are suggested (e.g. review Oliveira et al. 2006). The infinite allele mutation mode (IAM), where each mutation by chance creates a new allele and an allele pair of 130 bp and 100 bp is as different from each other as a pair of 131 bp and 130 bp alleles. Wright's (1949) F_{ST} and Nei's (1973) Gst are commonly used population differentiation estimates from allele frequencies applied under the IAM. The stepwise mutation model SMM, where mutations are believed to occur in stepwise manner at a constant rate (molecular clock) so that the alleles with similar repeat number may share more close ancestry than alleles differing markedly in repeat number (Slatkin 1995). The is the preferred model for SSRs, where the polymorphism in the number of microsatellite repeats is assumed to occur mainly due to DNA polymerase slippage during the DNA replication process (Eisen 1999, Ellegren 2004). For population differentiation assuming SMM, Slatkin (1995) suggested the R_{ST} as another fixation index that considers the differences in repeat number among genotypes. Still there is an uncertainty, perhaps depending on particular loci and species, where SMM or IAM could be more appropriate. For instance, R_{ST} under SMM is more sensitive to homoplasy and other violations such as non-step wise mutations, differential DNA repair mechanisms, effect of drift, where the IAM-based statistics is more robust and reliable (Balloux and Lugon-Moulin 2002, Landry et al. 2002). Therefore, there still is an interest to compare the efficiency of known SSR loci to detect population differentiation assuming diverse mutation models, algorithms on new material such as Scots pine populations in our study.

The objectives of our study were to estimate the efficiency of a multiplexed set of the nuclear microsatellite loci to estimate among population differentiation and genetic diversity by studying several Scots pine populations from Lithuania, Germany and Poland. The relationship between the polymorphism of the SSR loci and their efficiency for revealing the among population differentiation is discussed. We chose a subset of local populations together with a distant population, likely not connected via geneflow with the local genepools.

Material and methods

The following five populations were selected for the study to include geographically distant material when comparing the efficiency of the loci: Alpenkiefer, southern Germany (abbreviated as "Ger", lat. 47°57"N, long. 12°54"E, elevation 420 m a.s.l. BES), Suprasl, north-eastern Poland ("Pol", lat. 53°15"N, long. 23°23"E, elevation 181 m a.s.l.), Punia, southern Lithuania ("Pun", lat. 54°32'28.54"N, long. 24°4'24.82"E, elevation 78 m a.s.l.), Braziukai, central Lithuania ("Bra", lat. 54°53'25.13"N, long. 23°28'48.73"E, elevation 74 m a.s.l.) and Juodkrante, western sea-side Lithuania ("Juo", lat. 55°32'45.33"N, long. 21°6'49.03"E, elevation 35 m a.s.l.) (Figure 1).

The southern German population consists of 20 trees sampled as clones in a seed orchard located near Laufen (the coordinates of the seed orchard provided). The clones are plus trees selected in forest ranges of the Bavarian Alps. From 20 to 50 trees were sampled within natural stands to represent the Polish and Lithuanian populations, respectively. The sampled trees were of mature age and sampled systematically over 15–20-meter intervals within an area of 0.5 ha in stands of natural origin (Figure 2).

In total 190 individuals were genotyped. The wood samples for DNA extraction were collected by drilling with a 5-mm electric bore to the depth of up to 5 cm, placed into 2 ml plastic tubes and stored at –20°C until DNA extraction. DNA was extracted from the samples according to the modified ATMAB-method after Dumolin et al. (1995). The concentration and purity of the DNA samples obtained

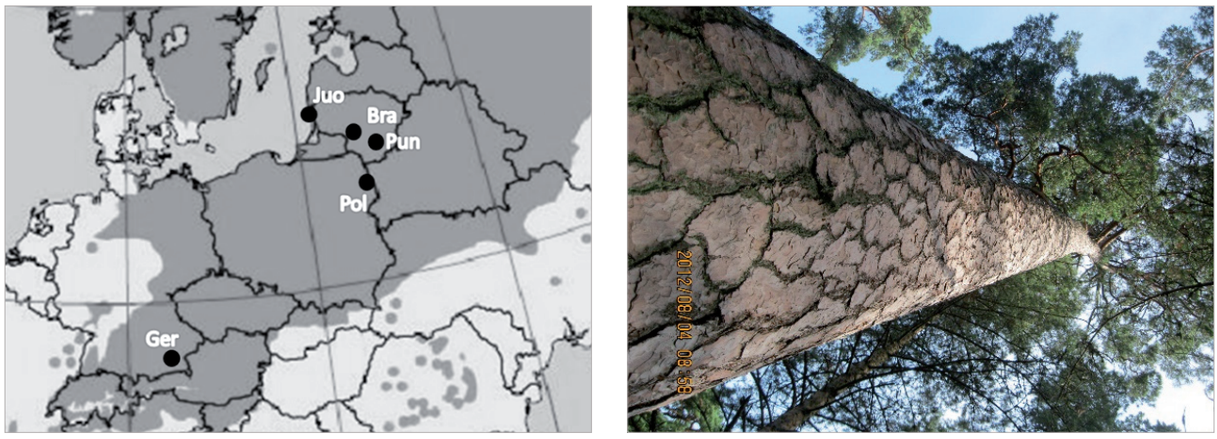


Figure 1. Left: Origin of the populations studied. The dark areas indicate the natural distribution range of Scots pine (EUFORGEN 2009). Right: A representative ca. 200-year-old tree from the sea-side Lithuanian Juodkrante (Juo) population with the specific "shallow" bark morphotype

Figure 2. Overmature trees of Scots pine sampled for the DNA study in Juodkrante in the Lithuanian sea-side spit Neringa. The oldest age class of ca. 200 years old was sampled

were fluorometrically quantified (Gene Quant Pro, Amersham Bioscience). Three multiplex-PCR reactions were performed using fluorescent labelled primers in a mixture of 10 µl total volume containing 1 X reaction buffer (Qiagen Multiplex PCR Master Mix), 0.5 to 4 µM of each primer and 1 µl of DNA (concentration 10 to 20 ng / µl). The program for amplification was optimized using gradient PCR-conditions. The final PCR program started with initial denaturation at 95°C for 15 min, followed by 30 cycles of 94°C for 30 sec, 57°C for 1.30 min, 72°C for 30 sec and a final elongation step at 60°C for 30 min (the thermal cycler "TProfessional Standard PCR Thermocycler") (The Biometra®). The length of the PCR fragments was determined by using an automated sequencer (CEQ GeXP Beckman-Coulter) and analysed by using an internal size standard (CEQ™ DNA Size Standard Kit – 400). Fragment length determination and allele assignment were carried out using the fragment analysis tool GeXP (Beckman-Coulter).

Eleven nuclear microsatellite loci were studied. The following primers for the microsatellite loci representing two to four nucleotide repeats were used: Psyl57, Psyl2, Psyl18, Psyl42, Psyl25, Psyl16 (Sebastiani et al. 2012), Spag7.14, Spac12.5, Spac11.4 (developed for *Pinus sylvestris*, Soranzo et al. 1998) and PtTX4011, PtTX4001 (developed for *Pinus taeda*, Elsik et al. 2000), combined into 3 multiplexes for the PCR and capillary electrophoresis as follows: PCR multiplex A: Psyl2, Psyl42, Psyl25, Spag7.14; PCR multiplex B: Psyl57, Psyl18, and PCR multiplex C: Psyl16, Spac12.5, Spac11.4, PtTX4011, PtTX4001.

After the allele scoring, the data set was checked for scoring errors due to stuttering, large allele dropouts with the approach estimating the excess of homozygotes implanted in the Microchecker software, ver. 2.2.3 (van Oosterhout et al. 2005). The null allele frequency per locus and populations were estimated by the maximum likelihood method implemented in Genepop 4.0 software.

Assuming the IAM the following allele identity-based fixation indices were calculated: Θ (theta) as an unbiased estimator of F_{ST} that corrects for error associated with incomplete sampling of populations (Weir and Cockerham 1984), (FSTAT, ver. 2.9.3.2, Goudet 1995). Assuming the SMM, the allele size-based R_{ST} fixation index was calculated with the FSTAT software. FSTAT employs jackknifing over populations to obtain the standard errors of the fixation indices separately for each locus (sampling from each population the number of individuals equal to sample size of the smallest population without replacement repeated 1,000 times). FSTAT also performs jackknifing and bootstrapping over loci to obtain the standard errors and confidence intervals for the multilocus estimates.

In addition two exact tests of population differentiation were computed: (a) the exact *G*-test (less biased under unequal sample sizes than the F_{ST}; Goudet et al. 1996) was used for testing for population differentiation assuming random mating within samples, where the proportion of randomized runs with the permuted values of larger differentiation than observed over 10,000 permutations are used to test for the significance of the differentiation (implemented in FSTAT software), and (b) the differentiation test by a contingency table approach to determine if groups of individuals have significant differences in allele frequencies for each locus (Raymond and Rousset 1995) implemented in PowerMarker software with 10,000 permutations (Liu and Muse 2005).

With more than two alleles, however, F_{ST} or G_{ST} cannot reach 1.0 even when no alleles are shared between the two populations, as there will always be some heterozygosity within populations (Meirmans and Hedrick 2011). Considering the limitations of the F_{ST} and G_{ST} fixation indices in detecting differentiation with diverse markers such as SSR, we computed the differentiation index, D_{est}, which measures of the fraction of allelic variation among populations as described by Jost (2008) by using the GenAlEx, ver. 6.5 software.

Unbiased gene diversity (expected heterozygosity) and allelic richness corrected for sample size (rarefaction based on minimum population sample size of 20 diploid individuals) were used as the estimates of the loci variability and both were calculated with FSTAT software (population estimate averages the values over the loci). Loci were also characterized with the F_{IS} fixation index testing of variation of individuals within sub-populations as an

estimate of inbreeding levels and an indication of the deviation from the HW equilibrium (FSTAT software). If F_{IS} is significantly positive or negative (significant deficit and excess of heterozygotes, respectively) or values close to 0 indicate neither outbreeding nor inbreeding so that the population is in HW equilibrium for this locus.

The population structure was assessed with the STRUCTURE ver. 2.3.4 software (Pritchard et al. 2000) as follows: the burn-in period of 100000, 100000 replications, 10 runs for each of 1 to 3 clusters, with the LOCPRIOR function, allele frequencies assumed to be correlated and no admixture model was assumed. Our recent study revealed weak differentiation among the Lithuanian Scots pine populations (Kavaliauskas et al. 2022). Therefore, we limited the *K* value to 3 clusters within our material. The most likely number of clusters (1 to 3) was estimated by submitting the STRUCTURE output to the STRUCTURE HARVESTER software, ver. 6.94 (Earl and Holdt 2012).

The genetic associations among the populations were studied based on the Goldstein et al. (1995) SMM based genetic distances calculated from the allele frequencies with of the PowerMarker ver. 3.25 software (Liu and Muse 2005), which generated 60000 bootstrapped trees (bootstrapping over loci). The procedure CONSENSE within the PHYLIP ver. 3.69 software (Felsenstein 1989) was used to obtain the consensus tree to be depicted an UPGMA and NJ dendrogram with the TREEVIEW software (Page 1996).

Relative migration among studied Scots pine populations were assessed using the approach implemented in *R*-package diveRsity (function divMigrate) (Keenan et al. 2013). The method is based on defining a pool of migrants for each combination of two samples in pairwise comparison and measure of genetic differentiation is then calculated for both the first and the second population compared to the pool, generating two directional measures of genetic differentiation (Sjöqvist et al. 2015, Sundqvist et al. 2016). The relative migration was calculated based on Nm (Alcala et al. 2014) and was used as a measure of genetic differentiation.

Results

Population differentiation

No evidence for scoring errors due to stuttering, nor for large allele dropout were found at all loci. The null allele frequency (NAF) varied at about 0.04, except for the almost monomorphic EST locus Psyl25, where NAF was as high as 0.9. All the loci, except Psyl25, were polymorphic, producing 6 to 35 alleles (Table 1). The Spac loci were the most variable with 16, 30 and 35 alleles. The PtTX loci were less variable with 10 to 17 alleles. The Psyl series loci (the EST-SSRs) were the least variable and the locus Psyl25 was basically fixed for the 215 bp allele (only 2 individuals contained other than the 215 bp allele; Figure 3).

Allele frequency of the highly variable loci Spag7.14 and Spac12.5 perfectly followed the Gaussian-shaped allele frequency distribution assumed by the SMM (Ellegren 2004; Figure 3). Most of the other loci also showed no marked deviation from the Gaussian distribution, except for the least polymorphic EST loci with several alleles such as Psyl25 (Figure 3). In general, there was a tendency for a lack of short alleles in the lower tail of the Gaussian-shaped distribution (Spac11.4, Psyl16, Figure 3).

The locus-wise rarified allelic richness revealed a similar trend as the observed number of alleles: the highest number of alleles was observed for the Spac loci, followed by PtTX and Psyl loci (Table 3 and S.M. Figure 1). Expected heterozygosity was positively associated with allelic richness (S.M. Figure 1). For most of the loci the F_{IS} index was positive close to zero indicating no marked HW deviation (Table 1). The locus PtTX4011, however, exhib-

Table 1. Population differentiation statistics compared between the loci

Locus	AN	p_{diff} G-test (FSTAT)	p_{diff} (Power marker)	R_{ST} FSTAT	F_{ST} (Θ) (se)	D_{est}	$p\ D_{est}$	F_{IS} FSTAT
Spag7.14	35	0.0028	0.0035	0.007	0.005 (0.003)	0.064	0.130	0.089
Psyl57	7	0.1245	0.0628	0.019	0.008 (0.009)	0.005	0.258	0.052
Psyl2	6	0.9000	0.8709	–0.005	–0.012 (0.003)	–0.007	0.987	0.146
Psyl18	6	0.431	0.3533	0.001	–0.001 (0.008)	0.001	0.236	0.176
Psyl42	5	0.2101	0.2247	–0.007	0.012 (0.011)	0.018	0.134	0.024
Psyl25	2	0.2463	0.6060	0.006	0.012 (0.006)	0.000	0.212	–0.013
Spac12.5	30	0.0948	0.0709	0.007	0.002 (0.002)	0.032	0.233	0.045
Spac11.4	16	0.0287	0.0577	–0.011	0.000 (0.003)	0.006	0.352	0.038
PtTX4011	10	0.0001	0.0004	0.031	0.011 (0.010)	0.027	0.084	0.279
PtTX4001	17	0.0639	0.0380	–0.003	0.006 (0.008)	0.026	0.056	0.015
Psyl16	10	0.0927	0.0969	0.014	0.001 (0.002)	0.009	0.322	0.112

Notes: AN – observed allele number; p_{diff} *G*-test is the probability for differentiation according to the exact *G*-test (Goudet et al. 1996) based on 10000 permutations implemented in FSTAT; p_{diff} is the probability for differentiation according to a contingency table approach to determine if groups of individuals have significant differences in allele frequencies for each locus (Raymond and Rousset 1995) implemented in PowerMarker software; F_{ST} (Θ, theta) is an unbiased estimator of F_{ST} that corrects for error associated with incomplete sampling of populations, (Weir and Cockerham 1984), (FSTAT); D_{est} is the Jost's estimate of differentiation (D_{est}) calculated following Meirmans and Hedrick eq 2 in Meirmans and Hedrick (2011); $p\ D_{est}$ is the significance of the D_{est} tested with 10000 bootstraps over individuals (GenAlEx ver. 6.5 software).

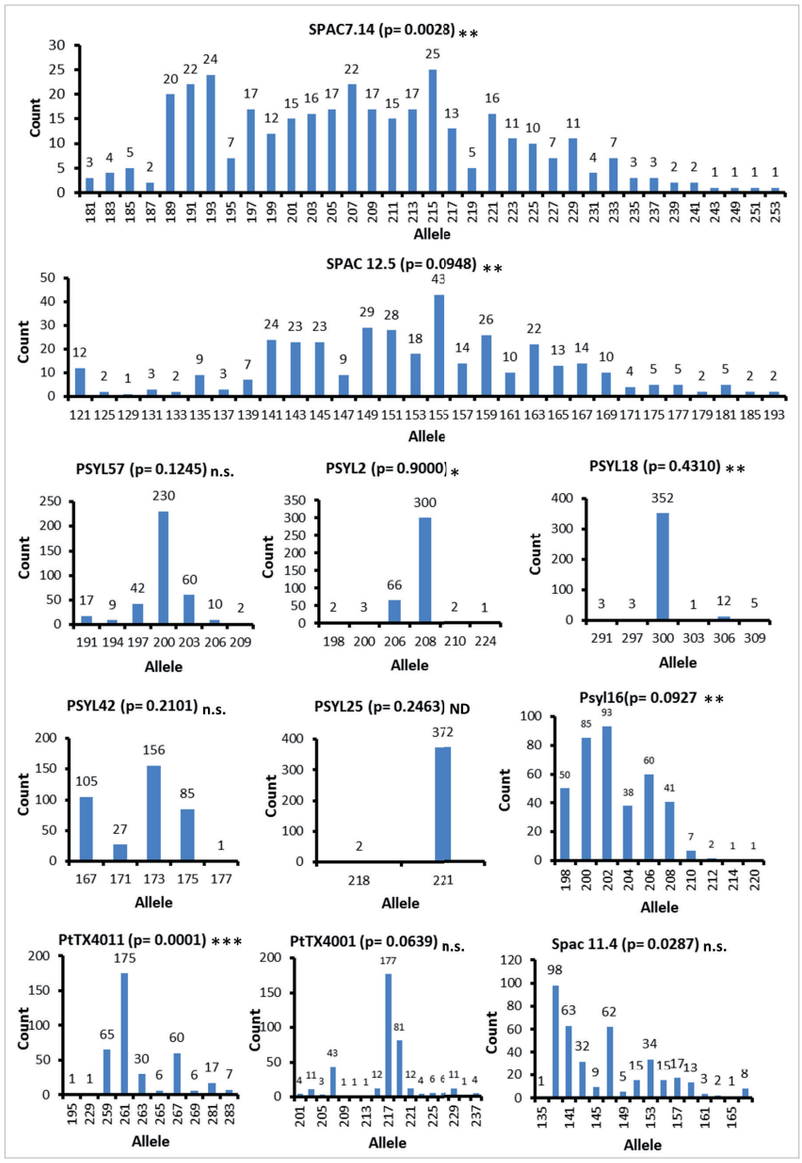


Figure 3. The allelic variability of the loci. The significance values for the exact population differentiation test (*G*-test in FSTAT) are given at each locus name

Note: The asterisk at the differentiation p values present results of the Hardy-Weinberg equilibrium deviation test by FSTAT (F_{IS} based), where n.s. = not significant deviation, * = 0.05–0.01, ** = 0.01–0.001; ND = not determined.

ited the highest positive F_{IS} value indicating a heterozygote deficit (Table 1, Figure 3).

For the least variable EST-SSR Psyl loci, the population differentiation was markedly lower than for the rest of the loci (*G*-test *p* value ranged from 0.0927 to 0.9000; Table 1). For the most variable Spac loci, the differentiation was significant or close to significant (0.0028 to 0.0927; Table 1). Though being of medium variability, the locus PtTX4011 had the greatest discriminating power with highly significant *p* values of 0.0001 to 0.0004 from both the differentiation tests (Table 1). There was good agreement between the two differentiation tests implemented in FSTAT and PowerMarker software packages (Table 1). Regarding the comparison of the discriminating power between the loci, the theta fixation index was less informative than the exact differentiation tests, because the theta values were rather uniform ranging at about 0.001–0.008 with high standard errors (Table 1). The D_{est} test returned higher values than F_{ST} for highly variable loci (Spac, PtTX) and similar values as F_{ST} for less variable Psyl loci (Table 1). Generally, the R_{ST} values were higher than the F_{ST} values (Table 1). The multilocus AMOVA revealed higher among population variance components for the R_{ST} index than for the F_{ST} index (Figure 4).

The relative migration network analysis (Figure 5) showed equally strong relative migration rates among the Lithuanian populations and slightly lowers genetic exchange with the Bavarian population. These results well corresponded well to the STRUCTURE clustering (Figure 6), where it revealed an unexpected genetic association between the sea-side Lithuanian population of Juodkrante and the Bavarian population. Directional relative migration rates below 0.5 were filtered out to emphasize the major gene flow and differentiation among the Scots pine populations.

Population structure

The STRUCTURE HARVESTER analysis returned the highest delta *K* value and the lowest standard deviation for the mean *ln* probability of the number of clusters (*K*) for the two-cluster structure (not shown). If added, the third cluster was equally shared among the individuals and provided no useful information for the population structure (not shown). All 10 STRUCTURE runs with the number of clusters equal to 2 returned similar inferred ancestry, as shown in Figure 6. The proportion of individual membership in the two clusters revealed a genetic association between the sea-side Lithuanian population of Juodkrante and the Bavarian population, both containing over 90% of the cluster 2 individuals (red bars in Figure 6). Whereas cluster 1 dominated in the remaining geographically close populations from north-eastern Poland and central Lithuania (green bars in Figure 6).

Table 3. The rarefied allelic richness given for each population and locus (F_{STAT} soft.)

Locus	Ger	Pol	Bra	Juo	Pun
Spag7.14	15.6	21.0	18.5	19.1	19.2
psyl57	4.8	5.6	5.4	5.4	5.0
psyl2	3.6	2.8	2.4	2.8	2.8
psyl18	3.7	1.8	2.5	3.0	2.5
psyl42	4.0	4.0	4.0	4.3	4.0
psyl25	1.0	1.0	1.0	1.6	1.0
Spac12	15.8	20.1	16.7	17.4	16.6
Spac11	10.4	10.1	9.4	10.1	10.0
PtTX40	4.7	5.4	5.4	8.3	4.8
PtTX40	12.7	6.2	6.1	8.3	9.0
Psyl16	6.6	7.7	7.0	7.1	5.9
Average	7.5	7.8	7.1	7.9	7.4

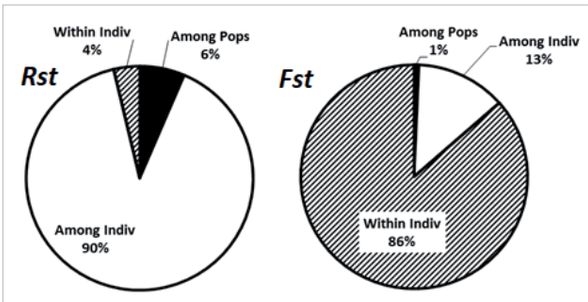


Figure 4. The partition of the variance based on the multilocus AMOVA carried out for F_{ST} and R_{ST} estimates with GenAlEx, ver. 6.5

With all loci included, the clustering based on the distance (which assumes mutations as the major force for population differentiation) assigned the Bavarian population into a separate node and revealed a close association between the Lithuanian population of Braziukai and the north-eastern Polish population. However, less than 50% of bootstraps indicate low reliability of this separate Bavarian population node (Figure 7). On the contrary, the association between Braziukai and the north-eastern Polish populations was strong, as indicated by over 90% of

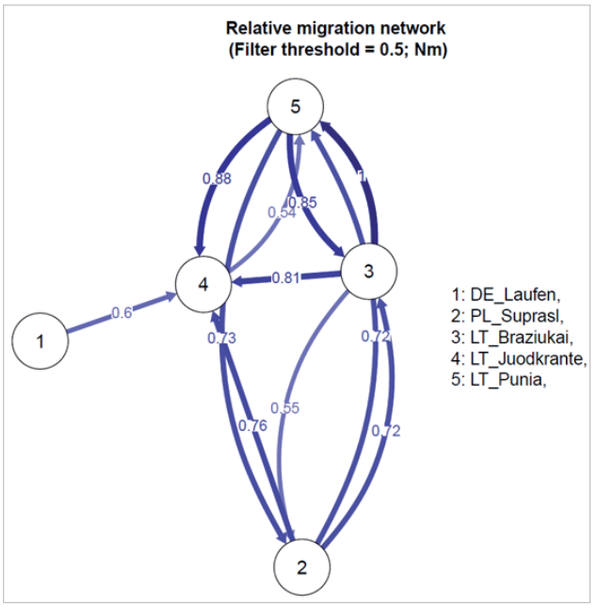


Figure 5. The relative migration network (based on *Nm*) (*R*-package diveRsity)

bootstrapped data merging these populations into a single node (Figure 7).

After exclusion of the loci with no significant population differentiation, the clustering based on the remaining seven loci returned similar results as the Bayesian inference: a node of Bavarian and the Lithuanian sea-side population of Juodkrante (71% of bootstraps) and a cluster containing the remaining populations from the north-eastern range (Figure 7). The UPGMA and NJ tree construction methods had no marked effect on the clustering outcome (Figure 7).

The main genetic diversity criteria were similar among the populations, except for slightly higher inbreeding coefficient and greater difference between the H_e and H_o values in the Braziukai population from central Lithuania (Table 2).

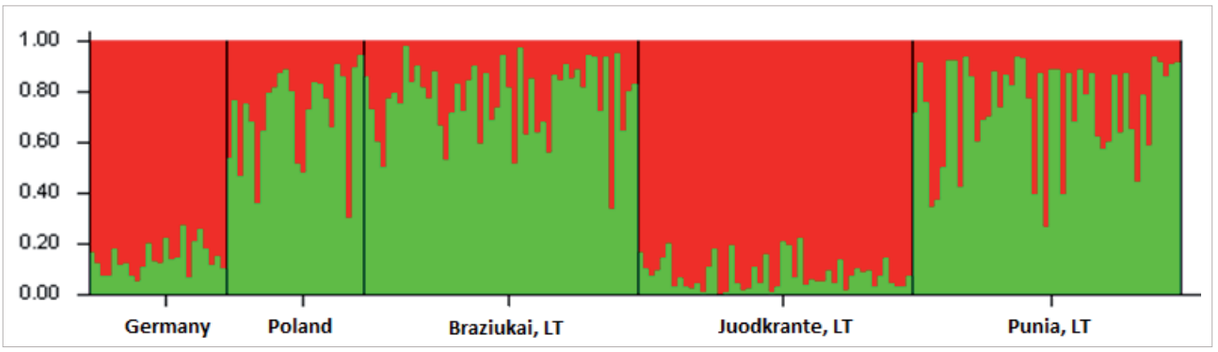


Figure 6. The plot of the inferred ancestry of individuals with the most likely structure of two clusters from the STRUCTURE software

Note: The colour of the bars indicates the proportion of membership of the individuals in each of the 2 clusters (Y axis). The populations are outlined by vertical lines and identified on the X axis).

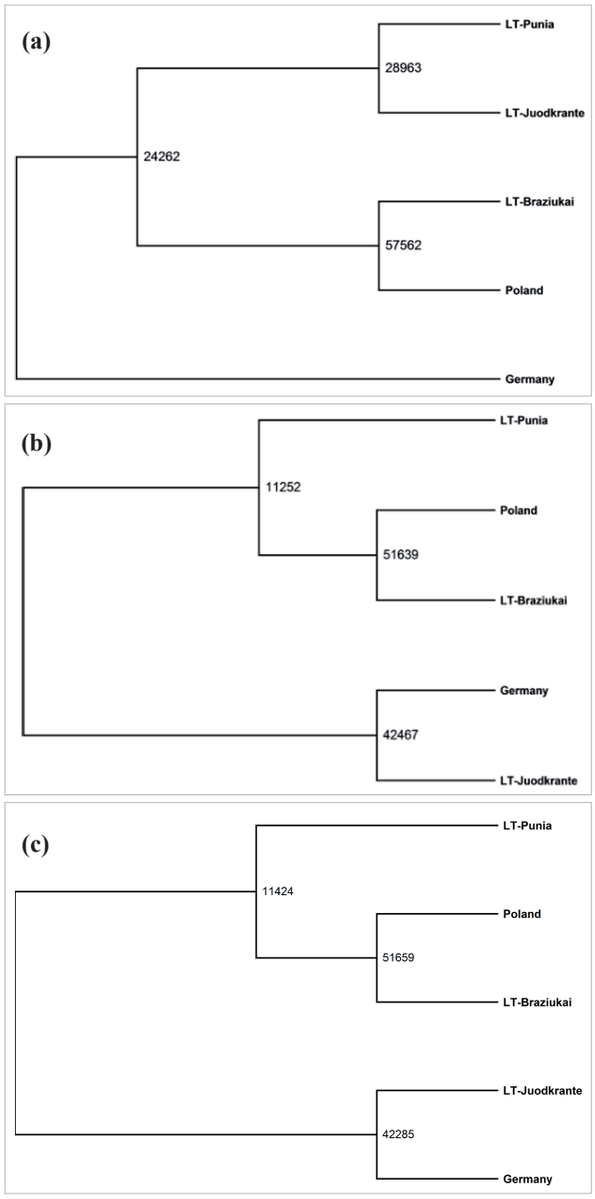


Figure 7. Comparison of dendrograms after clustering based on Goldstein et al. (1995) genetic distance with all the loci (upper left dendrogram (a)) and two dendrograms (on the right side (b and c)) with loci which showed significant population differentiation

Note: Significance of the nodes was tested by 60000 bootstraps and is given at the fork of each node as number of trees with the populations was joined into the node. The plots to the left show analysis with all the loci; the plots to the right only loci with significant population differentiation.

Table 2. The genetic diversity indices averaged over the loci. The standard errors of the multilocus means are given below each estimate (GenAlEx soft.)

Population	*N*	*Na*	*Ne*	*Ho*	*uHe*	*F*
Alpenkiefer, Bavaria (Ger)	**22**	**8.00**	**4.80**	**0.58**	**0.62**	**0.07**
		1.67	1.17	0.09	0.09	0.04
Suprasl, north-eastern Poland (Pol)	**22**	**8.36**	**5.48**	**0.55**	**0.59**	**0.03**
		2.21	1.81	0.10	0.10	0.04
Braziukai, central Lithuania (Bra, LT)	**50**	**9.09**	**5.49**	**0.53**	**0.61**	**0.16**
		2.42	1.72	0.09	0.10	0.05
Juodkrante, sea-side Lithuania (Juo, LT)	**50**	**10.45**	**5.80**	**0.58**	**0.62**	**0.05**
		2.45	1.80	0.09	0.10	0.04
Punia central Lithuania (Pun, LT)	**50**	**9.82**	**5.36**	**0.56**	**0.60**	**0.04**
		2.61	1.66	0.10	0.10	0.03

Notes: *N* – number of individuals; *Na* – observed number of alleles; *Ne* – No of Effective Alleles = 1 / (Sum pi^2); *Ho* – Observed Heterozygosity = No of Hets / *N*; *He* – Expected Heterozygosity = 1 – Sum *pi*^2; *uHe* – Unbiased Expected Heterozygosity = (2*N* / (2*N* – 1)) × *He*; *F* – Fixation Index = (*He* – *Ho*) / *He* = 1 – (*Ho* / *He*), where *pi* is the frequency of the i-th allele for the population and Sum *pi*^2 is the sum of the squared population allele frequencies.

Discussion and conclusions

Population comparison

The genetic association between the geographically distant populations from Bavaria and Juodkrante in sea-side Lithuania is unexpected, because the old growth stand sampled in Juodkrante is expected to represent natural regeneration from autochthonous pools surviving the desertification of the 18th century in the seaside spit Neringa. Nevertheless, this genetic association between the geographically distant populations from Bavaria and Juodkrante is strong, because all the trees within these two populations were assigned with over 80% likelihoods into the single STRUCTURE cluster number 2 (darker colour in Figure 6). Given such a uniform likelihood for each tree, the homoplasy is unlikely. The likelihood of a non-random association between these two populations is also strengthened by the finding that the UPGMA or NJ clustering structures turned to statistically significant after removing the loci with no significant differentiation among populations (Figure 7). The remaining three Baltic populations were genetically similar, which supports the geographical over a random differentiation background (Figures 6 and 7). It is likely that southern German and Baltic populations have different evolutionary backgrounds and geneflow over such large distances is unlikely (Naydenov et al. 2007, Buchovska 2013, Buchovska et al. 2013, Dering et al. 2017). Significant among-population differentiation based on nSSR is the least expected for geographically close populations such as in the Baltic region, including the Juodkrante population in our study (Kavaliauskas et al. 2022). Considering the considerations given above, a possible scenario is artificial establishment of at least a part of

the sampled stand in Juodkrante back in 1800's. The scale of such introductions in the seaside spit Neringa is unclear and only a few of such 200 years old stands are remaining in the spit Neringa. However, the scope of our study allows us to only raise the hypothesis of a higher likelihood for genetic associations between the seaside Lithuanian population of Juodkrante and the central European populations of Scots pine.

Similar allele numbers together with uniform H_e values indicate no marked differences in allelic diversity among the populations (Table 2). However, the relatively higher inbreeding coefficient in the Braziukai population indicates that the heterozygotes were created by a comparably lower number of alleles than in the other populations. This implies a relatively greater relatedness among the sampled trees in the Braziukai population. The main difference between the sampled stands is that the Braziukai stand was heavily thinned to ca. 200 trees per ha for natural regeneration purposes from seed trees. Consequently, there is a tendency for a greater relatedness among the superior trees left as seed trees for regeneration in the Braziukai stand than among randomly sampled trees in the remaining stands.

The loci comparison

The deviation from the HW equilibrium was stronger for the loci with a significant population differentiation (Figure 1). This may indicate the effects of divergent mutations or selection via the divergence hitchhiking (Goicoechea et al. 2012).

The allele conformity to the expected repeat size and proximation to the Gaussian – shaped distribution of the allele frequencies indicates neither marked genotyping nor serious sampling problems (Selkoe and Toonen 2006). The allele variability obtained in our study was similar as in the other studies using the same loci over a comparable geographical range of Scots pine (Soranzo et al. 1998, García-Gill et al. 2009). Especially, Spag7.14 with 35 alleles in our material is highly polymorphic for Scots pine, considering that Kuchma et al. (2011) found 39 alleles for Spag7.14 in a strongly mutagenic environment of the Chernobyl exclusion zone. We found 30 alleles for Spac12.5, the second most polymorphic locus. García-Gill et al. (2009) reported 47 alleles for Spac12.5 within a forest district of 25 ha and sample size of ca. 400 individuals in Sweden. Buchovska (2013) reported 39 Spac12.5 alleles for the material representing ca. 500 individuals from a large part of the Scots pine natural distribution range but having ca. 15 individuals per population. Kavaliauskas et al. (2022) found 35 alleles for Psyl12.5 in 20 Lithuanian Scots pine populations with 20 trees each. As expected, the genomic SSR loci (Spac; PtTX) showed markedly higher variability than the EST-SSR loci (the Psyl series; Li et al. 2004). Furthermore, in our recent study with a large within-population sample size of ca. 200 trees from each of 6 populations (938 trees in total) in Lithuania, we obtained 33 alleles for Spac12.5 (Danusevicius, submitted). Such observations for allele numbers that a single locus may exhibit with variable sample depths indicate vast allelic diversity reservoirs laying within populations. Is there a limit for mutation of the microsatellite motive repeats? Perhaps as for locus Spac12.5, Buchovska (2013) found 39 size variants in the Scots pine material covering most of the species range including the easternmost parts of Russia (compare 30 alleles of Spac12.5 in our study; Table 1). More investigations in this subject are of great interest.

Similar ranking among the populations as regards the observed and rarefied allele numbers indicate sufficient sample sizes for the estimate of population allelic diversity. Another feature of SSRs to be investigated in greater detail is the relationship between heterozygosity and allelic richness in such widely outcrossing species as Scots pine. Due to deviations from random mating, similar allele numbers can be combined towards a higher number of heterozygotes or homozygotes. This could be a specific feature of the cDNA SSRs. cDNA SSRs also exhibit a high transferability rate among pine species (Liewlaksaneeyanawin et al. 2004).

There may be a connection between the loci variability and the capacity to reveal the population differentiation so that the least variable loci were less efficient in revealing the differentiation (Figure 3, Table 1). One reason for such result is straightforward, variable markers provide a higher statistical power (both variation and more informative bootstrapping or permutation tests for significance) to estimate genetic structure (Epperson 2004). The loci with several alleles common in all populations certainly are least informative for differentiation. However, the highest variability does not mean the strongest differentiation, e.g. the Spac12.5 locus being marked more variable than PtTX4011 had lower differentiation power than PtTX4011 (Table 1, Figure 3). Relatively higher F_{ST} values for low diversity loci, especially for few allele cases, may be a good indicator when searching for population-specific markers, and useful for control of origin of forest reproductive material (Konnert and Behm 2006).

Clearly, the Spac loci series are among the most polymorphic and represent genomic microsatellites with high mutation rates, perhaps relatively more influenced by homoplasy than the other loci investigated in our study. This observation is supported by the findings that for the Spac12.5 locus, the high number of individuals from a small area returned a higher number of size variants (20 ha, ca. 96 trees and 47 alleles; García-Gill et al. 2009) than material representing most of the species range (ca. 2,000 km range, 500 individuals, 39 alleles; Buchovska 2013). The PtTX loci series may reveal differentiation owing to population specific alleles (Figure 4) and are good candidates for population differentiation as found by our and other Scots pine studies (Soranzo et al. 1998, Gonzlez-Martinez et al. 2004, García-Gill et al. 2009). Psyl loci showed low variability and low differentiation in our material but being

connected with the functional part of the genome, may be efficient in detecting population structure within a broad geographic range such as Eurasia (studied by Buchovska et al. 2013).

Most of the F_{ST} values were below 0.05 which is generally accepted as the lower differentiation threshold (Conner and Hartl 2004). Only for locus PtTX4011, the F_{ST} was close to moderate differentiation (Table 1). R_{ST} values were in general higher than F_{ST} ones indicating the importance of the SMM in revealing population differentiation at these loci. This was also confirmed by the multilocus R_{ST} with much greater percentage of variation among populations than F_{ST} (Figure 5). There also was a tendency for the R_{ST} values to be greater than the F_{ST} for the loci of medium variability but of high differentiation (Psy57 and PtTx4011; Table 1). Whereas for the most variable loci (Spag7.14 and Spac12.5), R_{ST} was not markedly different from the F_{ST} values. R_{ST} provides better estimates of population differentiation than F_{ST} for the material with low genetic exchange (Balloux and Goudet 2002), as in our case where gene flow between the southern German and the Baltic Scots pine populations is expected to be limited. The D_{est} value was higher for more variable loci as expected following the reasoning for the D statics by Jost (2008) and Meirmans and Hedrick (2011).

In conclusion, our study indicates that (a) there may be a genetic background for distinct morphotypes of Scots pine occurring without the absence of obvious natural reasons, and (b) for Scots pine, the association between the variability of the nuclear microsatellite loci and their efficiency in detecting population differentiation is not perfect and highly variable loci do not guarantee the greatest and evolutionary most meaningful population discrimination. Before carrying out large-scale population differentiation and phylogeographic studies with nuclear microsatellites in Scots pine, a pre-screening of the loci within the frames of a pilot investigation could markedly improve the cost efficiency and evolutionary significance of the results.

Acknowledgement

This work was supported by the European Social Fund under the grant agreement with the Research Council of Lithuania (LMTLT) Project No 09.3.3-LMT-K-712-19-0007 "Development of Competences of Scientists, other Researchers, and Students through Practical Research Activities".

Conflict of interest statement

The authors declare no conflict of interest.

Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

References

Alcala, N., Goudet, J. and Vuilleumier, S. 2014. On the transition of genetic differentiation from isolation to panmixia: what we can learn from GST and D. *Theoretical Population Biology* 93: 75–84. https://doi.org/10.1016/j.tpb.2014.02.003.

Balloux, F. and Goudet, J. 2002. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology* 11: 771–783. https://doi.org/10.1046/j.1365-294X.2002.01474.x.

Balloux, F. and Lugon-Moulin, N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11(2): 155–165. https://doi.org/10.1046/j.0962-1083.2001.01436.x.

Blankenship, S.M., May, B. and Hedgecock, D. 2002. Evolution of a perfect simple sequence repeat locus in the context of its flanking sequence. *Molecular Biology and Evolution* 19: 1943–1951. https://doi.org/10.1093/oxfordjournals.molbev.a004018.

Buchnan, J.C., Archie, E.A., van Horn, R.C., Moss, C.J. and Alberts, S.C. 2005. Locus effects and sources of error in noninvasive genotyping. *Molecular Ecology Notes* 5: 680– 683. https://doi.org/10.1111/j.1471-8286.2005.01002.x.

Buchovska, J., Danusevičius, D., Baniulis, D., Stanys, V., Šikšnianienė, J.B. and Kavaliauskas, D. 2013. The location of the northern glacial refugium of Scots pine based on mitochondrial DNA markers. *Baltic Forestry* 19(1): 2–12.

Buchovska, J. 2013. Geographical distribution patterns of DNA polymorphism of Scots pine in the eastern part of the species range. PhD Thesis. Lithuanian Forest Research Institute, Kaunas.

Conner, J.K. and Hartl, D.L. 2004. A Primer of Ecological Genetics. Sinauer Associates, Inc., Sunderland, MA, U.S.A. 304 pp. Available online at: https://livresbioapp.files.wordpress.com/2016/02/connorhartl_2004_aprimerofecologicalgenetics.pdf.

Danusevicius, D., Buchovska, J., Stanys, V., Šikšnianienė, J.B., Baliuckas, V. and Brazaitis, G. 2012. Chloroplast DNA polymorphism of an exotic *P. mugo* Turra population introduced to sea-side spit of Kursiu Nerija in Lithuania. *European Journal of Forest Research* 132(1): 137–150. https://doi.org/10.1007/s10342-012-0663-0.

Danusevičius, D., Buchovska, J. Stanys, V., Šikšnianiene, J.B., Marozas, V. and Bendokas, V. 2013. DNA marker based identification of spontaneous hybrids between *Pinus mugo* and *P. sylvestris* at the Lithuanian sea-side. *Nordic Journal of Botany* 31: 344–352. https://doi.org/10.1111/j.1756-1051.2012.01575.x.

Danusevičius, J. 2000. Pušies selekcija Lietuvoje [Selection of pines in Lithuania]. Monograph. Kaunas: Lithuanian Forest Research Institute, Lutute, 352 pp. (in Lithuanian with English summary).

Danusevicius, D., Kavaliauskas, D. and Fussi, B. 2016. Optimum sample size for SSR-based estimation of representative allele frequencies and genetic diversity in Scots pine populations. *Baltic Forestry* 22: 194–202.

Dering, M., Kosiński, P., Wyka, T.P., Pers-Kamczyc, E., Boratyński, A., Boratyńska, K., Reich, P.B., Romo, A., Zadworny, M., Żytkowiak, R. and 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 to Fennoscandia. *Diversity and Distributions* 23(5): 540–555. https://doi.org/10.1111/ddi.12546.

Dumolin, S., Demesure, B. and Petit, R.J. 1995. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theoretical and Applied Genetics* 91: 1253–1256. https://doi.org/10.1007/BF00220937.

Earl, D.A. and von Holdt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2): 359–361. https://doi.org/10.1007/s12686-011-9548-7.

Eisen, J.A. 1999. Mechanistic basis for microsatellite instability. In: Goldstein, D.B. and Schlötterer, C. (Eds.) Microsatellites – evolution and applications. Oxford University Press, Oxford (UK), p. 34–48.

Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* 5(6): 435–445. https://doi.org/10.1038/nrg1348.

Ellis, J.R. and Burke, J.M. 2007. EST-SSRs as a resource for population genetic analyses. *Heredity* (Edinb.) 99: 125– 132. https://doi.org/10.1038/sj.hdy.6801001.

Elsik, C.G., Minihan, V.T., Hall, S.E., Scarpa, A.M. and Williams, C.G. 2000. Low-copy microsatellite markers for *Pinus taeda* L. *Genome* 43(3): 550–555. https://doi.org/10.1139/g00-002.

Epperson, B.K. 2004. Multilocus estimation of genetic structure within populations. *Theoretical Population Biology* 65: 227–237. https://doi.org/10.1016/j.tpb.2003.11.003.

Epperson, B.K. 2005. Mutation at high rates reduces spatial structure within populations. *Molecular Ecology* 14: 703– 710. https://doi.org/10.1111/j.1365-294X.2005.02429.x.

Felsenstein, J. 1989. PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* 5: 164–166.

Finkeldey, R., Leinemann, L. and Gailing, O. 2010. Molecular genetic tools to infer the origin of forest plants and wood. *Applied Microbiology and Biotechnology* 85(5): 1251– 1258. https://doi.org/10.1007/s00253-009-2328-6.

Gaggiotti, O.E., Lange, O., Rassmann, K. and Gliddon, C. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8: 1513–1520. https://doi.org/10.1046/j.1365-294x.1999.00730.x.

García-Gill, M.R., Olivier, F., Kamruzzahan, S. and Waldmann, P. 2009. Joint analysis of spatial genetic structure and inbreeding in a managed population of Scots pine. *Heredity* 103: 90–96. https://doi.org/10.1038/hdy.2009.33.

Geburek, T. and Myking, T. 2018. Evolutionary consequences of historic anthropogenic impacts on forest trees in Europe. *Forest Ecology and Management* 422: 23–32. https://doi.org/10.1016/j.foreco.2018.03.055.

Goicoechea, P.G., Petit, R.J. and Kremer, A. 2012. Detecting the footprints of divergent selection in oaks with linked markers. *Heredity* 109: 361–371. https://doi.org/10.1038/hdy.2012.51.

Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L. and Feldman, M.W. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139(1): 463–471. https://doi.org/10.1093/genetics/139.1.463.

Gomez-Uchida, D. and Banks, M.A. 2005. Microsatellite analyses of spatial genetic structure in darkblotched rockfish (*Sebastes crameri*): is pooling samples safe? *Canadian Journal of Fisheries and Aquatic Sciences* 62: 1874–1886. https://doi.org/10.1139/f05-084.

Gonzlez-Martinez, S.C., Robledo-Arnuncio, J.J., Collada, C., Daz, A., Williams, C.G., Ala, R. and Cervera, M.T. 2004. Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines. *Theoretical and Applied Genetics* 109(1): 103–111. https://doi.org/10.1007/s00122-004-1596-x.

Goudet, J., Raymond, M., de Meeus, Th. and Roussett, F. 1996. Testing Differentiation in Diploid Populations. *Genetics* 144: 1933–1940. https://doi.org/10.1093/genetics/144.4.1933.

Goudet, J. 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.

Hedrick, P.W. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313–318. https://doi.org/10.1111/j.1558-5646.1999.tb03767.x.

Jansen, S. and Geburek, T. 2016. Historic translocations of European larch (*Larix decidua* Mill.) genetic resources across Europe – A review from the 17th until the mid-20th century. *Forest Ecology and Management* 379: 114–123. https://doi.org/10.1016/j.foreco.2016.08.007.

Jansen, S., Konrad, H. and Geburek, T. 2017. The extent of historic translocation of Norway spruce forest reproductive material in Europe. *Annals of Forest Science* 74(3): 1–17. https://doi.org/10.1007/s13595-017-0644-z.

Jansen, S., Konrad, H. and Geburek, T. 2019. Crossing borders – European forest reproductive material moving in trade. *Journal of Environmental Management* 233: 308– 320. https://doi.org/10.1016/j.jenvman.2018.11.079.

Jost, L. 2008. Gst and its relatives do not measure differentiation. *Molecular Ecology* 17(18): 4015–4026. https://doi.org/10.1111/j.1365-294X.2008.03887.x.

Kalinowski, S.T. 2002. How many alleles per locus should be used to estimate genetic distances? *Heredity* 88: 62–65. https://doi.org/10.1038/sj.hdy.6800009.

Kalinowski, S.T. 2005. Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* 94: 33–36. https://doi.org/10.1038/sj.hdy.6800548.

Kashi, Y., and Soller, M. 1999. Functional roles of microsatellites and minisatellites. In: Goldstein, D.B. and Schlötterer, C. (Eds.) Microsatellites: Evolution and Applications. Oxford University Press, Oxford (UK), p. 10–23.

Kavaliauskas, D., Fussi, B., Sirgėdienė, M., Kembrytė, R. and Danusevičius, D. 2021. Genetic diversity and parentage analysis for DNA marker-based forest reproductive material traceability system in Lithuania. In: Černiauskienė, J. (Ed.) Proceedings of the 10th International Scientific Conference 'Rural Development 2021', Issue 2021: Challenges for Sustainable Bioeconomy and Climate Change, Section: Multifunctional Approach for Sustainable Use of Bio-Resources. Vytautas Magnus University Press, Kaunas, p. 226–231. https://doi.org/10.15544/RD.2021.040.

Kavaliauskas, D., Danusevičius, D. and Baliuckas, V. 2022. New Insight into Genetic Structure and Diversity of Scots Pine (*Pinus sylvestris* L.) Populations in Lithuania Based on Nuclear, Chloroplast and Mitochondrial DNA Markers. *Forests* 13(8): 1179. https://doi.org/10.3390/f13081179.

Kembrytė, R., Danusevičius, D., Buchovska, J., Baliuckas, V., Kavaliauskas, D., Fussi, B. and Kempf, M. 2021. DNA-based tracking of historical introductions of forest trees: The case of European beech (*Fagus sylvatica* L.) in Lithuania. *European Journal of Forest Research* 140(2): 435–449. https://doi.org/10.1007/s10342-020-01341-0.

Kembrytė, R., Danusevičius, D., Baliuckas, V. and Buchovska, J. 2022. Phenology Is Associated with Genetic and Stem Morphotype Variation in European Beech (*Fagus sylvatica* L.) Stands. *Forests* 13(5): 664. https://doi.org/10.3390/f13050664.

Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W. and Prodöhl, P.A. 2013. diveRsity: An R package for the estimation and exploration of population genetics param-

eters and their associated errors. *Methods in Ecology and Evolution* 4(8): 782–788. https://doi.org/10.1111/2041-210X.12067.

Konnert, M. and Behm, A. 2006. Proof of identity of forest reproductive material based on reference samples. *Mitteilungen der Bundesforschungsanstalt der Forst- und Holzwirtschaft (BFH)* 221: 61–71.

Konnert, M., Fady, B., Gömöry, D., A'Hara, S., Wolter, F., Ducci, F., Koskela, J., Bozzano, M., Maaten, T. and Kowalczyk, J. 2015. Use and transfer of forest reproductive material in Europe in the context of climate change. In: European Forest Genetic Resources Programme (EUFORGEN), Bioversity International, Rome, Italy, Vol. XVI, 75 pp. Available online at: https://www.euforgen.org/fileadmin/templates/euforgen.org/upload/Publications/Thematic_publications/EUFORGEN_FRM_use_transfer.pdf.

Koskela, J., Vinceti, B., Dvorak, W., Bush, D., Dawson, I.K., Loo, J. and Ramamonjisoa, L. 2014. Utilization and transfer of forest genetic resources: a global review. *Forest Ecology and management* 333: 22–34. https://doi.org/10.1016/j.foreco.2014.07.017.

Krakau, U.K., Liesebach, M., Aronen, T., Lelu-Walter, M.A. and Schneck, V. 2013. Scots pine (*Pinus sylvestris* L.). In: Forest Tree Breeding in Europe. Springer, Dordrecht, p. 267–323.

Kuchma, O., Vornam, B. and Finkeldey, R. 2011. Mutation rates in Scots pine (*Pinus sylvestris* L.) from the Chernobyl exclusion zone evaluated with amplified fragment-length polymorphisms AFLPs and microsatellite markers. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* 725(1-2): 29–35. https://doi.org/10.1016/j.mrgentox.2011.07.003.

Landry, P.A., Koskinen, M.T. and Primmer, C.R. 2002. Deriving evolutionary relationships among populations using microsatellites and Dl2: all loci are equal, but some are more equal than others. *Genetics* 161(3): 1339–1347. https://doi.org/10.1093/genetics/161.3.1339.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. 2002 Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11(12): 2453–2465. https://doi.org/10.1046/j.1365-294X.2002.01643.x.

Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E. 2004. Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution* 21(6): 991–1007. https://doi.org/10.1093/molbev/msh073.

Liewlaksaneeyanawin, C., Ritland, C.E., El-Kassaby, Y.A. and Ritland, K. 2004. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* 109(2): 361–369. https://doi.org/10.1007/s00122-004-1635-7.

Liu, K. and Muse, S.V. 2005. Integrated analysis environment for genetic marker data. *Bioinformatics* 21(9): 2128–2129. https://doi.org/10.1093/bioinformatics/bti282.

Luikart, G. and England, P.R. 1999. Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution* 14(7): 253–256. https://doi.org/10.1016/S0169-5347(99)01632-8.

Meirmans, P.G. and Hedrick, P.W. 2011. Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources* 11: 5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x.

Myking, T., Rusanen, M., Steffenrem, A., Kjær, E.D. and 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(4): 325–337. https://doi.org/10.1093/forestry/cpw020.

Naydenov, K., Senneville, S., Beaulieu, J., Tremblay, F. and 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. https://doi.org/10.1186/1471-2148-7-233.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* 70(12): 3321–3323. https://doi.org/10.1073/pnas.70.12.3321.

Nielsen, E.E., Hansen, M.M., Ruzzante, D.E., Meldrup, D. and Grønkjær, P. 2003. Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology* 12: 1497–1508. https://doi.org/10.1046/j.1365-294X.2003.01819.x.

O'Reilly, P.T., Canino, M.F., Bailey, K.M. and Bentzen, P. 2004. Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology* 13: 1799–1814. https://doi.org/10.1111/j.1365-294X.2004.02214.x.

Oliveira, E.J., Pádua, J.G. and Zucchi, M.I. 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* 29(2): 294–307.

Olsen, J.B., Habicht, C., Reynolds, J., and Seeb, J.E. 2004. Moderately and highly polymorphic microsatellites provide discordant estimates of population divergence in sockeye salmon, *Oncorhynchus nerka*. In: Genetics of Subpolar Fish and Invertebrates. Developments in environmental biology of fishes, Vol. 23. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0983-6_21.

Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Bioinformatics* 12(4): 357–358.

Pearse, D.E. and Crandall, K.A. 2004. Beyond F_{ST}: analysis of population genetic data for conservation. *Conservation Genetics* 5: 585–602. https://doi.org/10.1007/s10592-003-1863-4.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945–959. https://doi.org/10.1093/genetics/155.2.945.

Putman, A.I. and Carbone, I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution* 4(22): 4399–4428. https://doi.org/10.1002/ece3.1305.

Ramanauskas, V. 1994. Paprastosios pušies formos Lietuvoje [Forms of Scots pine in Lithuania]. *Dendrologia Lithuaniae* 94: 5–8 (in Lithuanian).

Raymond, M. and Rousset, F. 1995. Genepop (version-1.2) – population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.

Ruzzante, D.E. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 15–27. https://doi.org/10.1139/f97-203.

Sebastiani, F., Pinzauti, F., Kujala, S.T., Gonzalez-Martinez, S.G. and Vendramin, G.G. 2012. Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. *Conservation Genetics Resources* 4(2): 231–234. https://doi.org/10.1007/s12686-011-9513-5.

Selkoe, K.A. and Toonen, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsat-

ellite markers. *Ecology Letters* 9(5): 615–629. https://doi.org/10.1111/j.1461-0248.2006.00889.x.

Shutyaev, A.M. and Giertych, M. 1997. Height Growth Variation in a Comprehensive Eurasian Provenance Experiment of Scots pine (*Pinus sylvestris* L.). *Silvae Genetica* 46(6): 332–349.

Sjöqvist, C., Godhe, A., Jonsson, P.R., Sundqvist, L. and Kremp, A. 2015. Local adaptation and oceanographic connectivity patterns explain genetic differentiation of a marine diatom across the North Sea–Baltic Sea salinity gradient. *Molecular Ecology* 24(11): 2871–2885. https://doi.org/10.1111/mec.13208.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139(1): 457–462. https://doi.org/10.1093/genetics/139.1.457.

Soranzo, N., Provan, J. and Powell, W. 1998. Characterization of microsatellite loci in *Pinus sylvestris* L. *Molecular Ecology* 7(9): 1260–1261.

Soto, A., Robledo-Arnuncio, J.J., Gonzalez-Martınez, S.C., Smouse, P.E. and Alıa, R. 2010. Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. *Molecular Ecology* 19(7): 1396–1409. https://doi.org/10.1111/j.1365-294X.2010.04571.x.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P. and Kleinhans, D. 2016. Directional genetic differentiation and relative migration. *Ecology and Evolution* 6(11): 3461–3475. https://doi.org/10.1002/ece3.2096.

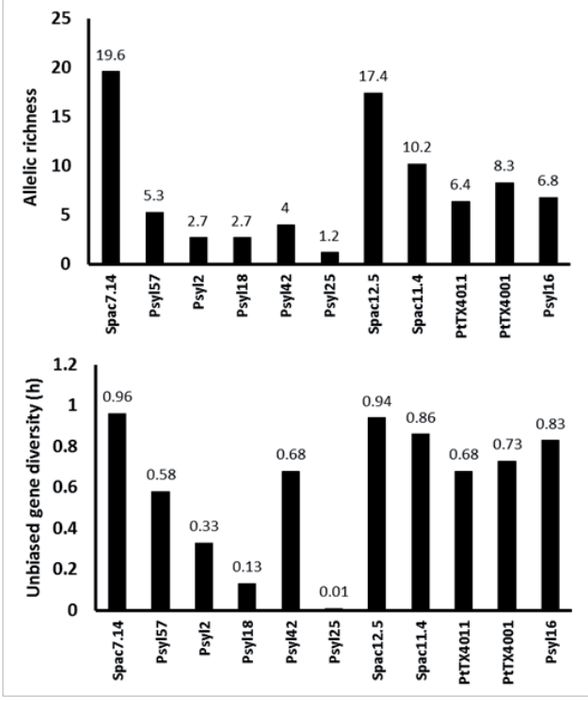
Strakauskaite, N. 2004. Klaipeda and the Curonian Spit Guide. Klaipėda, R. Paknio leidykla, 80 pp.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P.F. 2005. Micro-Checker User Guide. The University of Hull, Hull (UK).

Weir, B.S. an Cockerham, C.C. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38(6): 1358. https://doi.org/10.2307/2408641.

Wright, S. 1949. The genetical structure of populations. *Annals of Eugenics* 15(1): 323–353. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x.

Supplementary material



S.M. Figure 1. The rarefied allelic richness (upper plot) and unbiased gene diversity (lower plot) compared among the loci