Universitatea Transilvania din Brașov

INTERDISCIPLINARY DOCTORAL SCHOOL

Faculty of Silviculture and Forest Engineering

Marina SHELLER

GENETIC DIVERSITY OF SCOTS PINE (*PINUS SYLVESTRIS* L.) IN MIDDLE SIBERIA

ABSTRACT

Scientific supervisor

Prof. Dr. Alexandru Lucian CURTU

BRAȘOV, 2023

CONTENTS

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Prof. Alexandru Lucian Curtu, for his guidance and support during my PhD. Many thanks to my colleagues who have contributed to this study. Special thanks to Assoc. prof. dr. Elena Ciocîrlan for help in laboratory work; Endre György Tóth for help in statistical analysis; Pavel Mikhaylov, Sergey Kulakov, Nadezhda Kulakova, Natalia Melnichenko, Aleksey Ibe and Tatyana Sukhikh for help in collecting material for this PhD thesis research. I would also like to express my gratitude to the members of the advisory committee, Prof. dr. Neculae Șofletea and Assoc. Prof. dr. Adrian Victor Indreica for their helpful comments.

I would like to thank the Transilvania University of Brasov for supporting my education with the scholarship. The financial support provided by the Ministry of Education and Science of the Russian Federation (State Assignment "Fundamental principles of forest protection from entomo- and phyto-pests in Siberia» No. FEFE 2020-0014) is also acknowledged.

INTRODUCTION

Boreal forests are one of the largest ecosystems on Earth. Most of the boreal forests of the planet are located in Russia, where they represent a wide strip between the forest tundra in the north and the zone of mixed broad-leaved forests in the south. Coniferous trees predominate in boreal forests. Scots pine and larch form a light coniferous taiga, and fir, spruce and cedar – dark coniferous taiga (Markatyuk, 2012). Up to 75 percent of the taiga in Siberia is near to its natural state (WWF). Thus, in this region, where the natural structure of forest populations has still been preserved, it is possible to conduct unique studies aimed at preserving the genetic diversity of woody plants in the future. Conservation of forest genetic resources is of vital importance because they are an exceptional and irreplaceable resource for future generations. Genetic diversity is the key to ensuring the sustainability of forest ecosystems and the adaptation of forest species to climate change (FAO 2014). The higher the genetic potential, the stronger the impact of stressful situations a population can withstand and the more diverse the habitats it can inhabit (Padutov et al. 2008). Currently, the main factors causing a decrease in the genetic diversity of forests are fires, diseases, windfalls, floods, anthropogenic impacts. It is believed that the level of genetic diversity (heterozygosity) below 30% threatens the survival of the forest stand in changing climate (The State Coordination Program for the Development of Biotechnology in the Russian Federation 2012, Padutov V. E. 2001, Padutov et al. 2008). Therefore, the sustainable development of forest stands is possible only with the maximum preservation of genetic diversity.

Scots pine (*Pinus sylvestris* L.) is one of the keystone species in forest ecosystems of the boreal regions in Eurasia. It is of great ecological importance and is adapted to a variety of environmental conditions (Naydenov et al. 2007, Dering et al. 2017). Scots pine is one of the most commercially important tree species, which is valued worldwide for its wood (Mátyás et al. 2004). A significant part of the species` distribution range is located in Siberia, where pine forests are scattered as a complex spatial mosaic and only in the Angara river basin do they form a continuous zone. Due to natural disturbances, overexploitation and mismanagement, the area of pine forests in Russia has been decreased by two

million ha over the last ten years (Sheller et al. 2021). In this regard, the study of genetic diversity and structure of Scots pine forests in Russia is very relevant.

In recent decades, neutral DNA markers such as microsatellites (simple sequence repeats, SSRs) have been widely used in genetic studies. They are featured by a high level of polymorphism and they can be extremely useful in studying the population structure of coniferous tree species characterized by a low level of inter-population diversity (Bernhardsson et al. 2016, Șofletea et al. 2020, Dering et al. 2021, Kavaliauskas et al. 2022). In this thesis, we address the problem of assessment of genetic diversity and structure of Middle Siberian Scots pine populations based in nuclear and chloroplast DNA markers.

CHAPTER 1. BACKGROUND TO THE STUDY

1.1. Biological features, ecology and distribution of Scots pine (*Pinus sylvestris* L.)

Scots pine (*Pinus sylvestris* L.) is a widespread species in the pine family (*Pinaceae*). It is a large tree up to 25-35 m in height with a trunk diameter of up to 1 m. The crown is pyramidal, but over time the lower shoots dry up, fall off and the crown forms a spherical shape. The bark is reddish-brown, the upper layer of which is constantly peeling off in the form of thin plates. The needles are paired dark green with a bluish tinge, covered with a waxy coating. Male cones are yellow and female cones are reddish, localized on the tops of annual shoots, which are formed from April to May. The seeds sit on the seed scales and mature after fertilization for 1.5 years. The seed cones are dense green at first, but during maturation they gain a brown colour, their scales diverge and the seeds with wings fly apart (Pravdin 1964, Steven and Carlisle 1959).

Scots pine is one of the light-demanding species. Light stimulates the germination of Scots pine seeds. The weak shade tolerance of Scots pine is explained by the inability of its assimilatory apparatus to adapt to strong shading. Pine undergrowth under the forest canopy suffers from light starvation but when lightening, part of the undergrowth manages to adapt its assimilatory apparatus, after which it significantly accelerates the growth (Zelawsk and Zelawska 1967, Rysin and Savelieva 2008). Scots pine is considered as a thermophilic species. The dependence of Scots pine on the temperature regime is very clearly visible in the mountainous regions of Eastern Siberia and Transbaikalia, where it grows on the better-warmed southern slopes and is absent on the north facing slopes (Rysin and Savelieva 2008). Scots pine is a xerophytes species. The peculiarities of its assimilatory apparatus and root system allow it to grow on different soils with very different hydrological regimes. Scots pine is able to exist with a relatively low oxygen content in the soil, but up to certain limits. For instance, when the oxygen concentration in the soil decreases from 21% to 9%, the growth of the pine seedlings roots does not change significantly but at 4% it slows down by 20 times (Leyton and Rousseau 1958). Scots pine trees can grow on very poor soils, they are able to exist with a significant shortage and even with the complete cessation of the nutrients intake. Scots pine can grow on very acidic soils (pH = 3.0-3.5) and on chalky soils, but the optimal pH for this species is in the range of 4.5-5.6 (Ivanov 1970). With an

increase in pH to 7.0, the weight of pine seedlings decreases by almost half compared to the control (Levkina 1964).

Scots pine is a widespread tree in Eurasia, starting from Spain and Great Britain and further east to the Aldan River basin and the middle Amur River in Eastern Siberia. In the north, it grows up to Lapland, in the south it is found in Mongolia and China (Pravdin 1964). The area of Scots pine distribution in Russia is shown in Figure 1.1.

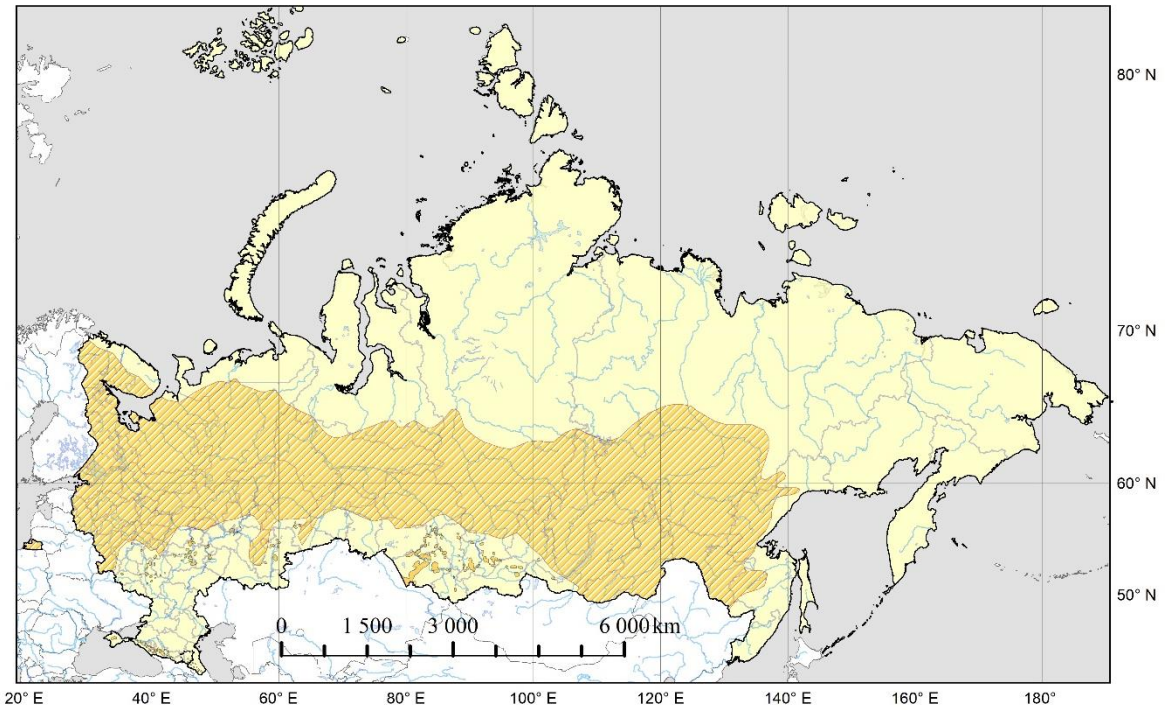


Figure 1.1. Scots pine distribution in Russia (Pravdin 1964). Orange areas represent natural presence of Scots pine.

In Russia, Scots pine is one of the main forest-forming tree species and it covers 15.6 % of Russian forests (FAO 2012). A significant part of its range is located in Siberia, where Scots pine forests form a complex spatial mosaic and only in the Angara River basin they form a continuous distribution range, occupying a variety of soils (Pimenov 2015).

Scots pine is often planted to consolidate sands and ravines, strengthen mountain slopes, create protective and roadside lanes. It is of great interest as an ornamental tree used in landscaping and landscape construction. Furthermore, it is a commercially important tree species, its wood is easily workable with good mechanical properties and has many uses, primarily as construction material and pulpwood (Mátyás et al. 2004).

1.2. Aim and objectives of the study

The aim of the study is to characterize the patterns of genetic diversity of Scots pine from different ecosystems and geographic regions at nuclear and chloroplast levels.

More specifically, the research objectives are the assessment of:

1. chloroplast DNA diversity in Scots pine populations from Middle Siberia;
2. nuclear DNA diversity within and among natural Scots pine populations from Middle Siberia;
3. nuclear DNA diversity of southern Middle Siberian mountain and foothill populations of Scots pine.

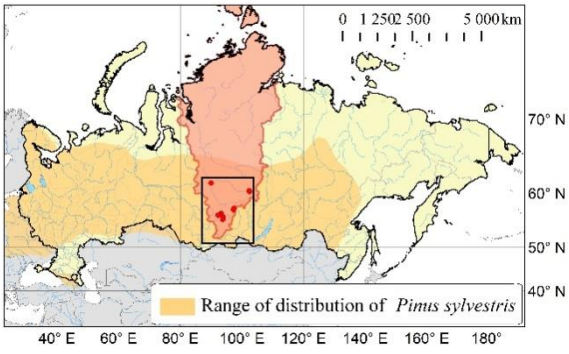
1.3. Material and methods

1.3.1. Study area

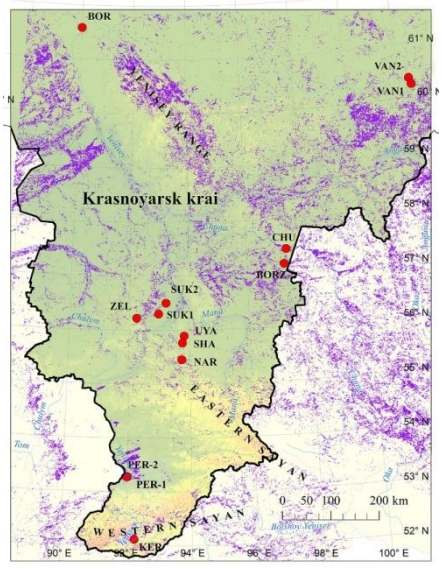
Middle Siberia is a huge region in terms of territorial coverage and unique in terms of natural conditions. A number of geographers associate the territory of this region mainly with the Central Siberian Plateau, determining its location between the Yenisey and Lena rivers to the basin of the Aldan River, the coastline of the Kara and Laptev seas, the slopes of the Eastern Sayan Mts. and the mountain belt of the Baikal region (Parmuzin 1964; Snytko and Konovalova 2005). According to Gerasimov and co-authors (1964), Middle Siberia also includes the part of the West Siberian Plain adjacent to the Yenisey River and the central part of the Altai-Sayan Mountains. In this study, we also adhere to latter vision of the territory of Middle Siberia. Middle Siberia can be considered as the most forested region of Russia. The European taiga is heavily cut down, in Western Siberia about half of the area of the forest zone is occupied by swamps, and to the east of the Lena River and Lake Baikal, the taiga zone up to the Pacific Ocean is crossed by treeless or sparsely forested mountain ranges (Gerasimov 1964). The study areas are located on the territories of three federal regions of Russia: Krasnoyarsk krai, the Republic of Khakassia and the Republic of Tuva.

1.3.1.1. Climate characteristics of Krasnoyarsk krai

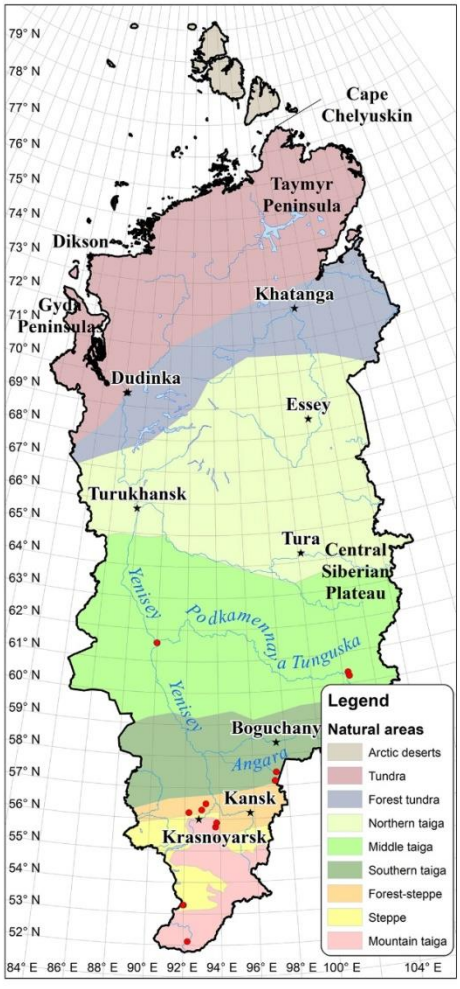
Krasnoyarsk krai occupies 2 366 797 km^2 and expands from 650 to 1460 km from west to east and 2990 km from north to south (Figure 1.2). The area of the forest of the Krasnoyarsk krai is 168.1 million hectares. Forests cover 69% of the territory of the region. The natural conditions of the region are very diverse due to its huge size. The south of the region is part of the Altai-Sayan mountain physico-geographical country, in the north of the region there is a vast North Siberian Lowland (Goryachko et al. 2010). The Yenisey River crosses most of the Krasnoyarsk krai. The West Siberian Plain occupies the left bank of the Yenisey River, and the East Siberian Plateau occupies the right bank.



(a)



(b)



(c)

Figure 1.2. (a), (b) Maps of native Scots pine populations from the Krasnoyarsk krai. Studied Scots pine populations are indicated by red dots. (b) Range of Scots pine distribution is shaded by purple (Shvidenko and Schepaschenko 2014). (c) Natural areas of Krasnoyarsk krai.

The territory of the Krasnoyarsk krai is located within three climatic zones: Arctic, subarctic and temperate. Mean annual temperature in the north is -10°C (Khatanga, Dudinka), in the center is -7°C (Turukhansk), in the south is -2.4°C (Boguchany) and -0.7°C (Kansk). Only in the Minusinsk basin temperatures are positive (1.2°C). Winter is cold and long. Mean January temperatures in the northern and central parts are from -27 (-29)°C (Dikson, Dudinka) to -36 (-39)°C (Tura, Essey); in the southern part are from -16°C (Krasnoyarsk) to -24°C (Boguchany) (Figure 1.2c). Mean July temperatures range from less than 10°C on the Arctic coast to 19°C in the south. The southern part is characterized by late spring and early autumn frosts, prolonged rains in the second half of summer. Annual precipitation amounts decrease from 550 mm in the west to 300 mm in the east; the maximum amount of precipitation (1000-1200 mm per year) falls on the windward slopes of the Sayan Mountains.

On the territory of the Krasnoyarsk krai there are several natural areas: Arctic deserts, tundra, forest tundra, taiga, forest-steppe, steppe (Figure 1.2c). The zone of Arctic deserts is represented on the extreme northern ledge of the Taymyr Peninsula - Cape Chelyuskin. There is a kingdom of mosses and lichens, among which there are flowering pioneer plants - Arctic poppy, saxifrage, tussock grass, etc. The largest part of the Taymyr and Gyda Peninsulas is occupied by a tundra zone, which is 600-700 km wide and is divided into Arctic, moss-lichen and shrub subzones. The forest tundra zone extends in the north of the West Siberian Plain with a strip of 150-200 km and in Middle Siberia the width of this zone reaches 900 km, merging with the zone of pre-tundra woodlands. The forest tundra is represented by Siberian larch and Siberian spruce is mixed in the south (Efits 2012). Almost 50% of the territory of the region is occupied by the taiga zone, which is divided into three subzones: northern taiga, middle taiga and southern taiga. To the south of the Arctic Circle, the northern taiga extends to 64°N on the Central Siberian Plateau and almost to 62°N west of the Yenisey River. To the east of the Yenisey River, shrub-moss and lichen forests of Siberian larch with an admixture of spruce predominate, and to the west of the Yenisey River moss- and lichen-shrub larch, spruce-larch and pine-larch forests with an admixture of cedar predominate. The middle taiga on the Central Siberian

Plateau reaches 58°N, and west of the Yenisey River it is spread to 60°N. Dark coniferous mixed forests dominate on the Yenisey Range, among which there are massifs of birch and aspen forests. In the Angara River basin and the valley of the Podkamennaya Tunguska, larch-pine and pine forests are presented. The southern taiga, extending to the west and east of the Yenisey Range, is characterized by an abundance of pine forests and floral diversity. The forest-steppes are represented by the Mariinsky-Achinsk forest-steppe (West Siberian type), Krasnoyarsk and Kansk forest-steppes (Central Siberian type). The forest-steppe zone is the most economically developed. Large areas of forest-steppe are occupied by arable land and haymaking. The vegetation cover of the southern mountainous part of the region is extremely diverse. The Eastern Sayan Mts. are characterized by dark coniferous forests, and the Western Sayan Mts. are occupied by mixed forests. Starting from an altitude of 1300-1400 m in the Kuznetsk Alatau Mts. and 1500-1700 m in the Sayan Mts., the subalpine belt stretches, and at an altitude of 1500- 2000 m the Alpine belt extends. Colorful subalpine and alpine meadows occupy mainly the southern slopes (Goryachko et al. 2010).

1.3.1.2. Climate characteristics of the Republic of Khakassia

Republic of Khakassia occupies 61 569 km^2 and expands 200 km from west to east and 460 km from north to south. Republic of Khakassia is located in the south of Siberia. Mountains occupy 80% of the territory (Figure 1.3).

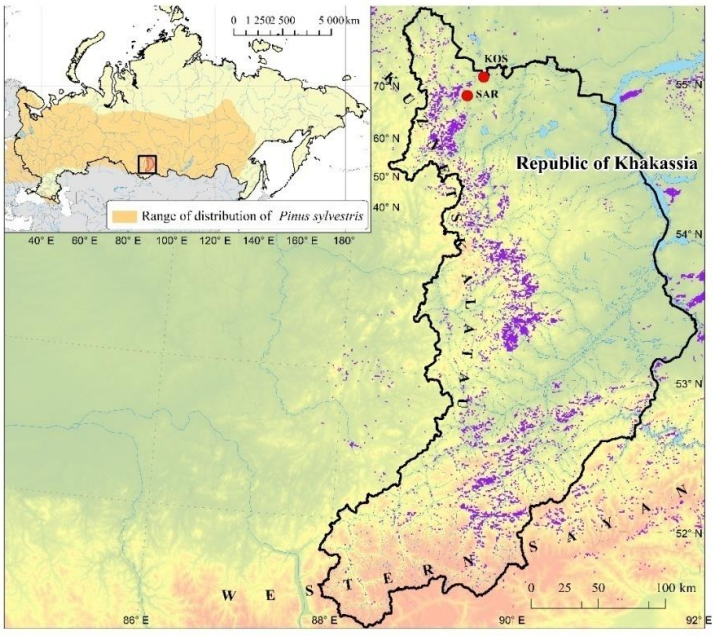


Figure 1.3. Map of native Scots pine populations from the Republic of Khakassia. Studied Scots pine populations are indicated by red dots. Range of Scots pine distribution is shaded by purple (Shvidenko and Schepaschenko 2014).

The territory of the region is heterogeneous and belongs to three large geographical areas: the Western Sayan Mts., the Kuznetsky Alatau Mts. and the Minusinskaya basin (Lushnikova et al. 2018). Hills with rocky outcrops and low ridges alternate with river valleys and lake basins. The climate is extremely continental with cold winters and hot summers in depressions. The average January temperature is from -19 to -21°C, in the foothills from -16 to -18°C. Summer is hot, with an abundance of sunny days. The average temperature in July is 18-19°C, in the mountains 12°C. The average annual precipitation ranges from 250 mm in the steppes to 600-700 mm in the mountains. The growing season is 155-165 days. In depressions, partly in the foothills and low mountains of Khakassia, steppes are common on typical black soils (26.4% of the territory). In the foothills and low mountains, there are expositional forest-steppes, in which larch-birch, birch-pine forests are widespread on gray forest soils. In the middle mountains, cedar-fir taiga with an admixture of larch prevails on mountain taiga brown, sod-podzolic soils. On the northern slopes of the Western Sayan Mts. there are larch-dark coniferous forests on mountain-taiga peat soils, on the southern slopes - larch herbaceous forests on mountain-forest black soils (Samoylova 2017).

1.3.1.3. Climate characteristics of the Republic of Tuva

Republic of Tuva occupies 168 604 km^2 and expands 630 km from west to east and 420 km from north to south (Figure 1.4). Republic of Tuva is located in the south of Siberia, in the upper reaches of the Yenisey River. Mountains occupy 82% of the territory (17% of them are highlands), depressions - 18%. The mountains with an altitude of 1200-1800 m prevail (Samoylova 2017).

The climate is extremely continental. Winters are harsh, long (from November to April), with many sunny days. The average temperature in January is from -28 to -34.9°C (in depressions). On the slopes, especially in the south, there is a temperature inversion. In depressions, the height of the snow cover is 10-15 cm, in the mountains 50-80 cm or more. Summers are warm, arid, hot in depressions,

moderately warm in the mountains. The average temperature in July is 16-18°C (max. 39°C). Annual precipitation ranges from 600 mm in the mountains to 138 mm in the depressions. The vegetation period is 150-160 days. The diversity of the soil and vegetation cover is associated with the intermediate position of Tuva between the Mongolian semideserts and the boreal regions of Siberia. In the south, semideserts are spread on brown desert and light chestnut soils. In depressions of the central part, there are steppes on black and dark chestnut soils. Scots pine forests (Balgazynsky forest) are represented in depressions on sandy sediments (Kuzhuget 2014). In the mountains, the altitudinal zonation of landscapes occurs. At an altitude of 1300-1500m, steppes or forest-steppes are replaced by mountain cedar-larch forests with an admixture of spruce, birch and aspen. Forests occupy 64.5% of the territory (Samoylova 2017).

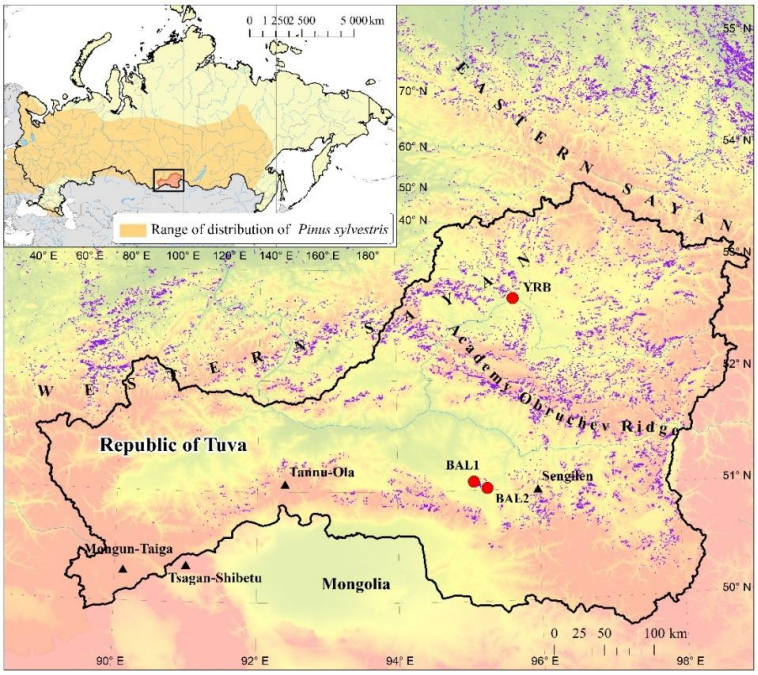


Figure 1.4. Map of native Scots pine populations from the Republic of Tuva. Studied Scots pine populations are indicated by red dots. Range of Scots pine distribution is shaded by purple (Tikhonova et al. 2018).

1.3.2. Sampled Scots pine populations

Nineteen native populations of Scots pine were chosen within the natural distribution range of the species in Middle Siberia (Figures 1.5). Scots pine is a dominant tree species in all studied populations. Five of them (BOR, VAN1, VAN2, CHU and BORZ) are located in taiga forest zone, three (SUK1, SUK2 and ZEL) grow in forest-steppe zone and eleven (UYA, SHA, NAR, KOS (SAR1), SAR (SAR-2), PER-1, PER-2, KER, YRB (TOD), BAL1 and BAL2) are distributed in Southern-Siberian mountain zone. Seven distant populations were included in the study: four from the Romanian Carpathians (POI, CHE, RET, LOT), one from the European part of Russia (SHA), one from West Siberia (ARO) and one from the Russian Far East (SVO).

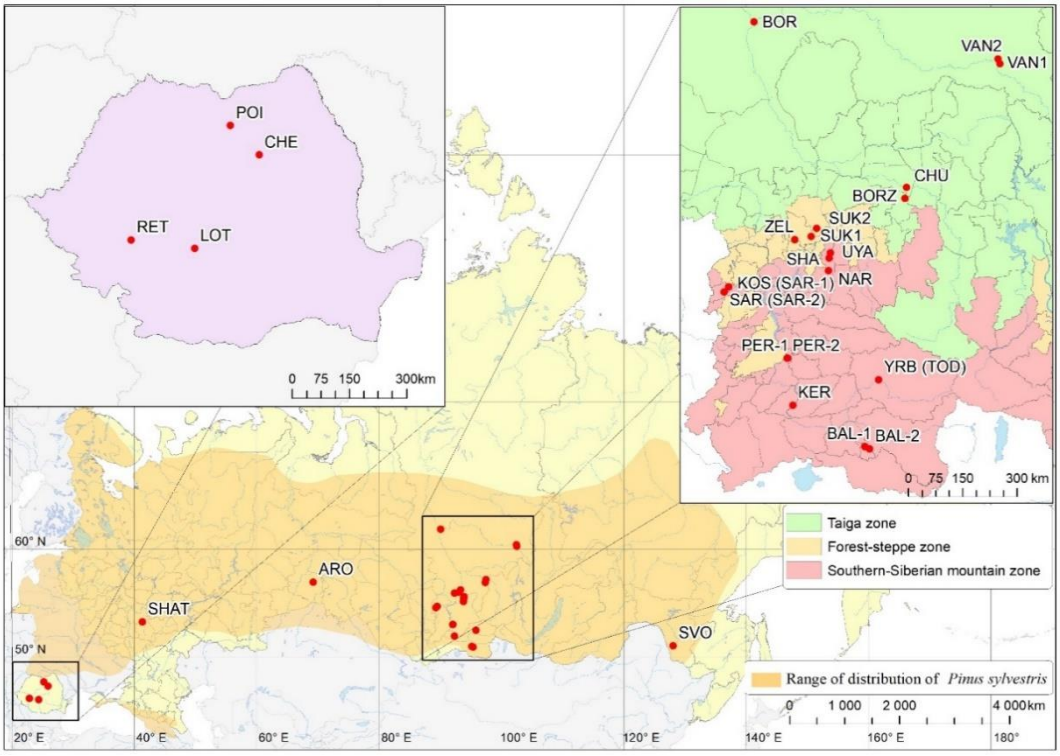


Figure 1.5. Location of the studied Scots pine populations in Middle Siberia, the European part of Russia, West Siberia, the Russian Far East and the Romanian Carpathians.

1.3.3. Molecular Analysis

Genomic DNA was extracted from dried needles with the CTAB method (Doyle and Doyle 1990). The quality and concentration of the extracted DNA were measured with Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific, USA), then diluted to a concentration of 10 - 20 ng/µl.

Ten chloroplast microsatellite markers were chosen for the genetic analysis: PCP45071, PCP36567, PCP48256, PCP41131, PCP30277, PCP26106, Pt1254, Pt15169, Pt71936, Pt87268 (Vendramin et al. 1996; Provan et al. 1998). The cpSSR loci were amplified in two PCR multiplex reactions in total volume of 10-12 µl using Qiagen Multiplex PCR Kits (Qiagen, Germany) under conditions recommended by the manufacturer. PCR amplification was performed in a Corbett thermal cycler (Corbett Research, Australia) with the following conditions: an initial denaturation of 15 min at 95°C, then 30 cycles of 15 s at 94°C, 1 min 30 s at 60°C (for the set of 6 PCP loci) and 58°C (for the set of 4 Pt loci), 1 min 30 s at 72°C and a final extension of 10 min at 72°C.

Ten nuclear SSR primer pairs were used in the genetic analysis: Psyl16, Psyl17, Psyl42, Psyl44, Psyl57 (Sebastiani et al. 2012); PtTX2146 (Auckland et al. 2002); PtTX4001 (Auckland et al. 2002); lw_isotig04195, lw_isotig04306, lw_isotig07383 (Fang et al. 2014). All primers were combined into three multiplex sets: set 1 consisted of Psyl44, Psyl57 and lw_isotig04306; set 2 comprised of Psyl16, PtTX2146 and lw_isotig07383; and set 3 contained Psyl17, Psyl42, lw_isotig04195 and PtTX4001. Reverse primers were labelled for use in GenomeLab GeXP genetic analyzer (Beckman Coulter, Fullerton, CA, USA). The chosen loci were amplified using Qiagen Multiplex PCR Kits (Qiagen, Hamburg, Germany) under conditions recommended by the manufacturer. PCR reactions were performed under the following conditions: an initial denaturation of 5 min at 95 °C, then 32 cycles of 30 s at 95 °C, 90 s at 58 °C, 50 s at 72 °C and a final extension of 7 min at 72 °C. Genotyping of individuals was performed using the GenomeLab GeXP software (Version 10.2, Beckman Coulter, Fullerton, CA, USA).

Amplified fragments were analyzed on a GenomeLab GeXP Genetic Analyser (Beckman Coulter, Fullerton, CA) genetic analyzer with an internal size standard. Fragment sizing was performed using the GenomeLab GeXP software (Version 10.2, Beckman Coulter, Fullerton, CA).

1.3.4. Statistical Analysis

POPGENE ver. 1.31 (Yeh et al. 1999) and GenAlEx v. 6.5 (Peakall and Smouse 2006) were used to estimate genetic diversity parameters: number of alleles (*Na*); number of effective alleles (*Ne*); inbreeding coefficient of an individual relative to the subpopulation (*F*is); inbreeding coefficient of an individual relative to the total population (*F*it); genetic differentiation coefficient (*F*st); observed heterozygosity (*Ho*); expected heterozygosity (*He*); Shannon`s Information index (*I*); Nei`s gene diversity index (*h*); Gene flow (*Nm*). The allelic richness (A_R) was computed in R (R Core Team. R 2013) using the "hierfstat" package (Goudet 2005).

Chloroplast DNA haplotypes were determined as a combination of the different microsatellite variants across ten cpSSR loci. HAPLOTYPE ANALYSIS ver. 1.05 (Eliades and Eliades 2009) was used to estimate the number of different haplotypes (*A*), number of private haplotypes (*P*), effective number of haplotypes (N_E), haplotype diversity (H_{CP}) and mean genetic distance between haplotypes (D^2sh).

BOTTLENECK software v.1.2.02 (Piry et al. 1999) was used to test for recent population bottlenecks on the basis of the stepwise mutation model (SMM). Statistical significance was determined by one-tailed Wilcoxon signed-rank test with 1000 iterations.

Population structure was analyzed using STRUCTURE v.2.3.4 with a Bayesian clustering approach (Pritchard et al. 2000). Testing twenty independent runs with K from 1 to 10, each run had a burn-in period of 100 000 iterations and 500 000 Monte Carlo Markov iterations, assuming admixture model (with LocPrior) with correlated allele frequencies. The studied populations were separated into groups by the Structure Harvester program (Earl and VonHoldt 2012) based on ΔK and mean L(*K*) values (Evanno et al. 2005). The average matrices of individual membership proportions for each population were estimated using CLUMPP v.1.1.2. (Jakobsson and Rosenberg, 2007).

Hierarchical analysis of molecular variance (AMOVA), implemented in GenAlEx v. 6.5 software, was used to determine the partitioning of the genetic variation among populations. The significance of differences was estimated using a permutation approach with 999 replications. Additionally, a pairwise *F*st matrix (Nei et al. 1983) and a a Principal Coordinates Analysis (PCoA) was conducted using the "ade4" package in R (Thioulouse et al. 2018) to compare genetic differentiation among populations.

SAMOVA software v.2.0 (Dupanloup et al. 2002) was used to identify groups of populations that are maximally differentiated from each other. Runs were conducted with the number of groups set from two to seven, performing 100 independent simulated annealing processes. The maximum F_{CT} value was chosen as the indicator of the best grouping.

Genetic Landscape Shape Interpolation analysis was carried out, as implemented in Alleles in Space (Miller 2005), to produce a surface plot that shows major genetic discontinuities, indicating probable contact areas between the detected genetic clusters. On the surface plot, positive peaks indicate areas with high genetic discontinuities (high genetic distances) and negative peaks represent of areas with genetic similarities (low genetic distances).

Potential barriers to gene flow among the studied populations were identified using Monmonier's maximum-difference algorithm (Monmonier 2010) implemented in BARRIER software v.2.2 (Manni et al. 2004). We generated 1000 *D* distance matrices (Nei's standard genetic distance corrected for sample size) in MSA software (Dieringer and Schlötterer 2003) by bootstrapping over the eight nSSR loci. The matrices were subsequently used to estimate possible species boundaries.

To test the correlation between geographical distances and genetic distances between population pairs a Mantel test was implemented. The Mantel test was performed in the "adegenet" package in R (Jombart and Ahmed 2011). Additionally, since geographic isolation can lead to either continuous clines of genetic differentiation, or to existence of distant patches, we used a 2-dimensional (2-D) kernel density estimator to the linearized F_{ST} values using the "MASS" package in R (Venables and Ripley 2013). The kernel density approach is aimed to reveal an underlying genetic structure that may help to explain observed correlation between the two distances.

To test the isolation-by-environment (IBE) hypothesis (Wang and Bradburd 2014), Euclidean climatic distances were calculated from recent (c. 1950–2000) climate data using 19 bioclimatic variables (Table S1) which were extracted from the global climate layer data using a grid size of 30 arc-seconds and downloaded from the WorldClim v.1.4 database (http://www.worldclim.org/). After, genetic, geographic, and the climatic distances were used in Mantel, partial-Mantel and MMRR (Multiple Matrix Regression with Randomization) regression analyses. The partial-Mantel test was conducted using the "vegan" package (Oksanen et al. 2022), while the MMRR carried out using the custom script of Wang

et al. (2013). The MMRR R script is deposited in the Dryad Data Repository under DOI:10.5061/dryad.kt71r)

1.4. Molecular genetic studies on Scots pine

In recent decades, a considerable amount of regional studies of the genetic diversity of Scots pine populations have been performed using mitochondrial, chloroplast and nuclear SSR (simple sequence repeats) markers. The DNA of these organelles is inherited differently. Cytoplasmic genomes of *Pinaceae* species have a special inheritance system, e.g. mitochondrial DNA is inherited maternally and, accordingly, its genetic flows are limited by the mobility of seeds. Due to the reduced mobility of seeds, mitochondrial markers are suitable for identifying the genetic structure of populations. However, the application of mitochondrial markers in relation to Scots pine is hampered by the case that the polymorphism detected in nad1 and nad7 mitochondrial regions is limited to the western part of Scots pine range while the populations from the northeast European Russia and Siberia revealed to be not polymorphic according to these fragments (Naydenov et al. 2007; Čelepirović et al. 2009; Buchovska et al. 2013; Dering et al. 2017; Semerikov et al. 2018).

Chloroplast DNA is inherited paternally, and the high mobility of pollen provides a more significant genetic flow. Chloroplast DNA markers allow to identify recently appeared DNA haplotypes and newly emerged genetic structure (Provan et al. 1998, Robledo-Arnuncio et al. 2005, Cheddadi et al. 2006, Semerikov et al. 2014, Bernhardsson et al. 2016, Dering et al. 2017).

Nuclear DNA markers provide opportunities for studying the history and mechanisms of evolution. However, there are some problems related to the application of nuclear DNA markers, such as recombination, selection, mutation, heterozygosity, PCR and sequencing difficulty, etc. (Zhang and Hewitt 2003). Even though nuclear microsatellites have proved to be useful to study phylogeographic and gene flow patterns in conifers to develop reliable new markers for conifer species is difficult due to their large genome size (estimated in 22.474 Mb for Scots pine) and the extensive repetitive nature of their DNA (Sebastiani et al. 2012; González-Martínez et al. 2004, 2010). For this reason, cross-amplification of earlier designed SSR markers is a cost-effective way to study highly polymorphic loci on several different species (Tóth et al. 2017). The studies based on nDNA in Scots pine confirmed that

nuclear SSR markers can be effectively used to investigate genetic diversity of Scots pine populations (Scalfi et al. 2009, Belletti et al. 2012, Naydenov et al. 2011, Lučić et al.2014, Pyhäjärvi et al. 2007). The knowledge of the Scots pine genetic diversity patterns is very important for basic and practical forestry. It can improve management of Scots pine genetic resources especially under changing climatic factors.

1.5. Outline of the research

The Doctoral thesis is presented in manuscript-style chapters. The focus of these chapters is as follows:

Chapter 2: Chloroplast DNA diversity in populations of P. sylvestris L. from Middle Siberia and the Romanian Carpathians.

In this study, we used ten paternally inherited chloroplast microsatellite loci to investigate the genetic diversity of nineteen Scots pine populations from Middle Siberia and the Romanian Carpathians. The results of the study showed high genetic diversity (H_{CP} = 0.91-1.00) in all the investigated populations. The cpSSR analysis yielded a total of 158 cpDNA haplotypes. The majority of the haplotypes (85%) were detected only once (unique haplotypes), and the rest were observed in two-to-nine individuals. Three common haplotypes were found between the Carpathian and the Siberian populations of Scots pine. Analysis of molecular variance (AMOVA) showed that only 3% of the variation occurred among populations from Middle Siberia and 6% of the variation existed among populations from the Carpathian Mountains. Overall, we found a weak geographic population structure in Scots pine from Middle Siberia and the Romanian Carpathians. The present study on genetic diversity in the Siberian and the Carpathian populations of Scots pine may contribute to the sustainable management and conservation of Scots pine genetic resources in Middle Siberia and the Romanian Carpathians.

Chapter 3: Genetic diversity and population structure of Scots pine (Pinus sylvestris L.) in Middle Siberia

Here, we assessed the genetic variation of 17 populations representing different parts of Scots pine range in Russia by using nuclear microsatellite markers (nSSR). Specifically, 14 populations were chosen within the natural distribution range of the species in Middle Siberia and three distant populations were sampled from the European part of Russia, the West Siberia and the Russian Far

East. All populations showed high values of genetic diversity (H_E = 0.514) and allelic richness A_R = 4.150. However, the easternmost population has shown the lowest level of genetic diversity (H_E = 0.433) and allelic richness (A_R = 3.505). Five genetic groups could be detected that correspond to: the European part of Russia, the south of Middle Siberia, the northwest of Middle Siberia, West and Middle Siberia, and the Russian Far East. However, the European population was the most genetically distinct one. The variation among Scots pine populations accounted for only 5% of the total variance. The highest level of genetic differentiation was found only between west-ernmost and easternmost populations (F_{ST} = 0.097). This data may contribute to a better understanding of the pattern of genetic diversity of Scots pine populations in Middle Siberia and help the conservation efforts of Scots pine genetic resources.

Chapter 4: Genetic legacy of southern Middle Siberian mountain and foothill populations of Scots pine (Pinus sylvestris L.): Diversity and differentiation.

Genetic diversity and structure of eight Scots pine populations located in southern Middle Siberia was studied. A high level of genetic diversity (H_E=0.518) was detected in the studied populations. No recent bottleneck effect, isolation by distance or isolation-by-environment were detected. Most genetic diversity was found within populations, while only 7% of genetic diversity occurred among populations. Both STRUCTURE analysis and UPGMA clustering showed two genetic groups. Two populations from the Minusink basin and a population from the Western Sayan Mts. formed the first group and the second group was composed of the other populations from Kuznetsk Alatau Mts., Central Tuva basin and Todzha basin. Our findings suggest that the studied Scots pine populations originate from different gene pools. The pattern of genetic diversity revealed by our study may be useful for the elaboration of conservation measures of genetic resources of Scots pine in southern Middle Siberia.

1.6. References

1. Auckland, L.D., Bui, T., Zhou, Y., Shepherd, M., Williams, C.G. (2002). Conifer microsatellite handbook. Texas A&M University, College Station TX, 57.
2. Belletti, P., Ferrazzini, D., Piotti, A., Monteleone, I., Ducci, F. (2012). Genetic variation and divergence in Scots pine (*Pinus sylvestris* L.) within its natural range in Italy. Eur J For Res 131, 1127–1138.

3. Bernhardsson, C., Floran, V., Ganea, S.L., García-Gil, (2016). Present genetic structure is congruent with the common origin of distant Scots pine populations in its Romanian distribution. For Ecol Manag 361,131-143.
4. Buchovska, J., Danusevičius, D., Baniulis, D., Stanys, V., Šikšnianienė, J.B., Kavaliauskas, D. (2013). The Location of the Northern Glacial Refugium of Scots Pine Based on Mitochondrial DNA markers. Baltic Forestry 19 (1), 2-12.
5. Čelepirović, N., Ivankovic, M., Gradecki-Postenjak, M., Nagy, L., Borovics, A., Agbaba, A.N., Littvay, T. (2009). Review of investigation of variability of nad1 gene intron B/C of mitochondrial genome in Scots pine (*Pinus sylvestris* L.). Period Biol 111, 453–457.
6. Cheddadi, R., Vendramin, G. G., Litt, T., François, L., Kageyama, M., Lorentz, S., Laurent, J.-M., de Beaulieu, J.-L., Sadori, L., Jost, A., and Lunt, D. (2006). Imprints of glacial refugia in the modern genetic diversity of Pinus sylvestris. Global Ecol. Biogeogr. 15, 271-282.
7. Dering, M., Kosiński, P., Wyka, T.P., Pers-Kamczyc, E., Boratyński, A., Boratyńska, K., Reich, P.B., 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.
8. Dering, M., Baranowska, M., Beridze, B., Chybicki, I.J., Danelia, I., Iszkuło, G., Kvartskhava, G., Kosiński, P., Rączka, G., Thomas, P.A. et al. (2021). The evolutionary heritage and ecological uniqueness of Scots pine in the Caucasus ecoregion is at risk of climate changes. Sci. Rep. 11, 22845.
9. Dieringer, D., Schlötterer, C. (2003). MICROSATELLITE ANALYSER (MSA): A platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes. 3, 167–169.
10. Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. Focus 12, 13–15.
11. Dupanloup, I., Schneider, S., and Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11, 2571–2581.
12. Earl, D. A., and vonHoldt, B. M. (2012). Structure harvester: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361.
13. Efits, O.A. (2012). Vegetation cover of the Krasnoyarsk krai. Siberian Federal University. 128.

14. Eliades, N.-G., Eliades, D.G. (2009). HAPLOTYPE ANALYSIS: Software for Analysis of Haplotypes Data. Distributed by the Authors. Forest Genetics and Forest Tree Breeding; Georg-Augst University Goettingen: Göttingen, Germany,18.
15. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14, 2611–2620.
16. Fang, P., Niu, S., Yuan, H., Li, Z., Zhang, Y., Yuan, L., Li, W. (2014). Development and characterization of 25 454 EST-SSR markers in Pinus sylvestris var. mongolica (Pinaceae). Appl. Plant Sci. 2, 455.
17. FAO (2012). The Russian Federation forest sector. Outlook study to 2030, Rome, Italy, 84.
18. FAO (2014) The state of the world's forest genetic resources. Commission on genetic resources for food and agriculture, Rome, Italy, 277.
19. Gerasimov, I.P., Preobrazhensky, V.S., Richter, G.D. (1964). Middle Siberia. Moscow, 480.
20. González-Martínez, S.C., Robledo-Arnuncio. J.J., Collada. C., Díaz, A., Williams, C.G., Alía, R., Cervera, M.T. (2004). Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines. Theor Appl Genet 109, 103–111.
21. González Díaz P. (2018). Development and maintenance of genetic diversity in Scots pine, *Pinus sylvestris* (L.). Doctoral thesis. University of Stirling. United Kingdom, 196.
22. Goryachko, M.D., Samoylova, G.S., Kuzminykh, S.V., Fedorova, V.I., Prokinova, A.N., Pavlinov, P.S. (2010). Krasnoyarsk krai. The Great Russian Encyclopedia, 5, 630.
23. Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes. 5, 184-186.
24. Ivanov, A.F. (1970). Growth of woody plants and soil acidity. Minsk, 218.
25. Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.
26. Jombart, T., Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27, 3070–3071.

27. Kavaliauskas, D., Danusevičius, D., 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, 1179.
28. Kuzhuget, S.K. (2014). Forest landscape - Balgazyn ribbon pine forest. Science Time 11, 163-166.
29. Levkina, T.I. (1964). On the question of the relation of seedlings of tree species to the reaction of the environment and liming of soils of forest nurseries. The renewal of forests in deforestation and the cultivation of seedlings in nurseries. Petrozavodsk. 179-202.
30. Leyton, K., Rousseau, L. (1958). Root growth of tree seedlings in relation to aeration. In: K. Thimann (ed.), The Physiology of Forest Trees, 467-475.
31. Lučić, A., Popović, V., Nevenić, M., Ristić, D., Rakonjac, L., Čirković –Mitrović, T., Mladenović –Drinić, S. (2014). Genetic diversity of Scots pine (*Pinus sylvestris* L.) populations in Serbia revealed by SSR markers. Arch Biol Sci 66, 1485–1492.
32. Lushnikova, O.S., Andreeva, A.A., Zalutsky, D.V., Lysogorsky, K.V., Adygaeva, N.N., Nepomnyashchy, V.V., Shurkina, V.V., Popov, A.A., Ankipovich, E.S., Zlotnikova, T.V., Shanina, E.V., Ivanova, T.V. (2018). State report. Ministry of Natural Resources and Ecology of the Republic of Khakassia, 242.
33. Markatyuk A. A. (2012). The importance of boreal forests located on territories of Eastern Siberia. Actual problems of the forest complex. Bryansk, 31, 40-41.
34. Mátyás, C., Ackzell, L., Samuel, C.J.A. (2004). EUFORGEN Technical Guidelines for Genetic Conservation and Use for Scots Pine (*Pinus sylvestris* L.); International Plant Genetic Resources Institute: Rome, Italy, 6.
35. Miller, M. P. (2005). Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. J. Hered. 96, 722–724.
36. Monmonier, M.S. (2010). Maximum-Difference Barriers: An Alternative Numerical Regionalization Method. Geogr. Anal. 5, 245–261.
37. Naydenov, K.D., Senneville, S., Beaulieu, J., Tramblay, 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.

38. Naydenov, K.D., Naydenov, M.K., Tremblay, F., Alexandrov, A., Aubin-Fournier, L.D. (2011). Patterns of genetic diversity that result from bottlenecks in Scots Pine and the implications for local genetic conservation and management practices in Bulgaria. New Forests 42, 179-193.
39. Nei, M., Tajima, F., Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J. Mol. Evol. 19, 153-170.
40. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2022). Package 'vegan': community ecology package.
41. Padutov, V.E. (2001). Genetic resources of pine and spruce of Belarus, monograph, Gomel, 144.
42. Padutov, V. E., Khotyleva, L.V., Baranov, O.Yu., Ivanovskaya, S.I. (2008). Genetic effects of transformation of forest ecosystems. Ecological Genetics VI (1), 3-11.
43. Parmuzin ,Yu. P. (1964). Middle Siberia. Moscow. 311.
44. Peakall, R., Smouse, P.E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
45. Pimenov, A.V. (2015). Biodiversity of the Scots pine (*Pinus sylvestris* L.) in contrasting ecotopes of the south of Siberia]. Biology Doctoral dissertation. Krasnoyarsk. 406.
46. Piry, S., Luikart, G., and Cornuet, J. M. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 90, 502–503.
47. Pravdin, L.F. (1964). Scots pine. Moscow. Science. 189.
48. Pritchard, J., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
49. Provan, J., Soranzo, N., Wilson, N.J., McNicol, J.W., Forrest, G.I., Cottrell, J., Powell, W. (1998). Gene-pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. The Royal Society 265, 1697-705.
50. Pyhäjärvi, T., García-Gil, M.R., Knürr, T., Mikkonen, M., Wachowiak, W., Savolainen, O. (2007). Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. Genetics 177, 1713–1724.
51. R Core Team (2013). R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria.

52. Robledo-Arnuncio, J.J., Collada, C., Alia, R., Gil, L. (2005). Genetic structure of montane isolates of Pinus sylvestris L. in a Mediterranean refugial area. Journal of Biogeography 32(4), 595–605.

53. Rysin, L.P., Savelyeva, L.I. (2008). Pine forests of Russia. Moscow. Association of Scientific Publications, 289.

54. Samoylova, G.S. 2017. The nature of Khakassia. The Great Russian Encyclopedia 2004-2017. Osipov, Yu.S. (Ed.) Moscow, 799.

55. 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.

56. Sebastiani, F., Pinzauti. F., Kujala, S. T., González-Martínez, S. C., Vendramin, G. G. (2012). Novel polymorphic nuclear microsatellite markers for Pinus sylvestris L. Conserv. Genet. Resour. 4, 231-234.

57. Semerikov, V.L., Semerikova, S.A., Dymshakova, O.S., Zatsepina, K.G., Tarakanov, V.V., Tikhonova, I.V., Ekart, A.K., Vidyakin, A.I., Jamiyansuren, S., Rogovtsev, R.V., et al. (2014). Microsatellite loci polymorphism of chloroplast DNA of Scots pine (*Pinus sylvestris* L.) in Asia and Eastern Europe. Russ. J. Genet. 50, 577–585.

58. Semerikov, V., Semerikova, S., Putintseva, Yu., Tarakanov, V., Tikhonova, I., Vidyakin, A., Oreshkova, N., Krutovsky, K. (2018). Colonization history of Scots pine in Eastern Europe and North Asia based on mitochondrial DNA variation. Tree Genetics & Genomes 14, 1-7.

59. Sheller, M., Ciocîrlan, E., Mikhaylov, P., Kulakov, S., Kulakova, N., Ibe, A., Sukhikh, T., Curtu, A.L. (2021). Chloroplast DNA Diversity in Populations of *P. sylvestris* L. from Middle Siberia and the Romanian Carpathians. Forests 12, 1757.

60. Shvidenko, A.Z., Schepaschenko, D.G. (2014). Carbon Budget of Russian Forests. Siberian Journal of Forest Science 1, 69-92.

61. Snytko, V.A., Konovalova, T.I. (2005). Stability and anthropogenic transformation of taiga geosystems in the South of Central Siberia. Siberian Ecological Journal 4, 651-661.

62. Șofletea, N., Mihai, G., Ciocîrlan, E., Curtu, A.L. (2020). Genetic Diversity and Spatial Genetic Structure in Isolated Scots Pine (*Pinus sylvestris* L.) Populations Native to Eastern and Southern Carpathians. Forests 11, 1047.

63. Steven, H.M., Carlisle, A. (1959). The native pinewoods of Scotland, 368.

64. The State Coordination Program for the Development of Biotechnology in the Russian Federation for 2011-2020. Approved by the Government of the Russian Federation on 24.04.2012. 67.
65. Thioulouse, J., Dray, S., Dufour, A.B., Siberchicot, A., Jombart, T., Pavoine, S. (2018). Description of Species Structures. In Multivariate Analysis of Ecological Data with ade4. Springer: New York, NY, USA, 97–117.
66. Tikhonova, I.V., Mukhortova, L.V., Danilova, I.V., Mikhailova, I.A. (2018). Assessment of Soil Suitability for Growth of Coniferous Forests Species in Tuva Republic. Russian Journal of Forest Science 3, 194-209.
67. Tóth, E.G., Köbölkuti, Z.A., Pedryc, A., Höhn, M. (2017). Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. J. For. Res. 28, 637-651.
68. Venables, W.N., Ripley, B.D. (2013). Modern Applied Statistics with S-PLUS; Springer Science & Business Media: New York, NY, USA, 5, 1–6.
69. Vendramin, G.G., Lelli, L., Rossi, P., Morgante, M. (1996). A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol. Ecol. 5, 111–114.
70. Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution 67, 3403–3411.
71. Wang, I. J., and Bradburd, G. S. (2014). Isolation by environment. Mol. Ecol. 23, 5649–5662.
72. World Wildlife Fund (WWF). http://www.wwf.ru/forests/eng/ (accessed on 24 June 2023).
73. Yeh, F.C., Yang, R.C., Boyle, T. (1999). POPGENE Version 1.32 Microsoft Windows-Based Freeware for Populations Genetic Analysis; University of Alberta: Edmonton, AB, Canada, 1999.
74. Zelawski, W., Zelawska, B. (1967). Some aspects of the effect of shade on growth of Scots pine (*Pinus sylvestris* L.) seedlings of various provenance. Ecol. Polska ser. A., 15.
75. Zhang, D.X., Hewitt, G.M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular ecology 12(3), 563-584.

Chapter 2

CHLOROPLAST DNA DIVERSITY IN POPULATIONS OF *P. SYLVESTRIS* L. FROM MIDDLE SIBERIA AND THE ROMANIAN CARPATHIANS

The results presented in this chapter were published in the *Forests* Journal with minor modifications (Sheller, M.; Ciocîrlan, E.; Mikhaylov, P.; Kulakov, S.; Kulakova, N.; Ibe, A.; Sukhikh, T.; Curtu, A.L. Chloroplast DNA Diversity in Populations of *P. sylvestris* L. from Middle Siberia and the Romanian Carpathians. Forests. 2021, 12, 1757. Q1, Impact Factor – 2.9 https://doi.org/10.3390/f12121757)

2.1. Introduction

Genetic diversity is the basis for biological stability, which allows species to evolve and to adapt to changing environmental conditions (FAO 2014). Knowledge of the genetic structure and the level of genetic variability of populations is relevant for the development of measures aimed at conservation of species-genetic diversity (Altukhov 2004).

Scots pine (*Pinus sylvestris* L.) is one of the keystone species in forest ecosystems of the boreal regions in Eurasia. It is of great ecological and economic importance and is adapted to a variety of environmental conditions (Naydenov et al. 2007, Dering et al. 2017). Scots pine is a monoecious, wind-pollinated and predominantly outcrossing conifer (Robledo-Arnuncio et al. 2005, Sebastiani et al. 2012, Şofletea et al. 2020). It usually forms extensive pure forests or mixed stands with birch and other conifers (Olsson 2019).

In this study, we used chloroplast microsatellite markers to characterize the level of genetic diversity of Scots pine populations in two geographic regions of its Eurasian natural range with different ecological settings and evolutionary history. Specifically, we aimed to address the following questions: (1) What is the pattern of genetic structure and diversity in Scots pine populations from Middle Siberia and the Romanian Carpathians? (2) How large is the degree of genetic differentiation between the two Scots pine distribution regions at chloroplast DNA level?

2.2. Material and methods

2.2.1. Plant Material

Nineteen native populations of Scots pine were chosen within the natural distribution range of the species in Middle Siberia and the Romanian Carpathians (Figure 1.5). Scots pine is a dominated tree species in all studied populations. Four of them (VAN2, VAN1, CHU and BORZ) are located in taiga forest zone, three (SUK2, SUK1 and ZEL) grow in forest-steppe zone and eight (UYA, SHA, NAR, KOS, SAR, YRB, BAL1 and BAL2) are distributed in Southern-Siberian mountain zone. Four relict populations of Scots pine (POI, CHE, RET, LOT) are located in the Southern and Eastern Carpathian Mountains. Ten adult trees were randomly chosen in each population. Consequently, the total number of analyzed individuals was 190. Needles collected from the trees were stored in silica gel until DNA extraction was carried out.

2.2.2. Molecular Analysis

Genomic DNA was extracted from dried needles with the CTAB method (Doyle and Doyle 1990). Ten chloroplast microsatellite markers were chosen for the genetic analysis: PCP45071, PCP36567, PCP48256, PCP41131, PCP30277, PCP26106, Pt1254, Pt15169, Pt71936, Pt7268 (Vendramin et al. 1996, Provan et al. 1998). The cpSSR loci were amplified in two PCR multiplex reactions in total volume of 10-12 µl using Qiagen Multiplex PCR Kits (Qiagen, Germany) under conditions recommended by the manufacturer. Amplified fragments were analyzed on a GenomeLab GeXP Genetic Analyser (Beckman Coulter, Fullerton, CA) genetic analyzer with an internal size standard. Fragment sizing was performed using the GenomeLab GeXP software (Version 10.2, Beckman Coulter, Fullerton, CA).

2.2.3. Statistical Analysis

Parameters of genetic diversity and structure were calculated using the following programs: POPGENE ver. 1.31 (Yeh et al. 1999), GenAlEx v. 6.5 (Peakall and Smouse 2006), HAPLOTYPE ANALYSIS ver. 1.05 (Eliades and Eliades 2009), STATISTICA software v.8 (Weiß 2007), STRUCTURE ver.2.3.4 (Pritchard et al. 2000).

2.3. Results

The cpSSR analysis of 190 individuals of Scots pine yielded a total of 158 haplotypes. The majority of the haplotypes (85%) were detected only once (unique haplotypes). The four Carpathian and three Siberian populations of Scots pine (CHU, ZEL and UYA) were characterized by the highest number of private haplotypes (nine), while the lowest number (five) of private haplotypes was recorded in two Siberian populations (SUK1 and BAL1 populations). Three common haplotypes (H86, H107, H118) were found in two Siberian (NAR and BAL2) and three Carpathian populations of Scots pine (POI, CHE and RET) (Table S2.1).

The values of Nei`s genetic distance among populations ranged from 0.0032 (NAR/KOS) to 0.2441 (VAN1/POI) (Table S2.2). UPGMA clustering showed that two groups were separated at the population level (Figure 2.1.). The first group consisted of fifteen Middle Siberian populations and the second group was composed of four Carpathian populations.

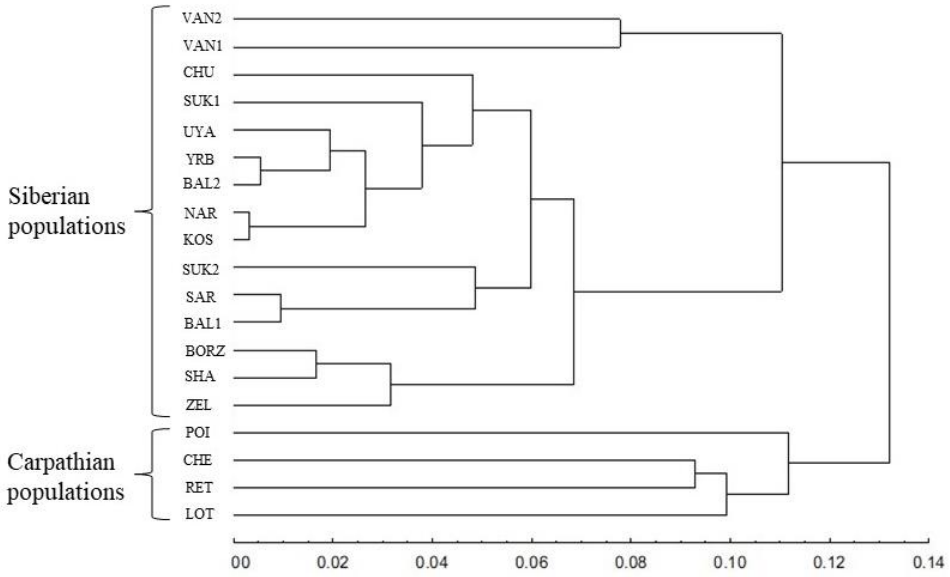


Figure 2.1. UPGMA dendrogram based on Nei's genetic distances among populations of Scots pine.

The hierarchical analysis of molecular variance (AMOVA) showed that the variation among two geographic regions (Middle Siberia and Carpathian Mountains) accounted for 5% of the total variance. The variance among populations within regions was 3% ($p < 0.01$). The AMOVA performed within the Siberian populations of Scots pine showed lower differentiation among populations (PhiPT = 3%, $p < 0.05$) compared to the Carpathian populations (PhiPT = 6%; $p < 0.01$).

The population structure analysis of the Siberian and Carpathian populations of Scots pine showed that ΔK was the highest when $K = 3$ (Figure 4). All the studied plants exhibited admixture from three genetic clusters. For $K = 3$, the Siberian populations had a larger membership in cluster 3 (blue color) than the Carpathian populations. The membership in cluster 1 (red color) was higher in case of Carpathian populations (Figure 2.2.).

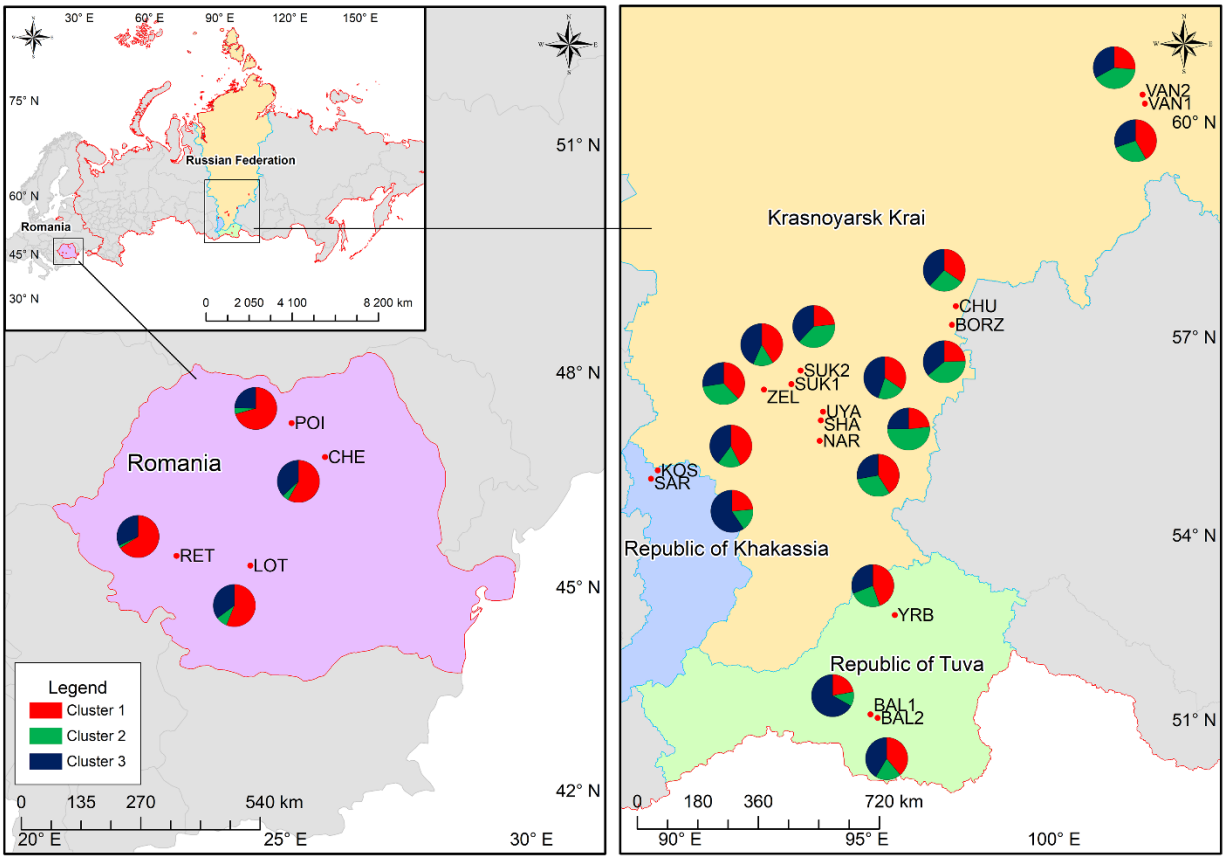


Figure 2.2. Geographical location of the studied Scots pine populations in Middle Siberia and the Romanian Carpathians. Populations of Scots pine are indicated by red dots. The sector maps show the mean cluster membership proportions of the analyzed individuals in each of the 19 Scots pine populations based on the structure at K = 3.

We analyzed the correlation between genetic distance and geographic distance for the studied populations using the Mantel test. The results showed a low but significant correlation between genetic differentiation and geographical distance among Siberian populations (R^2 = 0.1746, $p < 0.01$) and there was no significant correlation between genetic distance and geographic distance among Carpathian populations (R^2 = 0.0059, $p > 0.05$).

2.4. Discussion

Our results showed high genetic diversity (H_{CP} = 0.91-1.00) in all Scots pine populations studied. In Middle Siberia, two northernmost populations showed slightly higher haplotype diversity (H_{CP} = 1.00) than two southernmost populations (H_{CP} = 0.91 and H_{CP} = 0,98, respectively. The lowest haplotype number (seven) was detected in BAL1 population that refers to the southernmost pine forest of Northern Asia. Local fires, which occurred between 1988 and 2014, may be the cause of the reduction

in haplotype number observed in BAL1 population (Kuular and Namzyn 2015, Kuzhuget 2014). Decrease of population size, due to fires, logging, diseases and environmental pollution may lead to the decline in genetic diversity (Ellstrand and Elam 1993). Another explanation would be the relatively small sample size (10 individuals per population).

AMOVA showed a low level of genetic structure (PhiPT = 3%) among the fifteen Siberian populations, even though they are scattered over large distances of more than 1000 km. Based on STRUCTURE analysis three clusters were defined in the Siberian and the Carpathian populations of Scots pine where each tree was comprised of three genetic groups, showing a weak genetic structure between the two geographic regions, which is consistent with the grouping revealed by the UPGMA dendrogram.

Our findings, based on chloroplast DNA analysis, indicate that a high genetic diversity of individuals exists in Scots pine populations from Middle Siberia and the Romanian Carpathians and that, despite large geographic distances and barriers, there is limited genetic differentiation.

2.5. Conclusions

In the present study, ten cpSSR markers were employed to estimate genetic diversity within and among nineteen natural populations of Scots pine in Middle Siberia and the Romanian Carpathians. All populations showed high levels of genetic diversity. However, one of the southernmost Siberian populations has shown the lowest value of haplotype diversity. Accordingly, the performance of this population of Scots pine should be the focus of long-term study aimed at monitoring of population dynamics. Our study may contribute to the development of a strategy of sustainable management of Scots pine genetic resources in Middle Siberia and the Romanian Carpathians.

2.6. References

1. Altukhov, Y.P. (2004) Dynamics of gene pools under anthropogenic influences. Vestnik VOGiS 8, 40–59.
2. Dering, M., Kosiński, P., Wyka, T.P., Pers-Kamczyc, E., Boratyński, A., Boratyńska, K., Reich, P.B., 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. Diversity and Distributions. 23,540–555.

3. Doyle, J.J., Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus 12, 13–15.
4. Eliades, N-G., Eliades, D.G. (2009). HAPLOTYPE ANALYSIS: software for analysis of haplotypes data. Distributed by the authors. Forest Genetics and Forest Tree Breeding, Georg-Augst University Goettingen, Germany. 18.
5. Ellstrand, N.C., Elam, D.R. (1993). Population genetic consequences of small population size: Implications for plant conservation. Annu. Rev. Ecol. Syst. 24,217–242.49. Prishnivskaya, Ya.; Nassonova, E.; Vasileva, Yu.; Boronnikova, S. (2019). Selecting of Polymorphic Loci of Genome for Identification of Populations of *Pinus sylvestris* L. on East-Europe Plain. Bulletin of Science and Practice 5, 25-30.
6. Kuular, K.B. and Namzyn, S.A. (2015). Climatology and anthropogeny dynamics of Balgazyn ribbon pine forest. Interexpo Geo-Sib. 4,25–28.
7. Kuzhuget, S.K. (2014). Forest landscape - Balgazyn ribbon pine forest. Science Time 11, 163–166.
8. Naydenov, K.D., Senneville, S., Beaulieu, J., Tramblay, 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.
9. Olsson, J. (2019). Genetic diversity and hardiness in Scots pine from Scandinavia to Russia. Master thesis in Ecology. Umeå University, Sweden.
10. Peakall, R., Smouse, P.E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
11. Pritchard, J., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
12. Provan, J., Soranzo, N., Wilson, N.J., McNicol, J.W., Forrest, G.I., Cottrell, J., Powell, W. (1998). Gene-pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. The Royal Society 265,1697–705.
13. Robledo-Arnuncio, J.J., Collada, C., Alia, R., Gil, L. (2005). Genetic structure of montane isolates of *Pinus sylvestris* L. in a Mediterranean refugial area. Journal of Biogeography 32(4), 595–605.

14. Sebastiani, F., Pinzauti, F., Kujala, S.T., González-Martínez, S.C., Vendramin, G.G. (2012). Novel polymorphic nuclear microsatellite markers for Pinus sylvestris L. Conserv Genet Resour 4,231–234.
15. Șofletea, N., Mihai, G., Ciocîrlan, E., Curtu, A.L. (2020). Genetic Diversity and Spatial Genetic Structure in Isolated Scots Pine (*Pinus sylvestris* L.) Populations Native to Eastern and Southern Carpathians. Forests 11,1047.
16. Vendramin, G.G., Lelli, L., Rossi, P., Morgante, M. (1996). A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol. Ecol. 5,111–114.
17. Weiß, C.H. (2007). StatSoft, Inc., Tulsa, OK.: STATISTICA, Version 8. Asta Adv. Stat. Anal. 91, 339–341.
18. Yeh, F.C., Yang, R.C., Boyle, T. (1999). POPGENE Version 1.32 Microsoft Windows-Based Freeware for Populations Genetic Analysis; University of Alberta: Edmonton, AB, Canada, 1999.

2.7. Supplementary material

Table S2.1. Frequency of the 17 common haplotypes within Scots pine populations

No.	Haplotype*	Label	Population																	
			VAN2	VAN1	CHU	BORZ	SUK2	SUK1	ZEL	UYA	SHA	NAR	KOS	SAR	YRB	BAL1	BAL2	POI	CHE	RET	LOT
1	112/120/138/137/147/154/64/124/147/164	H41					0.1	0.1													
2	112/120/138/137/147/154/64/126/147/164	H42					0.1									0.1					
3	113/119/138/137/146/154/64/128/147/164	H59	0.1											0.1							
4	113/119/138/137/147/154/63/126/147/164	H62												0.1	0.1						
5	113/119/138/137/147/154/64/126/147/164	H64		0.1				0.1						0.1							
6	113/119/138/139/147/153/64/125/148/164	H68	0.1	0.1																	
7	113/120/138/137/146/154/64/125/147/164	H85				0.1						0.1									
8	113/120/138/137/146/154/64/126/147/164	H86															0.1			0.1	
9	113/120/138/137/147/153/65/125/147/164	H100									0.1				0.1						
10	113/120/138/137/147/154/64/124/148/164	H107										0.1							0.1		
11	113/120/138/137/147/154/64/126/147/164	H112	0.1				0.1	0.2		0.1			0.1			0.3					

12	113/120/138/137/147/154/64/127/146/165	H116							0.1						0.1						
13	113/120/138/137/147/154/64/127/148/164	H117									0.1	0.1									
14	113/120/138/137/147/154/64/128/147/165	H118										0.1						0.1			
15	113/120/138/137/147/155/64/126/148/164	H122			0.1	0.1											0.1				
16	113/120/138/138/147/153/64/125/148/164	H130		0.1									0.1								
17	113/120/139/137/147/154/64/127/149/166	H149					0.1	0.1													

* The second column shows the size variants at each of the ten cpSSR loci studied (PCP36567, PCP48256, PCP41131, PCP30277, PCP26106, PCP45071, Pt1254, Pt15169, Pt71936, and Pt87268, respectively).

Table S2.2. Nei`s genetic distances of 19 Scots pine populations

ID	VAN2	VAN1	CHU	BORZ	SUK2	SUK1	ZEL	UYA	SHA	NAR	KOS	SAR	YRB	BAL1	BAL2	POI	CHE	RET	LOT
VAN2	****																		
VAN1	0.0778	****																	
CHU	0.0916	0.1095	****																
BORZ	0.0729	0.1117	0.0643	****															
SUK2	0.1287	0.1956	0.053	0.0451	****														
SUK1	0.076	0.1335	0.0444	0.0608	0.0587	****													
ZEL	0.1036	0.0673	0.0506	0.0206	0.0688	0.0713	****												
UYA	0.0959	0.1541	0.058	0.05	0.06	0.0353	0.0473	****											
SHA	0.0909	0.1042	0.0828	0.0167	0.073	0.1325	0.0427	0.0989	****										
NAR	0.0531	0.0935	0.0491	0.0183	0.0571	0.0535	0.0171	0.0333	0.0636	****									
KOS	0.0068	0.0699	0.0395	0.0385	0.0792	0.011	0.0549	0.0318	0.0789	0.0032	****								
SAR	0.095	0.1642	0.0788	0.0565	0.0525	0.0536	0.0871	0.029	0.1048	0.0667	0.0606	****							
YRB	0.1192	0.1314	0.0539	0.0515	0.0822	0.057	0.0462	0.0261	0.0684	0.0332	0.0219	0.0601	****						
BAL1	0.1182	0.1973	0.0753	0.0564	0.0448	0.0389	0.1145	0.0294	0.1144	0.0779	0.0581	0.0095	0.0692	****					
BAL2	0.0929	0.1906	0.0431	0.0481	0.0849	0.0327	0.0859	0.0129	0.103	0.0265	0.0123	0.0491	0.0055	0.0334	****				
POI	0.1841	0.2441	0.1657	0.1345	0.1971	0.1549	0.1042	0.0485	0.1822	0.0662	0.1078	0.1351	0.0652	0.1483	0.0654	****			
CHE	0.145	0.1566	0.1389	0.1495	0.1884	0.139	0.1235	0.0926	0.1825	0.1072	0.1077	0.0936	0.1652	0.1264	0.1249	0.097	****		
RET	0.113	0.1777	0.1023	0.1662	0.1911	0.0823	0.1095	0.0666	0.1974	0.1152	0.0678	0.1436	0.1196	0.1458	0.1003	0.1035	0.0928	****	
LOT	0.1483	0.2324	0.1467	0.1627	0.1929	0.12	0.1363	0.0959	0.1849	0.1018	0.0718	0.1029	0.1106	0.117	0.0531	0.1348	0.095	0.1034	****

Chapter 3

GENETIC DIVERSITY AND POPULATION STRUCTURE OF SCOTS PINE (*PINUS SYLVESTRIS* L.) IN MIDDLE SIBERIA

The results presented in this chapter were published in the *Forests* Journal with minor modifications (Sheller, M.; Tóth, E.G.; Ciocîrlan, E.; Mikhaylov, P.; Kulakov, S.; Kulakova, N.; Melnichenko, N.; Ibe, A.; Sukhikh, T.; Curtu, A.L. Genetic Diversity and Population Structure of Scots Pine (*Pinus sylvestris* L.) in Middle Siberia. Forests 2023, 14, 119. Q1, Impact Factor – 2.9. https://doi.org/10.3390/f14010119)

3.1 Introduction

Genetic diversity is important for the long-term survival of species and plays a key role in their conservation (Spielman et al. 2004). It can contribute to the adaptability of species and increase chances that at least some individuals in a population are capable to withstand changing environmental conditions (Radu et al. 2014). Thus, to efficiently conserve the genetic diversity of a species, the level of genetic diversity should be defined (Graudal et al. 2020).

Scots pine is one of the most widespread tree species in the boreal forest of Eurasia. It has great ecological, economic and social importance (Şofletea et al. 2020, Sheller et al. 2021, Vasilyeva et al. 2021, Przybylski et al. 2022). In Russia, Scots pine forests are in the first place in terms of the timber production and clear cuttings are widely used (Markatyuk et al. 2013). At the same time, reforestation processes are passive, often significantly delayed and occur with an undesirable succession of tree species (Markatyuk et al. 2013). Due to natural disturbances (mainly forest fires), overexploitation and mismanagement, the area of pine forests in Russia has been decreased by two million ha over the last ten years (Parliamentary Newspaper 2022). Therefore, the study of Scots pine genetic resources in Russia is highly relevant for a sustainable use in breeding and conservation programs (Ilyinov and Raevsky 2016, Torbik et al. 2019).

In this study, we explored the genetic diversity and population structure of 17 Scots pine populations from different locations of the natural distribution range in Russia, using nuclear microsatellite

markers. We addressed the following questions: (i) What is the current level of genetic diversity and differentiation within and among Scots pine populations across Russia? (ii) Is there a population structure among the studied Scots pine populations?

3.2. Materials and methods

3.2.1. Plant material

A total of 17 natural populations of Scots pine were collected from the Russian distribution of the species. Among them, 14 populations were chosen within the natural distribution range of the species in Middle Siberia (BOR, VAN2, VAN1, CHU, BORZ, SUK2, SUK1, ZEL, NAR, KOS, SAR, SHAT, YRB, BAL1). Three distant populations were included in the study: one from the European part of Russia (SHAT), one from West Siberia (ARO) and one from the Russian Far East (SVO) (Figure 1.5). Geographic distances between populations ranged from 14 km (VAN2 and VAN1) up to 9656 km (SHAT and SVO). Initially, 510 individuals were sampled across Russia. Due to PCR failure, the number of studied trees was reduced to 406.

3.2.2. DNA isolation and microsatellite genotyping

DNA isolation was performed according to the CTAB method (Doyle and Doyle 1990). Initially, ten SSR primer pairs were used in the genetic analysis: Psyl16, Psyl17, Psyl42, Psyl44, Psyl57 (Sebastiani et al. 2012); PtTX2146 (Elsik et al. 2000); PtTX4001 (Zhou et al. 2002); lw_isotig04195, lw_isotig04306, lw_isotig07383 (Fang et al. 2014). All primers were combined into three multiplex sets: set 1 consisted of Psyl44, Psyl57 and lw_isotig04306; set 2 comprised of Psyl16, PtTX2146 and lw_isotig07383; and set 3 contained Psyl17, Psyl42, lw_isotig04195 and PtTX4001. Genotyping of individuals was performed using the GenomeLab GeXP software (Version 10.2, Beckman Coulter, Fullerton, CA, USA).

3.2.3. Data analysis

Parameters of genetic diversity and structure were calculated using the following programs: Micro-Checker software (Van Oosterhout et al. 2006), GenAlex v. 6.5 software (Peakall and Smouse 2006), R (R Core Team 2013), STRUCTURE software v.2.3.4 (Pritchard et al. 2000), CLUMPP v.1.1.2. (Jakobsson and Rosenberg 2007), BARRIER software v.2.2 (Manni et al. 2004).

3.3. Results

The mean number of alleles present per population ranged from 4.0 (SUK1) to 5.5 (BOR) with an overall mean of 4.6. Effective number of alleles (N_E) varied between 2.045 in SVO population and 2.869 in SHAT population, with an average of 2.494 per population. Shannon Information Index (I) ranged from 0.822 (SVO) to 1.133 (SHAT) population. The SVO population had the lowest values for allelic richness (A_R = 3.505) and the VAN2 population had the highest value (AR = 4.764). The expected heterozygosity (H_E) ranged from 0.433 (SVO) to 0.580 (SHAT). The inbreeding coefficient (F_{IS}) indicated an excess of homozygotes in BOR, VAN2, SUK1, NAR and KOS populations.

The hierarchical analysis of molecular variance (AMOVA) showed high molecular variance within populations (95%) and a low molecular variance among populations (5%). The Mantel test of isolation by distance yielded a significant correlation between genetic differentiation and geographical distance among populations (R^2 = 0.43, p = 0.01). A higher level of differentiation was found between two most distant populations (SHAT and SVO) (F_{ST} = 0.097, p = 0.001) and the number of migrants per generation (N_m) was estimated at 2.331.

The highest ΔK value of the Structure analysis was observed at K = 2; however, there was a second, lower ΔK peak at K = 5, indicating the presence of sub-clusters within the dataset (Figure 3.1A,B). The populations were first grouped according to K = 2, and then according to K = 5 (Figure 3.1C). For K = 2, one cluster clearly corresponds to the SHAT population (orange colour) and one to the 15 Siberian populations (blue colour). Individuals from the BOR population contained two genetic groups that were

highly admixed. The dataset contained sub-clusters that showed the most distant populations of SHAT, SVO, BAL, BOR, and ARO formed distinct clusters, while the remaining populations were highly admixed. For *K* = 5, SHAT, BOR, BAL and SVO populations were clearly separated from the remaining Scots pine populations (Figure 3.1D).

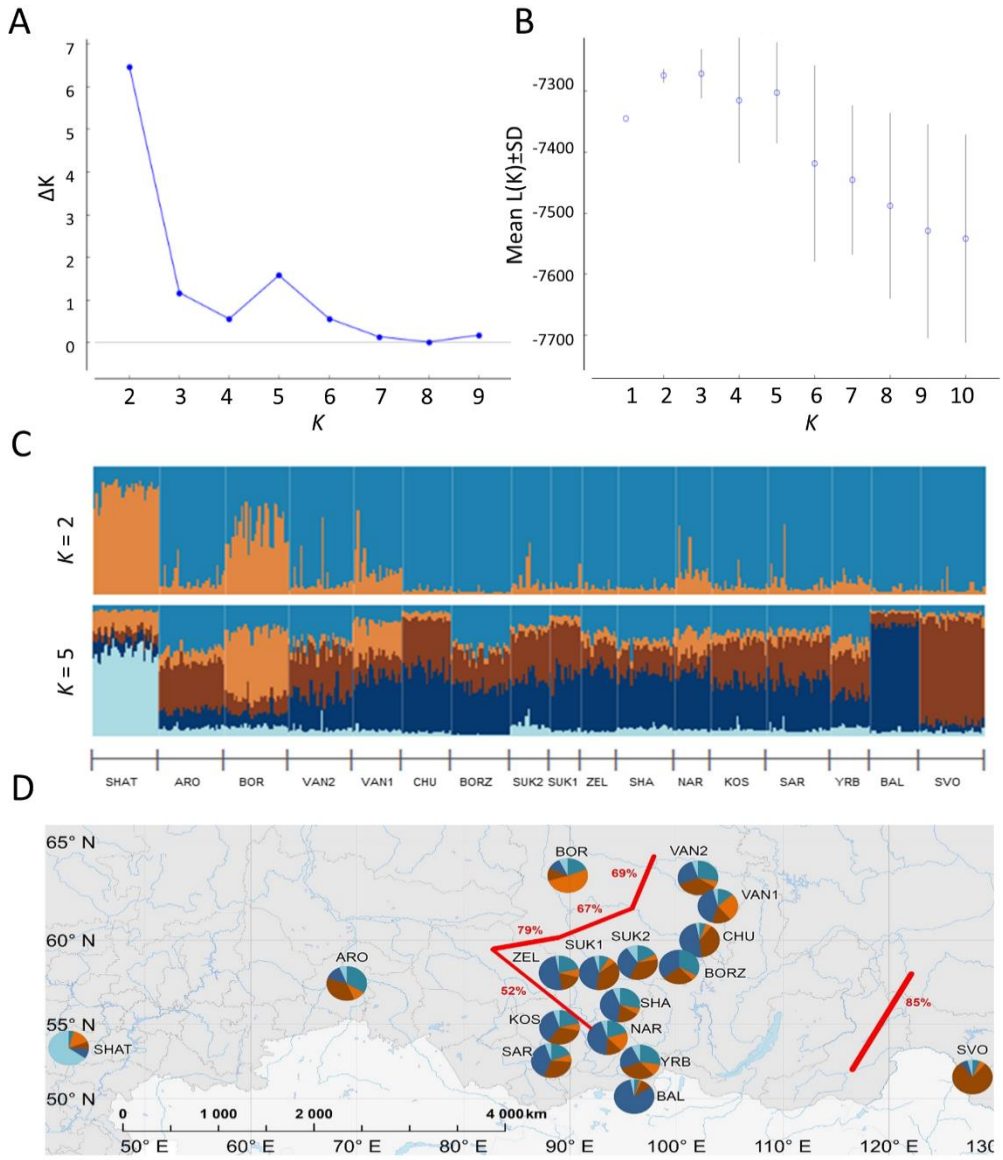


Figure 3.1. Results of population genetic structure analysis of 17 Scots pine populations. (A) Estimated population structure (*K* = 2 and *K* = 5). (B) Estimation of the best subpopulation numbers based on Δ*K* and Mean L(*K*)±SD values. (C) Barplot of genetic structure for *K* = 2 and *K* = 5. Light blue, blue, dark blue, orange and brown colors correspond to different clusters. (D) Geographic distribution of the five genetic clusters and the two barriers revealed by Barrier analysis (barriers are shown in red bold lines with bootstrap value).

Two barriers with bootstrap support between 50% and 90% were detected, using Barrier software (Figure 3.1D). The first barrier with the highest bootstrap value (85%) separates the easternmost population (SVO) from the remaining ones. The second barrier, with bootstrap support between 52% and 79%, delineates SHAT, ARO, BOR, KOS and SAR populations. All the other barriers between the populations were weak and showed a non-significant separation with <50% bootstrap support.

UPGMA analysis confirmed the STRUCTURE results for $K = 2$, separating the SHAT population from all the others (Figure 3.2A). It also showed similarities to $K = 5$ in which SHAT, BAL, SVO, BOR differentiated from the rest of the populations. Congruencies were also apparent on the heat map where the highest F_{ST} values were detected for those populations that were separated in the UPGMA (Figure 3.2B).

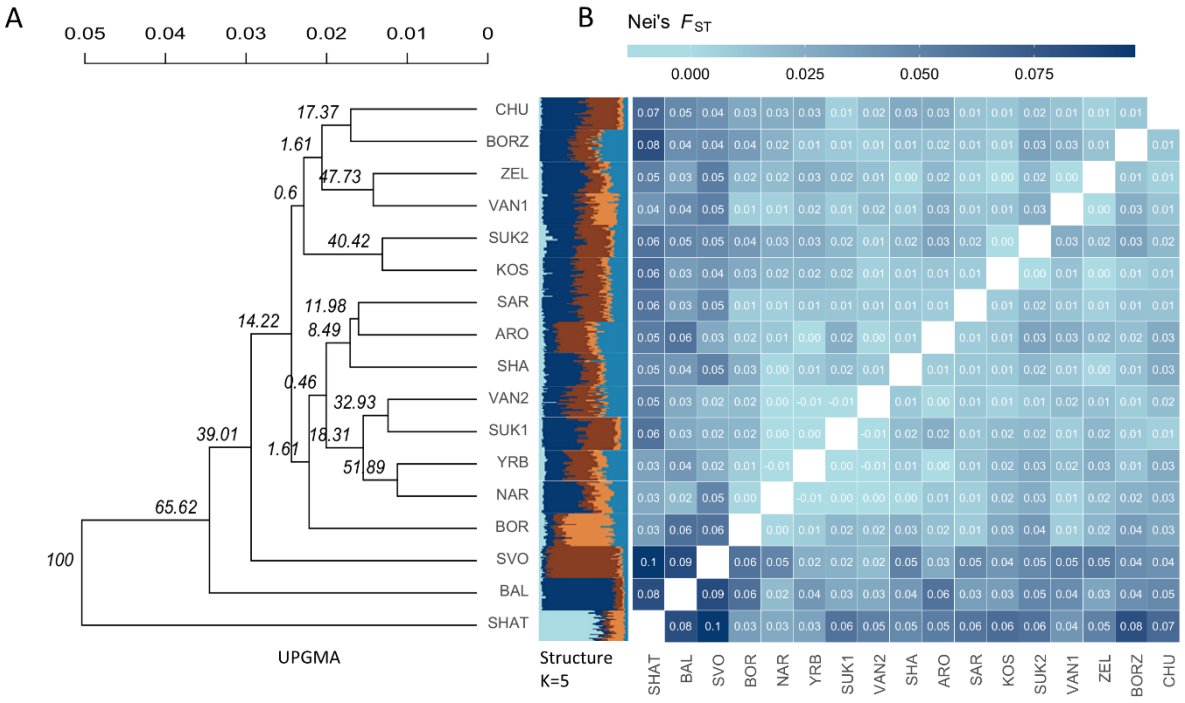


Figure 3.2. Bootstrapped UPGMA dendrogram (A), cluster assignments at $K = 5$ (light blue, blue, dark blue, orange and brown colors correspond to different clusters), and the heatmap of pairwise of F_{ST} between the 17 Scots pine populations investigated (the darker blue color means the higher level of genetic differentiation between populations) (B).

The first three axes of the Principal Coordinates Analysis (PCoA) accounted for 76.93% of the accumulated variability (Figure 3.3). The clusters distinguished by the PCoA analysis were found to be similar to the results of the Structure analysis at $K = 5$. At the first two axes, which explained 63.12%

of the total variation, the population from the European part of Russia (SHAT), as well as the easternmost (SVO) and southernmost (BAL) populations were separated. This pattern was confirmed by the second and third axes. The remaining populations from

the central part of the Siberian sampling range in the Asian part of Russia were not separated.

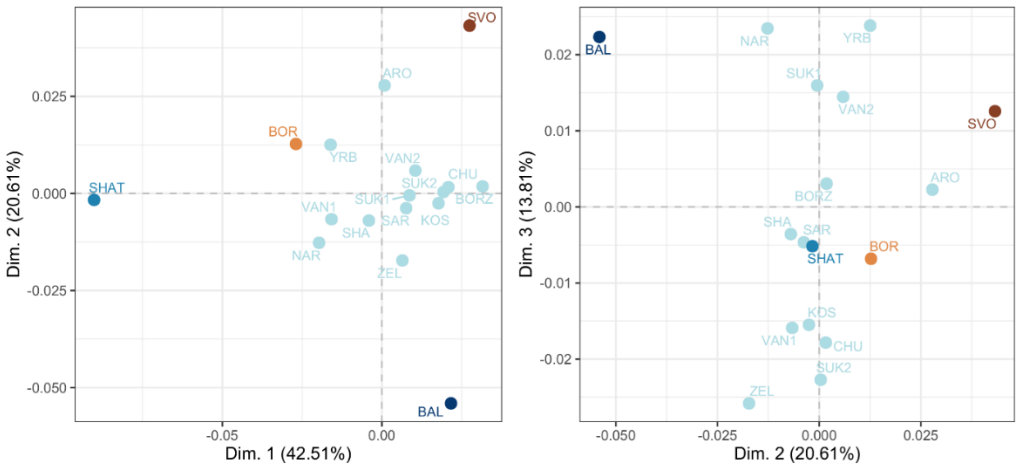


Figure 3.3. Two-dimensional plot of the three main principal components and their part of the total variance in % using Principal Coordinates Analysis (PCoA). The first and the second principal coordinates accounted for 63.12% of the total variance (A). The second and the third principal coordinates accounted for 34.42% of the total variance (B).

The Mantel test detected significant correlation between F_{ST} and geographical distances between the populations (R^2 = 0.429, p = 0.008), indicating that genetic differentiation among the populations significantly increases with geographic distance. However, the 2-D kernel density estimate indicated patches instead of a continuous cloud, suggesting genetic discontinuity. The SHAT population, as the most geographically distant population, showed discontinuity in genetic differentiation and spatial segregation.

3.4. Discussion

The genetic diversity and population structure of Scots pine populations from different parts of Russia were assessed based on the polymorphism of eight nuclear microsatellite markers. The highest genetic diversity was observed in SHAT population (H_E = 0.580) located in the European part of Russia while the lowest genetic diversity was detected in SVO population (H_E = 0.433) from the Russian Far East. Previous studies conducted on Scots pine based on mitochondrial and chloroplast DNA markers also

indicate a decrease in genetic diversity in Asian Scots pine populations compared to European ones (Semerikov et al. 2014; Semerikov et al. 2018). Semerikov et al. (2014) suggested that this fact might be the result of the loss of genetic variability due to multiple so-called "bottleneck" processes during the eastward distribution of the species.

AMOVA showed that only 5% of the total genetic variation occurred among populations, even if they are distributed at distances of up to 9656 km. Probably, free gene flow over large areas without any significant geographic barriers may have a homogenizing effect on the gene pool of Scots pine populations.

Using the software STRUCTURE, we found support for two and five genetic clusters. According to K = 2, all Siberian populations except for the BOR population were grouped into a single cluster while one European population (SHAT) was distinct. The presence of two genetic clusters in the studied Scots pine populations was also confirmed by UPGMA analysis. The clustering of studied Siberian populations into one group might indicate their common ancestry. However, based on the STRUCTURE results for K = 5, the studied Scots pine populations were separated into five distinct genetic groups, as follows: the European part of Russia (SHAT), the south of Middle Siberia (BAL), the northwest of Middle Siberia (BOR), West and Middle Siberia (13 populations) and the Russian Far East (SVO).

Our results indicate that despite the balanced diversity and substantial gene flow among Scots pine populations in Russia, there is evidence of genetic differentiation in the nuclear genomes. The main factor contributing to this is most likely the spatial segregation of populations due to the large geographic distances, which was confirmed by our significant Isolation-By-Distance hypothesis, as well as the inflated F_{ST} values of SHAT, BOR, BAL, SVO. This might be a possible mechanism shaping the present distribution of genetic diversity. It should also be considered that distant populations could have originated from different glacial refugia and, thus, correspond to different genetic lineages.

3.5. Conclusions

Our study showed that the Siberian populations of Scots pine harbor a large amount of genetic diversity, despite having a low level of genetic differentiation among its vast distribution area in the Eurasian part of its natural range. Genetic diversity found to decrease from west to east, and the

easternmost population has shown the lowest level of genetic diversity including allelic richness. Patterns of differentiation indicated separate genetic clusters of Scots pine in the European part of Russia, the south of Middle Siberia, the northwest of Middle Siberia, West and Middle Siberia and the Russian Far East, respectively. In these regions, besides the spatial segregation that indicated restrictions to gene flow, past demographic events also potentially affected neutral genetic variation. Our findings can be used in long-term monitoring of the state of Scots pine genetic resources in Russia and can provide guidance for future studies of population genetics. In addition, further studies encompassing more populations from these aforementioned regions could reveal the primary source of genetic variation and can provide insights into the adaptive genetic variation of Siberian Scots pine populations.

3.6. References

1. Doyle, J.J., Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus 12,13–15.
2. Elsik, C.G., Minihan, V.T., Hall, S.E., Scarpa, A.M., Williams, C.G. (2000) Low-copy microsatellite markers for Pinus taeda L. Genome 43, 550–555.
3. Fang, P., Niu, S., Yuan, H., Li, Z., Zhang, Y., Yuan, L. and Li, W. (2014). Development and characterization of 25 454 EST-SSR markers in *Pinus sylvestris* var. mongolica (*Pinaceae*). Appl. Plant Sci. 2, 455.
4. Graudal, L., Loo, J., Fady, B., Vendramin, G., Aravanopoulos, F.A., Baldinelli, G., Bennadji, Z., Ramamonjisoa, L., Changtragoon, S. and Kjær, E.D. (2020) Indicators of the genetic diversity of trees—State, Pressure, benefit and response. In State of the World's Forest Genetic Resources—Thematic Study; FAO: Rome, Italy, p. 92.
5. Ilyinov, A. A., Raevsky, B.V. (2016). The current state of *Pinus sylvestris* L. gene pool in Karelia. Sib. J. For. Sci. 5, 45–54.
6. Jakobsson, M., Rosenberg, N.A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.

7. Manni, F., Guerard, E., Heyer, E. (2004). Geographic Patterns of (Genetic, Morphologic, Linguistic) Variation: How Barriers Can Be Detected by Using Monmonier's Algorithm. Hum. Biol. 76, 173–190.
8. Markatyuk, A.A., Runova, E.M., Gavrilin, I.I., Vedernikov, I.B. (2013). Current state of East Siberian boreal forests in the aspect of natural *Pinus sylvestris* L. regeneration Systems. Methods. Technol. 1, 163–169.
9. Parliamentary newspaper. Available online: https://www.pnp.ru/ (accessed on 14 December 2022).
10. Peakall, R., Smouse, P.E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
11. Pritchard, J., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
12. Przybylski, P., Tereba, A., Meger, J., Szyp-Borowska, I., Tyburski, Ł. (2022). Conservation of Genetic Diversity of Scots Pine (*Pinus sylvestris* L.) in a Central European National Park Based on cpDNA Studies. Diversity 2022, 14, 93.
13. Radu, R.G., Curtu, L.A., Spârchez, G., Şofletea, N. (2014). Genetic diversity of Norway spruce [*Picea abie*s (L.) Karst.] in Romanian Carpathians. Ann. For. Res. 57, 19–29.
14. R Core Team (2013). R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria.
15. Sebastiani, F., Pinzauti, F., Kujala, S.T., González-Martínez, S.C., Vendramin, G.G. (2012). Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. Conserv Genet Resour 4,231–234.
16. Semerikov, V.L., Semerikova, S.A., Dymshakova, O.S., Zatsepina, K.G., Tarakanov, V.V., Tikhonova, I.V., Ekart, A.K., Vidyakin, A.I., Jamiyansuren, S., Rogovtsev, R.V., et al. (2014). Microsatellite loci polymorphism of chloroplast DNA of Scots pine (*Pinus sylvestris* L.) in Asia and Eastern Europe. Russ. J. Genet. 50, 577–585.

17. Semerikov, V., Semerikova, S., Putintseva, J., Tarakanov, V., Tikhonova, I., Vidyakin, A., Oreshkova, N., Krutovsky, K. (2018). Colonization history of Scots pine in Eastern Europe and North Asia based on mitochondrial DNA variation. Tree Genetics & Genomes. 14.

18. Sheller, M., Ciocîrlan, E., Mikhaylov, P., Kulakov, S., Kulakova, N., Ibe, A., Sukhikh, T., Curtu, A.L. (2021). Chloroplast DNA Diversity in Populations of *P. sylvestris* L. from Middle Siberia and the Romanian Carpathians. Forests. 12, 1757.

19. Șofletea, N., Mihai, G., Ciocîrlan, E., Curtu, A.L. (2020). Genetic Diversity and Spatial Genetic Structure in Isolated Scots Pine (*Pinus sylvestris* L.) Populations Native to Eastern and Southern Carpathians. Forests 11,1047.

20. Spielman, D., Brook, B.W., Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proc. Natl. Acad. Sci. 101, 15261–15264.

21. Torbik, D.N., Bedrickaya, T.V., Vlasova, M.M., Sinelnikov, I.G. (2019). Genetic diversity of natural populations of *Pinus sylvestris* L. In Proceedings of the Federal Budget Institution «Northern Research Institute of Forestry»; Demidova, N.A., Ed.; Arkhangelsk, Russia, pp. 91–99.

22. Van Oosterhout, C., Weetman, D., Hutchinson, W.F. (2006). Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Mol. Ecol. Notes 6, 255–256.

23. 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.

24. Zhou, Y., Bui, T., Auckland, L.D., Williams, C.G. (2002). Undermethylated DNA as a source of microsatellites from a conifer genome. Genome 45, 91–99.

Chapter 4

GENETIC LEGACY OF SOUTHERN MIDDLE SIBERIAN MOUNTAIN AND FOOTHILL POPULATIONS OF SCOTS PINE (*PINUS SYLVESTRIS* L.): DIVERSITY AND DIFFERENTIATION

The results presented in this chapter were published in the *Frontiers in Forests and Global Change* Journal with several improvements (Sheller, M.; Tóth, E.G.; Ciocîrlan, E.; Mikhaylov, P.; Tatarintsev A.; Kulakov, S.; Kulakova, N.; Melnichenko, N.; Ibe, A.; Sukhikh, T. and Curtu, A.L. Genetic legacy of southern Middle Siberian mountain and foothill populations of Scots pine (*Pinus sylvestris* L.): Diversity and differentiation. Front. For. Glob. Change 6:1152850. doi: 10.3389/ffgc.2023.1152850 Q1, Impact Factor – 3.2.)

4.1. Introduction

Genetic diversity is recognized as one of three basic elements of biodiversity (Hoban et al. 2020). It plays an important role in species adaption to changing climate, habitats, and biotic interactions (Spielman et al. 2004). The characterization of genetic diversity pattern within species and among populations is a fundamental requirement for the establishment of programs aimed at biodiversity preservation (Belletti et al. 2017).

Scots pine (*Pinus sylvestris* L.) is one of the most widespread conifers in the world, which has great economic and ecological importance (Floran et al. 2011). Within its vast distribution, Scots pine grows in various soil and climatic conditions. It forms over 20 geographical races and about 100 forms and varieties (Ekart et al. 2014). In the south of Siberia, pine forests are classified as especially valuable natural objects that stabilize environmental conditions. The maximum spectrum of adaptive variations for Scots pine is observed in this part of the region. Relict morpho- and genotypes of Scots pine can still be preserved in the south of Siberia (Pimenov 2015).

In this study, we used nuclear SSR markers to explore the genetic diversity and structure of Scots pine populations in southern Middle Siberia. Specifically, we aimed to: (a) assess the patterns of genetic diversity within and among populations, (b) test whether genetic differentiation is related to climatic variables.

4.2. Materials and methods

4.2.1 Plant material

Initially, 210 Scots pine individuals were sampled from eight populations (SAR-1, SAR-2, PER-1, PER-2, KER, TOD, BAL-1, BAL-2) located in southern Middle Siberia on a vast territory covering the mountain and foothill forests of the Western Sayan Mts., Kuznetsk Alatau Mts. and forest-steppe and steppe landscapes of the Minusinsk basin, Todzha basin and Central Tuva basin (Figure 1.5). Due to amplification failures, the number of studied individuals was reduced to 169.

4.2.2. DNA extraction and microsatellite analysis

Total genomic DNA was isolated according to the CTAB method (Doyle and Doyle, 1990). Seven nuclear microsatellite primers were selected: Psyl16, Psyl42, Psyl44, Psyl57 (Sebastiani et al. 2012); PtTX2146 (Elsik et al. 2000); lw_isotig04306, lw_isotig07383 (Fang et al. 2014). All primers were combined into two multiplex sets: set 1 consisted of Psyl44, Psyl57 and lw_isotig04306; set 2 comprised of Psyl16, Psyl42, PtTX2146 and lw_isotig07383. Reverse primers were labelled with a fluorescent dye. Polymerase chain reaction (PCR) was performed in a 10 μL reaction volume using Qiagen Multiplex PCR Kits (Qiagen, Germany) under the manufacturer`s instructions. The amplified fragments were run on a GenomeLab GeXP genetic analyzer (Beckman Coulter, Fullerton, CA) with an internal size standard. Genotyping was performed with the GenomeLab GeXP software (Version 10.2, Beckman Coulter, Fullerton, CA).

4.2.3. Data analysis

Parameters of genetic diversity and structure were calculated using the following programs: GenAlEx v. 6.5 software (Peakall & Smouse 2006), R (R Core Team 2013), Micro-Checker software (Van Oosterhout et al. 2006),

BOTTLENECK software v.1.2.02 (Piry et al. 1999), R (R Core Team 2013), STRUCTURE v.2.3.4 (Pritchard et al. 2000), SAMOVA software v.2.0 (Dupanloup et al. 2002).

To test whether climatic variation contributed to the patterns of genetic differentiation, i.e., to test the isolation-by-environment (IBE) hypothesis (Wang and Bradburd, 2014), three different approaches were taken. First, in addition to the genetic and geographic distances calculated for IBD, Euclidean climatic distances were calculated from recent (c. 1950–2000) climate data using 19 bioclimatic variables (Table S4.1) which were extracted from the global climate layer data using a grid size of 30 arc-seconds and downloaded from the WorldClim v.1.4 database.1 After, genetic, geographic, and the climatic distances were used in Mantel, partial-Mantel and MMRR (Multiple Matrix Regression with Randomization) regression analyses. The partial-Mantel test was conducted using the "vegan" package (Oksanen et al., 2022), while the MMRR carried out using the custom script of Wang (2013). The MMRR R script is deposited in the Dryad Data Repository under DOI:10.5061/dryad.kt71r.

4.3. Results

Using seven nuclear microsatellite loci, we identified a total of 49 alleles in the 169 individuals (Table S4.2). The mean number of alleles present per population varied from 3.571 (BAL-2) to 5.143 (PER-1 and SAR-1). Effective number of alleles (N_E) ranged between 2.191 in KER population and 2.910 in PER-2 population, with an average of 2.517 per population. Shannon Information Index (I) varied from 0.891 (BAL-1) to 1.118 (PER-2) population. The BAL-2 population had the lowest values for allelic richness (A_R = 3.571) and PER-1 population had the highest value (A_R = 4.764). The values of expected heterozygosity (H_E) ranged from 0.472 (BAL-1) to 0.565 (PER-2). The inbreeding coefficient (F_{IS}) values were between -0.161 and 0.229, but in general placed around zero in most of the populations. The value of the number of migrants per generation was high (N_m = 5.690) indicates high gene flow between populations. The ratio of observed and expected heterozygosity was balanced (mean H_O : H_E = 0.510 : 0.518).

The population stability analysis revealed no evidence for recent bottlenecks in the studied populations.

We performed AMOVA among and within Scots pine populations and the results showed that the genetic variation among populations was 7%, whereas most of genetic variation occurred within populations (93%, $p < 0.001$) (Table S4.3).

The matrix of pairwise F_{ST} values (Figure 4.1A) revealed that the highest differentiation apparent between KER, PER-1, PER-2 and SAR-1, SAR-2 and BAL-1 populations, while the lowest between KER, PER-1, PER-2 populations. Similarly, the UPGMA clustering indicated two groups. The first group consisted of PER-1, PER-2 and KER populations and the second group was composed of the five remaining Scots pine populations (Figure 4.1B).

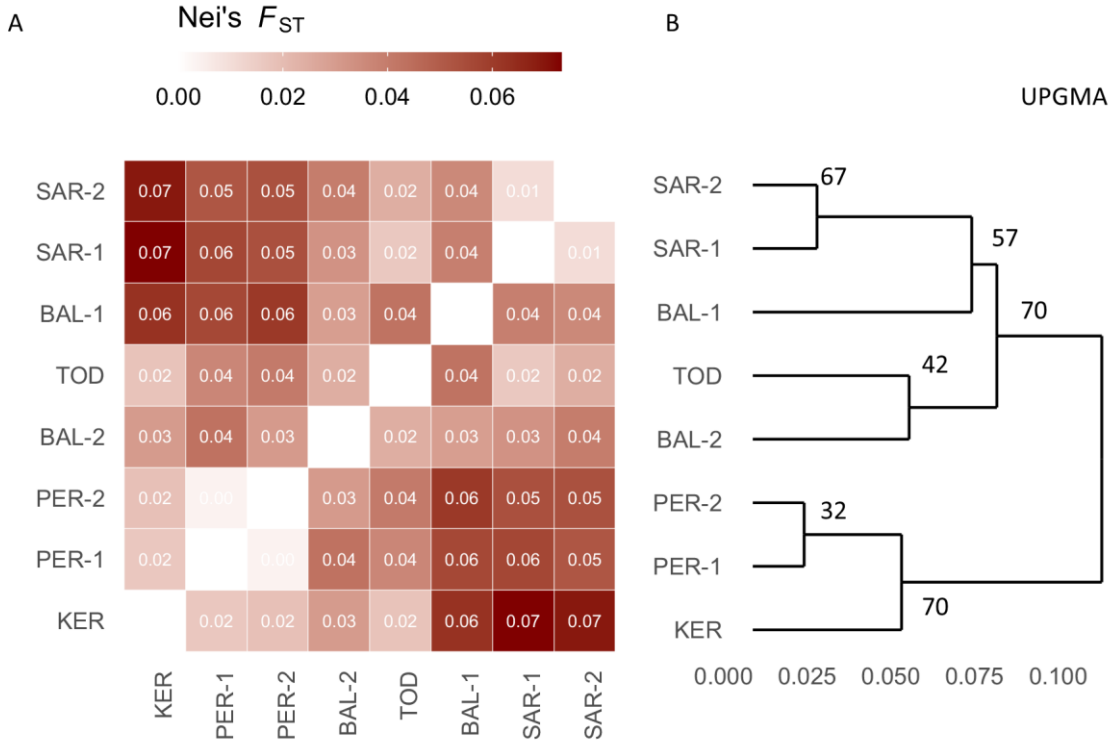


Figure 4.1. Heatmap of pairwise Fst (Nei, 1978) values and bootstrapped UPGMA tree of the eight Scots pine populations investigated.

The Bayesian STRUCTURE analysis revealed two Scots pine gene pools in the south of Middle Siberia (Figure 4.2), based on the Mean L(K) (±SD) and ΔK values. Group I included PER-1, PER-2 and KER populations, whereas Group II included all remaining populations.

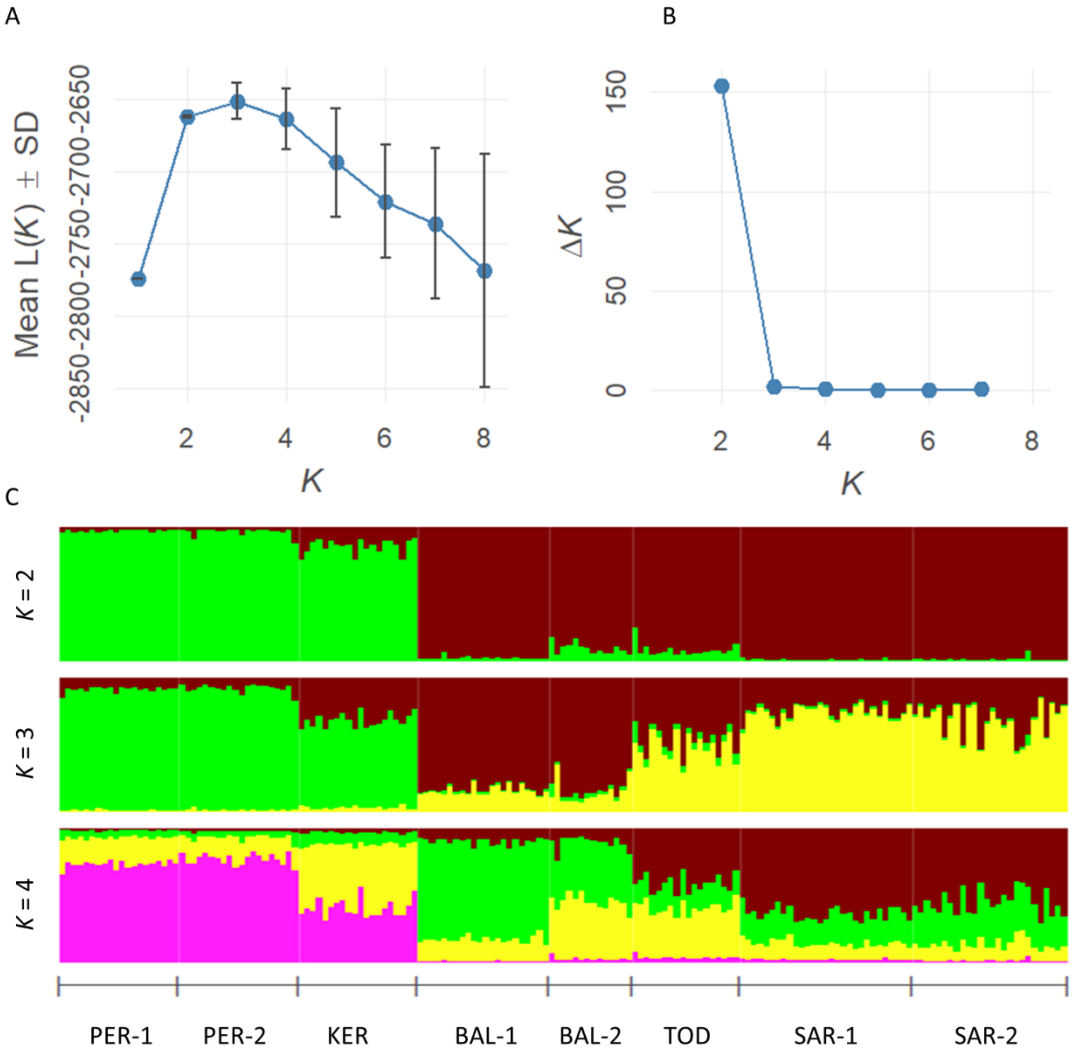


Figure 4.2. Estimation of the best subpopulation numbers based on Mean L(K) (±SD) and ΔK values (A) and (B). Genetic structural plot of eight Scots pine populations (C).

Principal Coordinates Analysis (PCoA) based on the F_{ST} values identified two major groups (Figure S4.1) at Dim. 1 vs. Dim. 2, explaining jointly 56.99% of the total variation for the nSSR markers. One group included populations from the Minusinsk basin (PER-1 and PER-2) and the Western Sayan Mts. (KER). The second group contained five populations from the Kuznetsk Alatau Mts. (SAR-1 and SAR-2), the Central Tuva basin (BAL-1 and BAL-2) and the Todzha Basin (TOD).

Genetic Landscape Spatial Interpolation has detected a significant barrier to gene flow in the form of a genetic discontinuity in the contact zone between the genetic lineages in this region (Figure 4.3).

The estimation of the contribution of genotypes in each population showed that the PER-1, PER-2 and KER populations contained a higher proportion of genotypes originated from the Minusinsk basin, compared to other samples from the Kuznetsk Alatau Mts. (SAR-1 and SAR-2), the Central Tuva basin (BAL-1 and BAL-2) and the Todzha basin (TOD).

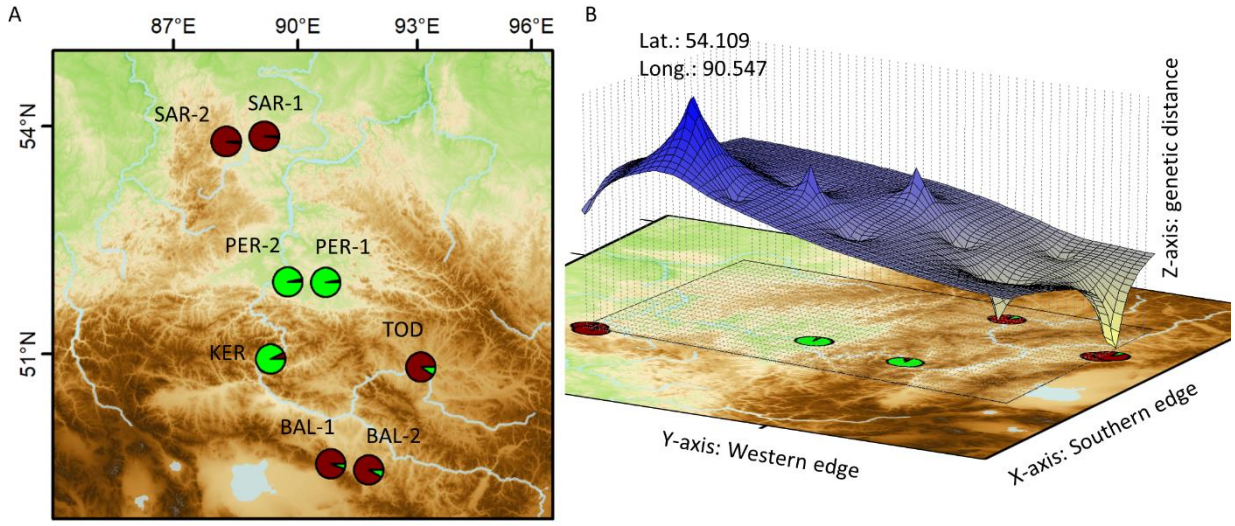


Figure 4.3. Spatial extent of the detected genetic clusters (A) and the genetic discontinuity revealed by the Genetic Landscape Shape Interpolation analysis (B).

Spatial analysis of molecular variance (SAMOVA) produced values of F_{CT} ranging from 0.0305 (K=3) to a maximal value of 0.03336 (K=6), which indicated number 6 to be the preferred number of genetically homogenous clusters for the whole dataset (Table S4.4). The six defined clusters contained the following populations: (I) PER-1 and PER-2; (II) BAL-1; (III) BAL-2; (IV) SAR-1 and SAR-2; (V) KER; (VI) TOD.

We further analyzed the correlation between genetic distance and geographic distance for the studied populations using the Mantel test. The results showed no correlation between genetic differentiation and geographical distance among Scots pine populations (R^2 = 0.043, p = 0.163). In addition, none of the matrix regression approaches (Mantel, partial-Mantel and MMRR) to investigate IBE, were able to find significant relationships between the genetic, geographic and climatic distances (Table S4.5).

4.4. Discussion

Our results showed that, despite the detectable effects of fragmentation, the level of genetic diversity in Scots pine populations in southern Middle Siberia is high (H_E = 0.518) and is similar to the other Scots pine populations in Middle Siberia (H_E = 0.514) (Sheller et al., 2023). The highest level of genetic diversity and allelic richness were detected in PER-2 and PER-1 populations (H_E = 0.565 and A_R=4.764 respectively) located in the Minusinsk basin. PER-1 and PER-2 populations represent geographically isolated pine forests, which are unique in their origin and ecological functions (Tatarintsev et al., 2015, Polyakova, 2008). In the study by Ekart et al. 2014 the highest values of heterozygosity (H_O and H_E) were also found in Scots pine population from the Minusinsk basin. Towards the south the pattern of genetic diversity changes, and considerably lower levels of genetic diversity were found. Expected heterozygosity and allelic richness in two southernmost populations, BAL-1 and BAL-2, were H_E = 0.472 and A_R=3.571, respectively. These populations are located in the Central Tuva basin and belong to Balgazyn relict pine forest. It is possible that the adverse effects of the extremely continental climate of Central Asia and the anthropogenic factors in the area of the Balgazyn pine forest, which is steadily shrinking, contributed to the decreased diversity.

Our estimate on total genetic variation occurring among populations showed a 7% applying AMOVA, which is considerably high, by taking into account the relatively short geographical distances among some of the populations (from 3 km to 593 km). In similar microsatellite studies, but on a larger geographical scales, Shuvaev et al. (2022) found a F_{ST} value of only 0.026 in Scots pine populations in Krasnoyarsk region (Middle Siberia) and Sheller et al. (2023) found a F_{ST} value of 0.097 among distant Scots pine populations in Russia. Based on allozyme analysis, Sannikov and Petrova (2012) revealed that the genetic differentiation of Scots pine populations in the southern part of the range (south of 52° - 53° latitude) in Central and Eastern Siberia is 2-4 times higher than in the contiguous forest zone. By decomposing our F_{ST} value among populations, based on pairwise estimates, majority of differentiation was found between populations located in the Kuznetsk Alatau Mts. (SAR-1 and SAR-2), the Minusinsk basin (PER-1 and PER-2), and the Western Sayan (KER).

Our Bayesian and the spatial clustering approaches (STRUCTURE and SAMOVA) were consistent with the F_{ST} estimates. The STRUCTURE analysis divided the studied populations into two main groups (K =

2), with PER-1, PER-2 and KER forming one group and the remaining five populations forming another group. The Minusinsk basin and the Western Sayan populations (PER-1, PER-2 and KER) wedged among the members of the other genetic group. However, towards the south they showed mixing, albeit at an extremely low proportion. This was evident also on our pairwise F_{ST} estimates, because lower F_{ST} values were typical here than towards the north, in the direction of the West Siberian Plain. We assume that this genetic differentiation across the landscape provided evidence of a contact zone of distinct genetic lineages, and a sharp boundary limiting gene flow, of a wedged population group.
It should be noted, as an interesting fact, that SAMOVA indicated six different groups, and dissected the population groups in the same order as F_{ST} decreases. First separated the Minusinsk basin and the Western Sayan populations (PER-1, PER-2 and KER), from the remaining ones, and so on. This shows that the primary barrier to restriction of gene flow is located here. Our Genetic Landscape Spatial Interpolation concurs with this, having identified a significant barrier to gene flow in the form of a genetic discontinuity in the contact zone between the genetic lineages in this region.

4.5. Conclusions

In our study, we assessed the genetic diversity and population structure of eight Scots pine populations in southern Middle Siberia using seven nuclear SSR markers. The study has revealed genetic heterogeneity of Scots pine populations near the southern boundaries of the species distribution in Middle Siberia. Despite fragmentation, the studied populations preserved high genetic diversity. The highest level of genetic diversity and allelic richness was detected in two populations from the Minusinsk basin while the lowest level of genetic diversity and allelic richness was found in two southernmost populations, which belong to a relict Balgazyn pine forest in the Central Tuva basin. However, to confirm a relictary status of Balgazyn pine forest additional studies should be carried out. Two clustering methods showed that the Minusinsk basin and the Western Sayan populations formed a distinct genetic group. The pattern of genetic diversity suggests a different origin of the studied Scots pine populations. However, further investigation is needed to study the evolutionary history of Scots pine populations in southern Siberia.

4.6. References

1. Belletti, P., Ferrazzini, D., Ducci, F., De Rogatis, A., Mucciarelli, M. (2017). Genetic diversity of Italian populations of *Abies alba*. Dendrobiology. 77, 147-159.
2. Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. Focus. 12, 13-15.
3. Dupanloup, I., Schneider, S., Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. Mol Ecol. 11, 2571-2581.
4. Ekart, A. K., Larionova, A. Y., Kravchenko, A. N., Tikhonova, I. V., Zatsepina, K. G., Tarakanov, V.V., Zhamyansuren, S. (2014). Genetic diversity and differentiation of Scots pine populations in Southern Siberia and Mongolia. Contemp. Probl. Ecol. 7(1), 52-59.
5. Elsik, C. G., Minihan, V. T., Hall, S. E., Scarpa, A. M., Williams, C. G. (2000). Low-copy microsatellite markers for *Pinus taeda* L. Genome. 43, 550-555.
6. Fang, P., Niu, S., Yuan, H., Li, Z., Zhang, Y., Yuan, L. and Li, W. (2014). Development and characterization of 25 454 EST-SSR markers in *Pinus sylvestris* var. mongolica (*Pinaceae*). Appl. Plant Sci. 2, 455.
7. Floran, V., Sestras, R. E., Garcia-Gil, M. R. (2011). Organelle Genetic Diversity and Phylogeography of Scots Pine (*Pinus sylvestris* L.). Not. Bot. Horti Agrobot. Cluj-Napoca 39, 317-322.
8. Hoban, S., Bruford, M., D'Urban, Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P. A., Paz-Vinas, I., Sjögren-Gulve, P., Segelbacher, G., Vernesi, C., et al. (2020). Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biol. Conserv. 248, 108654.
9. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2022). Package 'vegan': community ecology package.
10. Peakall, R. and Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. 6, 288-295.
11. Pimenov, A.V. (2015). Biodiversity of the Scots pine (*Pinus sylvestris* L.) in contrasting ecotopes of the south of Siberia. [dissertation/Doctoral thesis]. [Krasnoyarsk]: Sukachev Forest Institute of the Siberian Branch of the Russian Academy of Sciences.

12. Piry, S., Luikart, G., Cornuet, J. M. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 90, 502-503.
13. Polyakova, M. A. (2008). Vegetation of ribbon forests of the Minusinsk basin. [dissertation/PhD thesis]. [Novosibirsk]: Central Botanical Garden SB RAS.
14. Pritchard, J., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics. 155, 945-959.
15. R Core Team. (2013). R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria.
16. Sannikov, S. N., Petrova, I. V. (2012). Phylogenogeography and genotaxonomy of *Pinus sylvestris* L. populations. Ecology. 4, 252-260.
17. Sebastiani, F., Pinzauti. F., Kujala, S. T., González-Martínez, S. C., Vendramin, G. G. (2012). Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. Conserv. Genet. Resour. 4, 231-234.
18. Sheller, M., Tóth, E. G., Ciocîrlan, E., Mikhaylov, P., Kulakov, S., Kulakova, N., Melnichenko, N., Ibe, A., Sukhikh, T., Curtu, A. L. (2023). Genetic Diversity and Population Structure of Scots Pine (*Pinus sylvestris* L.) in Middle Siberia. Forests. 14, 119.
19. Shuvaev, D. N., Ibe, A. A., Shcherba, Yu. E., Sukhikh, T. V., Shilkina, E. A., Sheller, M. A., Usova, E. A., Lisotova, E. V., Repyakh, M. V., Stupakova, O. M. (2022). Molecular genetic characteristics of Scots pine populations in Krasnoyarsk region by panel of 15 nuclear microsatellite loci. Forestry. 5, 530-539.
20. Spielman, D., Brook, B. W., Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proc. Natl. Acad. Sci. 101, 15261-15264.
21. Tatarintsev, A. I., Kalenskaya, O. P., Bublikov, A. G. (2015). On the issue of root rot infestation of pine forests of the Minusinsk basin. Conifers of the Boreal zone, XXXIII 5(6), 240-247.
22. Van Oosterhout, C., Weetman, D., Hutchinson, W. F. (2006). Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Mol. Ecol. Notes. 6, 255–256.
23. Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution

67, 3403–3411.Wang, I. J., and Bradburd, G. S. (2014). Isolation by environment. Mol. Ecol. 23, 5649–5662.

4.7. Supplementary material

Table S4.1. Bioclimatic variables downloaded from WorldClim (http://www.worldclim.org/) and used in MMRR (Multiple Matrix Regression with Randomization) analysis.

Variable	Description
Bio 1	Annual Mean Temperature
Bio 2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
Bio 3	Isothermality (Bio 2/Bio 7) (* 100)
Bio 4	Temperature Seasonality (standard deviation *100)
Bio 5	Max Temperature of Warmest Month
Bio 6	Min Temperature of Coldest Month
Bio 7	Temperature Annual Range (BIO5-BIO6)
Bio 8	Mean Temperature of Wettest Quarter
Bio 9	Mean Temperature of Driest Quarter
Bio 10	Mean Temperature of Warmest Quarter
Bio 11	Mean Temperature of Coldest Quarter
Bio 12	Annual Precipitation
Bio 13	Precipitation of Wettest Month
Bio 14	Precipitation of Driest Month
Bio 15	Precipitation Seasonality (Coefficient of Variation)
Bio 16	Precipitation of Wettest Quarter
Bio 17	Precipitation of Driest Quarter
Bio 18	Precipitation of Warmest Quarter
Bio 19	Precipitation of Coldest Quarter

Table S4.2. Diversity statistics of the seven nuclear SSR loci across 169 Scots pine individuals.

Locus	N_A	F_{IS}	F_{IT}	F_{ST}
Psyl44	3	0.249	0.281	0.043
Psyl42	8	-0.055	0.063	0.111
Psyl57	5	0.095	0.121	0.029
Psyl16	7	0.013	0.054	0.042

lw_isotig04306	9	0.006	0.049	0.044
lw_isotig07383	5	0.058	0.084	0.028
PtTX2146	12	-0.004	0.056	0.060
Mean	7	0.052	0.101	0.051
SD	1.225	0.037	0.031	0.011

Note: number of alleles (N_A); inbreeding coefficient of an individual relative to the subpopulation (F_{IS}); inbreeding coefficient of an individual relative to the total population (F_{IT}); genetic differentiation coefficient (F_{ST}); ± standard deviation (SD).

Table S4.3. Hierarchical AMOVA of Scots pine populations.

Source	df	SS	MS	Est. Var.	%
Among Pops	7	72.238	10.320	0.308	7***
Within Pops	161	621.596	3.861	3.861	93
Total	168	693.834		4.168	100

Note: (df) degrees of freedom; (SS) sum of squares; (MS) mean of the squares; (Est. Var.) estimated variance of components; (%) percentage of total variance contributed by each component; *** = p <0.001

Table S4.4. Population groups identified by SAMOVA for eight Scots pine populations

Number of groups (K)	F_{CT}	Population grouping
K=2	0.03114	{PER-1/PER-2/KER} {BAL-1/BAL-2/SAR-1/SAR-2/TOD}
K=3	0.0305	{PER-1/PER-2/KER} {BAL-1/BAL-2} {SAR-1/SAR-2/TOD}
K=4	0.03132	{PER-1/PER-2/KER} {BAL-1} {BAL-2} {SAR-1/SAR-2/TOD}
K=5	0.03215	{PER-1/PER-2} {BAL-1} {BAL-2} {SAR-1/SAR-2} {KER/TOD}
K=6	0.03336	{PER-1/PER-2} {BAL-1} {BAL-2} {SAR-1/SAR-2} {KER} {TOD}
K=7	0.03216	{PER-1/PER-2} {BAL-1} {BAL-2} {SAR-1} {SAR-2} {KER} {TOD}
K=8	-	-

Note: F_{CT}: index of pairwise genetic differentiation

Table S4.5. Results of standard Mantel tests, partial-Mantel tests and Multiple Matrix Regression with Randomization (MMRR) analyses.

Test	Parameters		R	β	p
Mantel	Gen vs. Geo		0.217	-	0.131^{ns}
	Gen vs. Clim		0.267	-	0.091^{ns}
partial-Mantel	Gen vs. Geo (Clim)		-0.033	-	0.543^{ns}
	Gen vs. Clim (Geo)		0.163	-	0.203^{ns}
MMRR	Gen vs. Geo + Clim		0.268	Geo: -5.691	0.870^{ns}
				Clim: 4.178	0.377^{ns}

*significance calculated with 999 permutations: ns; not significant

Gen, genetic distance (F_{ST}); Geo, geographic distance; Clim, climatic distance

partial-Mantel tests: X ~ Y(Z) is the correlation between X and Y matrices, controlling for Z

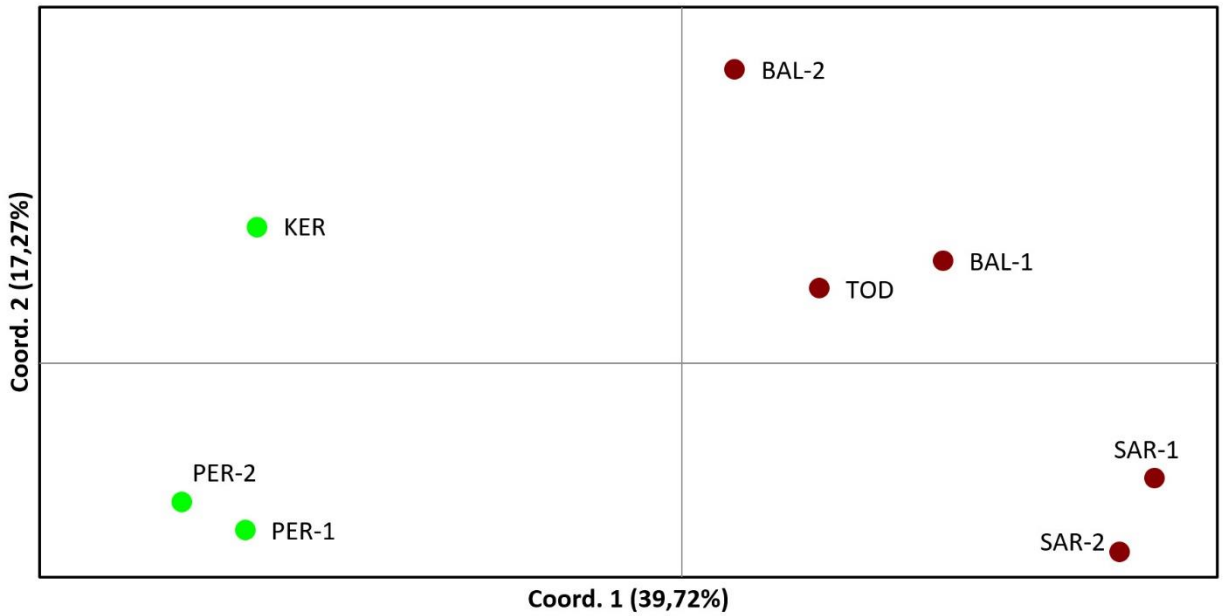


Figure S4.1. Principal Coordinates Analysis (PCoA) and their part of the total variance in percentages (%), for Dim 1. vs. Dim. 2., for the eight analyzed populations.

Chapter 5. CONCLUSIONS. ORIGINAL CONTRIBUTIONS. DISSEMINATION OF RESULTS

5.1. Conclusions

The presented PhD thesis focuses on the ecologically and economically important tree species, Scots pine (*Pinus sylvestris* L.), one the main forest forming tree species in Middle Siberia. The aim of this research was to assess the genetic diversity of natural Scots pine populations in Middle Siberia at chloroplast and nuclear levels. Our results showed high levels of genetic diversity across the Middle Siberian populations of Scots pine for both nuclear and chloroplast microsatellite markers. Our main findings are outlined below.

1. The study of cpDNA diversity of 19 natural populations of Scots pine in Middle Siberia and the Romanian Carpathians showed high levels of genetic diversity in all of the populations studied. Three common haplotypes were found between the Carpathian and the Siberian populations of Scots pine, which can be explained by very efficient long-distance gene flow or common ancestry. A weak geographic population structure in Scots pine from Middle Siberia and the Romanian Carpathians was detected.

2. The nSSR polymorphism of Scots pine populations across Russia showed high values of genetic diversity and allelic richness. Genetic diversity found to decrease from west to east, and the easternmost population has shown the lowest level of genetic diversity including allelic richness. Five genetic groups could be detected that correspond to: the European part of Russia, the south of Middle Siberia, the northwest of Middle Siberia, West and Middle Siberia, and the Russian Far East. However, the European population was the most genetically distinct one.

3. The study of eight Scots pine populations in southern Middle Siberia showed a high level of genetic diversity and occurrence of two genetic clusters in the studied populations. Two populations from the Minusink basin and a population from the Western Sayan Mts. Formed the first group and the second group was composed of the other populations from Kuznetsk Alatau Mts., Central Tuva basin and

Todzha basin. Our findings suggest that the studied Scots pine populations originate from different gene pools.

5.2 Original contributions

This study is one of the first SSR marker based studies on Scots pine (*Pinus sylvestris* L.) population genetic in Middle Siberia at the cpDNA and nDNA genome level. It focuses on Scots pine populations originating from various ecosystems and geographic regions and compares genetic diversity patterns across different parts of Scots pine distribution range. The data on the genetic diversity and differentiation of Scots pine populations in Middle Siberia can be used in the development of conservation programs of Scots pine genetic resources and further population genetic studies. Furthermore, the data obtained can be integrated in an existing reference database of Scots pine genetic resources for the identification of geographical origin of the species.

5.3. Dissemination of results

The results of the research carried out within this doctoral thesis were disseminated through three scientific articles and presented at four international conferences.

Scientific publications based on the material of the thesis:

1. **Sheller**, M.; Ciocîrlan, E.; Mikhaylov, P.; Kulakov, S.; Kulakova, N.; Ibe, A.; Sukhikh, T.; Curtu, A.L. Chloroplast DNA Diversity in Populations of *P. sylvestris* L. from Middle Siberia and the Romanian Carpathians. Forests 2021, 12, 1757. https://doi.org/10.3390/f12121757
2. **Sheller**, M.; Tóth, E.G.; Ciocîrlan, E.; Mikhaylov, P.; Kulakov, S.; Kulakova, N.; Melnichenko, N.; Ibe, A.; Sukhikh, T.; Curtu, A.L. Genetic Diversity and Population Structure of Scots Pine (*Pinus sylvestris* L.) in Middle Siberia. Forests 2023, 14, 119. https://doi.org/10.3390/f14010119
3. **Sheller**, M.; Tóth, E.G.; Ciocîrlan, E.; Mikhaylov, P.; Tatarintsev A.; Kulakov, S.; Kulakova, N.; Melnichenko, N.; Ibe, A.; Sukhikh, T. and Curtu, A.L. Genetic legacy of southern Middle Siberian

mountain and foothill populations of Scots pine (*Pinus sylvestris* L.): Diversity and differentiation. Front. For. Glob. Change 6:1152850. doi: 10.3389/ffgc.2023.1152850

Participation in international scientific conferences:

1. **Sheller** M., Ciocîrlan E., Mikhaylov P., Kulakov S., Kulakova N., Ibe A., Sukhikh T., Curtu A. L. Genetic diversity of Scots pine (*Pinus sylvestris* L.) populations from Southern mountain forest zone in Middle Siberia. 7th International Meeting "Conservation of Forest Genetic Resources" 20-22 September 2022, Pushkino, Russia.
2. **Sheller** M., Ciocîrlan E., Mikhaylov P., Kulakov S., Kulakova N., Ibe A., Sukhikh T., Curtu A. L. Genetic diversity of *Pinus sylvestris* L. populations in Middle Siberia by Chloroplast Microsatellite Markers. Proceedings of the International Scientific and Practical Conference "Conservation and Rational Use of Biological Resources in the System of Sustainable Forest Management". 27-29 September 2022, Gomel, Republic of Belarus.
3. **Sheller** M., Ciocîrlan E., Mikhaylov P., Kulakov S., Kulakova N., Ibe A., Sukhikh T., Curtu A. L. Genetic diversity of Siberian Scots pine (*Pinus sylvestris* L.) populations. International Symposium "Forest and Sustainable Development". 14-15 October 2022, Brasov, Romania.
4. **Sheller** M., Tóth E.G., Ciocîrlan E., Mikhaylov P., Tatarintsev A., Kulakov S., Kulakova N., Melnichenko N., Ibe A., Sukhikh T. and Curtu A.L. Genetic diversity of Scots pine (*Pinus sylvestris* L.) populations in the south of Middle Siberia. VII International Scientific Conference "Plant Genetics, Genomics, Bioinformatics and Biotechnology (PlantGen 2023)". 10-15 July 2023, Kazan, Republic of Tatarstan, Russia.

SHORT SUMMARY

Genetic diversity plays an important role in species adaption to changing climate, habitats, and biotic interactions. It helps to maintain ecosystem functions, stability and services. The characterization of genetic diversity pattern within species and among populations is a fundamental requirement for the establishment of programs aimed at biodiversity preservation. In this PhD thesis we address the issue of studing the genetic diversity in natural Scots pine (*Pinus sylvestris* L.) populations in Middle Siberia. Scots pine is one of the main forest-forming tree species in Middle Siberia and is of great ecological and economic importance. We assessed neutral genetic variation of Scots pine at the cpSSR and nSSR loci. The study of cpDNA diversity of nineteen natural populations of Scots pine in Middle Siberia and the Romanian Carpathians showed high levels of genetic diversity in all the populations studied. Three common haplotypes were found among the Siberian and the Carpathian populations, which can be explained by long-distance gene flow or common ancestry. The nSSR polymorphism of Scots pine populations across Russia showed high values of genetic diversity and allelic richness. Patterns of differentiation indicated separate genetic clusters of Scots pine in the European part of Russia, the south of Middle Siberia, the northwest of Middle Siberia, West and Middle Siberia and the Russian Far East. The study of eight Scots pine populations in southern Middle Siberia revealed the occurrence of of two genetic clusters. Two populations from the Minusink basin and a population from the Western Sayan Mts. formed the first cluster and the second cluster was composed of the other populations from Kuznetsk Alatau Mts., Central Tuva basin and Todzha basin. Our findings suggest that the studied Scots pine populations originate from different gene pools. The patterns of genetic diversity revealed by our study may be useful for futher studies of Scots pine in Siberia and for the elaboration of conservation measures of genetic resources of Scots pine in Middle Siberia.