B-SAP Markers Derived from the Bacterial *KatG* Gene Differentiate Populations of *Pinus sylvestris* and Provide New Insights into Their Postglacial History

Piotr Androsiuk, Roman Zielinski & Kornelia Polok

The aim of the studies was to evaluate the efficiency of the *KatG* gene based B-SAP markers as a tool to distinguish morphologically diversified and geographically distant Scots pine populations and to track the routes of migrations. The 19 populations growing in the IUFRO 1982 provenance experiment and representing the natural distribution of the species in Europe were scored using 103 B-SAP loci. Among them 26% loci were polymorphic. The level of polymorphism was associated with the location of primers on the *KatG* template. The diversity was low, $H_e = 0.086$, and deposited mostly among populations. Seven unique markers were found that identified populations and likely they were associated with morphology. The overall genetic identity was relatively low, $I = 0.933$ ($D = 0.069$). The block of six B-SAP markers discriminated populations into two groups in agreement with their geographic origin and thereby further described as the North and the South. The North group was uniform with genetic diversity, $H_e = 0.026$ and the overall genetic distance $D = 0.022$. Presumably, it migrated from refugia in the Alps via France, northern Germany and Denmark, to Scandinavia and Russia. The South group was heterogeneous with $H_e = 0.063$ and $D = 0.047$. This group migrated from the Carpathians via Slovakia to Germany and Poland. The Balkans and Asian refugia did not take part in recolonization of Europe. The block of six B-SAP/*KatG* markers can be recommended for tracking postglacial history of Scots pine.

1   Introduction

Scots pine (*Pinus sylvestris* L.) is an important component of the boreal forests across Europe and Asia where it shows strong dominance on the xeric slopes and sandy soils. Besides having major economic values, it also influences hydrological, fire regimes, provides food and create habitat for animals, and plays a significant role in determining climate. The distribution of *P. sylvestris* was mostly shaped in the Holocene when the species retreated to its northern limits due to climatic cooling as well as anthropogenic activity (Willis et al. 1998). Scots pine ranges from Scottish Highlands to northern Mediterranean Basin through Atlantic and Pacific coast of eastern Siberia with extreme environments such as tundra, bogs and mountains and belongs to the most widespread species of any pine. It is not surprising therefore, that it forms a variety of geographical races differing in growth rate, foliage colour, needle characters, winter survival, wood gravity, susceptibility to diseases and many others (Wright et al. 1966, Ruby 1967, Alía et al. 2001). Many of these traits vary continuously with latitude and elevation, and are correlated with climatic conditions (Oleksyn et al. 1998, Andersson and Fedorkov 2004). Along with this unusual morphological diversity, the genetic diversity within populations is also very high. With more than 90% of polymorphic loci, gene diversity (H) ranging from 0.282 for isoenzymes up to 0.740 in a case of microsatellites and more than four enzymatic alleles per locus (Goncharenko et al. 1994, Karhu et al. 1996), *P. sylvestris* belongs to the most variable organisms.

High genetic diversity makes Scots pine ideal material for breeding and conservation studies. However, effective manipulation often depends on molecular tools for marker-assisted selection of economically or adaptively important quantitative traits. Unfortunately, the great genetic variation of *P. sylvestris* is generally observed within populations while differences between populations are small. Populations from distant geographical locations have the same gene pool with the same alleles, frequencies of which are also similar as measured on both izoenzyme (Goncharenko et al. 1994, Hertel and Schneck 1999) and DNA levels, using different molecular markers such as

RFLP, RAPD, SSR, rDNA and low copy DNA, (Karhu et al. 1996). Diagnostic alleles unique for a population are rarely found. The uniformity of populations from the wide geographical range are exemplified by low Nei's genetic distance coefficients ranging from 0.005 to 0.056 (Goncharenko et al. 1994, Sannikov et al. 2005). Similarly, correlations between morphological characters and genetic markers have rarely been found. Among a few examples of such relationships are the *Lap-A* allele typical of thicker Scots pine trees (Blumenrother et al. 2001) and seven DNA markers specific to the *turfosa* phenotype characterized by a short stature, curve log and "umbrella" like crown and inhabiting Polish peat-bogs (Polok et al. 2005a)

In general, however predicting the differentiation of morphological or adaptive traits based on universal molecular markers is a kind of a standstill. First, these traits are quantitative and thus encoded by many genes. Second, such predictions can only be possible if a relatively tight linkage exists between a marker and a trait of interest. A third complication is the very large Scots pine genome consisting for the most part of repetitive sequences. Consequently, the great majority of molecular markers target non-coding sequences. Thus, revealed patterns of variation do not reflect morphological differences among populations. A somehow different proposal is to survey in the taxa under considerations a growing number of sequences identified in plant genome projects and known to be related with morphological or adapted traits in related species. For example, the *pal1* gene from *Pinus taeda* encoding phenylalanine ammonia-lyase - a key enzyme in the secondary metabolism, important in wood formation, ozone tolerance and defence against pathogens, was studied in four *P. sylvestris* populations from Finland, Russia and Spain (Dvornyk et al. 2002). Unfortunately, the low nucleotide diversity (0.0056) and low divergence between geographically distant populations have confirmed earlier molecular observations.

In the present studies, we propose to employ new molecular markers based on the bacterial *KatG* gene as a tool to differentiate European populations of Scots pine. The bacterial *KatG* gene encodes the catalase-peroxidase belonging to the class I of the plant peroxidases superfamily.

The family evolved from a common prokaryotic ancestor through multiple duplications, conversions and translocations. However, proteins still share the sequence similarity up to 76% (Guan and Scandalios 1996, Zamocky 2004). Some members of the family are conserved within many groups while the others have evolved after the taxa differentiation. Therefore, a number of universal pairs of primers can generate fingerprints differentiating taxa on different levels of evolutionary relationships (Zielinski and Polok 2005). The *KatG* based approach has been effective in resolving phylogenetic relationships within the genus *Lolium* and related members of "Core Pooids". The homology of generated fragments to plant peroxidases has been confirmed by the tight linkage between the *Per3* locus encoding peroxidase and *KatG*-based markers on *L. perenne* x *L. multiflorum* genetic map (Polok 2007). Sequencing of selected B-SAP (Bacteria Specific Amplification Polymorphism) fragments revealed from 49% to 100% identity to plant peroxidases at both nucleotide and protein levels (Polok and Zielinski 2009). The revealed markers inherit as dominant and they are ascribed as B-SAP because primers are designed on bacterial sequences. The B-SAP markers are especially useful for species delimitation as exemplified by plenty of species specific markers differentiating three species of *Polygonatum* (Szczecinska et al. 2006), four cryptic species of *Aneura pinguis* (Baczkiewicz et al. 2008) and five *Sphagnum* species (Sawicki and Zielinski 2008). Among pines, B-SAP markers were used to assess genetic identity among *Pinus cembra* populations from the Tatras (Chmiel et al. 2008) as well as between *P. sylvestris* and *Pinus mugo* (Zielinski and Polok 2005). A further point of evidence lies in the linkage between some *KatG*-based B-SAP markers and QTLs involved in domestication of *L. perenne* and *L. multiflorum* (Polok 2007). This relationship makes chance to differentiate pine populations contrasting in adaptive and morphological features. No less important is easy use, and relatively low costs of the system in comparison with more sophisticated ones such as AFLP or SSAP. Thus, our aim was to evaluate *KatG*-based B-SAP markers as a tool to distinguish morphologically diversified Scots pine populations of distant geographic origin. We studied the level of variation within the species and how the B-SAP grouping correlates with morphological and geographical diversity as well as we searched for population specific markers. We discussed the implication of results to the studies on genetic diversity of Scots pine and its postglacial history.

2 Material and Methods

2.1 Plant Material

In total 19 populations of *P. sylvestris* representing the natural distribution of the species along the transect running across Europe were analysed (Fig. 1). The trees originated from the area between 40°N–60°N latitude, 4°E–33°E longitude and 40 m –1400 m altitude (Table 1). Material was provided by the Institute of Dendrology, Polish Academy of Sciences (Jacek Oleksyn), where plants grew in the IUFRO 1982 provenance experiment at the Zwierzyniec experimental forest near Kórnik, Poland (52°15´N, 17°04´E). The trial is static in that genetic diversity is fixed. Trees are planted in a randomized block design with seven blocks and three to seven replications of each provenance. The identity of the stands with the population of origin is maintained through disabling the tree reproduction within the frame of the trial. To secure the physical stability of the stand, a careful silvacultural practices have been used plus the individual identity has been kept. Significant losses of trees have not been observed. Moreover, to ensure that the samples are representative of the populations of origin, to avoid the risk of hybridization and adaptation that may influence offspring, needles were collected from original trees at the age of 24 years. These trees originated from seeds sampled from native populations according to the IUFRO instruction. Needles from ten trees per population were sampled. After collection, needles were washed using deionized water, then white spirit (Shell) to remove the resin, and finally surface sterilized with 95% ethanol and frozen in –20 °C until use.
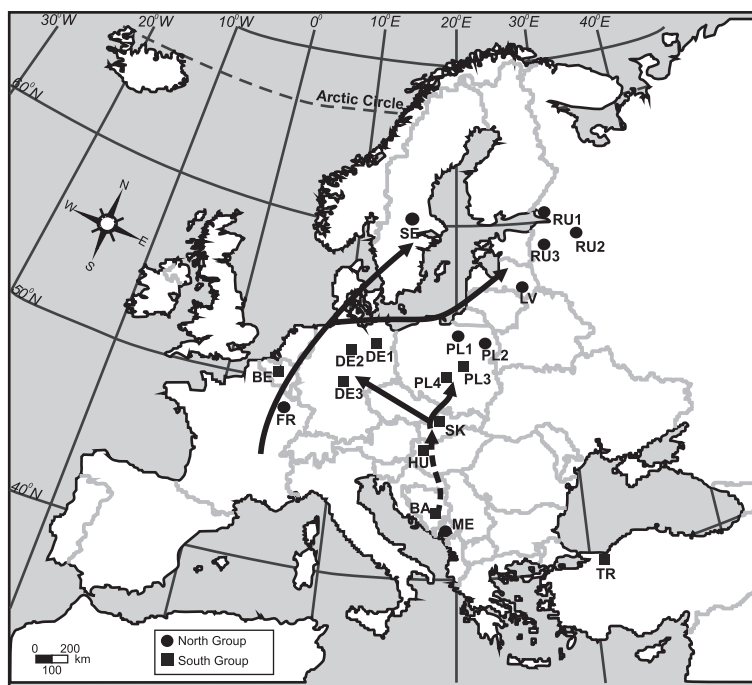
Fig. 1. Geographical location of studied *Pinus sylvestris* populations. Acronyms refer
to the populations as described in the Table 1. Solid arrows indicate migration
routes based on B-SAPs, a dotted arrow indicates an alternative migration route
from Balkan refugia.

Table 1. Populations of *Pinus sylvestris* analysed for B-SAP polymorphism and their affiliation based on a block
of six B-SAP markers [a].

Acronym[b]	Country	Population origin	Coordinates	Elevation [m a.s.l.]	Group[a]
SE	Sweden	Sumpberget, Hedemora	60°11´N, 15°52´E	185	North
RU1	Russia	Roshchinskaya Dacha	60°15´N, 29°54´E	80	North
RU2	Russia	Kondezhkoe	59°58´N, 33°30´E	70	North
RU3	Russia	Serebryanskoe	58°50´N, 29°07´E	80	North
LV	Latvia	Silene	55°45´N, 26°40´E	165	North
PL1	Poland	Milomlyn, Pilawki	53°34´N, 20°00´E	110	North
PL2	Poland	Suprasl, Krasne	53°12´N, 23°22´E	160	North
PL3	Poland	Spala, Malomierz	51°37´N, 26°40´E	160	South
PL4	Poland	Rychtal, Sadogora	51°08´N, 17°55´E	190	South
DE1	Germany	Neuhaus, Wollenz	53°02´N, 13°54´E	40	South
DE2	Germany	Betzhorn	52°30´N, 10°30´E	65	South
DE3	Germany	Lampertheim	50°00´N, 10°00´E	100	South
BE	Belgium	Ardennes, Groenendaal	50°46´N, 04°26´E	110	South
FR	France	Haguenau, Alsace, Plain	48°49´N, 07°46´E	150	North
SK	Slovakia	Zahorie	48°46´N, 17°03´E	160	South
HU	Hungary	Pornoapati	47°20´N, 16°28´E	300	South
ME	Montenegro	Maocnica, Pljevla	43°10´N, 19°30´E	1200	North
BA	Bosnia	Prusacka Rijeka	44°06´N, 17°21´E	885	South
TR	Turkey	Catacik	40°00´N, 31°30´E	1400	South

[a] B-SAP block comprises *KatG2-8, KatG9-6, KatG9-7, KatG9-9, KatG2-11* and *KatG10-8;* [b] Country codes according to ISO3166.

2.2 DNA Isolation

Total DNA was extracted from 1.5 g of needles adopting the modified CTAB procedure (Polok 2007). Briefly, needles ground in liquid nitrogen were transferred to tubes with 2 ml of CTAB buffer (2% CTAB, 100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% β-mercaptoethanol) preheated to 60 °C. Then, following reagents, all preaheated to 60 °C, were added – 1 ml of 20% CTAB, 1 ml of 10% PVP, 330 μl of 30% sarcosyl. Samples were incubated at 60 °C for 2 h. DNA was precipitated after three chloroform-isoamyl alcohol (24:1) extractions, then DNA was washed and dissolved in sterile, deionized H_2O. RNA was removed with RNA-ase at a final concentration 200 μg/ml. The quality of DNA was verified on 1% agarose while the purity was assessed spectrophotometrically and it ranged between 89% and 95%. The DNA content of the samples ranged from 290 μg to 959 μg.

2.3 B-SAP Markers Based on the *KatG* Gene

Twelve pairs of primers were designed using the 4801 bp *KatG* gene sequence from *Mycobacterium tuberculosis,* in the NCBI file, accession N° X68081.1. The well-known *KatG* structure, the presence of both conservative and highly variable domains, and usefulness in strain identification were behind the reasons of using this gene. Pairs of primers were distributed every 250–300 bp from 1514 to 4272 bp, so the coding part of the *KatG* gene (1797–4201) was targeted (Table 2). The respective protein NCBI file, N°CAA48213 was used to describe the putative function of fragments flanking by each primer pair. The

Table 2. Characteristics of B-SAP primers based on the *M. tuberculosis KatG* gene and PCR conditions.

Primer pair	Sequence (5'→3')[a]	5' position on *KatG* template in the NCBI file [bp]	Annealing temperature [°C]	Number of PCR cycles [n]
*KatG*1	**F:** GACTACGCCCAACAGCTCC	1 514	53	30
	R: GCGATAACCCCGCAAGACC	1 781		
*KatG*2	**F:** GCGGGGTTATCGCCG ATG	1 787	53	35
	R: GCCCTCGACGGGGTATTTC	2 056		
*KatG*3	**F:** AACGGCTGTCCCGTCGTG	2 030	53	30
	R: GTCGTGGATGCGGTAGGTG	2 311		
*KatG*4	**F:** CAGGTCCGCCCATGAGAG	2 453	53	30
	R: TCGACTTGACGCCCTGACG	2 753		
*KatG*5	**F:** CGACAACGCCAGCTTGGAC	2 386	53	35
	R: GGTTCACGTAGATCAGCCCC	2 653		
*KatG*6	**F:** GCAGATGGGGCTGATCTACG	2 647	53	35
	R: ACCTCGATGCCGCTGGTG	2 917		
*KatG*7	**F:** GCTGGAGCAGATGGGCTTG	2 854	52	30
	R: ATCCACCCGCAGCGAGAG	3 122		
*KatG*8	**F:** GTCACTGACCTCTCGCTG	3 094	50	30
	R: CGCCCATGCGGTCGAAAC	3 395		
*KatG*9	**F:** GCGAAGCAGATTGCCAGCC	3 408	50	30
	R: ACAGCCACCGAGCACGAC	3 608		
*KatG*10	**F:** CAAAGTGTCCTTCGCCGACC	3 586	51	35
	R: CACCTACCAGCACCGTCATC	3 847		
*KatG*11	**F:** TGCTCGACAAGGAGAACCTG	3 807	53	35
	R: TCCGAGTTGGACCCGAAGAC	4 067		
*KatG*12	**F:** TACCAGGGCAAGGATGGTAG	4 010	53	30
	R: GCAAACACCAGCACCCCG	4 272		

[a]F – forward primer, R – reverse primer.

PCR reaction was optimized in earlier studies (Zielinski and Polok 2005, Polok 2007). Each PCR reaction was performed in a 20 µl volume containing 20 mM $(NH_4)_2SO_4$, 100 ng DNA, 50 mM Tris-HCl, pH 9.0 at 25 °C, 1.5 mM $MgCl_2$, 2 µl of Enhancer with betaine (Epicentre Technology), 200 µM dNTPs, 1.0 µM primers, 0.75 U of *Tfl* polymerase (Epicentre Technology) and 100 ng of template DNA. The PCR programme for amplification of B-SAP markers based on the *KatG* gene was 94 °C for 3 min followed by 30–35 cycles at 94 °C for 1 min, 50–53 °C for 1 min, 72 °C for 1.5 min. Parameters for each pair of primers are shown in Table 2. The final step was 5 min at 72 °C. PCR reactions were done twice for each sample as well as using different DNA isolates of the same sample. PCR products were loaded on 1.5% (w/v) agarose gel containing 0.5 µg/ml ethidium bromide, separated in 1 x TBE buffer (Tris-Borate-EDTA) at 100 V constant power, visualized under UV light (312 nm), photographed with Olympus Camera and stored as .jpg files.

2.4 Data Analysis

All bands that could be reliably read were treated as independent loci, and scored either present (1) or absent (0). The polymorphic products were referred as e.g., *KatG1-2* when referring to a band revealed by a set of *KatG* primers N°1 and the second product from the anode. Basic genetic parameters were calculated using the POPGENE 1.32 software (Yeh et al. 2000). They include mean number of alleles per locus (N_a), effective number of alleles (N_e), percentage of polymorphic loci (P), Nei's gene diversity statistics i.e., gene diversity (H_e), total diversity (H_T) and average gene diversity (H_s). Diversity was also measured in categorical data using Shannon's diversity (H_o) as well as by F_{ST} statistics. The deviation of F_{ST} from zero was calculated by t-test. For a relatively large number of alleles the distribution of parameters is approximately normal, and the ordinary statistics can be used for comparisons of different populations. The statistical significance of population genetic parameters was tested by analysis of variance with LSD test in STATISTICA 9.0 software. Genetic identities

and distance were determined with the Nei and Li formula (Nei and Li 1979). The matrix of dissimilarities was used in the cluster analysis with the UPGMA (Unweighted pair-group using arithmetic mean) for the amalgamation. The Euclidean distances derived from the matrix of similarities were used in the multidimensional scaling in the STATISTICA 9.0. Pairwise geographic distances were computed as great-circle distances from the geographic co-ordinates and then transformed to natural logarithms. The distance matrices (genetic and geographic) were compared by the Mantel test with 10 000 permutations using the MANTEL software and significance level, P = 0.05 (Cavalcanti 2008).

3 Results

3.1 Efficiency of B-SAP Markers Based on the *KatG* Gene

Twelve pairs of primers generated in total 103 reproducible bands of mobility ranging from 200 bp to 2000 bp (Table 3). The number of bands observed per a primer pair ranged from 6 to 11 with an average of 8.6 bands. Among them 27 bands (26%) were polymorphic. As expected, differences were observed between *KatG* pairs of primers used. The pair *KatG*2 detected the highest number of polymorphic bands (five bands) while the *KatG*8 with one band was at the opposite extreme. Three pairs of primers, *KatG*4, *KatG*6 and *KatG*7 detected no polymorphic bands. Similarly, various pairs of primers revealed various degrees of genetic diversities. Gene diversities, H_s ranged from 0.031 to 0.137 with a mean value of 0.047. The highest diversity was attributed to *KatG*5 and *KatG*11. Total diversity ranged from 0.038 to 0.174 with an average 0.114. The markers generated by *KatG*2, *KatG*5 and *KatG*9 (H_T from 0.161 to 0.174) showed the highest contribution to the total variation and were responsible for dividing a total population into two groups (Fig. 2).

Interestingly, these differences between levels of genetic variation revealed by *KatG* primer pairs seem not to have been random, but rather associated with the primers' location on the *KatG*
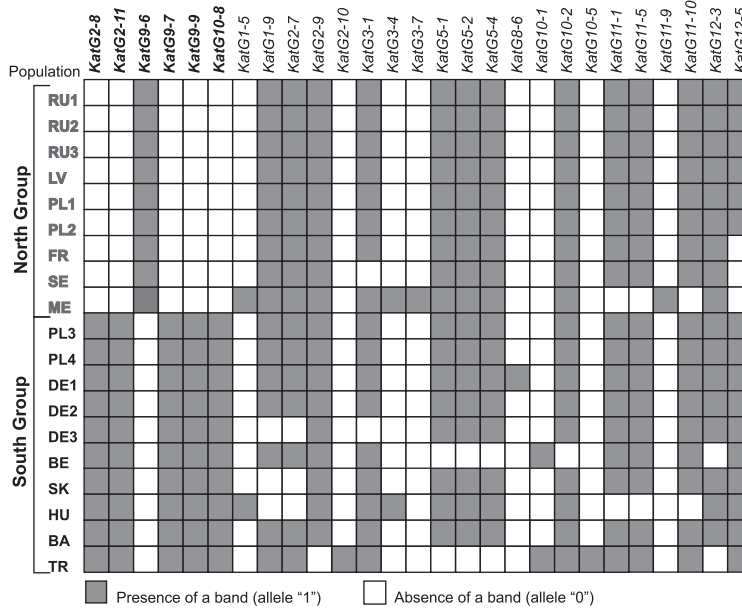
Fig. 2. Summary of B-SAP fingerprints revealed by primers complementary to the *KatG* gene. Bold indicates six B-SAP markers dividing populations into two groups.

Table 3. Variation generated by B-SAP markers in *Pinus sylvestris*. N_a = mean number of alleles, P = percentage of polymorphic loci, H_S = average gene diversity, H_T = total gene diversity.

Primer pair	Putative function of *KatG* region[a]	Fragment size range [bp]	N° of bands	N° of polymorphic bands	N_a	P	H_S	H_T
*KatG*1	N	350–1900	10	2	1.20	20	0.035	0.050
*KatG*2	N, CDS	570–1800	11	5	1.45	45	0.042	0.161
*KatG*3	N, CDS, PX	200–1950	8	3	1.38	38	0.074	0.075
Mean for *KatG*1–*KatG*3			9.7	3.3	1.34a	34	0.050a	0.095a
*KatG*4	CDS, PX, A, B	250–1650	9	0	1.00	0	0.000	0.000
*KatG*5	CDS, PX	350–1800	6	3	1.50	50	0.124	0.174
*KatG*6	CDS, PX, A, B	300–1750	9	0	1.00	0	0.000	0.000
*KatG*7	CDS, PX, A		9	0	1.00	0	0.000	0.000
*KatG*8	CDS	350–1450	7	1	1.14	14	0.031	0.038
*KatG*9	CDS, PX, A, B	400–1800	9	3	1.33	33	0.000	0.166
Mean for *KatG*4–*KatG*9			8.2	1.2	1.16b	16	0.026b	0.063b
*KatG*10	C, CDS, PX	350–1300	9	4	1.44	44	0.039	0.039
*KatG*11	C, CDS, PX, A	400–1800	10	4	1.40	40	0.137	0.137
*KatG*12	C, CDS, PX	750–2000	6	2	1.33	33	0.082	0.082
Mean for *KatG*10–*KatG*12			8.3	3.3	1.40a	39	0.088c	0.086a
Mean for *KatG*1–*KatG*12			8.6	2.3	1.26	26.4	0.047	0.114
Total			103	27		26.2		

[a] N – N-terminal part of putative protein, C – C-terminal part, CDS – coding region, PX – regions homologous to plant peroxidases, A – active site, B – binding pocket. Statistically significant values between means for *KatG* regions at P=0.05 are indicated by different letters.

gene template (Table 3). In general, statistically higher values of parameters were obtained for primers designed on both the ends of the *KatG* template. For instance, *KatG*1, *KatG*2 and *KatG*3 pairs located on the N terminal part of a putative protein and those on the C-terminal part, *KatG*10, *KatG*11, *KatG*12, produced on average 3.3 (37%) of polymorphic loci, which is twofold more than in a case of primers from the "middle" part of the gene (1.2 polymorphic loci, 16%). All primers detecting no polymorphism were located in the "middle" part of the *KatG* template. Likely, mutations responsible for changes in banding patterns are more frequent in variable N and C terminal parts, not determining functional or structural properties of *KatG* derivatives. The lack of active sites and binding pockets in fragments of the *KatG* template flanked by *KatG*1–*KatG*3 (N-terminal) and *KatG*10–*KatG*12 (C-terminal) primers confirmed the view. By contrary, such structural properties were predicted in the majority of "middle" fragments of the *KatG* gene, flanked by *KatG*4–*KatG*9. Furthermore, active sites were predicted in *KatG* parts flanked by all three primer pairs producing monomorphic loci and two of them also flanked regions with a binding pocket.

3.2 B-SAP Polymorphism in *Pinus sylvestris*

The relatively low overall proportion of polymorphic bands (26%) confirmed the conservative character of the B-SAP system. Genetic diversities measured both by Nei's diversity ($H_e = 0.086$) and the Shannon index ($H_o = 0.128$) were also lower than expected for highly variable species such as *P. sylvestris* (Table 4). However, the majority of this variation was deposited among populations as estimated from the high F_{ST} value (0.473). This entails the possibility of identifying markers specific to one or more populations. Notwithstanding the majority of *P. sylvestris* populations shared the same bands, seven unique alleles were found (Fig. 2, Fig. 3). Three of them were typical of the Turkish population (TR), two of the Montenegro population (ME) and one each in the Belgian (BE) and the German (DE1) populations. In addition, 11 markers were observed in two or three populations (semi-diagnostic) thus, in combination they allowed population discrimination. For example, at *KatG1-5* and *KatG3-4* loci a band was amplified only in the Hungarian (HU) and the Montenegro (ME) populations. At another locus, *KatG3-1* the band was lacking only in the German (DE3) and the Swedish (SE) populations.

Another line of evidence that B-SAP markers are effective in population discrimination was provided by a block of six loci that divided populations into two groups mostly in agreement with their geographic origin and thereby further described as the North and the South (Fig. 2). The diagnostic loci involved all polymorphic loci identified by *KatG*9 primers, two by *KatG*2 primers, *KatG2-8, KatG2-11* in addition to one revealed by *KatG*10, *KatG10-8* (Fig. 3). Both B-SAP groups differed significantly with respect to most genetic variation parameters (Table 4).

Table 4. Genetic variation parameters summarized for total *P. sylvestris* population and two groups, North and South, identified based on B-SAP markers. N_a = mean number of alleles, N_e = effective number of alleles, P = percentage of polymorphic loci, H_e = gene diversity, H_o = Shannon's diversity index, F_{ST} = fixation index, I = mean similarity coefficient (Nei and Li 1979). For I values, "total" means a mean identity coefficient calculated based on all pair-wise comparisons, "North-South" means the I value between the North and the South group of populations.

Group	N_a	N_e	P	H_e	H_o	F_{ST}	I
Total	1.26	1.15	26.2	0.086	0.128	0.473	0.933
North	1.09	1.05a	8.7	0.026a	0.039a	0.694a	0.978a
South	1.17	1.11b	17.5	0.063b	0.093b	0.258b	0.954b
North-South	-	-	-	-	-	-	0.903

Statistically significant values at P = 0.05 for the North and South groups are indicated by different letters.
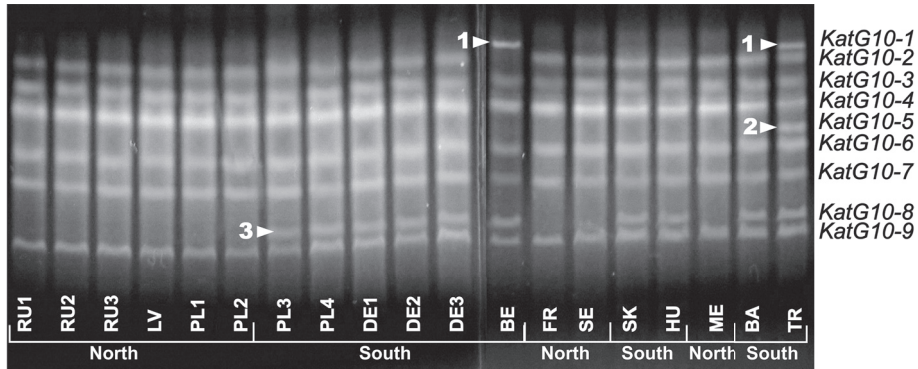
Fig. 3. Amplification patterns revealed by *KatG*10 primers in populations of *P. sylvestris*. Arrows and numbers correspond to 1 – semi-diagnostic alleles, 2 – unique alleles, 3 – alleles included in a block of six markers dividing populations into two groups, North and South.

The geographical separation between the North and South group run right across the Polish territory (Fig. 1) in the way that populations from the North-East (PL1, PL2) belonged to the North group while from Central Poland to the South group (PL3, PL4). The first group was described as the North due to mostly northeastern origin of comprising populations, i.e. from Russia, Latvia and North-Eastern Poland. This group was characterized by the lack of amplification products at all but *KatG9-6* diagnostic loci mentioned above (Fig. 2). Classification of the French and Montenegro populations into the North group was somehow unexpected. The North group was uniform in its B-SAP fingerprints with the exception of populations from France (FR), Sweden (SE) and Montenegro (ME). Twofold lower values of genetic variation parameters (N_e, P, H_e, H_o) in relation to the South group confirmed this view. Conversely, ten populations classified as the South group had bands in five diagnostic loci and were relatively variable as exemplified by twofold higher genetic variation parameters, unique markers and marker combinations enabling to recognize the majority of populations. These populations originated from Central Poland, Germany, Slovakia, Belgium, Bosnia and Turkey. Somehow intermediate was the Hungarian population, which despite the B-SAP pattern typical of the South group, still shared some alleles with the Montenegro population from the North group.

3.3 Genetic Similarity and Distance of *P. sylvestris* Populations, Correlation with Geographic Distance

Notwithstanding the conservative character of B-SAP markers, the average genetic similarity among populations, I=0.933 (D=0.069) was quite low as for *P. sylvestris*. As expected the Nei's identities varied among populations with the lowest value (I=0.777, D=0.252) between the Montenegro (ME) and the Turkey (TR) populations while the highest reaching I=1.00 was typical of all Russian, Latvian and two Polish populations (PL1, PL2). Considering the North and South group separately, Nei's coefficients confirmed geographical structuring and higher uniformity of the former (Table 4, Fig. 4). Genetic similarities among populations from the North group ranged from 0.922 (D=0.080) to 1.000 (D=0.0) with the average of 0.978 (D=0.022) while the respective values for the South group were lower, from 0.874 (D=0.135) to 1.000 with an average of 0.954 (D=0.047).

Cluster analysis and multidimensional scaling showed clearly the ability of the B-SAP markers to discriminate populations into two groups, classified as the North and the South. In the UPGMA dendrogram (Fig. 4) these two groups formed separate clusters joined at I=0.903 (D=0.102). Besides the North and South major groups, smaller clusters were found within them. In the North group, indistinguishable populations
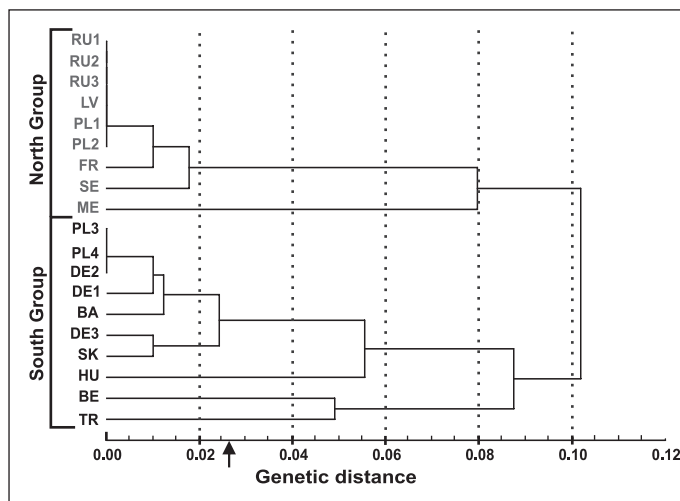
11

Fig. 4. UPGMA grouping of *P. sylvestris* populations based on B-SAP/*KatG* markers.

from Russia, Latvia and Poland formed a cluster with French (I=0.990, D=0.01) and Swedish (I=0.981, D=0.020) populations and then joined with the Montenegro population, which however, formed a separate branch. Somehow, distinct position of this population was also demonstrated in the plot based on three principal coordinate axes in multidimensional scaling (Fig. 5). The South group was represented in the UPGMA tree (Fig. 4) by a larger, main cluster of two Central Polish (PL3, PL4), all German (DE1, DE2, DE3) in addition to Slovakia (SK) and Bosnia (BA) populations. The Hungarian population formed a separate branch, however merged with the main cluster. Belgian (BE) and Turkish (TR) populations remained independent of that cluster. Likewise, multidimensional scaling (Fig. 5) unequivocally demonstrated a distinct position of three populations originated from Hungary (HU), Belgium (BE) and Turkey (TR).

Significant correlation between genetic and geographical distances (r=0.289, p=0.009) provided an evidence that B-SAP markers differentiate populations with agreement to their geographic origin. In general, the more geographically distant populations the bigger genetic differences were found. Stronger correlation was typical of the North group (r=0.428, p=0.02), in which all
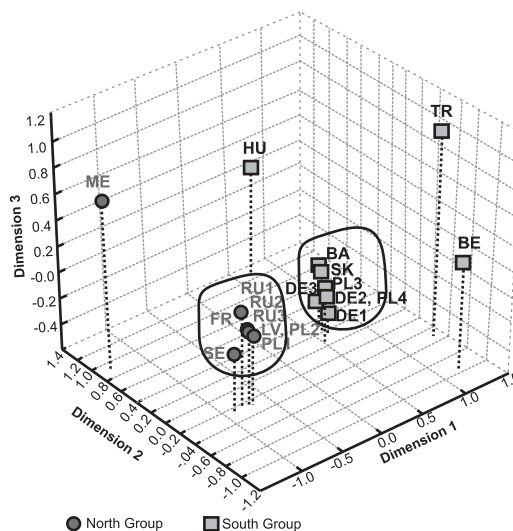

Fig. 5. Multidimensional scaling of *P. sylvestris* populations based on B-SAP/*KatG* markers.

pairwise distances were also significant. Slightly weaker relationships in the South group (r=0.384, p=0.04) resulted from the lack of significant correlations between Belgian, Slovak and Hungarian populations and the rest seven populations comprising the South group.

4 Discussion

Spatial diversity, mating system along with seed dispersal and gene flow favour the creation of genetic variation in *P. sylvestris*. This is emphasized by very high genetic diversity on the isoenzyme level ranging from $H_e = 0.211$ to $H_e = 0.365$ (Ledig 1998). With this respect, the low values of H ($H_e = 0.086$, $H_s = 0.047$, $H_o = 0.128$) obtained in our studies by means of B-SAP markers are contrary. Low B-SAP polymorphism (26%) also confirmed by considerably uniform fingerprints is somehow surprising. In principal, DNA markers are expected to reveal higher molecular diversity than do isoenzymes. Accordingly, codominant RFLPs and microsatellite markers might show even twofold higher variability than allozymes as demonstrated by respective H_T values, 0.54; 0.74 and 0.34 (Karhu et al. 1996). Similar results were obtained using mitochondrial genes (Pyhäjärvi et al. 2008) and dominant nuclear markers such as ISSR (Labra et al. 2006) and RAPDs (Naugzemys et al. 2006). On the other hand, the B-SAP marker system is known to be conservative and especially recommended for species delimitation (Polok 2007). The more distant species the higher discriminatory power of the B-SAPs is observed. It is demonstrated by 11 *KatG* loci that differentiate species from two sections of the genus *Polygonatum*. However, only a single marker distinguishes species within the same section and no intraspecific polymorphism is observed (Szczecinska et al. 2006). Similarly in pines, surveys employing *P. cembra* demonstrate as well uniform B-SAP/*KatG* fingerprints in populations from the Polish and Slovak Tatras (Chmiel et al. 2008). At this point, the low genetic variation in *P. sylvestris* revealed by B-SAP markers based on the *KatG* gene is in agreement with the previous data. The plausible explanation of the phenomenon is the putative origin of at least some *KatG* markers from sequences related with enzymatic functions. In such loci mutations might not be neutral. Remarkably, primers designed on *KatG* fragments with active or binding sites revealed lower or no variation in Scots pine. Tight linkage between *KatG* markers and enzymatic loci on genetic maps provides another support for the hypothesis that *KatG* markers originate from functional genes not repetitive sequences. For example in grasses from the genus *Lolium* several *KatG* markers are tightly linked with *Per1* and *Per3* loci encoding peroxidases (Polok 2007). Further, the linkage between *KatG* markers and other five enzymatic loci on the *L. multiflorum* x *L. perenne* genetic map entails the possibility of similar linkage between important genes and *KatG* markers in other species including *P. sylvestris*. Then, although *KatG* markers themselves would be neutral, selective forces at other loci reduce variation due to a hitchhiking effect.

Despite the *KatG* based B-SAP markers reveal low variation in *P. sylvestris*, they prove to be effective in population differentiation as demonstrated by both genetic similarities (distances) and diagnostic markers. For instance, the distance equal to 0.252 between the Montenegro and Turkish populations as well as the mean genetic distance among all populations analysed ($D = 0.069$) belong to the highest observed so far. For example, Eurasian populations differ from $D = 0.005$ to 0.046 (Goncharenko et al. 1994), Swiss populations from 0.0042 to 0.0084 (Neet-Sarqueda 1994) and Scandinavian, Mediterranean and Transcaucasian from 0.045 to 0.056 (Sannikov et al. 2005). It has been speculated that the greater amounts of DNA in pines in general, may favour recombination hence population specific allele combinations are not captured. Low divergence of populations estimated from nucleotide diversity at the *pal1* locus indicates the large effective population size among responsible factors (Dvornyk et al. 2002). In contrast, a view emerging from RAPD studies is that the Scots pine nuclear genome can be a source of hundreds of markers serving as a road for studying pine diversity (Polok et al. 2005b, Naugzemys et al. 2006). The small sampling size involving three to seven populations is a certain limitation of the cited studies. Thus, the recent B-SAP appraisal employing 19 populations along a transect running across Europe is an important evidence for diagnostic power of nuclear DNA markers in population studies of Scots pine. Nevertheless, the present B-SAP results also underline the need for caution if sequences derived from a single gene are used because primers designed on different gene fragments can reveal different levels of variation in a species, thus having different utilities. For example, only primers designed on both ends

of the *KatG* gene may be useful in population identification.

The diagnostic power of B-SAP markers resides in cumulative information provided by population specific fingerprints. Firstly, they enable, to divide *P. sylvestris* into two groups, the North and the South according to geographic origins and secondly, they identify the majority of populations within groups. Populations from France (FR), Germany (DE1), Poland (PL1) and Sweden (SE) are known to be fixed for a single mitotype (Sinclair et al. 1999) but they are distinguishable using the combination of B-SAP markers. Thus, in contradiction to common observations of low diagnostic power of nuclear markers, the B-SAP approach proved to be useful in diversification of morphologically distinct populations. Unlike frequently used "random" markers such as RFLP, RAPD, AFLP, SSR, which scan the entire genome, the B-SAP assay preferentially tags sequences having functional properties or at least B-SAPs are linked to them. The B-SAP markers are predominantly mapped in genic regions with several examples of tight linkage with QTLs controlling morphological characters (e.g., leaf colour, plant height) in *L. multiflorum* x *L. perenne* (Polok 2007). Scots pine populations distant geographically and differing in morphology are likely dissimilar in the genomic region around QTLs but seemingly monomorphic at all other regions. If any B-SAP marker is within a QTL region, it will differentiate populations in agreement with morphological variation. By contrary, markers unlinked to QTLs are expected to be randomly distributed across populations. The "linkage hypothesis" has strong roots in occurrence of private B-SAP alleles in Belgian (BE), Montenegro (ME), and Turkish (TR) populations, i.e., these known for distinct height and branching (Stephan and Liesebach 1996). Moreover, the Belgium population characterized by the lack of a band in *KatG10-2*, is also exceptional with respect to straight stem in the majority of trees. Similarly, a marker revealed by *KatG10* is fixed in typical forms of *P. sylvestris* (i.e., with straight stem) inhabiting the Gazwa peat bog in North-Eastern Poland. The marker is not observed in "*turfosa*" forms (i.e., with curved log and umbrella like crown). Molecular differences are correlated with differences in needle morphology and anatomy

(Zielinski et al. 2009). It should be stressed however, that the "linkage hypothesis" needs further insights in mapping studies.

A counterproposal is that differences in B-SAP patterns are rather genetic footprints of the Quaternary refugia and postglacial migrations. This view emerges from distribution of populations falling into two B-SAP lineages, the North and South based on the block of six markers. During the Last Glacial Maximum (LGM), 21000 years ago, Scots pine was restricted to refuge areas across Southern Europe, from where it started to expand about 16000 years ago reaching northernmost maximum 7800 years ago (Willis et al. 1998). Therefore, pan-European populations are commonly regarded as offsets of glacial populations located in South. Unfortunately, little is known about the origin of populations that colonized northern Europe. It can not be exluded that ancestral populations might be more genetically diverged. Genetic differentiation initiated within refugia might then be widespread through specific routes of colonization. Heterogeneous origin of European Scots pine populations is obvious from variants in the intron 1 of *nad7* of mtDNA that signals distinctive evolution of northeastern Europe populations (Naydenov et al. 2007). Recent macrofossils and pollen data confirm existence of glacial refugia in more northern locations – south parts of Central and Eastern Europe (Willis and van Andel 2004). From there Scots pine might colonize northern Europe. A first glimpse at the B-SAP record marked on the European map (Fig. 1) entails two plausible routes of migrations directed from the South toward North-East (North group) and North-Central (South group) Europe with a border in Poland. The B-SAP based colonization history indicates alternative routes of Scots pine migration in Europe. Such routes were previously postulated based on morphology (Urbaniak 1998) as well as they are compatible with dual migration of *P. sylvestris* into Fennoscandia (Sinclair 1999). However, several cautionary points should be made about such B-SAP based routes.

What remains contentious is a dichotomy of Polish populations located just 300 km away (PL1–PL2; PL3–PL4) but classified to different groups due to B-SAP patterns. The majority of Polish pine forests are not primeval. The presence

of two *P. sylvestris* types in Poland would then indicate different origin of seeds/plants used in setting up forest plantations. Especially, German seeds were used in Central and West Poland what explains similarity of two Polish populations (PL3, PL4) to German ones (DE1, DE2). On the other hand, the rationale for the North-Eastern affiliation of PL1 and PL2 is a known border for relict pines that occurred easternmost during the LGM (Svenning et al. 2008). Likewise, RAPD patterns are consistent with eastern origin of the Milomlyn population (PL1), which is clustered with the relict ecotype inhabing peat bogs in North-Eastern Poland (Polok et al. 2005a). The dichotomy similar to that of Polish populations is also observed for populations from Belgium (South group) and France (North Group). Apart from the fact they are close enough to hybridize, they belong to different B-SAP groups and no signs of intercrossing have been observed. However, it can not be excluded that the Belgian population is not representative. Natural pine forests were degraded in Belgium with maximum in the 12th and 13th century. Then Scots pine was artificially planted in the first half of the 19th century. Thus, it is likely that seeds used for planting, in fact originated from populations belonging to the South Group. To resolve this problem more dense sampling is needed but unfortunately, Belgium provenances are rarely represented in studies on genetic diversity of Scots pine.

A second point of concern involves the Balkans refugia as the origin of Scots pine groups identified in the present studies. In a case of the South group, the similarities between populations from Bosnia (BA) and West-Central Europe are compatible with the presence of the Balkan refugium, from where Scots pine migrated along the Carpathian range. Current data suggest that another refugium was in Hungary (Willis and van Andel 2004) although an intermediate character of the Hungarian population may account for isolation for a long period. Presumably, during the LGM trees formed there opened and patchy stands. The conditions in the Hungarian Lowlands, Danube valley and Moravian were less favourable in comparison with the North-Western Carpathian valleys, where the climate was continental but still relatively humid, and favourable for growing rather a dense forest canopy (Jankovska and Pokorny 2008). The Carpathians could serve as a large-scale forest refugium, from where Scots pine colonized Slovakia, Germany and Poland. Scots pine is known to survive and grow reasonable well upon permafrost. Indeed, the mitochondrial haplotypes also suggest the refugial areas in Central Europe (Pyhäjärvi et al. 2008). The dual migration of the South group towards Germany and Poland correlates well with the distribution of mtDNA haplotypes in intron1 of *nad 7* (Naydenov et al. 2007). Such spread is also in agreement with colonization routes inferred for grasshopper, beech, oaks and Norway spruce. Much more controversial is the origin of the North group from the Balkans due to a number of private alleles in the Montenegro population. Similarities between French and North-eastern populations (SE, PL, RU) point at the alternative route of migration from the French Alps via northern Germany, Denmark, to Scandinavia and Russia. Early pine populations that colonized the French Alps and the Massif Central probably originated from the circum-Alpine refugia (Cheddadi et al. 2006). This route is also evident from morphological studies (Urbaniak 1998). The migration model becomes more plausible if we consider a large area of a dry land, likely a rich habitat ("Doggerland") stretching across the present Netherlands, Germany and Denmark in Mesolithic (Weninger et al. 2008). Similarly, about 11000–8700 BP the Baltic Sea region was a dry land with lakes, rivers and forests dominated by *P. sylvestris* on sandy soils (Schmolcke et al. 2006).

The important question remains why two neighbouring populations from the Balkans, just 200 km away (BA and ME) belong to different B-SAP lineages and differ in a number of unique alleles. One explanation is that these lineages had already existed in an ancestral population before the LGM. For example, the Turkish population originating from a refugium in Asia Minor did not take part in recolonisation of Europe because it possesses a number of private alleles not found in any other population (Naydenov et al. 2007, Pyhäjärvi et al. 2008). Nevertheless, it still shares the B-SAP pattern of the South group. An intermediate position of the Hungarian population as well as the distinctiveness of the Belgian population in terms of unique B-SAP alleles also states for the ancient origin of the northern and southern B-SAP types.

On the other hand, the number of unique alleles in the Montenegro population unequivocally entails limitation in gene flow accompanied by diversifying selection processes associated with high environmental heterogeneity typical of mountainous areas such as the Balkans.

A final point to consider is the lack of recombination between six B-SAP markers that produce two fixed B-SAP blocks. Obviously all six markers could be produced from the same sequence and consequently they represent a single gene or a group of tightly linked sequences. Another explanation may involve insertion of a transposon into the *KatG* derivatives. An insertion into a non-coding sequence could be neutral. If a target was a coding sequence, the transposon insertion should have not been lethal. Evidences of only slight modification of morphological characters depending on a place of transposon insertion are known in maize (Zhang et al. 2006). In grapes red, rose or white colour depend on the full retrotransposon, solo LTR or no insertion (Kobayashi et al. 2004). However, the present studies address the usefulness of the marker system in Scots pine diversification and the reasons underlying the B-SAP polymorphism await explanation in more detailed studies employing sequencing and insertion tracking.

In summary, the B-SAP marker system based on the *KatG* gene proved to be an effective tool in differentiation of *P. sylvestris* populations and phylogeographic studies. Presumably, fingerprints result from both diversification of ancestral populations and adaptation to changing environments visible as possible linkage with QTLs controlling morphological features. The B-SAP marker system, also due to its simplicity is a milestone raising the possibility to adopt a universally standardized scheme helpful in classifying populations and elucidating the postglacial history of Scots pine. Especially, when interpreted in conjunction with information from mtDNA and macrofossils.

Acknowledgements

References

Alía, R., Moro-Serrano, J. & Notivol, E. 2001. Genetic variability of Scots pine (Pinus sylvestris) provenances in Spain: growth traits and survival. Silva Fennica 35(1): 27–38.

Andersson, B. & Fedorkov, A. 2004. Longitudinal differences in Scots pine frost hardiness. Silvae Genetica 53: 76–80.

Baczkiewicz, A., Sawicki, J., Buczkowska, K., Polok, K. & Zielinski, R. 2008. Application of different DNA markers in studies on cryptic species of Aneura pinguis (Jungermannipsida, Metzgeriales). Cryptogamie Bryologie 29: 3–21.

Blumenrother, M., Bachman, M. & Müller-Starck, G. 2001. Genetic characteristics and diameter growth of provenances of Scots pine (Pinus sylvestris L.). Silvae Genetica 50: 212–222.

Cavalcanti, M.J. 2008. MANTEL – test for association between two symmetric distance matrices with permutation iterations. Departamento de Zoologia. Universidade. do Estatdo do Rio de Janeiro. Available at: http://maurobio.ifobio.net.

Cheddadi, R., Vendramin, G.G., Litt, T., Francois, L., Kageyama, M., Lorentz, S., Laurent, J-M., de Beaulieu, J-L., Sadori, L., Jost, A. & Lunt, D. 2006. Imprints of glacial refugia in the modern genetic diversity of Pinus sylvestris L. Global Ecology and Biogeography 15: 271–282.

Chmiel, J., Milewski, T. & Polok, K. 2008. Protection of the native Swiss stone pine (Pinus cembra L.) gene pool in the Tatras in the light of molecular studies. Roczniki Bieszczadzkie 16: 215–232. (In Polish with English summary).

Dvornyk, V., Sirvio, A,. Mikkonen, M. & Savolainen, O. 2002. Low nucleotide diversity at the pal1 locus in the widely distributed Pinus sylvestris. Molecular Biology and Evolution 19(2): 179–188.

Goncharenko, G.G., Silin, A.E. & Padutov, V.E. 1994. Allozyme variation in natural populations of Eurasian pines. III. Population structure, diversity,

differentiation and gene flow in central isolated populations of Pinus sylvestris L. in eastern Europe and Siberia. Silvae Genetica 43: 119–132.

Guan, L. & Scandalios, J.G. 1996. Molecular evolution of maize catalases and their relationships to other prokaryotic and eukaryotic catalases. Journal of Molecular Evolution 42: 570–579.

Hertel, H. & Schneck, V. 1999. Genetic and phenotypical variation of Scots pine (Pinus sylvestris L.) populations due to seed origin and environmental conditions at experimental sites. Forest Genetics 6(2): 65–72.

Jankovska, V. & Pokorny, P. 2008. Forest vegetation of the full glacial period in the Western Carpathians (Slovakia and Czech Republic). Preslia 80: 307–324.

Karhu, A., Hurme, P., Karjalainen, M., Karvonen, P., Karkkainen, K., Neale, D. & Savolainen, O. 1996. Do molecular markers reflect patterns of differentiation in adaptive traits in conifers? Theoretical and Applied Genetics 93: 215–221.

Kobayashi, S., Goto-Yamamoto, N. & Horochika, H. 2004. Retrotransposon-induced mutations in grape skin colour. Science 304: 982.

Labra, M., Grassi, F., Sgorbati, S. & Ferrari, C. 2006. Distribution of genetic variability in southern populations of Scots pine (Pinus sylvestris L.) from the Alps to the Apennines. Flora 201: 468–476.

Ledig, T. 1998. Genetic variation in Pinus. In: Richardson, D.M. (ed.). Ecology and biogeography of Pinus. Cambridge University Press. p. 251–280.

Naugzemys, D., Zvingila, D., Aucina, A. & Rancelis, V. 2006. Comparison of DNA polymorphism in seedlings of Pinus sylvestris L. from different populations by RAPD markers. Biologija 1: 30–35.

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): 1–12.

Neet-Sarqueda, C. 1994. Genetic differentiation of Pinus sylvestris L. and Pinus mugo aggr. populations in Switzerland. Silvae Genetica 43: 207–215.

Nei, M. & Li W-H. 1979. Mathematical model for studying variation in terms of restriction endonucleases. Proceedings of the National Academy of Science of USA 76: 5269–5273.

Oleksyn, J., Tjoelker, M.G. & Reich, P.B. 1998. Adaptation to changing environment in Scots pine populations across a latitudinal gradient. Silva Fennica 32(2): 129–140.

Polok, K. 2007. Molecular evolution of the genus Lolium L. Studio Poligrafii Komputerowej "SQL", Olsztyn. 317 p.

— & Zielinski, R. 2009. Evolution of peroxidases in selected plant species. In Libiakowa, G., Gajdosova, A. & Hajduch, M. (eds). Plant proteomics for Europe. COST FA0603. Institute of Plant Genetics and Biotechnology, Nitra. p. 58.

—, Pisarek, W., Urbaniak, L., Korzekwa, K. & Zielinski, R. 2005a. Genetic characteristics of Pinus sylvestris f. turfosa Woerl from the Gazwa reserve, north eastern Poland. In: Prus-Glowacki, W. & Pawlaczyk, E. (eds.). Variability and evolution: new perspectives. Wydawnictwo Naukowe UAM, Poznan. p. 225–239.

—, Urbaniak, L., Korzekwa, K., Androsiuk, P., Ciągło, S., Kubiak, K. & Zielinski, R. 2005b. Genetic similarity of Pinus sylvestris populations on the base of DNA markers. In: Prus-Glowacki, W. & Pawlaczyk, E. (eds). Variability and evolution: new perspectives. Wydawnictwo Naukowe UAM, Poznan. p. 253–267.

Pyhäjärvi, T., Salmela, M.J. & Savolainen, O. 2008. Colonization routes of Pinus sylvestris inferred from distribution of mitochondrial DNA variation. Tree Genetics & Genomes 4: 247–254.

Ruby, J.L. 1967. The correspondence between genetic, morphological, and climatic variation paterns in Scots pine. Silvae Genetica 16: 50–56.

Sannikov, S.N., Petrova, I.V., Sannikova, N.S. & Filippova, T.V. 2005. Gradient genogeographic analysis of Pinus sylvestris L. populations in Europe. Russian Journal of Ecology 36: 377–382.

Sawicki, J. & Zielinski, R. 2008. Phylogenetic relationships between five *Sphagnum* species of the section *Acutifolia* based on DNA markers. Čas Slez Muz Opava (A) 57: 63–80.

Schmolcke, U., Endtmann, E., Klooss, S., Meyer, M., Michaelis, D., Rickert, B.-H. & Roβler, D. 2006. Changes of sea level, landscape and culture: A review of the south-western Baltic area between 8800 and 4000 BC. Palaeogeography, Palaeoclimatology, Palaeoecology 240: 423–438.

Sinclair, W.T., Morman, J.D. & Ennos, A. 1999. The postglacial history of Scots pine (Pinus sylvestris L.) in western Europe: evidence from mitochondrial DNA variation. Molecular Ecology 8: 83–88.

Stephan, B.R. & Liesebach, M. 1996. Results from the IUFRO 1982 Scots pine (Pinus sylvestris L.) provenance experiment in southwestern Germany. Silvae Genetica 45: 342–349.

Svenning, J-C., Normand, S. & Kageyama, M. 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. Journal of Ecology 96: 1117–1127.

Szczecinska, M., Sawicki, J., Polok, K., Holdynski, Cz. & Zielinski, R. 2006. Comparison of three Polygonatum species from Poland based on DNA markers. Annales Botanici Fennici 45: 379–388.

Urbaniak, L. 1998. Geographic diversity of Scots pine (Pinus sylvestris L.) from the Eurasian territory based on needle anatomy and morphology. Wydawnictwo Naukowe UAM, Poznan. 141 p. (In Polish with English abstract).

Weninger, B., Schulting, R., Bradtmoller, M., Clare, L., Collard, M., Edinborough, K., Hilpert, J., Joris, Ol., Niekus, M., Rohling, W.J. & Wagner, B. 2008. The catastrophic final flooding of Doggerland by the Storegga Slide Tsunami. Documenta Praehistorica 25: 1–24.

Willis, K. & van Andel, T.H. 2004. Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. Quaternary Science Reviews 23: 2369–2387.

Willis, K.J., Bennet, K.D. & Birks, H.J. 1998. The late Quaternary dynamics of pines in Europe. In: Richardson, D.M. (ed.). Ecology and biogeography of Pinus. Cambridge University Press. p. 107–121.

Wright, J.W., Pauley S.S. & Polk, R.B. 1966. Performance of Scots pine varieties in the North Central Region. Silvae Genetica 15: 101–110.

Yeh, F.C., Yang, R.C., Boyle, T.B.J., Ye, Z.H. & Mao, J.X. 2000. POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada. Available at: http://www.ualberta.ca/~fyeh/index.htm.

Zamocky, M. 2004. Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. European Journal of Biochemistry 271: 3297–3309.

Zhang, J., Zhang, F. & Peterson, T. 2006. Transposition of reversed Ac element ends generates novel chimeric genes in maize. PloS Genetics 2: 1535–1540.

Zielinski, R., Osmęda, J., Urbaniak, L. & Polok, K. 2009. Molecular characteristics of an ancient Pinus sylvestris (Pinaceae) ecotype from the Gązwa peat bog reserve (NE Poland). In: Mirek, Z. & Nikiel, A. (eds.). Rare, relict and endangered plants and fungi in Poland. W. Szafer Institute of Botany, Polish Academy of Sciences, Kraków. p. 561–567.

— & Polok, K. 2005. Molecular evolution and taxonomy of plants. In: Prus-Glowacki, W. & Pawlaczyk, E. (eds). Variability and evolution: new perspectives. Wydawnictwo Naukowe UAM, Poznan. p. 37–56.

Total of 44 references