SYLWAN 165 (7): 509-518, July 2021
https://doi.org/10.26202/sylwan.2021046
Journal homepage: https://sylwan.lasy.gov.pl

ORIGINAL PAPER

A set of 20 single nucleotide polymorphism markers as a tool in genetic diversity study of Scots pine

Anna Tereba[(1)]✉, Agata Konecka[(2)], Kateryna Fyałkowska[(1)]

[(1)] Department of Forest Ecology, Forest Research Institute, 3 Braci Leśnej St., 05-090 Sękocin Stary, Poland

[(2)] Department of Silviculture, Institute of Forest Sciences, Warsaw University of Life Sciences, ul. Nowoursynowska 159, 02-776 Warszawa, Poland

ABSTRACT

Molecular techniques involved in the analysis of species population variability are based mainly on the analysis of microsatellite markers. However, ongoing development of genetic techniques increasingly allows this type of research to be carried out using variation in genome SNPs, mainly using NGS technology. This significantly expands the data pool on variability of genes responsible for a number of features, e.g. population variability related to adaptation to changing climatic conditions. In this study, 20 SNP (single nucleotide polymorphism) were selected from various genes related to, for example, abscisic acid responsive protein or dehydrins. Polymorphism of these markers was analyzed with traditional Sanger sequencing methods for 96 adult trees from populations in central Poland. Haplotype variation, h, was high (0.9902) and the average number of nucleotide differences between haplotypes was 10.625. The mean observed heterozygosity (Ho) was 0.311 and the mean expected heterozygosity (He) was equal to 0.360. The obtained results confirm the usefulness of the proposed non-standardised method of analysis of genetic variability as an alternative to other common techniques.

KEY WORDS

molecular polymorphism, *Pinus sylvestris* L., SNPs analysis

Introduction

Scots pine (*Pinus sylvestris* L.) is one of the dominant species in Poland and one of the main forest tree species in northern and central Europe. It occurs in forests where it is the dominant species, but also is often found in mixed species stands. It is a species of great economic importance, with commercial forests often grown in monocultures. Scots pine is highly adaptable to changing environmental conditions. A number of ecotypes have been characterized and it possesses strong clinal variability in, among other traits, resistance to low temperature and drought (Wachowiak, 2015). Formation of ecotypes is connected with development of different phenotypic characteristics (morphological, physiological and ecological) that influence Scots pine variability observed in the environment (Oleksyn *et al*., 1999; Barzdajn *et al*., 2016). For example, the research by Rehfeldt *et al*. (2002) showed that seedlings growing in the northern areas have higher wood productivity compared to southern populations and are characterized by higher survival rates. It

✉Tel. +48 22 7150318, e-mail: A.Tereba@ibles.waw.pl

Received: 5 May 2021; Revised: 21 July 2021; Accepted: 28 July 2021; Available online: 5 November 2021

has also been observed that in the oceanic climate the trees have thick branches and flexible trunks, while in the colder zones have thin branches and straight trunks (Govindarajulu, 2014). Tree-ring formation of Scots pine is conditioned by the environmental factors. Trees from higher latitudes maintain the orientation for accelerated growth at the beginning of a season, which is followed by a rapid transition to latewood formation and slowdown in growth earlier, than in mid-latitude trees (Savva and Vaganov, 2006). What is more, summer precipitation was shown to be the main factor in controlling ring-width (Misi *et al*., 2019).

Molecular techniques based on DNA polymorphism have been used for more than 20 years to study genetic diversity, ecology and adaptation processes in forest trees. Many researchers have used genetic data to deduce features of population genetics that could not be gained from traditional ecological studies. For example, genetic markers can be used to estimate parameters of genetic similarity/diversity, inbreeding, past geographical processes (e.g., bottlenecks) or testing the correlation of genetic isolation with geographical distance (Gugerli *et al*., 2001; Dzialuk *et al*., 2014; Antonecchia *et al*., 2015). Until recently, these types of studies were limited to standard use of microsatellite panels to identify polymorphic genetic markers. For studies of molecular diversity, an alternative approach appeared in the form of SNP panels, instead of microsatellite markers (Santure *et al*., 2010; Seifert *et al*., 2012; Plomion *et al*., 2016). However, the most common use of SNP markers has been to analyze genotype – phenotype relationships, by examining single nucleotide polymorphisms in relation to genes for adaptation, for example: the dehydrin family of genes (Wachowiak *et al*., 2009) and genes related to wood formation and quality (González-Martínez *et al*., 2007; Tian *et al*., 2012). Single nucleotide polymorphisms can be used in population and ecological studies with great success (Hess *et al*., 2011; Stolting *et al*., 2013).

Single-nucleotide polymorphism analysis may, in the near future, provide new directions in population studies due to several important attributes. Among others it is worth mentioning the high frequency of SNPs in the genome, with fairly well described mutation mechanisms and low error rates (Weller *et al*., 2006). Additionally, obtaining a large number of SNPs does not pose great difficulty due to their occurrence in the entire genome, whereas the number of microsatellite markers may be limited depending on the species. What is more, SNP markers show high stability and the facility of standardization (Werner *et al*., 2004). Studies comparing microsatellite markers and SNP suggest that microsatellite markers are much informative in assignment tests, while SNP can be used in admixture analysis and could be a good substitute for microsatellite markers (Liu *et al*., 2005; Narum *et al*., 2008; Hauser *et al*., 2011; Hess *et al*., 2011). Recently, the development of next generation sequencing methods made thousands of SNP available in non-model organisms (Santure *et al*., 2010). The amount of research incorporating SNP analysis based on NGS technology, which requires substantial funding and advanced bioinformatics tests, generates large datasets (Santure *et al*., 2010; Liu *et al*., 2014; Plomion *et al*., 2016). However, technology is available allowing the analysis of SNP markers that does not require the involvement of NGS techniques or high cost.

This study was carried out to verify usefulness of 20 SNP markers in genetic diversity study of *Pinus sylvestris* populations in central Europe. We also aimed to develop an alternative to popular microsatellite markers, while maintaining the low costs of the analyses.

Materials and methods

Plant material and DNA extraction. Pine shoots were sampled from 96 trees in stands from 4 forest stands in central Poland: Nagoszewka (27 samples), Sierzchów (20 samples), Olsztynek (25 samples) and Oleszyce (24 samples), which were located at 52°46'N 21°48'E; 51°58'N 21°06'E; 53°33'N 20°22'E; and 50°17'N 22°57'E, respectively.

Extractions of genomic DNA from samples were made using a DNA isolation kit (Macherey-Nagel®, Germany), with minor modification. Lysis buffer PL2 was used at a volume of 600 µl in the lysis step, the volume of buffer used for precipitation after the PL3 lysis step was 150 µl; and PC binding buffer was increased to 900 µl.

SNP IDENTIFICATION AND GENOTYPING. The genome of *Pinus spp.* is sequenced and available for investigation (Zimin *et al*., 2014). One can even find information published concerning the level of SNP polymorphism for *Pinus sylvestris* (Cullingham *et al*., 2013; Semerikov *et al*., 2015). To design a multiplex for SNP identification, based on the primer extension method, we first selected 45 genetic markers with the highest potential number of SNPs (Wachowiak *et al*., 2009, 2011; Cullingham *et al*., 2013). Sixteen sequences for successfully amplified markers were sequenced to screen for variation using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Sci.) on an ABI 3500 capillary sequencer (Applied Biosystems), according to manufacturer's protocols. After multiple sequence alignment using BioEdit 7.0.9 (Hall, 1999), SNP sites were selected to maximize the number of genes and the possibility of designing primers for multiplex reactions. Primers were designed in the Primer 3 program (Koressaar and Remm, 2007; Untergasser *et al*., 2012), as predicted for SNP analysis by primer extension and nucleotide termination method (single-base extension and termination) (Pastinen *et al*., 1997), using the ABI Prism SNaPshot Multiplex Kit. The SNP was identified by SNP genotyping and confirmed by sequencing. Characteristics of primers and SNPs are presented in Table 1.

To perform the SNP identification reaction, the purified PCR products (using primers and conditions for PCR reaction from Cullingham *et al*. (2013) and Wachowiak *et al*. (2011) were pooled in two different mixes (multiplex A: PCR products of genes abaR, C35213, dhy2PP; multiplex B: PCR products of genes: C55378, C55401, C85506) for each individual and used as a template in the primer extension reaction. For the multiplex reaction, SNP primers were combined, generating 10 SNPs per reaction, with each primer attached with a (T)n tail, allowing reading length to be assessed. The contents of the multiplex reaction solution was: 3 µl of pooled PCRs, 5 µl of SnapShot Multiplex Kit (Thermo Fisher Scientific), 1 µl of mixed primers for reactions A and B (concentration 0.2 µM) and 1 µl of water. The reaction conditions were: 96°C – 10 sec, 50°C – 5 sec, 60°C – 30 sec for 25 cycles. After the reaction, 1 unit of alkaline phosphatase (SAP) was added and the contents were incubated for one hour at 37°C; enzyme inactivation was carried out at 75°C for 15 min. Samples prepared in this way were analyzed in the ABI 3500 capillary sequencer (Applied Biosystems), and SNP polymorphisms were read using GeneMapper ver. 5 (Thermo Fisher Scientific Inc.).

SNP ANALYSIS. Basic parameters of molecular variation, such as haplotype variation, h (Nei, 1973), were estimated using DnaSP ver 5.10.01 (Librado and Rozas, 2009). In addition, the expected heterozygosity for individual loci (Nei, 1987) and the average number of nucleotide differences were calculated using Arlequin ver. 3.1. (Excoffier *et al*., 2007). Using the same software, the mean number of alleles, genetic distance (Fst) between forest stands, the observed (Ho) and expected (He) heterozygosity, were calculated, as well as deviation from Hardy-Weinberg equilibrium. Using SNPAnalyzer 2.0 (Yoo *et al*., 2008), we calculated the frequency of alleles and identified haplotypes and their frequency. Haplotypes were reconstructed from diploid data using the EM algorithm (Expectation Maximization algorithm with iterative expectation and maximization steps) (Excoffier and Slatkin, 1995), with minimum occupancy rate=0.01 and the degree of convergence = 0.01. In order to verify the occurrence of disequilibrium between SNP marker pairs in the SNPAnalyzer 2.0 program, the square of Pearson's correlation coefficient (r^2) was calculated as a measure of the deviation of the observed haplotype frequency from

Table 1.

Characterization of 20 SNP markers with information about mix designation; annotation; primer sequence, length and direction; SNP position, type of polymorphism and basic diversity parameters

Mix	SNP name	Protein function	Primer Direction	Primer sequence	SNP position	Primer length	Substitution	H_o	H_e	P HWE
A	abaR2	Abscisic acid responsive protein	F	T(10)CCGACGGAACACCCC	218	25	A/G	0.531	0.469	0.272
A	abaR3	Abscisic acid responsive protein	F	T(13)GCACCAAAGTCATGGGC	146	30	T/C	0.217	0.244	0.376
A	C352131	Aspartyl protease family protein	F	T(14)GATGATAAATATGTGGCTGCT	172	35	T/C	0.362	0.437	0.103
A	C352132	Aspartyl protease family protein	F	T(20)GGTTCTCTTGCAGTTTCACC	224	40	T/C	0.309	0.277	0.450
A	dhy2PP1	Dehydrin	F	T(28)CTTCACCGCACAAACAG	192	45	T/C	0.204	0.293	0.007
A	Dhy2PP2	Dehydrin	F	T(29)GGTTGGTTGATAATGTAGTGG	287	50	T/C	0.269	0.293	0.474
A	Dhy2PP3	Dehydrin	F	T(30)CTCTAGCTCTGTAAGTCTCTTTCTC	229	55	G/T	0.258	0.314	0.099
A	Dhy2PP4	Dehydrin	F	T(39)TCTCKCTTTCTTACTGTGTGA	246	60	G/T	0.046	0.412	0.000
A	Dhy2PP5	Dehydrin	R	T(46)ATCGCTCGACTGCAATTAT	298	65	T/C	0.441	0.488	0.396
A	Dhy2PP6	Dehydrin	F	T(55)CCAGCCAAATGCCCA	96	70	G/A	0.174	0.212	0.112
B	C553781	Transcription factor jumonji domain-containing protein	F	GATTGTCAAGGAAAGAATGG	64	20	G/A	0.344	0.286	0.065
B	C554012	Transcribed locus	F	T(14)CTAACATAGGAAAATTGGTCA	380	35	G/A	0.179	0.245	0.018
B	C554013	Transcribed locus	F	T(25)GCGCCTCAAACCTGC	52	40	T/C	0.372	0.448	0.110
B	C554014	Transcribed locus	F	T(24)TCCTGCTTTTAGTTCTGCTTT	145	45	T/C	0.189	0.237	0.067
B	C554015	Transcribed locus	F	T(30)ATTTGCATGCCATTTAGATA	357	50	G/T	0.453	0.587	0.022
B	C855061	Transcribed locus	R	T(33)TGAATTTGATTTGTTGTGTTGT	181	55	G/A	0.457	0.482	0.669
B	C855062	Transcribed locus	F	T(40)TAGAAAAAGGGCAATGGAAT	292	60	G/A	0.436	0.486	0.392
B	C855063	Transcribed locus	F	T(44)GACATTTGTACCCCAGAACTT	55	65	T/C	0.511	0.470	0.510
B	C855064	Transcribed locus	R	T(48)AAGGGAGAYCAATAAAATCAAC	113	70	T/C	0.045	0.044	1.000
B	C855065	Transcribed locus	R	T(54)GAAGGGATAGATAAAGGGAGA	127	75	T/C	0.435	0.488	0.390

H_0 – observed heterozygosity, H_e – expected heterozygosity, P HWE – exact p value for H-W equilibrium test

the corresponding allele frequency, according to the formula used in the software. The analysis parameters were: minimum haplotype frequency=0.01, cut-off level for MAF=0.05, cut-off for factor r^2=0.8. Several SNPs that were in strong linkage disequilibrium were pull together into one LD block using Gabriel's method (Gabriel *et al*., 2002).

Results and discussion

For the 20 SNP loci evaluated, the observed heterozygosity (Ho) ranged from 0.045 to 0.531, averaging 0.311; the expected heterozygosity (He) varied from 0.044 to 0.469 (averaging 0.360). Only one SNP (dhy2PP4) showed significant deviation from Hardy-Weinberg equilibrium (Table 1). The calculated haplotype variation, h, was high (0.9902 ±0.002) and the average number of nucleotide differences between haplotypes was 10.625 ±4.877. Heterozygosity of SNPs was 0.531 ±0.270. The frequencies of alleles (nucleotides) in particular markers are presented in Table 2. In multiplex reaction A, most balanced frequencies were calculated in dhy2PP6 (G – 0.58 and A – 0.41), while much lower frequencies of the alleles were calculated for the dhy2PP6 and abaR3 markers (A – 0.12, T – 0.14). In multiplex reaction B significantly lower frequency showed allel T – 0.137 (locus C554014), allel G (frequency=0.142; locus C554012) and allel G (frequency=0.135; locus C554015). The 3 analyzed SNPs (C554013, C855061, C855063) similarly showed a balanced frequency in the ratio of about 0.6 (0.60 – 0.66) to about 0.3 (0.33 – 0.39) of both nucleotides. For the SNP C855064, almost 100% of allele showed A, only three individuals identified the T allele and one the G allele. The most frequently observed haplotype (with frequency of 8% among the analyzed samples) was the sequence with the order of bases – GCTTCTTTGGGACCCTGCAA (each nucleotide represents a particular SNP for the analyzed markers). With a slightly lower turnover (around 4% frequency), were three haplotypes,

Table 2.
Frequency of nucleotides occurrences in particular SNPs in both multiplexes

Marker	SNP			
	A	C	G	T
abaR2	0.370		0.630	
abaR3		0.859		0.141
C352131		0.319		0.681
C352132		0.165		0.835
dhy2PP1		0.823		0.177
dhy2PP2		0.177		0.823
dhy2PP3			0.194	0.806
dhy2PP4			0.287	0.713
dhy2PP5	0.414		0.586	
dhy2PP6	0.120		0.880	
C553781	0.172		0.828	
C554012	0.858		0.142	
C554013		0.665		0.335
C554014		0.863		0.137
C554015		0.542	0.135	0.317
C855061		0.601		0.399
C855062	0.590		0.410	
C855063		0.372		0.628
C855064	0.978		0.00	0.02
C855065	0.413		0.587	

and the remaining haplotypes each occurred at less than 1.5% frequency. Due to the specificity of SNP markers as well as the diploid nature of the data, reconstruction of haplotypes in an individual context requires the use of the Excoffier *et al*. algorithm. (1995); nevertheless, this method allows an estimation of the assignment of haplotypes in 7 cases, even with 100% matching accuracy. On the other hand, over half of the cases showed accuracy of matching at above 70%. The lowest match accuracy was estimated for haplotypes for which no full SNP sequence was identified, the lack of data being the main reason for obtaining low matching values (=0.125). The haplotype with the highest frequency was assigned to 14 samples based on the EM algorithm, but there was no case in which it occurred more than once in pairs with the same second haplotype. The analysis of the basic parameters of genetic differentiation for forest stands showed a mean number of alleles close to 2, which results directly from the characteristics of the SNP markers. – Both the observed and expected heterozygosity values were in the range from 0.343 to 0.396 in all populations, and the p-value for H-W was not statistically significant for all four forest stands (Table 3). There were no statistically significant genetic distance between forest stands. The analysis of the linkage disequilibrium between the SNP markers showed significant correlation between the 6 cases in which $r^2 > 0.8$. One of the SNP couplings was observed in the dhn2PP marker between 2 and 3 SNP ($r^2=0.898$), further in the C55401 marker ($r^2=0.957$), while the remaining 4 within the C85506 marker sequence had r^2 values from 0.812 to 0.933. The occurrence of linkage disequilibrium in the observed allele frequency is not an isolated incident in the case of the same coding gene. In the analyzed nucleotide polymorphisms, however, there was no LD occurring between SNP markers belonging to different genes. A graphical representation of the LD between SNPs is shown in Figure.

Liu *et al*. (2005), pointed that, when comparing different types of markers, SNPs of highest informativeness work better than the same number of microsatellites especially with a small number of markers apply to data analysis. Studies carried out by Rosenberg *et al*. (2003) and Liu *et al*. (2005) conclude that the appropriate number of SNP markers can be more informative for population structure inference, due to the fact that the markers with the highest informativeness are added to the analysis first.

In the research on the genetic variability of Scots pine microsatellite markers are the basic tools (Semerikov *et al*., 2014; Kalko and Kotova 2018), but the use of single nucleotide variability for this species is also carried out in the context of studying the degree of polymorphism or adaptation in a changing environment (González-Martínez *et al*. 2007; Wachowiak *et al*. 2009). On the other hand, the use of the sanger method in the analysis of DNA polymorphism for forest-forming tree species is not common in research (Seifert *et al*., 2012), and it is an important methodological alternative for this type of study. In the European beech study, the authors used a panel of 17 SNPs gene-related with drought stress response (Seifert *et al*., 2012). The obtained mean values of observed heterozygosity were 0.324, together with no significant deviations from the Hardy-Weinberg equilibrium allowed the authors to summarize the developed genetic

Table 3.

Basic genetic diversity parameters for four stands

Stand	Mean Nb of alleles	H_o	H_e	P HWE
Nagoszewka	2.118	0.353	0.396	0.414
Sierzchów	2.111	0.350	0.386	0.505
Oleszyce	2.056	0.343	0.347	0.568
Olsztynek	2.071	0.345	0.378	0.548

H_o – observed heterozygosity, H_e – expected heterozygosity, P HWE – exact *p* value for H-W equilibrium test

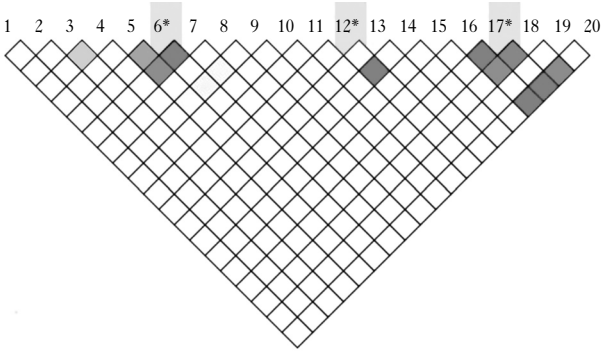

Fig.

Pattern of linkage disequilibrium for 20 SNP sites strong grey colour and * in the linkage disequilibrium pattern graph means that there exists strong pairwise linkage disequilibrium between adjacent SNPs, and the area enclosed by pale grey shows LD block

markers as a useful genomic tools to further investigate drought stress tolerance in population in *F. sylvatica*. What was confirmed by the team of Cuervo-Alarcon in 2018 using, inter alia, these SNPs for study genetic variation in European beech population from Switzerland (Cuervo-Alarcon *et al*., 2018).

Conclusions

In the paper, the authors proposed a new methodology for studying genetic diversity, using a standard capillary sequencer (the Sanger sequencing method) for SNP analysis. The markers used, along with the analysis method performed for this study, provide useful genetic tools for the study of pine genetic variability. The development of research techniques is a significant challenge in the context of research on adaptability to changing environmental conditions. Adverse climate changes will be a challenge in the near future and non-neutral DNA variability studies will be a significant addition to the body of knowledge in the field of broadly understood 'climate change studies'.

Authors' contributions

A.T. – the research concept, laboratory and statistical analysis, manuscript preparation; A.K. – sample collection, laboratory analysis, manuscript corrections; K.F. – manuscript corrections

Conflicts of interest

The authors declare no conflicts of interest.

Funding

The study was supported by the Ministry of Science and Higher Education of Poland Nr 241 401.

References

Antonecchia, G., Fortini, P., Lepais, O., Gerber, S., Legér, P., Scippa, G.S., Viscosi, V., 2015. Genetic structure of a natural oak community in central Italy: Evidence of gene flow between three sympatric white oak species (*Quercus, Fagaceae*). *Annals of Forest Research*, Volume 58, pp. 205-216. DOI: 10.15287/afr.2015.415.

Barzdajn, W., Kowalkowski, W., Chmura, D.J., 2016. Variation in growth and survival among European provenances of *Pinus sylvestris* in a 30-year-old experiment. *Dendrobiology*, Volume 75, pp. 67-77. DOI: 10.12657/denbio.075.007.

Cuervo-Alarcon, L., Arend, M., Müller, M., Sperisen, C., Finkeldey, R., Krutovsky, K.V., 2018. Genetic variation and signatures of natural selection in populations of European beech (*Fagus sylvatica* L.) along precipitation gradients. *Tree Genetics & Genomes*, Volume 14, 84. DOI: 10.1007/s11295-018-1297-2.

Cullingham, C.I., Cooke, J.E.K., Dang, S., Coltman, D.W., 2013. A species-diagnostic SNP panel for discriminating lodgepole pine, jack pine, and their interspecific hybrids. *Tree Genetics & Genomes*, Volume 9, pp. 1119-1127. DOI: 10.1007/s11295-013-0608-x.

Dzialuk, A., Chybicki, I., Gout, R., Mączka, T., Fleischer, P., Konrad, H., Curtu, A.L., Sofletea, N., Valadon, A., 2014. No reduction in genetic diversity of Swiss stone pine (*Pinus cembra* L.) in Tatra Mountains despite high fragmentation and small population size. *Conservation Genetics*, Volume 15, pp. 1433-1445. DOI: 10.1007/s10592-014-0628-6.

Excoffier, L., Laval, G., Schneider, S., 2007. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, Volume 1, pp. 47-50.

Excoffier, L., Slatkin, M., 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Molecular Biology and Evolution*, Volume 12, pp. 921-927. DOI: 10.1093/oxfordjournals.molbev.a040269.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J., Altshuler, D., 2002. The structure of haplotype blocks in the human genome. *Science*, Volume 296, pp. 2225-2229. DOI: 10.1126/science.1069424.

González-Martínez, S.C., Wheeler N.C., Ersoz E., Nelson C. D., Neale D. B. 2007. Association Genetics in *Pinus taeda* L. I. Wood Property Traits. *Genetics*, Volume 175, pp. 1399-1409. DOI: 10.1534/genetics.106.061127.

Govindarajulu, A., 2014. Adaptive variation in extent and timing of growth of Scottish Scots pine (*Pinus sylvestris* Linn.). *Journal of Biodiversity & Endangered Species*, Volume 2, 125. DOI: 10.4172/2332-2543.1000125.

Gugerli, F., Senn, J., Anzidei, M., Madaghiele, A., Buchler, U., Sperisen, C., Vendramin, G.G., 2001. Chloroplast microsatellites and mitochondrial nad1 intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembr*a), Siberian stone pine (*Pinus sibiric*a), and Siberian dwarf pine (*Pinus pumila*). *Molecular Ecology*, Volume 10, pp. 1489-1497. DOI: 10.1046/j.1365-294x.2001.01285.x.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, Volume 41, pp. 95-98.

Hauser, L., Baird, M., Hilborn, R., Seeb, L.W., Seeb, J.E., 2011. An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resources*, Volume 11, suppl. 1, pp. 150-161. DOI: 10.1111/j.1755-0998.2010.02961.x.

Hess, J.E., Matala, A.P., Narum, S.R., 2011. Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River Basin. *Molecular Ecology Resources*, Volume 11, suppl. 1, pp. 137-149. DOI: 10.1111/j.1755-0998.2010.02958.x.

Kalko, G.V., Kotova, T.M., 2018. The microsatellite markers for estimation of genetic diversity of Scots pine. *Proceedings of the Saint Petersburg Forestry Research Institute*, Volume 3-4, pp. 17-30. DOI: 10.21178/2079-6080.2018.3-4.17.

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics*, Volume 23, pp. 1289-1291. DOI: 10.1093/bioinformatics/btm091.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, Volume 25, pp. 1451-1452. DOI: 10.1093/bioinformatics/btp187.

Liu, J.-J., Sniezko, R.A., Sturrock, R.N., Chen, H., 2014. Western white pine SNP discovery and high-throughput genotyping for breeding and conservation applications. *BMC Plant Biology*, Volume 14, 380. DOI: 10.1186/s12870-014-0380-6.

Liu, N., Chen, L., Wang, S., Oh, C., Zhao, H., 2005. Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. *BMC Genetics*, Volume 6, S26. DOI: 10.1186/1471-2156-6-S1-S26.

Misi, D., Puchałka, R., Pearson, C., Robertson, I., Koprowski, M., 2019. Differences in the climate-growth relationship of Scots Pine: A case study from Poland and Hungary. *Forests*, Volume 10, pp. 243. DOI: 10.3390/f10030243.

Narum, S.R., Banks, M., Beacham, T.D., Bellinger, M.R., Campbell, M.R., Dekoning, J., Elz, A., Guthrie, C.M., Kozfkay, C., Miller, K.M., Moran, P., Phillips, R., Seeb, L.W., Smith, C.T., Warheit, K., Young, S.F., Garza, J.C., 2008. Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. *Molecular Ecology*, Volume 17, pp. 3464-3477. DOI: 10.1111/j.1365-294x.2008.03851.x.

Nei, M., 1973, Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, Volume 70, pp. 3321-3323. DOI: 10.1073/pnas.70.12.3321.

Nei, M., 1987. Genetic Variation Within Species. In: M. Nei, ed. *Molecular Evolutionary Genetics*. New York: Columbia University Press, pp. 176-207. DOI: 10.7312/nei-92038-009

Oleksyn, J., Reich, P.B., Tjoelker, M.G., Chałupka, W., 1999. Differential Above- and Below-ground Biomass Accumulation of European Pinus sylvestris Populations in a 12-year-old Provenance Experiment. *Scandinavian Journal of Forest Research*, Volume 14, pp. 7-17. DOI: 10.1080/02827589908540804.

Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., Syvanen, A.C., 1997. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Research*, Volume 7, pp. 606-614. DOI: 10.1101/gr.7.6.606.

Plomion, C., Bartholome, J., Lesur, I., Boury, C., Rodriguez-Quilon, I., Lagraulet, H., Ehrenmann, F., Bouffier, L., Gion, J.M., Grivet, D., de Miguel, M., de Maria, N., Cervera, M.T., Bagnoli, F., Isik, F., Vendramin, G.G., Gonzalez-Martinez, S.C., 2016. High-density SNP assay development for genetic analysis in maritime pine (*Pinus pinaster*). *Molecular Ecology Resources*, Volume 16, pp. 574-587. DOI: 10.1111/1755-0998.12464.

Rehfeldt, G.E., Tchebakova, N.M., Parfenova, Y.I., Wykoff, W.R., Kuzmina, N.A., Milyutin, L.I., 2002. Intraspecific responses to climate in *Pinus sylvestris*. *Global Change Biology*, Volume 8, pp. 912-929. DOI: 10.1046/j.1365-2486.2002.00516.x.

Rosenberg, N.A., Li, L.M., Ward, R., Pritchard, J.K., 2003. Informativeness of Genetic Markers for Inference of Ancestry. *The American Journal of Human Genetics*, Volume 73, pp. 1402-1422. DOI: 10.1086/380416.

Santure, A.W., Stapley, J., Ball, A.D., Birkhead, T.R., Burke, T., Slate, J. 2010. On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology*, Volume 19, pp. 1439-1451. DOI: 10.1111/j.1365-294X.2010.04554.x.

Savva, J.V., Vaganov, E.A., 2006. Genetic and Environmental Effects Assessment in Scots Pine Provenances Planted in Central Siberia. *Mitigation and Adaptation Strategies for Global Change*, Volume 11, pp. 269-290. DOI: 10.1007/s11027-006-1026-2.

Seifert, S., Vornam, B., Finkeldey, R., 2012. A set of 17 single nucleotide polymorphism (SNP) markers for European beech (*Fagus sylvatica* L.). *Conservation Genetics Resources*, Volume 4, pp. 1045-1047. DOI: 10.1007/s12686-012-9703-9.

Semerikov, V.L., Putintseva, Y.A., Oreshkova, N.V., Semerikova, S.A., Krutovsky, K.V., 2015. Development of New Mitochondrial DNA Markers in Scots Pine (*Pinus sylvestris* L.) for Population Genetic and Phylogeographic Studies. *Genetika*, Volume 51, pp. 1386-1390.

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., Kalchenko, L.I., 2014. Microsatellite Loci polymorphism of chloroplast DNA of the pine tree (*Pinus sylvestris* L.) in Asia and Eastern Europe. *Genetika*, Volume 50, pp. 660-669.

Stolting, K.N., Nipper, R., Lindtke, D., Caseys, C., Waeber, S., Castiglione, S., Lexer, C., 2013. Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology*, Volume 22, pp. 842-855. DOI: 10.1111/mec.12011.

Tian, J., Du, Q., Mengqi, C., Zhang, D., 2012. Allelic Variation in *PtGA20Ox* Associates with Growth and Wood Properties in *Populus* spp. *Plos One*, Volume 7, e53116. DOI: 10.1371/journal.pone.0053116.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Research*, Volume 40, e115. DOI: 10.1093/nar/gks596.

Wachowiak, W., 2015. Genetic relationships between Polish and reference populations of Scots pine (*Pinus sylvestris L.*) in Europe based on nucleotide polymorphism study at nuclear loci. *Sylwan*, Volume 159, pp. 53-61. DOI: 10.26202/sylwan.2014032.

Wachowiak, W., Balk, P.A., Savolainen, O., 2009. Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris L.*). *Tree Genetics & Genomes*, Volume 5, pp. 117. DOI: 10.1007/s11295-008-0188-3.

Wachowiak, W., Salmela, M.J., Ennos, R.A., Iason, G., Cavers, S., 2011. High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris L.*) in Scotland. *Heredity*, Volume 106, pp. 775-787. DOI: 10.1038/hdy.2010.118.

Weller, J.I., Seroussi, E., Ron, M., 2006. Estimation of the number of genetic markers required for individual animal identification accounting for genotyping errors. *Animal Genetics*, Volume 37, pp. 387-389. DOI: 10.1111/j.1365-2052.2006.01455.x.

Werner, F.A.O., Durstewitz, G., Habermann, F.A., Thaller, G., Kramer, W., Kollers, S., Buitkamp, J., Georges, M., Brem, G., Mosner, J., Fries, R., 2004. Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. *Animal Genetics*, Volume 35, pp. 44-49. DOI: 10.1046/j.1365-2052.2003.01071.x.

Yoo, J., Lee, Y., Kim, Y., Rha, S.Y., Kim, Y., 2008. SNPAnalyzer 2.0: a web-based integrated workbench for linkage disequilibrium analysis and association analysis. *BMC Bioinformatics*, Volume 9, 290. DOI: 10.1186/1471-2105-9-290.

Zimin, A., Stevens, K.A., Crepeau, M.W., Holtz-Morris, A., Koriabine, M., Marcais, G., Puiu, D., Roberts, M., Wegrzyn, J.L., de Jong, P.J., Neale, D.B., Salzberg, S.L., Yorke, J.A., Langley, C.H., 2014. Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics*, Volume 196, pp. 875-890. DOI: 10.1534/genetics.113.159715.

STRESZCZENIE

Panel 20 markerów polimorfizmu pojedynczego nukleotydu jako narzędzie w badaniu zróżnicowania genetycznego sosny zwyczajnej

W analizach struktury genetycznej populacji standardowo stosowane są markery mikrosatelitarne SSR. Źródłem polimorfizmu tych markerów jest różnica w długości powtórzeń tandemowych zlokalizowanych w genomach. Szereg cech związanych z budową i funkcjonowaniem genomu gatunku sosna zwyczajna (*Pinus sylvestris* L.) sprawia, że analiza tego typu markerów może być trudnym wyzwaniem. Zastosowanie alternatywnych markerów molekularnych, opartych na zmienności pojedynczych nukleotydów (SNP), stanowi niezbędne uzupełnienie technik stosowanych dotychczas w analizach molekularnych. W niniejszej pracy zaprojektowano dwie reakcje multipleksowe, pozwalające na jednoczesną analizę 20 miejsc polimorficznych SNP, w obrębie 6 sekwencji DNA dla wytypowanych markerów. Testowana metodyka polega na wykorzystaniu w reakcji PCR startera wydłużonego znakowanym nukleotydem, co pozwala na jednoczesną analizę do 10 miejsc SNP w jednej reakcji PCR, niezależnie od liczby sekwencji genów, w obrębie których analizowane są SNP. Uzyskane wyniki wskazują, że liczba analizowanych SNP jest wystarczająca do badania zmienności genetycznej populacji (dla 96 drzew – zróżnicowanie haplotypowe h=0,9902 ±0,002; średnia liczba różnic nukleotydowych=10,625 ±4,877). Dla 20 analizowanych loci SNP obserwowana heterozygotyczność (Ho) zawierała się w przedziale od 0,045 do 0,531, średnio 0,311; oczekiwana heterozygotyczność (He) wynosiła od 0,044 do 0,469 (średnio 0,360). Tylko jeden SNP (dhy2PP4) wykazał istotne odchylenie od równowagi Hardy'ego-Weinberga (tab. 1). Zarówno obserwowane, jak i oczekiwane wartości heterozygotyczności mieściły się w zakresie od 0,343 do 0,396 we wszystkich populacjach, a wartość p dla H-W nie była istotna statystycznie dla wszystkich czterech drzewostanów (tab. 3). Analiza nierównowagi sprzężeń między markerami SNP wykazała istotną korelację w 6 przypadkach, w których r^2>0,8 (Fig.).

Zaproponowana metodyka, polegająca na zastosowaniu detekcji fluorochromów w identyfikacji polimorfizmu, jest tożsama z wynikiem uzyskanym metodą sekwencjonowania DNA (Sangera), przy jednoczesnym znacznym obniżeniu kosztów analiz laboratoryjnych. Ma to istotne znaczenie we wdrażaniu metod molekularnych do badań ekologicznych, morfologicznych czy fizjologicznych.