Flora 220 (2016) 142–149

Contents lists available at ScienceDirect

Flora

journal homepage: www.elsevier.com/locate/flora

Substructuring of Scots pine in Europe based on polymorphism at chloroplast microsatellite loci

Błażej Wójkiewicz [a,*], Witold Wachowiak [a,b]

[a] Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, 62-035 Kórnik, Poland
[b] Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 29 February 2016
Accepted 7 March 2016
Edited by Karsten Wesche
Available online 11 March 2016

Keywords:
*cp*DNA
Genetic variation
Phylogeography
Population structure
Pinus sylvestris

ABSTRACT

Scots pine (*Pinus sylvestris* L.) is a forest forming tree species of great ecological and economic importance in the Northern Hemisphere. Assessment of genetic relationships between populations is needed for a better understanding of demographic and evolutionary processes of the species. We analyzed genetic variation and population structure of 24 native populations of Scots pine across its broad European distribution using a set of 13 polymorphic chloroplast simple-sequence repeats loci (*cp*SSRs). The number of alleles per locus ranged from 3 to 11 providing a set of 458 different haplotypes (haplotype diversity, $He = 0.982$). Genetic differentiation in most between-population comparisons was low (F_{ST} ranged from 0.00 to 0.08, R_{ST} ranged from 0.00 to 0.19) as expected for highly outcrossing and wind pollinated tree species. However, Bayesian clustering analysis revealed at least three distinct genetic clusters corresponding to geographical origin of the analyzed populations. Our results indicate genetic substructuring across the distribution range of Scots pine, and also point to the existence of populations of most likely different history that maintain their genetic identity.

1. Introduction

The current distribution of genetic diversity of plant species in Europe has been largely shaped during range shifts and population migration after the last glacial maximum (25–18,000 years ago) (Petit et al., 2003). Paleontological and phylogeographic data indicate that most of the northern regions of Europe were colonized from southern refugial areas in the Iberian Peninsula and Balkans, whereas Apennine Peninsula populations were more isolated due to the presence of the Alps as a main barrier (Taberlet et al., 1998). On the other hand, this simple south to north recolonization model may not hold for boreal forest tree species. Several lines of evidence in cold tolerant tree species such as silver birch (*Betula pendula*) and Norway spruce (*Picea abies*) highlight the potential role of populations from cryptic and higher latitude refugial areas in expansion to their present distribution ranges (Maliouchenko et al., 2007; Palme et al., 2003; Parducci et al., 2012; Tollefsrud et al., 2008; Wachowiak et al., 2011; Willis and van Andel, 2004). However, due to generally low phylogeographic resolution of most genetic markers used so far, the demographic history of conifer tree species especially in Central and Northern Europe remains unclear.

Gene flow is a major factor that shapes patterns of genetic variation between populations. In pine trees, the haploid chloroplast genome (*cp*DNA) is paternally inherited and transmitted by pollen. This mode of transmission has a homogenizing effect on genetic structure. Therefore, genetic signatures of population differentiation at neutral *cp*DNA markers could indicate their different origin and demographic history. Analysis of the geographical distribution of *cp*DNA variants can provide insights about locations of refugial areas, population history, genetic structure and level of gene flow between populations (Semerikov et al., 2014). Information about genetic relationships between populations is especially important for assessment of background genetic variation in local adaptation studies, as well as for forest management and conservation programs.

In the present study we focused on Scots pine (*Pinus sylvestris* L.) that has the largest Euro-Asian distribution range of all conifers. Previous studies of genetic diversity of Scots pine at isozymes or nucleotide sequence variation at genomic regions indicated low differentiation between populations (Pyhäjärvi et al., 2007; Wachowiak et al., 2009; Prus-Głowacki et al., 2012). Analysis of chloroplast microsatellite regions (SSRs) indicated no significant differences between Scotland and mainland European populations (Provan et al., 1998). Low differentiation at nuclear microsatellites was also observed between mountain isolates of Scots pine in Mediterranean refugial areas (Robledo-Arnuncio et al., 2005).

* Corresponding author.
E-mail address: bwojkiew@man.poznan.pl (B. Wójkiewicz).

http://dx.doi.org/10.1016/j.flora.2016.03.005
0367-2530/

However, as the previous chloroplast microsatellite studies were confined only to some areas of the species range, they could not have tested for broader relationships between populations in Europe.

In the present study we focused on Scots pine populations using a large collection of individuals (n = 676) sampled across the species distribution range in Europe. We genotyped the samples at thirteen variable chloroplast simple-sequence repeats (SSRs). Our aim was to assess genetic variation of the species across its European distribution and characterize relationships between Scots pine populations. We tested for signatures of population substructuring in relation to the species putative post glacial history across the study area. We provide information about background population differentiation and discussed our result in comparison to those obtained with other genetic markers at neutral loci. Such information on population structures is needed for studies of natural selection.

2. Material and methods

2.1. Plant material, DNA extraction and microsatellite genotyping

Needles from 676 individuals of Scots pine were collected from 24 populations (Table 1). The studied locations represent a wide distribution range of the species in Europe (Fig. 1). DNA was extracted following the standard CTAB protocol (Dumolin et al., 1995). Thirteen pairs of chloroplast microsatellites primers including Pt15169, Pt26081, Pt30204, Pt36480, Pt71936, PCP1289, PCP102652, PCP26106, PCP30277, PCP36567, PCP41131, PCP45071, PCP87314 were used in this study (Provan et al., 1998, 1999; Vendramin et al., 1996). DNA amplification by polymerase chain reaction (PCR) was carried out in two multiplex reactions for Pt and PCP primer sets, respectively. Individual-wise PCR reactions were conducted in a total volume of 10 μl containing 5 μl of Qiagen Multiplex Master Mix (2×), 0.2 μl of primer mix (20 μM), 1ul of Q-Solution (5×), 0.8 μl RNase-Free water, and 3 μl of template DNA (approximately 10-20 ng). The following PCR amplification conditions were used: initial denaturation at 95 °C for 15 min; 28 cycles of denaturation at 94 °C for 15 s, annealing at 58 °C for 1:30 min for multiplex 1, and at 60 °C for 1:30 min for multiplex 2, extension at 72 °C for 1:30 min and final extension at 72 °C for 10 min. The fluorescently labeled PCR products, along with a size standard (GeneScan 500 LIZ) were separated on a capillary sequencer ABI 3130 (Life Technologies, USA). The allele sizes were determined using GeneMapper software (ver. 4.0; Life Technologies, USA).

2.2. Data analysis

2.2.1. Allele and haplotype variation

The distribution of size variants among alleles in a given locus was checked using GenAlEx ver. 6 software (Peakall and Smouse, 2006). Haplotypes were determined as a combination of different microsatellite variants across the *cp*SSR loci. We used Haplotype Analysis ver. 1.05 software (Eliades and Eliades, 2009) to estimate haplotype frequencies and the following genetic diversity parameters: number of detected haplotypes in each population (A), number of private haplotypes (P), effective number of haplotypes (N_e), haplotypic richness (R_h; El Mousadik and Petit, 1996), genetic diversity (H_e) and mean genetic distance between individuals (D^2_{sh}; Goldstein et al., 1995). Regression analysis was used to infer the relationships between the proportion of private haplotypes in each population and their geographical location. The analyses were conducted using JMP® statistical software (SAS Institute Inc., Cary, NC, 1989–2007).

2.2.2. Genetic differentiation between populations

To test for genetic differentiation between populations, F_{ST} and R_{ST} statistics were estimated using Analysis of Molecular Variance (AMOVA) implemented in Arlequin ver. 3.5 software (Excoffier and Lischer, 2010). The inter-haplotype distance metric defined as the number of different *cp*SSR regions (for F_{ST}) and the sum of squared size differences (for R_{ST}) were used. Locus by locus analysis was performed to check the contribution of individual loci to differentiation between populations. Statistical significance of the results was evaluated using 10,000 random permutations of the samples between populations. Putative hybrid individuals (4 samples) carrying the size variants characteristic for *Pinus mugo* were excluded from the analysis (see Section 3 below).

To evaluate the influence of the stepwise mutation model (SMM) on differentiation among populations, the R_{ST} and permuted R_{ST} (pR_{ST}) values were compared using the test proposed by Hardy et al. (2003) and implemented in the SpaGeDi ver. 1.3d software (Hardy and Vekemans 2002). Comparison of these statistics allows identification of phylogeographic structure ($R_{ST} > pR_{ST}$). Moreover, we looked at the impact of isolation-by-distance (IBD) on the observed patterns of among-population differentiation (Rousset, 1997). To do so, we compared the matrix of pairwise geographic (logarithmic scale) and pairwise genetic distance (measured as $F_{ST}/(1\text{-}F_{ST})$) using the Mantel correlation test (Mantel, 1967). Statistical significance of the correlation was calculated for all populations and sets of populations located along latitudinal and longitudinal transects using 1000 permutations as implemented in GenAlEx ver. 6.

2.2.3. Population clustering and haplotype phylogenetic relationships

Population grouping and phylogenetic network analyses were done based on 8 loci which had significant impact on the among-population genetic differentiation (see Section 3 below). The genetic structure of populations was analyzed by Principal Coordinates Analysis (PCoA) based on Nei's genetic distance (Nei, 1972) using GenAlEx ver. 6 software. The genetic relationships between populations were further explored using Bayesian clustering methods. Generally, such methods provide an estimate of the number of distinct populations and their genetic structure. In this study, the population clustering was performed based on algorithms implemented in BAPS ver. 6.0 (Corander et al., 2008) and GENELAND ver. 3.2.4 software (Guillot, 2008). Following the results of the Bayesian clustering approaches, the populations were assigned to geographical groups to estimate their genetic differentiation using AMOVA (10,000 permutations) as implemented in Arlequin ver. 3.5 software. Moreover, the pairwise F_{ST} values were estimated between these groups of populations. Statistical significance of the results was evaluated using 10,000 random permutations.

Finally, a phylogenetic network of the established *cp*SSR haplotypes was constructed based on the median-joining method to check the haplotype relationships. We looked for signs of geographical affiliation of haplotypes (Bandelt et al., 1999) using maximum-parsimony analysis as implemented in NETWORK ver. 4.6.1.2 software (Fluxus Technology Ltd, http://www.fluxusengineering.com).

3. Results

3.1. Allele and haplotype variation

3.1.1. Allelic patterns found in Scots pine populations

All thirteen *cp*SSR loci were polymorphic providing a total of 46 size variants (3–11 per locus). Twenty one alleles were private for particular populations from Turkey (T1, T4), Greece (G), Spain (H1, H2, A) and Scotland (SC). Loci Pt30204, Pt71936, PCP30277 and