SZENT ISTVÁN EGYETEM
KERTÉSZETTUDOMÁNYI KAR, BUDAPEST

PHYLOGEOGRAPHY AND ADAPTIVE GENETIC VARIATION OF SCOTS PINE (*PINUS SYLVESTRIS* L.) POPULATIONS FROM THE CARPATHIANS AND THE PANNONIAN BASIN

PHD THESIS

ENDRE GYÖRGY TÓTH

SUPERVISOR: DR. MÁRIA HÖHN

BUDAPEST

2017

PH.D. SCHOOL

Name:	Doctoral School of Horticultural Science
Field:	Crop Sciences and Horticulture
Head of the Ph.D. school:	Prof. Dr. Éva Zámboriné Németh, Head of Department of Medicinal and Aromatic Plants SZENT ISTVÁN UNIVERSITY, Faculty of Horticultural Science
Supervisor:	Assoc. Prof. Dr. Mária Höhn, Head of Department of Botany and Botanical Garden of Soroksár SZENT ISTVÁN UNIVERSITY, Faculty of Horticultural Sciences

The applicant met the requirement of the Ph.D. regulations of the SZENT ISTVÁN UNIVERSITY and the thesis is accepted for the defense process.

.....	.....
Dr. Éva Zámboriné Németh	Dr. Mária Höhn
Head of Ph.D. School	Supervisor

1. INTRODUCTION AND OBJECTIVES

Scots pine (*Pinus sylvestris* L.) is a long-lived coniferous species of the Pinaceae family which occupies a continuous range as the dominant tree species of the Eurasian taiga communities (Pravdin 1969). On the southern and western edge of its wide distribution Scots pine has survived in small peripheral populations considered to be remnants from the Quaternary when pines were more widespread in Eurasia. As well, Scots pine has survived during the LGM in several southern latitude refugia across Europe in highly disperse isolated populations.

Along the Carpathian Mountain range Scots pine is distributed in island-like peripheral populations (Fekete and Blattny 1913), but there are also scattered natural populations sustained in mixed forest stands, with broad-leaved species in the western Pannonian Basin, at the foothills of the Alps (Pócs 1960, Fekete *et al.* 2014). Postglacial climate warming of the Holocene, forced Carpathian Scots pine populations to immigrate into edaphically specialized habitat types, including humid, cool peatbogs and sunny, dry, rocky outcrops (Giertych and Mátyás 1991, Matías and Jump 2012). In these habitats populations are strongly influenced by the interplay of genetic drift, gene flow and natural selection. Ongoing processes are also affected by the historical demographical events (Eckert *et al.* 2008).

Former study by Bernhardsson *et al.* (2016), has reported that populations from the Carpathians and the Pannonian Basin present low level of differentiation among the populations and the impact of Holocene population fragmentation. Despite combined pollen, macrofossils and organelle DNA analysis that could detect glacial refugia in the Pannonian Basin along the Danube, previous molecular studies performed in the region reported lack of geographic structure both with mtDNA and cpDNA within the Carpathian Mountains (Cheddadi *et al.* 2006, Bernhardsson *et al.* 2016). Similarly, no variation and no phylogeographic structure in mitochondrial DNA was found in provenance trials conducted in the region by Čelepirović *et al.* (2009). Historical demographic evaluations based on molecular genetic data were not conducted within the Carpathian Mountains. According to the fossil records, Scots pine has a complex spatio-temporal history in Central-Eastern Europe during the Holocene, influenced mainly by oscillations of the climate and also affected by human activities (Feurdean *et al.* 2007).

Studies that aimed to elucidate adaptive genetic variation of *P. sylvestris* have found overall moderate level of diversity and differentiation in continental European populations. Additionally, signals of negative selection and effects of historical demography on nucleotide

variation have been detected (Pyhäyärvi *et al.* 2007, Wachowiak *et al.* 2009, 2011, Kujala and Savolainen 2012). Sequence variation studies, involving candidate gene loci, are missing from the Carpathian region, while patterns of nucleotide diversity and divergence were only studied in western and southern Europe. Therefore, the adaptive potential of these peripheral populations are yet unknown.

The overall objectives of this research were as follows:

- Highlight the current population structure of the selected peripheral Scots pine populations native to Central-Eastern Europe, most of which formerly were not included in molecular studies.
- Identify genetic relationships, degree of diversity and divergence and infer gene flow between the studied stands.
- Describe historical demographical processes (expansions-contraction) and circumscribe putative refugia within the studied region that might have existed in the time of the Pleistocene.
- Assess the nucleotide diversity, divergence at candidate gene loci to infer adaptive nucleotide variation of peripheral populations as signs of local adaptation.

2. MATERIAL AND METHODS

2.1. Plant material

Plant material originates from Central-Eastern Europe, from the area of the Carpathian Mountains and the Pannonian Basin. Additionally, a population was included from the northern part of the distribution range, from Estonia (core population) and another from the south, from the Balkan Peninsula (Bulgaria).

Altogether 20 natural and autochthonous populations were sampled from the highly fragmented distribution range of the species. In total, 421 individuals were analysed with nuclear and chloroplast simple-sequence repeats (SSR).

To test neutral and adaptive genetic variation, i.e. genetic variation under natural selection, 10 populations (96 individuals) were used from the Hungarian Pannonian Basin and from the Slovakian Western Carpathians and the Romanian Eastern Carpathians.

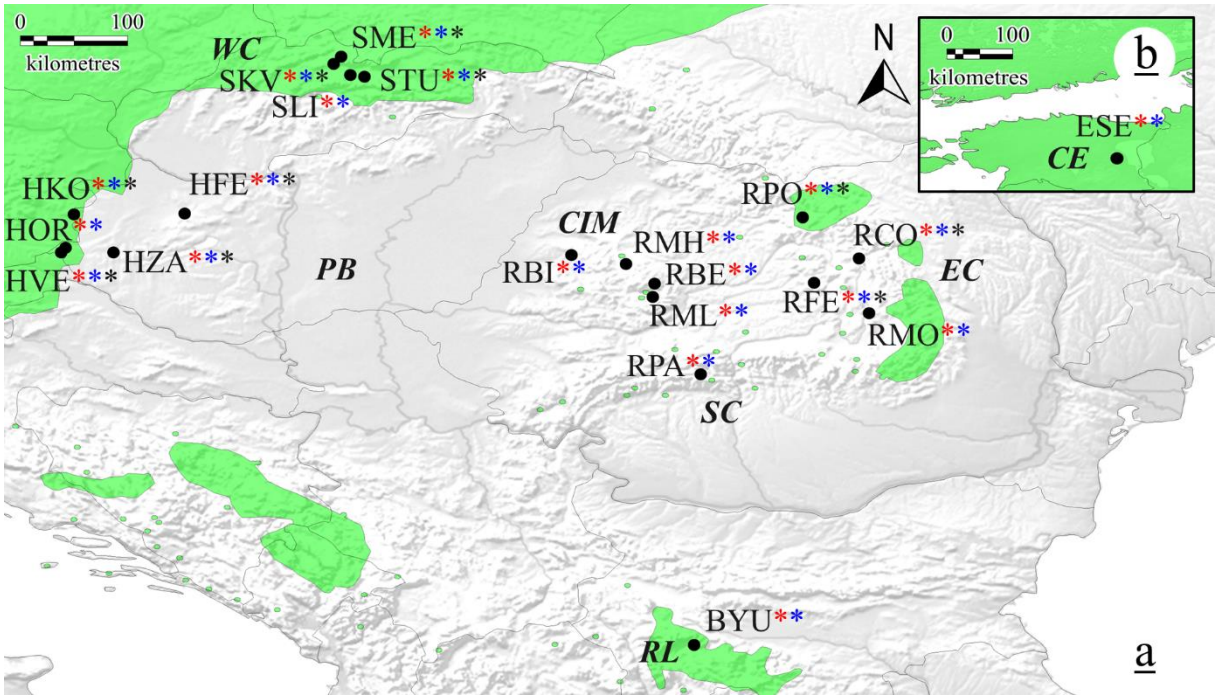

Fig. 1: Geographic location of the studied Scots pine populations. (a): Central-Eastern Europe, (b): Estonia. The acronyms stand for the population code. Asterisks indicate types of marker (*; cpSSR, *; nSSR, *; SNP) applied to evaluate the population. The natural distribution of Scots pine is marked in green according to the EUFORGEN database, with modifications by the authors.

Plant samples were analysed with eight nuclear SSR (nSSR) markers (SPAG 7.14, SPAC 11.4 from Soranzo *et al.* (1998) and psyl16, psyl17, psyl19, psyl36, psyl42, psyl57 from Sebastiani *et al.* (2012)) and with four chloroplast SSR (cpSSR) markers (Pt-30204, Pt-15169, Pt-45002, Pt-26081 from Vendramin *et al.* (1996)) by Polymerase Chain Reaction (PCR) and Fragment Length Analysis (FLA) to study variability and divergence. Eleven water dehydration stress response candidate genes (*abaR, ccoaomt, chcs, dhn3, dhn7, dhy2PP, erd3* from Wachowiak *et al.* (2011) and *cpk3, pal-1, ppap12, rd21A-like* from González-Martínez *et al.* (2006)) were amplified with PCR and sequenced to examine sequence variation by single-nucleotide polymorphism (SNP).

2.2. Statistical analysis

Standard population genetic diversity indices were calculated with GenAlEx v.6.5 (Peakall and Smouse 2006). For estimating haplotype diversity indices Haplotype Analysis software (Eliades and Eliades 2009) was used. To test for correlation between geographical (kilometres) and genetic (Pairwise Nei) distances at both nSSR and cpSSR markers, Mantel test (Mantel 1967) was carried out in the Isolation-by-Distance Web Service (IBDWS) 3.16 (Jensen *et al.* 2005). Analysis of molecular variance (AMOVA) implemented in Arlequin v.3.5 software (Excoffier and Lischer 2010) was used to determine the partition of the genetic variation within and among populations. Genetic discontinuities corresponding to the change in genetic variation among populations were identified with Barrier 2.2 (Manni *et al.* 2004) on both nuclear and chloroplast SSR datasets. We used BAPS 6.0 (Corander and Mattinen 2006) to conduct hierarchical clustering analyses of the cpSSR dataset. A Bayesian clustering approach implemented in STRUCTURE 2.3.4 (Pritchard *et al.* 2000) was used to infer groups or subpopulations based on the nSSR dataset. To detect any recent severe reduction in effective population size or possible expansion events in Scots pine populations, BOTTLENECK 1.2.02 was used on nSSR dataset (Cornuet and Luikart 1996). Analysis of the historical demographic changes was studied with Approximate Bayesian Computation implemented in DIYABC v2.0 software (Cornuet *et al.* 2014).

Basic sequence variation statistics of the candidate gene dataset was evaluated with DnaSP 5.10.1 (Librado and Rozas 2009). Nucleotide diversity indicator theta pi (θ_π) was determined with Arlequin v.3.5 software. To test that all mutations (SNPs) are selectively neutral, hence randomly evolving (Kimura 1983), and to identify deviations from this neutral expectation (loci under selection) Tajima's D (Tajima 1989) test was applied. Additionally, Fu

and Li's D and F (1993) statistical tests were also evaluated. We also tested if single SNPs are deviated from neutral expectations (outlier analysis) by detecting loci under selection with distinct approaches implemented in Arlequin v.3.5, in Lositan (Antao *et al.* 2008) and in BayeScan 2.1 (Foll and Gaggiotti 2008). Decay of linkage disequilibrium (LD) was measured in DnaSP. STRUCTURE was also used to infer groups or subpopulations in the candidate gene dataset. AMOVA and pairwise F_{ST} (matrix) was calculated with Arlequin v.3.5 software.

3. RESULTS

Haplotype analysis of the chloroplast microsatellite (cpSSRs) revealed 4 to 13 size variants per locus. These size variants are combined into 141 haplotypes. Haploid diversity was balanced along the populations. Mean number of private alleles was the highest in SME. The twenty most-frequent haplotypes show evident divergence in haplotype frequencies between the Eastern Carpathian populations on one hand, and the Transylvanian Central-Island Mountains (Apuseni) and Southern Carpathian populations on the other. AMOVA presented Φ_{PT}=0.081 ($p< 0.001$) value. The Mantel test revealed significant negative correlation between geographic and genetic distances among populations. The Bayesian approach of population structuring estimated with BAPS identified three main clusters (K=3), which formed geographically distinct groups **(Fig. 2)**. Two major groups were detected: Populations from Western Hungary clustered together with those of the Eastern Carpathians and the Bulgarian population (1), while populations of the Western Carpathians (Tatras) and the Transylvanian Central-Island Mountains (Apuseni) grouped together with the Southern Carpathian and Estonian populations (2). SME population from the Tatras was separated as a distinct group. The main barriers to gene flow as detected with Barrier analysis on the cpSSR dataset delimited the Eastern Carpathians populations with highest bootstrap support of 99.5% (data not shown).

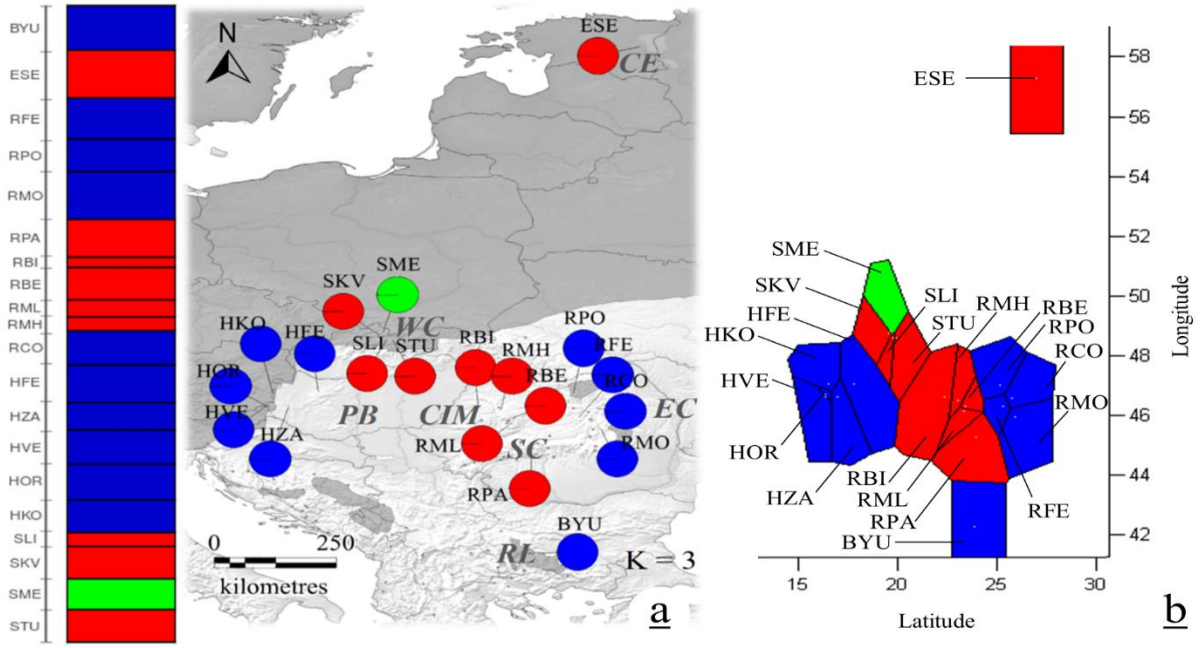

Fig. 2: Estimated spatial population structure K=3 (a and b) according to BAPS analysis. The acronyms stand for the population code. The natural distribution of Scots pine is marked in grey according to the EUFORGEN database, with modifications by the author. RL: Rila Mountains, CE: Central Estonian Plain, PB: Pannonian Basin, WC: Western Carpathians, EC: Eastern Carpathians, SC: Southern Carpathians and CIM: Central-Island Mountains (Apuseni), respectively.

Nuclear microsatellites (nSSRs) presented moderate level of intra-population variability. All loci largely conformed to the Hardy–Weinberg Equilibrium (HWE) and showed no significant deviations. AMOVA showed high molecular variance within individuals (90%) and relatively low molecular variance among individuals (3%), while among populations only 7% variation was observed. Overall F_{ST} was 0.071 ($p < 0.001$). STRUCTURE analysis revealed two genetic groups **(Fig. 3)** that are: (1) the Western Hungarian populations and the Southern Carpathian population, (2) the Eastern Carpathians population with the Bulgarian population (BYU) and the northernmost population from Estonia (ESE). Barrier analysis identified major genetic discontinuities with high bootstrap support (from 70.9% to 100%) around the Eastern Carpathians, separating these populations from the rest. The BOTTLENECK analysis showed no evidence of significant excess or deficit of heterozygosity in most populations. The Approximate Bayesian Computation (ABC) analysis, indicated that Sc6 (the 6th tested demographic scenario) an "admixture model" has the highest posterior probability (0.5670, 95% C.I.: 0.5307–0.6033). Results estimated an ancient divergence at t2 time (divergence of Pop1 and Pop2) 3560 (95% C.I.: 1040–8810) generations ago, and a recent admixture event and then a divergence (of Pop1-2-3) at t1 time at 236 (95% C.I.: 29.4–876) generations ago. Growth in effective population size was detected for all evaluated populations **(Fig. 4)**.

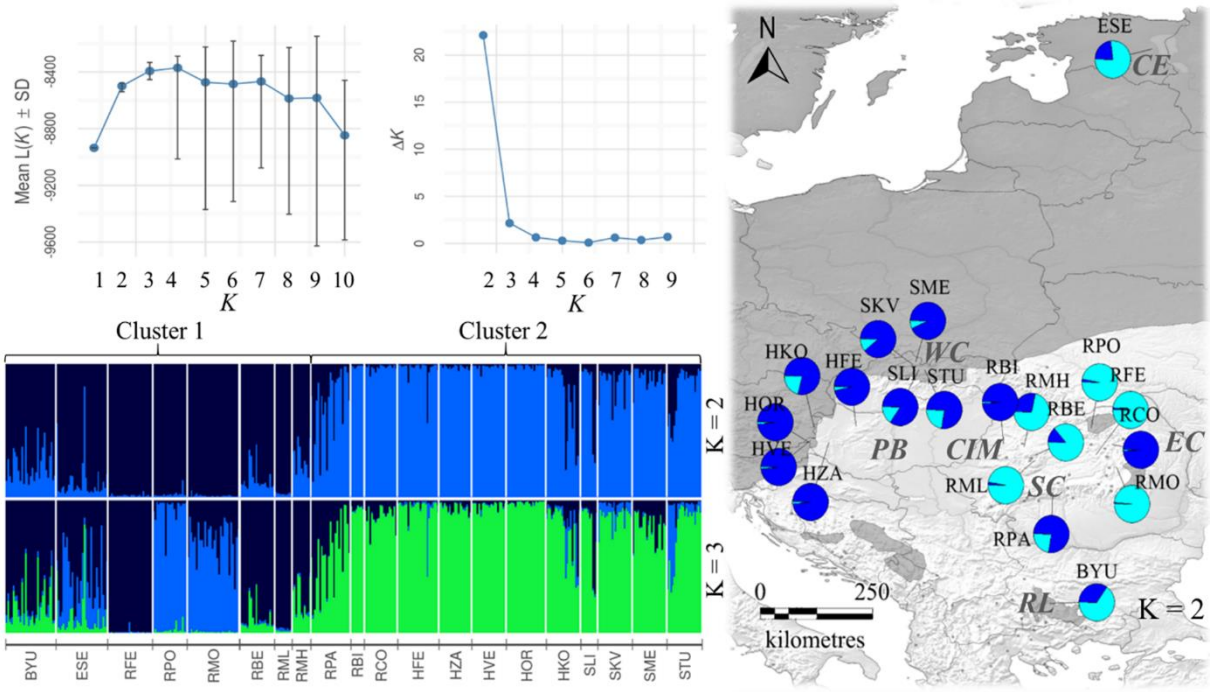

Fig. 3: Estimated population structure K=2 and K=3 of assignment analysis performed in STRUCTURE. K=2 plotted on a geographic map. Most likely membership in a population is presented by the colour of the individual's thin line. The acronyms stand for the population code. The natural distribution of Scots pine is marked in grey according to the EUFORGEN database, with modifications by the author. RL: Rila Mountains, CE: Central Estonian Plain, PB: Pannonian Basin, WC: Western Carpathians, EC: Eastern Carpathians, SC: Southern Carpathians and CIM: Central-Island Mountains (Apuseni), respectively.

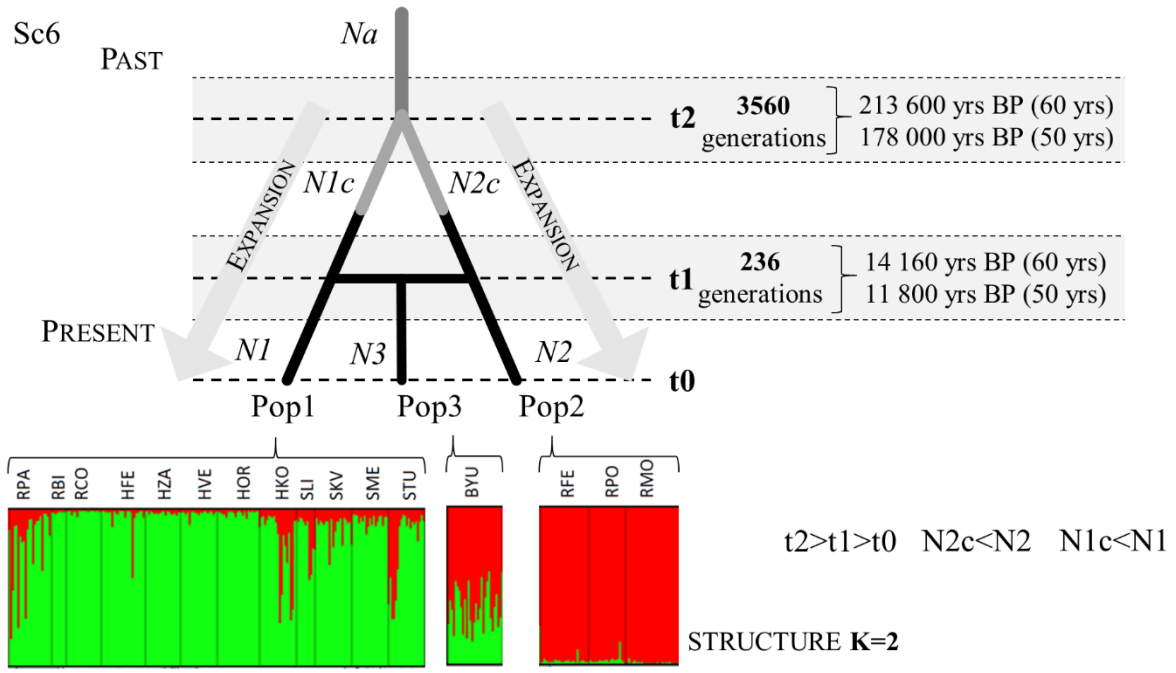

Fig. 4: The best supported demographic scenario with the highest posterior probability (0.5670, 95%CI: 0.5307–0.6033) resulted on the Central-Eastern European Scots pine populations in DIYABC. In the detected scenario t# represent the estimated time expressed in generation time, N# is the effective size of population Pop1, Pop2 and Pop3 respectively. N2c and N1c indicate expansion of ancient population expressed in effective population size, as like Na < N#c < N#.

The 11 candidate gene sequences related to draught stress response provided a set of 232 polymorphic (segregating) sites **(Table 1)**. The highest number of substitution sites was found at *dhy2PP* and *rd21A* loci and in the SME and STU populations. Most of the population private substitution sites were detected in SME which was about twice as many as in other populations. The average nucleotide diversity (π) over all loci was π=0.0056 and Theta pi ($\theta\pi$) nucleotide diversity was $\theta\pi=2.15\times10^{-3}$. An average haplotypic diversity (Hd) of 0.733 and an average number of 22.63 haplotypes (*Nh*) was present in the whole sample set. The population SME showed the highest number of population specific haplotypes (33 in total). Decay of linkage disequilibrium (LD) was detected in only seven candidate gene loci, of which 4 loci (*abaR*, *dhn7*, *dhy2PP*, *pal1*) presented a slow decay (R^2=0.01-0.02) at a distance about 100-400 bp and 3 loci (*ccoaomt*, *erd3*, *rd21A*) presented a faster, but still relatively slow decay (R^2=0.04-0.09) at a distance about 100-400, -600 and -800 bp, respectively. Test of Analysis of Molecular Variance (AMOVA) including all populations and the test of all candidate loci resulted an overall F_{ST}=0.063 (p<0.001), meaning that high molecular variance (93.63%) resides within populations, while among populations only a 6.36% variation was observed. STRUCTURE indicated that the most likely number of groups of populations was K=4, whereas high level of admixture was revealed. Two distinct clusters were also detected including only one population:

the Eastern Carpathian RFE and the Western Carpathian SME population. The analysis of deviations from neutral expectations (Tajima's D, Fu and Li's D and F statistics) resulted mostly negative values with statistical significance at *erd3* and *rd21A* loci. The value of *dhy2PP* was also significant, but only in the Fu and Li's D test. Outlier SNP analysis detected 13 SNP loci under selection (out of the 232) at different significance level. These loci reside in 8 candidate gene and alter 6 non-coding, 1 synonymous and 6 non-synonymous nucleotide change. Seven loci were determined as potentially being under negative (balancing or purifying) selection and 6 loci as to be under positive (diversifying) selection.

Table 1: The loci included in the nucleotide diversity analysis and the identified polymorphic sites.

			Base pairs screened				
Locus	***n***	***ns***	**Total**	**Coding region**	**Non-coding region**	***S***	***Sing***
abaR	95	190	378	321	57	17	5
ccoaomt	93	186	549	300	249	14	2
chcs2	95	190	341	231	110	16	2
cpk3	95	190	599	306	293	19	3
dhn3	61	122	268	216	52	11	4
dhn7	82	164	358	279	79	23	8
dhy2PP	96	192	454	357	97	34	13
erd3	95	190	679	474	205	14	9
pal1	96	192	459	276	183	15	4
ppap12	96	192	238	237	1	25	6
rd21A	80	160	1005	624	381	44	18
Total	984	1968	5328	3621	1707	232	74

n: number of samples analysed by locus
ns: number of haploid sequences
S: number of polymorphic (segregating) sites identified
Sing: total number of singleton mutations

4. CONCLUSIONS

Although the two SSR marker types exhibit particular genetic patterns, there were also some congruencies in the detected spatial genetic structure of populations. BAPS and STRUCTURE analysis delimited Western Carpathian populations that proved to be different from those inhabiting the eastern range of the Carpathians. The differentiation in the genetic pattern along the Carpathians has been recognized in earlier studies of other conifer species also. In the case of *Picea abies* or *Abies alba*, mitochondrial minisatellite regions and nSSRs delimited lineages of the Western and Eastern Carpathians and accordingly, these patterns suggest different origin of populations from distinct glacial refugia (Tollefsrud *et al.* 2008, Liepelt *et al.* 2009, Gömöry *et al.* 2012). Moreover, the haplotype pattern of Scots pine revealed in our study is highly congruent with those observed by Höhn *et al.* (2009) for *Pinus cembra*, where the populations of the Western and the Eastern Carpathians were spatially separated on the basis of chloroplast SSR variation. It is likely that the Carpathian populations of Scots pine harbour genetic material originating from at least two separate refugia, dating back to the Pleistocene. This was also highlighted in other alpine species by Tribsch and Schönswetter (2003). One refugium might have been situated around the Eastern Alps and the Hungarian Plain with the Danube region (Cheddadi *et al.* 2006), and the other might have existed in the Eastern Carpathians, where high abundance of fossil pollen remains was reported (Feurdean *et al.* 2011).

Genetic diversity revealed in our study was higher compared to that found in the Romanian-Hungarian populations by Bernhardsson *et al.* (2016). Despite this fact, chloroplast population diversity indices show detectable signs of segregation and fragmentation of these isolated populations, which can be the effect of restricted gene flow on a regional scale. Similarly, genetic discontinuity was also detected in both datasets with Barrier analysis. We did not found signs of inbreeding in most of the populations studied, although small and isolated populations are more vulnerable to inbreeding (Ellstrand and Elam 1993). In accordance, BOTTLENECK analysis provided evidence that the Carpathian populations studied are not influenced by a recent genetic bottleneck. Signs of a bottleneck event require multiple generations to appear. Furthermore, the effect of a bottleneck can vary based only on the reduction size and on the duration period (Busch *et al.* 2007, Peery *et al.* 2012). It is most plausible that populations that today are isolated have undergone a recent fragmentation and isolation event.

Our demographic (DIYABC) analysis supported an admixture scenario, in which the two main detected gene pools have separated at the same time, rather than hierarchically split. Additionally, a gradual demographic expansion of the populations were detected. This might have caused admixture of diverged populations. If we assume 50-60 years as generation time, the first divergence time (t2), from the ancestral population, falls within 178 ka and 213.6 ka BP and the admixture event (t1) from 11.8 ka to 14.1 ka. There are strong evidences that *Pinus* (diploxylon) species were dominating from the mid-Pleistocene's transition to glacial to early Holocene interglacial period. Deep pollen cores from the Tenaghi Philippon peatland in Greece, showed an overall increase of *Pinus* pollen (Milner *et al.* 2013). Recent study by Sadori *et al.* (2015) from lake Ohrid (western Balkan region, Albania) also highlighted the high abundance of *Pinus* pollen at this time period. Accordingly, *Pinus* pollen concentration remained high during the mid-Pleistocene to the LGM. Our estimated admixture event for Scots pine, might have happened between 11.8 ka to 14.1 ka BP, when Scots pine displayed a vast expansion, at the end of Younger Dryas and early Holocene, when *Pinus* (diploxylon) pollen percentages were at their maximum (Feurdean *et al.* 2011). A strong reduction of conifers and expansion of deciduous species has started in the Late Glacial/Holocene transition period (Feurdean *et al.* 2012). This decline in *Pinus* pollen abundance has been detected at several sites along the Carpathian mountain range (Tantau *et al.* 2003, 2006, Feurdean and Bennike 2004, Feurdean *et al.* 2007), suggesting, that after the expansion and admixture, populations have contracted causing fragmentation and reduction of population sizes.

Sequence diversity evaluated at a set of drought stress related candidate gene loci provided evidence for selection. Our results agreed that 13 significant outlier SNPs (true outliers) are under natural selection either located in coding or non-coding regions. Outliers suggest that there is a strong signal for selection acting to maintain genetic diversity (negative selection) and potentially to counteract the effect of genetic erosion. Our neutrality measures point to an excess of low-frequency polymorphism, which likely acts against local adaptation in our peripheral populations. Moreover negative values can also refer to hitchhiking effects resulting from a selective sweep or are the result of historical demographic events (either spatial or demographic expansion) formerly revealed in case of forest tree species at drought stress response genes (Grivet *et al.* 2009, 2011, Homolka *et al.* 2013). Effects of historical demography (population size fluctuations effect on sequence variation) were identified by former studies on dehydrin genes and on other cold-related candidate genes for coniferous species, including *Pinus sylvesris* and *Pinus mugo* by Wachowiak *et al.* (2009, 2013, 2014). The decay of intragenic linkage disequilibrium (LD) was overall slow. LD potentially is the

result of complex interactions of several factors like population size, selection process, structure and history and depends on sequence length and nucleotide variation (Krutovsky and Neale 2005, Lalagüe *et al.* 2014). Relatively slow decay was estimated for Scottish populations of *P. sylvestris,* whereas LD decay was almost three times slower than that in the mainland populations (Wachowiak *et al.* 2011). Nucleotide diversity was much higher than detected earlier for *P. sylvestris* (π=0.0014, Dvornyk *et al.* 2002). Values presented similar or even higher values, like in case of Scottish populations or populations from Central Europe and Spain (Wachowiak *et al.* 2011, 2014). Our set of candidate genes for plant response to environmental stress preserved high variation, accordingly are not affected by genetic erosion due to historical isolation, fragmentation. STRUCTURE analysis separated SME and RFE populations from each other and from the remaining highly admixed populations. The position of SME population is potentially assumed to be the consequence of hybrid individuals of *Pinus rhaetica* (*P. sylvestris* x *Pinus mugo*) and their specific genomic composition. The distinct clustering of RFE, both in our non-coding SSR and candidate gene study might be due to several factors, including introduced individuals (alien genetic material), effect of large-scale genetic drift or peculiar historical demographical events.

5. SUMMARY

During the Pleistocene Scots pine (*Pinus sylvestris* L.) was widespread in Eurasia, being present even in the outskirts of the glaciated territories. Scots pine also survived in several southern latitude refugia across Europe in highly disperse isolated occurrences. Putative refugia of the most frequent genetic lineage was detected from Central Continental Europe. Presumably Scots pine originated not just from the (Sub) Mediterranean areas like the Balkan Peninsula, but also from around the eastern Alps and the surroundings of the Danube plain (western Pannonian basin).

Results showed that the most differentiated region is the Romanian Carpathians, i.e. the Eastern Carpathians and the Transylvanian Central-Island Mountains (Apuseni Mts.). Our findings constitute evidence in support of the hypothesis that there were Pleistocene refugia in this region, also evidenced by some recent palynological records.

Demographic history highlighted that from the ancestral population two lineages have diverged (potentially from the two detected refugia) and later due to the favourable climatic conditions of the mid-Pleistocene, populations underwent an expansion leading to an admixture event between 11.8–14.1 ka BP in the early Holocene.

The analysis of nucleotide variation at candidate loci revealed that peripheral populations have maintained high genetic diversity (large nucleotide polymorphism) despite being fragmented and isolated. This strong signal suggests that selection is actually acting to maintain genetic diversity and to counteract the effect of genetic erosion.

6. NEW SCIENTIFIC ACHIEVEMENTS

1. **Non-coding (chloroplast and nuclear) microsatellite markers revealed two distinct genetic lineages and overall geographic structuring of Scots pine populations from the Carpathian Mountains and the Pannonian Basin.**

2. **In accordance to former studies, despite fragmentation and isolation, high genetic diversity was preserved in the natural Scots pine populations of the Carpathians and the Pannonian Basin.**

3. **Historical demographic expansion and recent fragmentation was detected for Scots pine populations inhabiting the Carpathians and the Pannonian Basin, based on non-coding nuclear microsatellite markers.**

4. **Neutral genetic variation was evidenced at candidate gene loci indicating long-term maintenance of genetic diversity in peripheral Scots pine populations.**

7. REFERENCES

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC bioinformatics*, 9(1), 323.

Bernhardsson, C., Floran, V., Ganea, S. L., & García-Gil, M. R. (2016). Present genetic structure is congruent with the common origin of distant Scots pine populations in its Romanian distribution. *Forest Ecology and Management*, 361, 131-143.

Busch, J. D., Waser, P. M., & De Woody, J. A. (2007). Recent demographic bottlenecks are not accompanied by a genetic signature in banner- tailed kangaroo rats (*Dipodomys spectabilis*). *Molecular Ecology*, 16(12), 2450-2462.

Čelepirović, N., Ivanković, M., Gradečki-Poštenjak, M., Nagy, L., Borovics, A., Agbaba, S. N., & Littvay, T. (2009). Review of investigation of variability of nad1 gene intron B/C of mitochondrial genome in Scots pine (*Pinus sylvestris* L.). *Periodicum Biologorum*, 111(4), 453-457.

Cheddadi, R., Vendramin, G. G., Litt, T., François, 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*. *Global Ecology and Biogeography*, 15(3), 271-282.

Corander, J., & Marttinen, P. (2006). Bayesian identification of admixture events using multi-locus molecular markers. *Molecular Ecology*, 15(10), 2833-2843.

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144(4), 2001-2014.

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J. M., & Estoup, A. (2014). DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30(8), 1187-1189.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, 17(5), 1170-1188.

Eliades, N. G., Eliades, D.G. (2009). Haplotype Analysis: software for analysis of haplotype data. Georg-August University Göttingen, Germany. Available at: http://www.uni-goettingen.de/en/134935.html. [Accessed: 05.10.2016].

Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual review of Ecology and Systematics*, 24(1), 217-242.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564-567.

Fekete, G., Molnár, Z., Magyari, E., Somodi, I., & Varga, Z. (2014). A new framework for understanding Pannonian vegetation patterns: regularities, deviations and uniqueness. *Community Ecology*, 15(1), 12-26.

Fekete, L., Blattny, T. (1913). Az erdészeti jelentőségú fák és cserjék elterjedése a Magyar Állam területén (The distribution of forest trees and shrubs in Hungary). *Erdészeti Lapok,* 52:743-758.

Feurdean, A., & Bennike, O. (2004). Late Quaternary palaeoecological and palaeoclimatological reconstruction in the Gutaiului Mountains, northwest Romania. *Journal of Quaternary Science,* 19(8), 809-827.

Feurdean, A., Björkman, L., & Wohlfarth, B. (2012). A paleoecological reconstruction of the Late Glacial and Holocene based on multidisciplinary studies at Steregoiu site (Gutai Mts., Romania). *Studia UBB Geologia,* 46(2), 125-140.

Feurdean, A., Tanţău, I., & Fărcaş, S. (2011). Holocene variability in the range distribution and abundance of *Pinus, Picea abies,* and *Quercus* in Romania; implications for their current status. *Quaternary Science Reviews,* 30(21), 3060-3075.

Feurdean, A., Wohlfarth, B., Björkman, L., Tantau, I., Bennike, O., Willis, K. J., Farcas, S., & Robertsson, A. M. (2007). The influence of refugial population on Lateglacial and early Holocene vegetational changes in Romania. *Review of Palaeobotany and Palynology,* 145(3), 305-320.

Feurdean, A., Wohlfarth, B., Björkman, L., Tantau, I., Bennike, O., Willis, K. J., Farcas, S., & Robertsson, A. M. (2007). The influence of refugial population on Lateglacial and early Holocene vegetational changes in Romania. *Review of Palaeobotany and Palynology,* 145(3), 305-320.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics,* 180(2), 977-993.

Fu, Y.-X., & W.-H. Li. (1993). Statistical tests of neutrality of mutations. *Genetics,* 133:693-709.

Giertych, M., & Mátyás, C. (1991). *Genetics of Scots pine*. Amsterdam, Elsevier. 3:19-76.

Gömöry, D., Paule, L., Krajmerová, D., Romšáková, I., & Longauer, R. (2012). Admixture of genetic lineages of different glacial origin: a case study of *Abies alba* Mill. in the Carpathians. *Plant Systematics and Evolution,* 298(4), 703-712.

González-Martínez, S. C., Ersoz, E., Brown, G. R., Wheeler, N. C., & Neale, D. B. (2006). DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics,* 172(3), 1915-1926.

Grivet, D., Sebastiani, F., Alía, R., Bataillon, T., Torre, S., Zabal-Aguirre, M., Vendramin G. G., & González-Martínez, S. C. (2011). Molecular footprints of local adaptation in two Mediterranean conifers. *Molecular Biology and Evolution,* 28(1), 101-116.

Grivet, D., Sebastiani, F., González- Martínez, S. C., & Vendramin, G. G. (2009). Patterns of polymorphism resulting from long- range colonization in the Mediterranean conifer Aleppo pine. *New Phytologist,* 184(4), 1016-1028.

Höhn, M., Gugerli, F., Abran, P., Bisztray, G., Buonamici, A., Cseke, K., Hufnagel, L., Quintela-Sabarís, C., Sebastiani, F., & Vendramin, G. G. (2009). Variation in the chloroplast DNA of Swiss stone pine (*Pinus cembra* L.) reflects contrasting post-glacial history of populations from the Carpathians and the Alps. *Journal of Biogeography,* 36(9), 1798-1806.

Homolka, A., Schueler, S., Burg, K., Fluch, S., & Kremer, A. (2013). Insights into drought adaptation of two European oak species revealed by nucleotide diversity of candidate genes. *Tree Genetics & Genomes,* 9(5), 1179-1192.

Jensen, J. L., Bohonak, A. J., & Kelley, S. T. (2005). Isolation by distance, web service. *BMC Genetics*, 6(1), 13.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge, Cambridge University Press. pp. 117-137.

Krutovsky, K. V., & Neale, D. B. (2005). Nucleotide diversity and linkage disequilibrium in cold-hardiness-and wood quality-related candidate genes in Douglas fir. *Genetics*, 171(4), 2029-2041.

Kujala, S. T., & Savolainen, O. (2012). Sequence variation patterns along a latitudinal cline in Scots pine (*Pinus sylvestris*): signs of clinal adaptation? *Tree Genetics & Genomes*, 8(6), 1451-1467.

Lalagüe, H., Csilléry, K., Oddou-Muratorio, S., Safrana, J., de Quattro, C., Fady, B., González-Martínez, S. C., & Vendramin, G. G. (2014). Nucleotide diversity and linkage disequilibrium at 58 stress response and phenology candidate genes in a European beech (*Fagus sylvatica* L.) population from southeastern France. *Tree Genetics & Genomes*, 10(1), 15-26.

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452.

Liepelt, S., Cheddadi, R., de Beaulieu, J. L., Fady, B., Gömöry, D., Hussendörfer, E., Konnert, M., Litt, T., Longauer, R., Terhürne-Berson, R., & Ziegenhagen, B. (2009). Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.)—a synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, 153(1), 139-149.

Manni, F., Guerard, E., & Heyer, E. (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*, 76(2), 173-190.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2 Part 1), 209-220.

Matías, L., & Jump, A. S. (2012). Interactions between growth, demography and biotic interactions in determining species range limits in a warming world: the case of *Pinus sylvestris*. *Forest Ecology and Management*, 282, 10-22.

Milner, A. M., Müller, U. C., Roucoux, K. H., Collier, R. E., Pross, J., Kalaitzidis, S., Christanis, K., & Tzedakis, P. C. (2013). Environmental variability during the Last Interglacial: a new high- resolution pollen record from Tenaghi Philippon, Greece. *Journal of Quaternary Science*, 28(2), 113-117.

Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288-295.

Peery, M. Z., Kirby, R., Reid, B. N., Stoelting, R., Doucet-Bëer, E., Robinson, S., Vásquez-Carrillo, C., Pauli, N. J., & Palsbøll, P. J. (2012). Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, 21(14), 3403-3418.

Pócs, T. (1960). Die zonalen Waldgesellschaften Südwestungarns. (The zonal forest communities of southwestern Hungary) *Acta Botanica Hungarica*, 6, 75-105.

Pravdin, L. F. (1969). *Scots pine: variation, intraspecific taxonomy and selection*. London, Annarbor Humphrey Science Publishers Ltd., p. 208.

Pritchard, J. K., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

Pyhäjärvi, T., Garcia-Gil, M. R., Knürr, T., Mikkonen, M., Wachowiak, W., & Savolainen, O. (2007). Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics*, 177(3), 1713-1724.

Sadori, L., Koutsodendris, A., Masi, A., Bertini, A., Combourieu-Nebout, N., Francke, A., Kouli, K., Joannin, S., Mercuri, A. M., Panagiotopoulos, K., Peyron, O., Torri, P., Wagner, B., Zanchetta, G., & Donders, T. H. (2015). Pollen-based paleoenvironmental and paleoclimatic change at Lake Ohrid (SE Europe) during the past 500 ka. *Biogeosciences Discussions*, 12(18), 15461-15493.

Sebastiani, F., Pinzauti, F., Kujala, S. T., González-Martínez, S. C., & Vendramin, G. G. (2012). Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. *Conservation Genetics Resources*, 4(2), 231-234.

Soranzo, N., Provan, J., & Powell, W. (1998). Characterization of microsatellite loci in *Pinus sylvestris* L. *Molecular Ecology*, 7(9), 1260-1261.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585-595.

Tantau, I., Reille, M., de Beaulieu, J. L., & Farcas, S. (2006). Late Glacial and Holocene vegetation history in the southern part of Transylvania (Romania): pollen analysis of two sequences from Avrig. *Journal of Quaternary Science*, 21(1), 49-61.

Tantau, I., Reille, M., de Beaulieu, J. L., Farcas, S., Goslar, T., & Paterne, M. (2003). Vegetation history in the Eastern Romanian Carpathians: pollen analysis of two sequences from the Mohoş crater. *Vegetation History and Archaeobotany*, 12(2), 113-125.

Tollefsrud, M. M., Kissling, R., Gugerli, F., Johnsen, Ø., Skrøppa, T., Cheddadi, R., van der Knaap, W. O., Latałowa, M., Terhürne-Berson, R., Litt, T., Geburek, T., Brochmann, C. & Sperisen, C. (2008). Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology*, 17(18), 4134-4150.

Tribsch, A., & Schönswetter, P. (2003). Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon*, 52(3), 477-497.

Vendramin, G. G., Lelli, L., Rossi, P., & Morgante, M. (1996). A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. *Molecular Ecology*, 5(4), 595-598.

Wachowiak, W., Balk, P. A., & Savolainen, O. (2009). Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes*, 5(1), 117.

Wachowiak, W., Boratyńska, K., & Cavers, S. (2013). Geographical patterns of nucleotide diversity and population differentiation in three closely related European pine species in the *Pinus mugo* complex. *Botanical Journal of the Linnean Society*, 172(2), 225-238.

Wachowiak, W., Salmela, M. J., Ennos, R. A., Iason, G., & Cavers, S. (2011). High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland. *Heredity*, 106(5), 775-787.

Wachowiak, W., Wójkiewicz, B., Cavers, S., & Lewandowski, A. (2014). High genetic similarity between Polish and North European Scots pine (*Pinus sylvestris* L.) populations at nuclear gene loci. *Tree Genetics & Genomes*, 10(4), 1015-1025.

PUBLICATIONS CONNECTED TO THE DISSERTATION

Papers in impact factored journals:

Tóth, E. Gy., Vendramin, G. G., Bagnoli, F., Cseke, K., & Höhn, M. (2017). High genetic diversity and distinct origin of recently fragmented Scots pine (*Pinus sylvestris* L.) populations along the Carpathians and the Pannonian Basin. *Tree Genetics & Genomes*, 13(2), 47.

Tóth, E. Gy., Köbölkuti, Z. A., Pedryc, A., & Höhn, M. (2017). Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. *Journal of Forestry Research*, 1-15.

Other papers:

Tóth E., Köbölkuti Z. A., Pedryc A., Höhn M. (2015). Az SNP markerek és alkalmazhatóságuk fenyőfajok molekuláris genetikai vizsgálatainál. *Kertgazdaság*, 47(2), 79-87.

Conference papers (abstracts):

Höhn, M., Köbölkuti, Z. A., **Tóth, E. Gy.** (2017). A Fenyőkúti tőzegláp erdeifenyves populációjának különleges helyzete molekuláris markerek vizsgálata alapján. IV.Transylvanian Horticulture and Landscape Studies Conference. Marosvásárhely. Romania.

Tóth, E., Höhn, M. (2016). A Kárpátok és a Pannon medence izolált erdeifenyő (*Pinus sylvestris* L.) populációinak molekuláris genetikai struktúrája és demográfiai története. XI. Aktuális flora- és Vegetációkutatás a Kárpát-medencében (11th International Conference „Advances in research on the flora and vegetation of the Carpato-Pannonian region"). Budapest, Hungary. pp. 7-8, ISBN 978-963-9877-25-2

Tóth, E., Höhn, M. (2016). Adaptive potential of fragmented Scots pine populations from the Carpathians. 5th Workshop of the AForGen network. Sacel. Romania.

Tóth, E., Vendramin, G. G., Bagnoli, F., Köbölkuti, Z. A., Höhn, M. (2016). Filogeográfiai és adaptációs mintázat a közép-kelet-európai Erdeifenyő (*Pinus sylvestris* L.) populációkban. MTA DAB, Molekuláris biodiverzitás: taxonómia, filogenetika, filogeográfia konferencia. Debrecen, Hungary.

Tóth, E., Höhn, M. (2015). Further progresses in population genetic study of Carpathian pine populations: the case of Scots pine 4rd Workshop of the AForGen network. Bavaria, Berchtesgaden, Germany.

Tóth, E., Vendramin, G. G., Bagnoli, F., Cseke, K., Höhn, M. (2015). Genetic structure of relict and isolated Scots pine (*Pinus sylvestris* L.) populations from Central-Eastern Europe: signals of recent population fragmentation? IUFRO Tree Biotechnology Conference 2015, Florence, Italy.

Tóth, E., Andrzej, P., Höhn, M. (2014). Diversity pattern of natural *Pinus sylvestris* L. populations from South-Western Hungary revealed by cpSSR markers. Horticulture in quality and culture of life conference. Lednice, Czech Republic.

Tóth, E., Cseke K., Höhn, M. (2014). Phylogeographic evaluation of *Pinus sylvestris* L. populations from East-Central Europe. 3rd Workshop of the AForGen network. Fafleralp canton, Valis, Switzerland.

Tóth, E., Andrzej, P., Höhn, M. (2014). Peremhelyzetű, közép-kelet európai erdeifenyő (*Pinus sylvestris* L.) populációk genetikai variabilitásának értékelése mikroszatellit markerekkel (Evaluation of genetic variability of situated Middle-Eastern-European Scots pine populations with microsatellite markers). X. Aktuális flora- és Vegetációkutatás a Kárpát-medencében. Sopron, Hungary.

Tóth, E., Andrzej, P., Höhn, M. (2013). Morphological and molecular evaluation of native Scots pine (*Pinus sylvestris* L.) populations from south-eastern Europe. Biogeography of the Carpathians Conference. Krakow, Poland.

Tóth, E., Hufnagel, L., Höhn, M. (2012). Chloroplast microsatellite study reveal lower variability in *Pinus sylvestris* L. populations inhabiting sunny, dry habitats. Interdisciplinarity in Geoscience in the Carpathian Basin (IGCB) conference. Suceava, Romania.