DENDROBIOLOGY

2023, vol. 90, 76–85

https://doi.org/10.12657/denbio.090.006

Sebastian Szczepański, Bartosz Łabiszak, Witold Wachowiak*

Development of a SNaPshot assay for the genotyping of organellar SNPs in four closely related pines

Received: 16 May 2023; Accepted: 30 September 2023

Abstract: Mitochondrial (*mt*DNA) and chloroplast (*cp*DNA) polymorphisms are valuable resources to study past demographic changes, phylogenetics and evolution, especially in forest tree species, where these genomes are haploid and uniparentally transferred. The organellar markers were usually scored separately using direct sequencing or PCR-based approaches, which can be time-consuming and expensive, especially in large-scale population genetics research. In this study, we developed an efficient and cost-effective SNaPshot assay for genotyping preselected *mt*DNA and *cp*DNA polymorphism in four closely related pine species including Scots pine (*Pinus sylvestris* L.) and three taxa from the *Pinus mugo* complex. We validated the method by genotyping the samples derived from 12 populations of the species from their wide geographical distribution range in Europe. The results proved high accuracy of the method with a genotyping success rate of 99.7%. The set of assayed markers shows significant genetic variation. By using multiplex SNaPshot assay, we provided an efficient and sensitive molecular tool for intra- and interspecific genetic analyses. The presented protocol is useful for fast and relatively cheap SNP genotyping of organelle genome of closely related pine species. The assayed SNPs allow studying the species discrimination and detailed investigations of their population history and structure. Given its numerous benefits and efficient genotyping rate, the SNaPshot method appears to be a valuable and practical resource for studying the genetic makeup of forest tree species. Particularly, it proves to be advantageous for population genetics.

Keywords: genetic diversity, *mt*DNA, hybridization, Pinus, SNP genotyping, organelle DNA

Addresses: S. Szczepański, B. Łabiszak, W. Wachowiak, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland;
SS https://orcid.org/0000-0001-9701-0771, e-mail: sebastian.szczepanski@amu.edu.pl;
BŁ https://orcid.org/0000-0002-2548-9186
W. Wachowiak, Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, 62-035 Kórnik, Poland;
https://orcid.org/0000-0003-2898-3523
* corresponding author

Introduction

Genetic variation is fundamental for species persistence and plays a critical role in shaping the diversity of life. Understanding the extent and distribution of genetic variation within and among species is essential for elucidating their evolution and adaptation, and for informing conservation efforts aimed at preserving biodiversity. Advances in genomic technologies have greatly expanded our ability to study genetic variation in natural populations. However,

analyzing genetic diversity at the population level remains a complex and challenging task, especially for species with large, structurally complex genomes, such as many important forest tree species (Cao et al., 2022; Chung et al., 2023; Neale et al., 2017).

Tree populations often exhibit high level of genetic diversity within populations, but low differentiation between individuals within the same population as a result of many factors including species history, distribution range or mating system (Petit & Hampe, 2006). To accurately study the genetic diversity and population history of tree species across a wide geographic range, a tradeoff must be made between the resolution of the available genetic markers and the number of individuals sampled from each population. This decision is strongly influenced by the cost of genotyping. Given the current biodiversity crisis and its significant threat to forest ecosystems, it is crucial to develop and use reliable and cost-efficient methods for population genetic analysis (Ceballos et al., 2015; Urban, 2015). These methods are aimed at describing and monitoring the genetic resources of tree species, which is essential for understanding and managing their populations.

The *Pinus* genus comprises over 100 species of trees that are widely distributed across the world and adapted to a range of climatic and soil conditions. They are one of the most significant forest-forming tree species in the Northern Hemisphere, with both ecological and economic importance (Farjon, 2018). Due to their numerous ecosystem services, it is critical to gain a better understanding of their evolution, adaptation, and thus response to environmental changes.

Polymorphisms in mitochondrial DNA (*mt*DNA) and chloroplast DNA (*cp*DNA) genomes are often used to study phylogenetics and evolutionary history of many species (Avise et al., 1987; Hewitt, 1999; Morris & Shaw, 2018; Naydenov et al., 2007; Palmer, 1992; Soranzo et al., 2000; Taberlet et al., 1998; Tóth et al., 2017). They are especially useful in the case of forest tree species, because of their haploidy, uniparental transmission through seeds and pollen and lack of sexual recombination. Mitochondrial DNA markers are particularly useful in studying pines because they are inherited maternally and dispersed by seeds over short distances. Due to their mode of inheritance and distribution, *mt*DNA markers provide higher resolution in population structure analysis as compared to pollen-mediated markers (including chloroplast and nuclear DNA). As a result, they offer a valuable perspective on the genetic history of these species, reflecting past demographic changes and retaining patterns of demographic structure over time (Jaramillo-Correa & Bousquet, 2005; Łabiszak et al., 2019; Polezhaeva et al., 2010; Semerikov & Lascoux, 2003; Senjo et al., 1999; Tollefsrud et al., 2015; Tóth et al., 2017)

Due to its relatively simple structure and size (~150k bp), chloroplast DNA has been sequenced for many pine species, providing complete *cp*DNA genomes and easy access to polymorphisms (Asaf et al., 2018; Sokolowska et al., 2021; Yu et al., 2022). The difficulty in discovering new mitochondrial DNA (*mt*DNA) markers is primarily due to the large size of conifers mitochondrial genome (Jackman et al., 2015; Jackman et al., 2020), the presence of multiple repeated regions in their complex structure, and their relatively low sequence evolution rate (Guo et al., 2016; Smith, 2016). However, advances in sequencing technologies allowed the development of novel genomic resources in non-model plants, including the sequence of a large fragment of the mitochondrial genome in pines (Donnelly et al., 2017). Based on the discovered polymorphisms, Łabiszak et al. (2019) developed a set of *mt*DNA markers that proved to be useful in population genetics studies of closely related pine species (*Pinus mugo* complex - *P. mugo*, *P. uliginosa*, *P. uncinata*). However, genotyping of the markers using PCR-based Restriction Fragment Length Polymorphism (PCR-RFLP) required time-consuming protocols of amplification, restriction enzyme digestion and electrophoresis of each marker separately, significantly limiting their applicability to large-scale studies.

To overcome this challenge, we present a method of genotyping a set of SNPs and indel markers from mitochondrial and organellar genomes in the complex of four pine species. This method can be applied to score genetic variation across many samples in large scale population genetic studies. The developed assay allows for the simultaneous analysis of multiple SNPs and is based on a commercial SNaPshot™ Multiplex Kit (Thermo Fisher) that combines the markers in multiplex reactions. The approach can be utilized to genotype predefined polymorphism in any species. Similar assays, containing 15-18 mitochondrial SNPs, were developed previously, but their usage was mainly restricted to forensic or anthropological studies, and not to broad-range plant population genetic investigations (Hu et al., 2016; Weiler et al., 2016). We validated our SNaPshot method using samples from a broad distribution range demonstrating its versatility and applicability in *Pinus sylvestris* and three taxa from the *P. mugo* complex. Our approach provides a useful tool to explore large-scale patterns of genetic variation including gene flow and genetic structure at both the intra- and interspecific levels. Compared to previously used genotyping approaches in studied pine species such as PCR-RFLP and Sanger sequencing, our method is faster, cheaper and more effective for scoring the markers.

Materials and methods

DNA extraction and PCR amplification

Samples from 147 trees derived from 12 natural populations of four pine species: *Pinus sylvestris*, *P. mugo*, *P. uliginosa* and *P. uncinata*, were included in the study (Table S1). Genomic DNA was extracted from the needles using Genomic Mini AX Plant extraction kit (A&A Biotechnology, Poland). The quantity of DNA was measured by Qubit 4 fluorometer, using the Broad Range (BR) Assay Kit and DNA was diluted to the working concentration of 40 ng/µl. PCRs were carried out in two multiplex reactions, each with seven markers (Table 1). The first multiplex contained PR5, PR7, PR15, PR19, PR20, PR21 and PR24 regions (Donnelly et al., 2017). The second one contained PR25, PR29, PR30, PR31 and PR32 (Donnelly et al., 2017) – names are consistent with those in Łabiszak et al. (2019), *nad*1 (Soranzo et al., 2000) mitochondrial fragments and *trn*L-*trn*F intergenic region of the chloroplast DNA (Taberlet et al., 1991; Wachowiak et al., 2000). The PCR reaction mixture for both multiplexes contained 1µl of Solis Biodyne HOT FIREPol® Blend Master Mix (5x) polymerase, 0.5 µl of primer mix (5 µM of each primer), 3 µl of Milli-Q water and 1 µl of sample DNA (40 ng/µl). PCRs included the following steps: initial denaturation (95 °C, 15 min), 32 cycles of denaturation (95 °C, 30 s), annealing (57 °C, 1:30 min), extension (72 °C, 1:30 min) and final extension (72 °C, 10 min). PCR products were separated in 1.5% agarose gels (1x TAE buffer) resulting in 5 (Multiplex I) and 4 (Multiplex II) clearly visible fragments as some of the amplification products were of similar size.

Development of SNaPshot assay

The SNaPshot method involves a process of amplification and purification of the DNA amplicons before using a specific primer to target a sequence next to the SNP site. PCR products in each multiplex were purified with enzymes exonuclease I and phosphatase. Each reaction contained 0.5 µl of phosphatase (1 U/µl, Thermo Fisher), 0.05 µl of exonuclease I (20 U/µl, Thermo Fisher), 0.25 µl of exonuclease buffer (10x, Thermo Fisher), 1.7 µl of Milli-Q water and 5 µl of PCR product. Total reaction volume of 7.5 µl was then heated to 37 °C for 60 min and to 80 °C for 15 min. The purified products were used in the SNaPshot reaction carried out according to SNaPshot™ Multiplex Kit protocol and involving extension of the primer targeted at each SNP marker position by a single base using fluorescently labeled dideoxynucleotides. Total volume of 5 µl of SNaPshot assay contained 0.5 µl SNaPshot™ Multiplex Ready Reaction Mix (5x), 0.1 µl SNaPshot primers mix (Table 2; 10 µM each primer), 0.5 µl sequencing buffer (5x), 2.4 µl Milli-Q water and 1.5 µl purified PCR product. Reaction conditions involved initial denaturation (96 °C, 10 s), 35 cycles of denaturation (96 °C, 10 s), annealing (50 °C, 5 s) and extension (60 °C, 30 s). After initial tests we modified the conditions for Multiplex I, extending the annealing time to 30 s in order to improve the quality of PR5 read. Final product was purified with 2.5 µl of phosphatase solution (0.25 µl phosphatase (1 U/ µl), 2.25 µl Milli-Q water) in 37 °C for 15 minutes and 80 °C for 15 minutes. This is an important step as failure to remove unincorporated labelled nucleotides can cause extraneous fluorescence, leading to incorrect

Table 1. List of markers divided into two multiplex reactions with primer names, sequences and SNP variants

	Marker	Forward primer	Reverse primer	SNP variant
Multiplex I	PR5[1]	PR5F ATTCCTGTGCTTGGTTGGGA	PR5R GGCGCTTACCCACACACTTA	[T/G]
	PR7[1]	PR7F TGAGTTCGTTGACCGCGTAA	PR7R TCAGGCGAGCTTGTGCTTTA	[C/A]
	PR15[1]	PR15F CATCCTCTCCTCTCGATGGC	PR15R GCTTTTGGCTTGGTGCGAAT	[T/G]
	PR19[1]	PR19F CGGAGCGAGGTGAAGAAACT	PR19R GCGAGAAGCAGTAGTGGGTT	[T/G]
	PR20[1]	PR20F GTTCCTACGATCCAGCCAGG	PR20R ACCATGGATTCTTCGGACGG	[C/A]
	PR21[1]	PR21F TCCGATGATGAGGTGGAGGT	PR21R AGTTGAAGGCAGGAAGGTCG	[T/G]
	PR24[1]	PR24F TGCATTCTGGCTGGCTTTCT	PR24R GGCGTCGATAGACTCGGTTT	[T/G]
Multiplex II	PR25[1]	PR25F GGCATGTCCGCTATGGAAGT	PR25R AGGCTCCGGAAGTACCTGT	[T/G]
	PR29[1]	PR29F GGTTGGTTGATCCATCCGGT	PR29R CCGGCTTGGGTACGTCTTTT	[T/G]
	PR30[1]	PR30F ACTTACATTGACCGGCGGAT	PR30R CACACATCTAGGGCACAGGG	[T/G]
	PR31[1]	PR31F TGCGACCTGTGAATGGATGT	PR31R CGGCGGTTCTAGCCTTGATT	[T/G]
	PR32[1]	PR32F ACCCTCCTTCAACTGATGCG	PR32R CCTCAACCAACCGTCAGTCA	[T/G]
	*nad*1[2]	nad1F TTAATCAAAAGGTCCGGAG	nad1R TGAAGTGACTCGCACTACTG	[C/G]
	*trn*L-*trn*F[3]	trnF ATTTGAACTGGTGACACGAG	trnL CGAAATCGGTAGACGCTACG	[C/A]

[1] Donnelly et al., 2017; [2] Soranzo et al., 2000; [3] Taberlet et al., 1991.

Table 2. List of primers used in SNaPshot reactions

	Primer name	Length	Sequence	SNP variant
Multiplex I	PR_5sF	20	CCTTTCTATATGAGAATATT	[T/G]
	PR_7sF	26	CTCTTATCCGATCCGAATATTATTTT	[C/A]
	PR_15sF	32	AACAACAGAAGCAAGGAAGGAATCAGCCAGAA	[T/G]
	PR_19sF	38	AAGGTGAAGCTAGTGTCACTGAGACTTATTAACTTATT	[T/G]
	PR_20sF	44	TGTTGCTGCCATACCCTTTCATGAGGTTTCTCTCTGCTGATAGC	[C/A]
	PR_21sF	50	GTCCTTTTGGTTCTTCGCTGATCAGCAAGATCTAATCTCTCTTCGGGGTT	[T/G]
	PR_24sF	56	TTTCTAAATATGTTTGAAGTGAATGCATCATAGCTGAGCTGGACAATAAGTGTTTT	[T/G]
	PR_25sF	20	AAAGGAGGCTGTAGGTAGGA	[T/G]
Multiplex II	PR_29sF2	18	AAAAGCAGGTGGGTTGGA	[T/G]
	PR_30sF	32	CTGGTTGGTTCCATTAAGGCCTTACTCCATGA	[T/G]
	PR_31sF	38	TTCGTTTCCTAATGACGACCAGACTGAGGTAGTTAATT	[T/G]
	PR_32sF	44	AGTGAGTGACTCCGTCCCTGGGAAATCGAATATCATATAAAATA	[T/G]
	*nad*1BC_sF	50	TCTTTTTTACTTACTTTAGAGGATGCGTAAGCACGCTCGACTGTTAAGGA	[C/G]
	*cp_trn*LF_sF	56	TATTTTCGATCTGGAAGTCACTAATATGATAAAAATGGACTGCAATTGAATAATTT	[C/A]

or failed genotyping results (Pati et al, 2004). Finally, 1 µl of purified SNaPshot product was added to 9 µl of formamide with GeneScan™ 120 LIZ™ dye Size Standard and denatured for 5 minutes in 95 °C. Capillary electrophoresis was carried out on the ABI Prism 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA). The chromatograms, reflecting the primer size used in the SNaPshot assay and extended by a labeled base, were viewed and analyzed in Peak Scanner™ Software v1.0.

Data analysis

Mitochondrial DNA polymorphisms were scored and the 13 SNP variants concatenated to form 13 nucleotide long haplotypes. The polymorphism at the *trn*L-*trn*F region, that discriminates *cp*DNA genome of *P. sylvestris* vs. the taxa from the *P. mugo* complex (*P. mugo*, *P. uliginosa*, *P. uncinata;* Wachowiak et al., 2000; Jasińska et al., 2010; Wachowiak et al., 2016) was analyzed separately. The number of haplotypes and haplotype diversity across populations and species was assessed in DnaSP v6 software (Librado & Rozas, 2009; Rozas et al., 2017). One individual was excluded from the analysis because it had missing data for three out of the thirteen markers used to determine haplotypes, probably due to poor quality/quantity of DNA. No other missing data was present in the analysis.

Haplotype network was made in PopART 1.7 software (Leigh & Bryant, 2015) using the Median Joining Network method (Bandelt et al., 1999). The map showing the distribution of the haplotypes was generated in R (R Core Team, 2021) using packages: foreign, tidyverse (Wickham, 2011, Wickham et al., 2019), raster (Hijmans and van Etten, 2012), rgdal (Keitt et al., 2010) and ggplot2 (Wickham, 2016). Haplotype numbers, number of singletons, haplotype diversities and Nei unbiased genetic distance (uNei) (Nei, 1973) were calculated in GenAlEx 6.5 (Peakall & Smouse, 2006; 2012). In order to visualize the genetic relationships between samples PCoA based on uNei distances was performed in ggplot2 R package.

Results and discussion

In this research, we deliver a novel tool for population genetic studies of polymorphisms in the organellar genome of closely related pine species – *P. sylvestris*, *P. mugo*, *P. uliginosa* and *P. uncinata*. The technique is time efficient and relatively low-cost, making it a valuable alternative to other methods of DNA genotyping of pre-selected genetic markers. Although the standard PCR-based Restriction Fragment Length Polymorphism (PCR-RFLP) method is usually known to be relatively inexpensive and frequently used in DNA polymorphism analysis of targeted genomic regions (Hashim & Al-Shuhaib, 2019; Wolf et al., 1999), it is time-consuming and may provide inconclusive results that require additional validations. This method was previously employed for genotyping the same set of SNPs (Łabiszak et. al 2019, Zaborowska et. al 2019, Wachowiak et al., 2023) and was estimated to be roughly four times more expensive and much slower than SNaPshot method due to time consuming steps involving separate amplification of each *mt*DNA region analysed, digestion with restriction enzymes targeted on the SNP position and their agarose gel electrophoresis. Similarly, direct sequencing of targeted amplicons is laborious and expensive, especially when the genotyped SNPs are distributed across many genomic

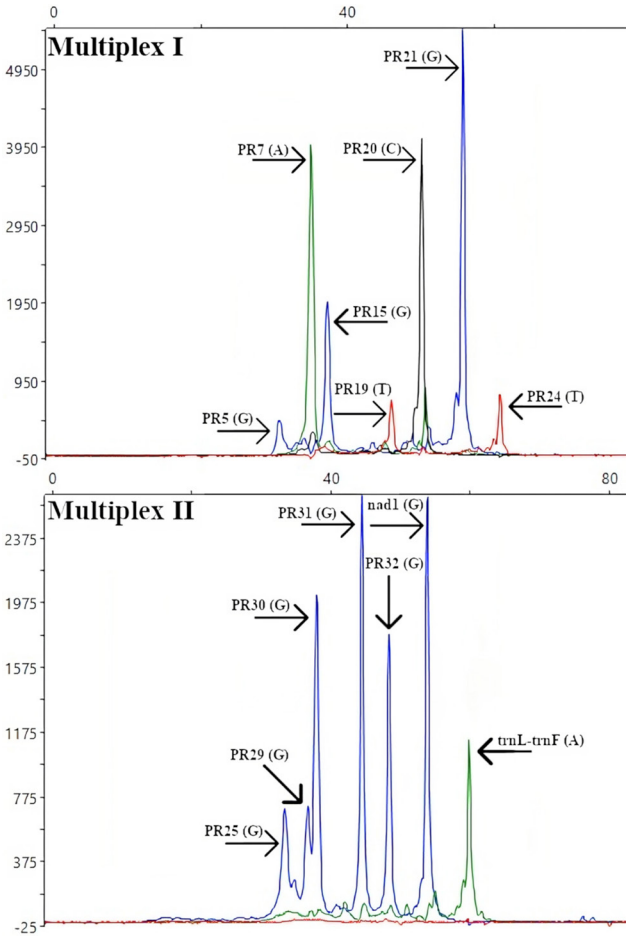

Fig. 1. Representative electropherograms of the two SNaPshot multiplex reactions scored for 14 SNPs color-coded as blue (G), green (A), red (T) and black (C)

regions. Additionally, our method provides consistent results across runs and is unaffected by the differences in laboratory equipment used during genotyping. Unlike microsatellite markers (SSR, Single Sequence Repeats) which depend on the exact length of products, our method relies on the relative sizes of the markers. Specifically, the markers are designed in a way that their sizes increase by 6 bp between markers (Table 2). Possible variations in product sizes (around ±1-3 bp) due to different equipment used, if present, should be consistent across all runs and thus cannot impact the proper scoring of genotypes.

In our study, all sites were successfully genotyped providing reliable scores for each of the assayed markers (Fig. 1). The assayed SNPs showed high efficiency in all studied pine species including 36 *P. sylvestris,* 35 *P. mugo*, 39 *P. uliginosa*, and 36 *P. uncinata* samples. The markers showed intra- and interspecific genetic variation (Fig. 2). Only two markers PR24 and PR32 were monomorphic within the studied samples. However, they showed variation in *P. sylvestris* distribution range not covered in the validation panel of our study (Wachowiak et al., 2023). Therefore, the SNaPshot assay we demonstrate here is a much simpler and cheaper genotyping method than previously mentioned alternatives since it involves the analysis of multiplexed markers and can be applied to most SNPs data. As we obtained complete genotypes for 146 of 147 studied samples, the developed assay is highly efficient providing a higher success rate (over 99%) than in similar studies based on other genotyping techniques using 12-14 markers, ranging from 80% to 94.4% (e.g. Boratyńska et al., 2021; Sobierajska et al., 2020; Wachowiak et al., 2022). Therefore, it seems a valuable method of choice for population genetic studies based on genotyping of informative SNPs markers.

The assayed SNPs allowed for the discovery of 12 different mitochondrial haplotypes (H1–H12) for the investigated species (Fig. 2, Table S2). *Pinus sylvestris*, *P. mugo* and *P. uliginosa* had the same number of haplotypes (5), whereas *P. uncinata* had 3 haplotypes. Haplotype diversity for all samples was 0.849. *P. uliginosa* had the highest and *P. uncinata* had the lowest unbiased haplotype diversity (uh = 0.329 and uh = 0.042, respectively) (Fig. 3A). The number of haplotypes varied among studied populations. We found the highest number of haplotypes in *P. uliginosa* population from Węglowiec (WLB, 4 haplotypes), which resulted also in the highest unbiased haplotype diversity uh = 0.311 (Fig. 3B–C). However, the separation of the individual population was not that clear, as some haplotypes were shared between species.

Significantly, the mitochondrial markers assayed could clearly discriminate the four pine species in the PCA analysis (Fig. 4A). This is also apparent at population level, although there *P. uliginosa* populations are closer to populations of other pines (Fig. 4B). The pattern we observe at the intraspecific level could be understood as mean haplotype composition within species. The mean is susceptible to the presence of outliers, and it is reflected in the pattern observed in PCoA, especially in the case of the population within *P. uliginosa*. The population from Wielkie Torfowisko Batorowskie (Bat) has only one distinctive haplotype present in all individuals that elevates mean distance between studied species in the PCoA analysis. As this haplotype is also common in Finnish populations of *P. sylvestris*, both "Bat" and "F1" populations are placed close to each other on the PCoA plot. Furthermore, the chloroplast DNA *trn*L-*trn*F marker (Wachowiak et al., 2000) confirmed differentiation between species (Fig. 5). Variant A (V_A) is characteristic of *Pinus mugo* complex (*P. mugo, P. uliginosa, P. uncinata*), whereas variant C (V_C) occurs only in *Pinus sylvestris* (Fig. 5). Although this marker differentiates only *Pinus sylvestris* from the taxa of the *P. mugo* complex, it is still useful in the pine species hybridization studies. As there are no reports of any sympatric population that would involve all four pine

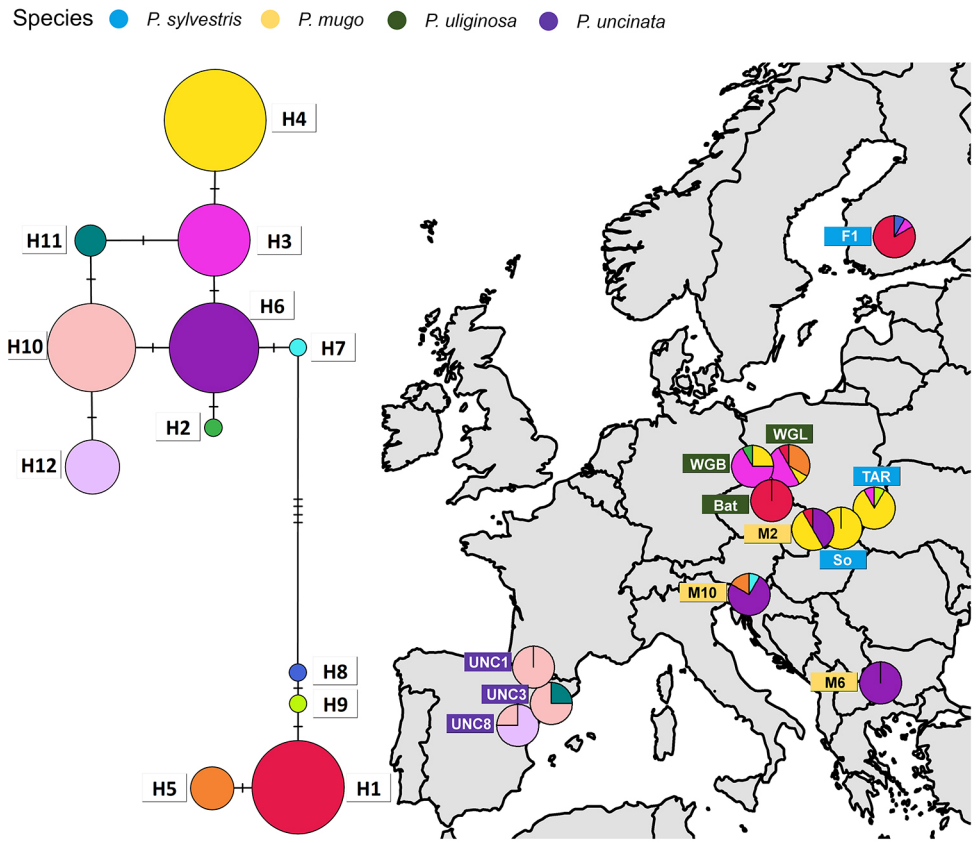

Fig. 2. Median-joining haplotype network (H1–H12) and the distribution map of haplotype frequencies in 12 populations of *P. sylvestris*, *P. mugo*, *P. uliginosa* and *P. uncinata* in Europe. Circle size corresponds to a frequency of a haplotype in the general population. Lines on branches represent a singular mutation event. Detailed information about populations is provided in Table S1 (Supplementary Material)

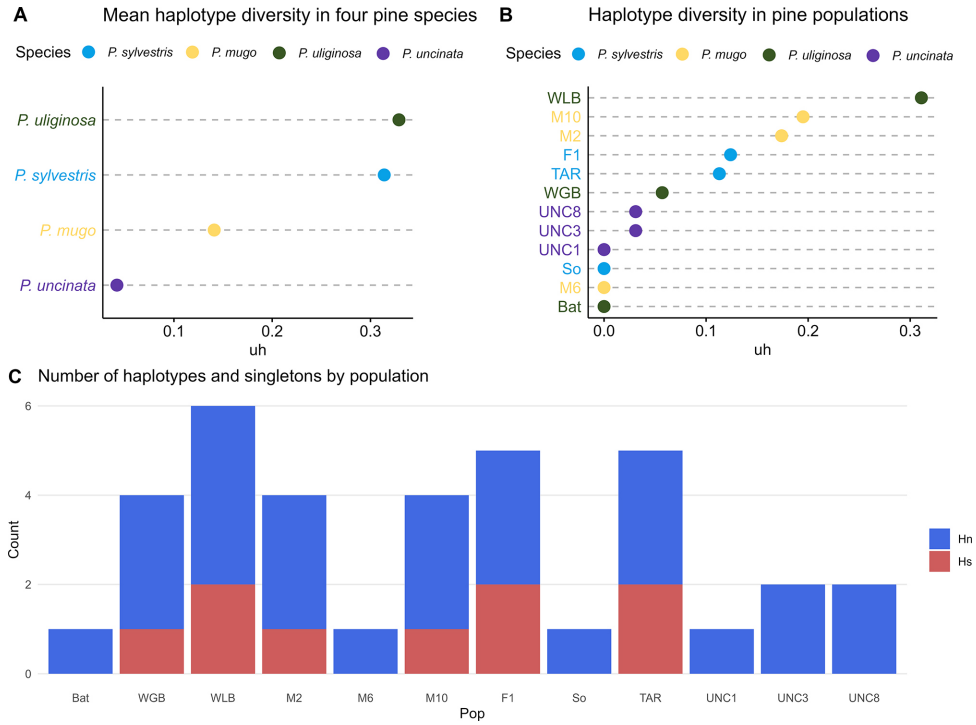

Fig. 3. Distribution of genetic diversity measured as haplotype diversity in four pine species (A), and in 12 pine populations (B). Number of haplotypes (Hn) and singletons (Hs) by population (C). Detailed information about populations is provided in Table S1 (Supplementary Material)

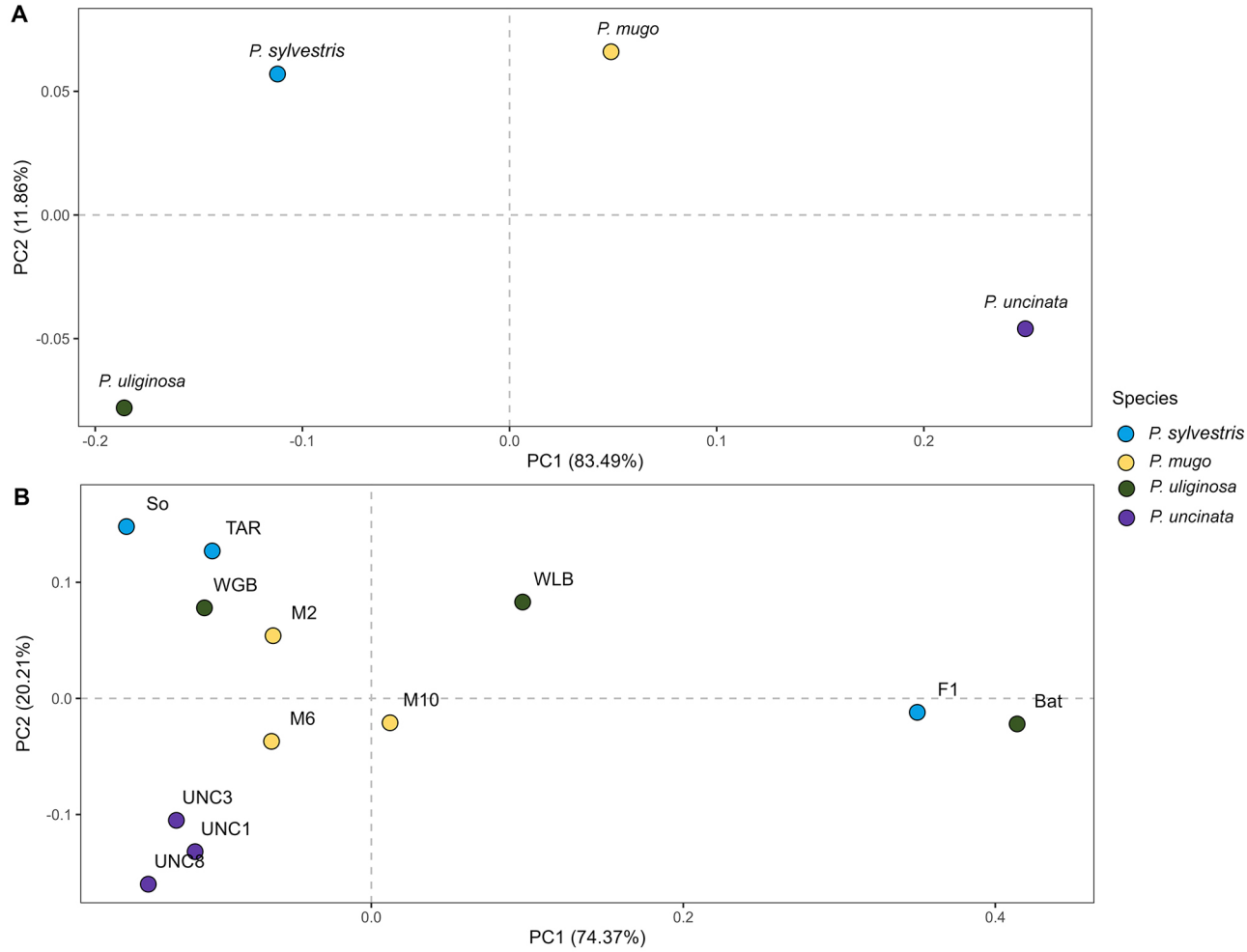

Fig. 4. Results of PCoA analysis at both species and population levels (A and B, respectively). Detailed information about populations is provided in Table S1 (Supplementary Material)

species, the marker could be used to track chloroplast DNA in contact zones in which *P. sylvestris* occurs together with other representative of taxa from the *P. mugo* complex. Consistent with the scenario of chloroplast capture during hybridization (Gernardt et al., 2018), we could expect hybrid trees of Scots pine morphology with V_A chloroplast or vice versa. Given the sympatric occurrences of *P. uliginosa* and *P. sylvestris* at Węgliniec and Węglowiec reserve, the *trn*L-*trn*F marker was added to Multiplex II to possibly detect such hybrids, but no such individuals were found. As the SNP markers assayed in our research have demonstrated high levels of genetic variability both within and between species, they are a valuable tool for studying population structure, hybridization and migration patterns of the investigated pine species.

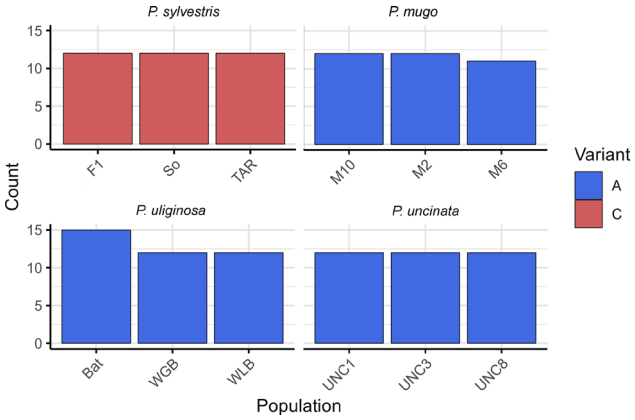

Fig. 5. Distribution of chloroplast DNA *trn*L-*trn*F marker variants among studied populations in four pine species. Detailed information about populations is provided in Table S1 (Supplementary Material)

Conclusions

We described here a protocol of fast and relatively cheap SNP genotyping of the organelle genome of closely related pine species. The assayed SNPs allow for intra- and interspecific investigation in studies focused on species discrimination and detailed investigations of their population history and structure. By simultaneously genotyping multiple SNPs, our

approach provides an efficient and cost-effective way of analyzing genetic variation at the population level. Additionally, by analyzing multiple SNPs, it provides a more comprehensive understanding of the genetic diversity within and between populations, which has significant implications for conservation biology and management practices. Considering its numerous advantages and high genotyping rate, the SNaPshot approach seems a very useful and convenient tool for population genetics studies of forest tree species that require a relatively small number of genotyped markers.

Acknowledgements

This work was financially supported by the Polish National Science Centre (Grant No. 2020/39/B/NZ9/00051). We acknowledge support from Adam Mickiewicz University in Poznań and Institute of Dendrology, Polish Academy of Sciences.

References

Asaf S, Khan AL, Khan MA, Shahzad R, Lubna, Kang SM, Al-Harrasi A, Al-Rawahi A & Lee IJ (2018) Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. PLoS ONE 13: e0192966. doi:10.1371/journal.pone.0192966.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA & Saunders NC (1987) Intraspecific phylogeography: The Mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489–522. doi:10.1146/annurev.es.18.110187.002421.

Bandelt HJ, Forster P & Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37–48. doi:10.1093/oxfordjournals.molbev.a026036.

Boratynska K, Gołąb Z, Łabiszak B, Niemczyk W, Sobierajska K, Ufnalski K, Wachowiak W & Boratynski A (2021) Are there any traces of *Pinus uliginosa* in the Stołowe Mountains outside the Wielkie Torfowisko Batorowskie and Błędne Skały?. Acta Societatis Botanicorum Poloniae 90: 904. doi:10.5586/asbp.904.

Cao HX, Vu GTH & Gailing O (2022) From genome sequencing to CRISPR-Based genome editing for climate-resilient forest trees. International Journal of Molecular Sciences 23: 966. doi:10.3390/ijms23020966.

Ceballos G, Ehrlich PR, Barnosky AD, Garcia A, Pringle RM & Palmer TM (2015) Accelerated modern human-induced species losses: Entering the sixth mass extinction. Science Advances 1: e1400253. doi:10.1126/sciadv.1400253.

Chung MY, Merila J, Li JL, Mao KS, Lopez-Pujol J, Tsumura Y & Chung MG (2023) Neutral and adaptive genetic diversity in plants: An overview. Frontiers in Ecology and Evolution 11: 1116814. doi:10.3389/fevo.2023.1116814.

Donnelly K, Cottrell J, Ennos RA, Vendramin GG, A'Hara S, King S, Perry A, Wachowiak W & Cavers S (2017) Reconstructing the plant mitochondrial genome for marker discovery: a case study using Pinus. Molecular Ecology Resources 17: 943–954. doi:10.1111/1755-0998.12646.

Farjon A (2018) Conifers of the world. Kew Bulletin 73: 8. doi:810.1007/S12225-018-9738-5.

Gernandt DS, Aguirre Dugua X, Vázquez-Lobo A, Willyard A, Moreno Letelier A, Pérez de la Rosa JA, Piñero D & Liston A (2018) Multi-locus phylogenetics, lineage sorting, and reticulation in *Pinus* subsection *Australes*. American Journal of Botany 105: 711–725. doi:10.1002/ajb2.1052.

Guo WH, Grewe F, Fan WS, Young GJ, Knoop V, Palmer JD & Mower JP (2016) *Ginkgo* and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. Molecular Biology and Evolution 33: 1448–1460. doi:10.1093/molbev/msw024.

Hashim HO & Al-Shuhaib MBS (2019) Exploring the potential and limitations of PCR-RFLP and PCR-SSCP for SNP detection: A review. Journal of Applied Biotechnology Reports 6: 137–144. doi:10.29252/JABR.06.04.02.

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68: 87–112. doi:10.1111/j.1095-8312.1999.tb01160.x.

Hijmans R & van Etten J (2012) raster: Geographic analysis and modeling with raster data. R package version 2.0-12. http://CRAN.R-project.org/package=raster.

Hu CT, Yan JW, Chen F, Zhang QX, Wang HD, Yin CY, Fan HT, Hu LL, Shen CM, Meng HT, Zhang YD, Wang H & Zhu BF (2016) Genetic analysis of 15 *mt*DNA SNP loci in Chinese Yi ethnic group using SNaPshot minisequencing. Gene 576: 105–108. doi:10.1016/j.gene.2015.09.071.

Jackman SD, Warren RL, Gibb EA, Vandervalk BP, Mohamadi H, Chu J, Raymond A, Pleasance S, Coope R, Wildung MR, Ritland CE, Bousquet J, Jones SJ, Bohlmann J & Birol I (2015) Organellar Genomes of white spruce (*Picea glauca*): Assembly and Annotation. Genome Biology and Evolution 81: 29–41. doi:10.1093/gbe/evv244.

Jackman SD, Coombe L, Warren RL, Kirk H, Trinh E, MacLeod T, Pleasance S, Pandoh P, Zhao Y, Coope RJ, Bousquet J, Bohlmann J, Jones SJM & Birol I (2020) Complete Mitochondrial Genome of a

Gymnosperm, sitka spruce (*Picea sitchensis*), indicates a complex physical structure. Genome Biology and Evolution 12: 1174–1179. doi:10.1093/gbe/evaa108.

Jaramillo-Correa JP & Bousquet J (2005) Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. Genetics 171: 1951–1962. doi:10.1534/genetics.105.042770.

Jasińska AK, Wachowiak W, Muchewicz E, Boratyńska K, Montserrat JM & Boratyński A (2010) Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. Botanical Journal of the Linnean Society 163: 473–485. doi:10.1111/j.1095-8339.2010.01065.x.

Keitt TH, Bivand R, Pebesma E & Rowlingson B (2011) rgdal: bindings for the Geospatial Data Abstraction Library. R package version 0.7-1, URL http://CRAN. R-project. org/package= rgdal.

Leigh JW & Bryant D (2015) popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6: 1110–1116. doi:10.1111/2041-210x.12410.

Librado P & Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452. doi:10.1093/bioinformatics/btp187.

Łabiszak B, Zaborowska J & Wachowiak W (2019) Patterns of *mt*DNA variation reveal complex evolutionary history of relict and endangered peat bog pine (*Pinus uliginosa*). AoB Plants 11: plz015. doi:10.1093/aobpla/plz015.

Morris AB & Shaw J (2018) Markers in time and space: A review of the last decade of plant phylogeographic approaches. Molecular Ecology 27: 2317–2333. doi:10.1111/mec.14695.

Naydenov K, Senneville S, Beaulieu J, Tremblay F & Bousquet J (2007) Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. BMC Evolutionary Biology 7: 233. doi:10.1186/1471-2148-7-233.

Neale DB, Martinez-Garcia PJ, De La Torre AR, Montanari S & Wei XX (2017) Novel insights into tree biology and genome evolution as revealed through genomics. Annual Review of Plant Biology 68: 457–483. doi:10.1146/annurev-arplant-042916-041049.

Nei M (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America 70: 3321–3323. doi:10.1073/pnas.70.12.3321.

Palmer JD (1992) Mitochondrial DNA in plant systematics: applications and limitations: Molecular systematics of plants (ed. by PS Soltis, DE Soltis & JJ Doyle) Springer US, Boston, MA, pp. 36–49.

Peakall R & Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288–295. doi:10.1111/j.1471-8286.2005.01155.x.

Peakall R & Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537–2539. doi:10.1093/bioinformatics/bts460.

Polezhaeva MA, Lascoux M & Semerikov VL (2010) Cytoplasmic DNA variation and biogeography of *Larix* Mill. in northeast Asia. Molecular Ecology 19: 1239–1252. doi:10.1111/j.1365-294X.2010.04552.x.

R Core Team R (2020) R: A language and environment for statistical computing.

Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE & Sanchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution 34: 3299–3302. doi:10.1093/molbev/msx248.

Semerikov VL & Lascoux M (2003) Nuclear and cytoplasmic variation within and between Eurasian *Larix* (Pinaceae) species. American Journal of Botany 90: 1113–1123. doi:10.3732/ajb.90.8.1113.

Senjo M, Kimura K, Watano Y, Ueda K & Shimizu T (1999) Extensive mitochondrial introgression from *Pinus pumila* to *P. parviflora* var. *pentaphylla* (Pinaceae). Journal of Plant Research 112: 97–105. doi:10.1007/PL00013867.

Smith DR (2016) The past, present and future of mitochondrial genomics: have we sequenced enough *mt*DNAs? Briefings in Functional Genomics 15: 47–54. doi:10.1093/bfgp/elv027.

Sobierajska K, Wachowiak W, Zaborowska J, Labiszak B, Wojkiewicz B, Sekiewicz M, Jasinska AK, Sekiewicz K, Boratynska K, Marcysiak K & Boratynski A (2020) Genetic consequences of hybridization in relict isolated trees *Pinus sylvestris* and *the Pinus mugo* complex. Forests 11: 1086. doi:1086 10.3390/f11101086.

Sokolowska J, Fuchs H & Celinski K (2021) New insight into taxonomy of european mountain Pines, *Pinus mugo* Complex, based on complete chloroplast genomes sequencing. Plants-Basel 10: 1331. doi:10.3390/plants10071331.

Soranzo N, Alia R, Provan J & Powell W (2000) Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. Molecular Ecology 9: 1205–1211. doi:10.1046/j.1365-294x.2000.00994.x.

Taberlet P, Gielly L, Pautou G & Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17: 1105–1109. doi:10.1007/BF00037152.

Taberlet P, Fumagalli L, Wust-Saucy AG & Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7: 453–464. doi:10.1046/j.1365-294x.1998.00289.x.

Tollefsrud MM, Latałowa M, van der Knaap WO, Brochmann C & Sperisen C (2015) Late Quaternary history of North Eurasian Norway spruce (*Picea abies*) and Siberian spruce (*Picea obovata*) inferred from macrofossils, pollen and cytoplasmic DNA variation. Journal of Biogeography 42: 1431–1442. doi:10.1111/jbi.12484.

Tóth EG, K öb ölkuti ZA, Pedryc A & H öhn M (2017) Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. Journal of Forestry Research 28: 637–651. doi:10.1007/s11676-017-0393-8.

Urban MC (2015) Accelerating extinction risk from climate change. Science 348: 571–573. doi:10.1126/science.aaa4984.

Wachowiak W, Lesniewicz K, Odrzykoski I, Augustyniak H & Prus-Głowacki W (2000) Species specific *cp*DNA markers useful for studies on the hybridisation between *Pinus mugo* and *P. sylvestris*. Acta Societatis Botanicorum Poloniae 69: 273–276. doi:10.5586/asbp.2000.036.

Wachowiak W, Zukowska WB, Wojkiewicz B, Cavers S & Litkowiec M (2016) Hybridization in contact zone between temperate European pine species. Tree Genetics & Genomes 12: 48. doi:10.1007/s11295-016-1007-x.

Wachowiak W, Perry A, Zaborowska J, González-Martínez SC & Cavers S (2022) Admixture and selection patterns across the European distribution of Scots pine, *Pinus sylvestris* (*Pinaceae*). Botanical Journal of the Linnean Society 200: 416–432. doi:10.1093/botlinnean/boac016.

Wachowiak W, Zukowska WB, Perry A, Lewandowski A, Cavers S & Labiszak B (2023) Phylogeography of Scots pine in Europe and Asia based on *mt*DNA polymorphisms. Journal of Systematics and Evolution 61: 315–327. doi:10.1111/jse.12907.

Weiler NEC, de Vries G & Sijen T (2016) Development of a control region-based *mt*DNA SNaPshot (TM) selection tool, integrated into a mini amplicon sequencing method. Science & Justice 56: 96–103. doi:10.1016/j.scijus.2015.11.003.

Wickham H (2011) ggplot2. Wiley interdisciplinary reviews-computational statistics 3: 180–185. doi:10.1002/wics.147.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K & Yutani H (2019). "Welcome to the tidyverse." Journal of Open Source Software 4: 1686. doi:10.21105/joss.01686.

Wolf C, Rentsch J & H übner P (1999) PCR-RFLP analysis of mitochondrial DNA: A reliable method for species identification. Journal of Agricultural and Food Chemistry 47: 1350–1355. doi:-DOI 10.1021/jf9808426.

Yu T, Jia ZY, Dayananda B, Li JQ, Guo XL, Shi L, Yuan XW & Gao Y (2022) Analysis of the chloroplast genomes of four *Pinus* species in Northeast China: Insights into hybrid speciation and identification of DNA molecular markers. Journal of Forestry Research 33: 1881–1890. doi:10.1007/s11676-021-01432-7.