SZENT ISTVÁN UNIVERSITY

FACULTY OF HORTICULTURAL SCIENCE

DEPARTMENT OF BOTANY

ADAPTIVE VARIATION OF SELECTED SCOTS PINE (*PINUS SYLVESTRIS* L.) POPULATIONS FROM THE CARPATHIANS AND THE PANNONIAN BASIN INFERRED FROM MORPHOLOGICAL, ANATOMICAL AND GENOMIC DATA

PH.D. THESIS

ZOLTÁN ATTILA KÖBÖLKUTI

SUPERVISOR: DR. MÁRIA HÖHN

BUDAPEST

2018

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.

....
    Head of Ph.D. School   Supervisor

1. INTRODUCTION AND OBJECTIVES

Scots pine *(Pinus sylvestris* L.) on its extremely large Eurasian distribution range has adapted to a wide variety of climates. In Europe, the species grows on different substrates and soil types: peat bogs, rocky substrates and also young glacial deposits, siliceous and acidic soils, frequently with deep litter and raw humus layer (Steinbeck 1966). The species' southern limit is more closely determined by edaphic conditions like soil moisture, southern populations representing the highly fragmented range of species' distribution. Peripheral populations cannot become established beyond these previously mentioned conditions because of the restrictive, unfavourable habitats (Bridle and Vines 2007). However the dynamics of populations on these latitudinal margins are critically important in the response of the species to expected climate change (Iverson et al. 2004).

It is accepted that on the periphery populations have maintained substantial genetic variation because of adaptation to different selective pressures and the reduced gene flow (Lenormand 2002). These occurences are considered of great importance for the long-term conservation of species (Hampe and Petit 2005). Resident genotypes of these peripheral populations have a higher relative fitness in their local habitat than genotypes originating from other habitats. Although well adapted to local environments with high phenotypic plasticity, the composition and structure of peripheral populations are determined also by historical events, long term geological changes, forcing populations to *in situ* adaptation (Losos 1996).

The specific migration history, natural selection over extended periods of time has left traces on genetically determined morphological and anatomical traits of populations and thus can be concluded that they display not only "geographic marks", but also genetic differences and differentiation at the phenotypic level. By this way, morphological studies offer a feasible method to compare these populations evolved under different conditions.

Beside the studies on variation of growth and survival, based on morphological, anatomical observations and provenance studies, during the last decades, the advancement in the field of molecular biology has facilitated the development of a range of molecular genetic markers. Furthermore, the rapid development of next-generation sequencing technology provided an opportunity to develop novel genomic tools. The lower cost and greater sequence yield has allowed the identification of candidate genes, functional genomic level data with genome characteristics. These data that have high functional information content, often correspond to genes with known or predicted functions, and so have proven to be unavoidable in comparative genomics (Vera et al.

2008) and population genomic studies of genetic variation associated with adaptive traits. The analysis of gene content is also very helpful for developing SNP markers, a useful tool for understanding the adaptive response in stressful environmental conditions and against pathogens.

The overall objectives of this research were as follows:

• Detect the level of phenotypic differentiation based on cone morphology and needle anatomy in selected peripheral populations of *Pinus sylvestris* L. in the Pannonian Basin and the Carpathian Mountains.

• Identify, if there is any concordance with pollen-based historical data about the common origin of these populations.

• Discern possible groups of populations by significant morphological and anatomical differentiation in response to different type of habitat.

• Quantify variation in seed traits and germination power among and within marginal populations, considering the type of habitat.

• Identify candidate genes with role in adaptation and develop novel SNP markers to assess the nucleotide diversity at these candidate gene loci, with the aim to infer adaptive responses.

2. MATERIAL AND METHODS

2.1. Plant material

The Scots pine plant material originates from Central-Eastern Europe, from the area of the Carpathian Mountains and the Pannonian Basin. For cone morphological and needle anatomical study 16 natural populations of *Pinus sylvestris* were sampled between 2011 – 2015.

The study of seed morphology and germination was completed on less populations (10) than the cone morphological and needle anatomical survey. All of these marginal populations within the natural range of the species originated from specific habitat types, such as raised bogs, dry rocky surfaces, or mixed forests on specific substrates with low nutrient content.
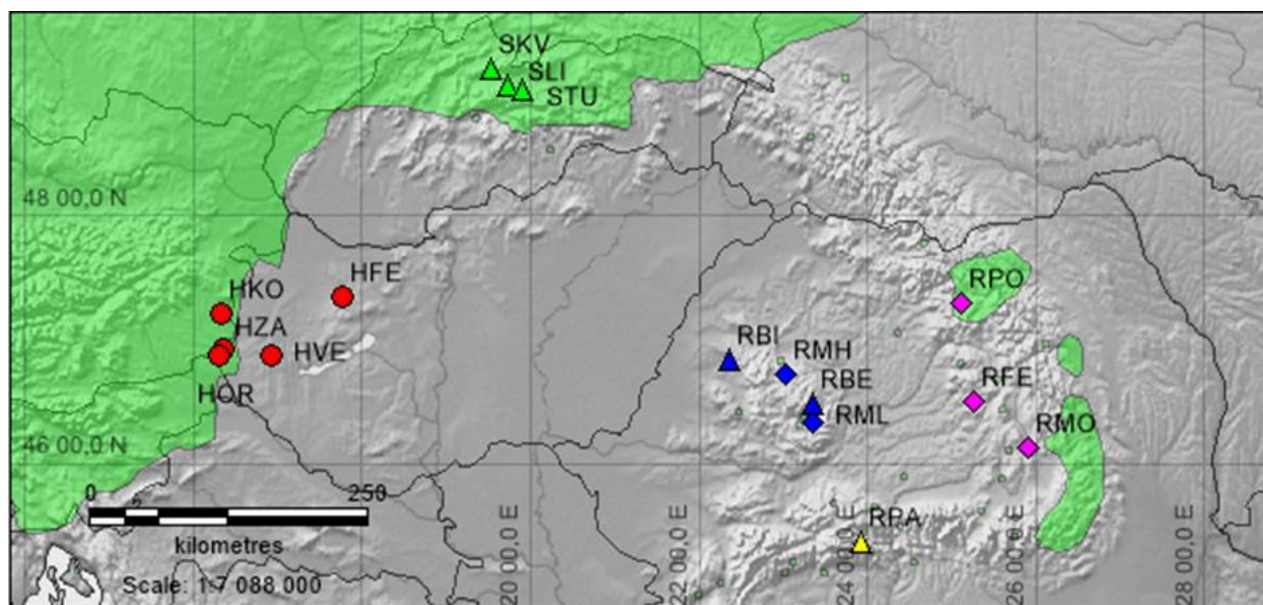

Figure 1. Sampled populations of *Pinus sylvestris*. The colors indicate the geographical affiliation (red: Pannonian basin (PB), green: Northern Carpathians (NC), blue: Central-Island Mountains/ Apuseni Mt. (CIM), pink: Eastern Carpathians (EC) and yellow: Southern Carpathians (SC)). The shape of icons (rectangulars: peat bogs, triangles: rock surfaces, circles: mixed forests) represents the type of the habitat. In green is highlighted the Scots pine distribution area based on Euforgen map (http://www.euforgen.org/species/pinus-sylvestris). For population abbreviation see the Abbreviation list of the sampled populations.

For the germination study, a progeny trial was established in the Botanical Garden of Soroksár, in spring 2016, with 15 seeds/mother tree sampled from four mother trees per population.

The marker development was completed after a previous pre-analysis in *Pinus cembra* L., at the Genetic Department, BFW, Vienna, Austria. The Swiss stone pine samples for this purpose were sampled from a single tree in Obergurgel, Austria, from two cones in different developmental stages and a needle sample (Jahn and Heinze, unpublished). The primers designed on *de novo* identified gene

sequences were tested in the laboratory at first on 84 Swiss stone pine DNA samples, selected at random, from six sites along the Austrian Alps, then on Scots pine samples originated from the Carpathian area, originating from three different types of habitat: peat bog [Mohos (RO)]; rocky surface [Kvacany (SK)] and beechpine mixed forest [Csörötnek (HU)], populations formerly included in the morphological and anatomical survey, as well as in a microsatellite analysis (Tóth et al. 2017).

2.2. Data collection

In reference to the studied traits, 12 cone and seed morphological, eight needle anatomical characters were measured and in total 12 ratios were calculated. Germination was scored once weekly between 5th of April to 19th of June 2016, followed by calculation of six germination associated parameters (International Seed Testing Association-ISTA 1985).

For marker development on the *Pinus cembra* transcriptome, a data mining was performed, from literature where candidate genes with possible role in the adaptation process were selected. Then a BLAST database with 5187 EST and protein sequences was filled, downloaded from NCBI database (http://www.ncbi.nlm.nih.gov). *De novo* assembled contigs were searched using BLASTN and BLASTX toolkits against the database. By this method, sequences were annotated, renamed by the coded protein, and selected for primer design. All previously tested markers on Swiss stone pine, to evaluate their functionability and transferability, were tested in the laboratory on Scots pine DNA samples. The PCR products that resulted from the CleanSweep process were sequenced using the forward primers, in one direction. To determine whether the previously annotated candidate genes can indeed be associated to the putative enzyme or transcription factor and to find out to what extent and in which manner non-synonymous mutations or indels affect the nature of the encoded protein, BLASTX search against the NCBI database was completed with each amplified sequence of every coding marker (amplified from samples from different type of habitat).

2.3. Statistical analysis

To investigate the cone traits and needle anatomical parameters statistical analysis was carried out on a dataset containing 4448 measured parameters. Methods were chosen according to Jasińska *et al.* (2014), Marcysiak (2006), Staszkiewicz (1961), Turna and Güney (2009). Multivariate ANOVA (MANOVA), discriminant analysis, and the Mantel test were performed. Maximal-minimal values, arithmetic means, and standard deviations were calculated and analyzed for all populations and

population-groups. The one-way multivariate analysis of variance (MANOVA) test with geographical position or habitat type as a factor was used, followed by variable-wise between-subjects effects analysis, to evaluate the significance of differences among populations for particular characteristics. We applied discriminant function analysis to predict a categorical dependent variable and determine whether a certain set of variables was effective in predicting category membership.

Statistical analysis was carried out on 3240 sampled morphological data to investigate the seed traits and the relationship between seed variables and germination associated parameters. Morphological variation was analyzed with IBM SPSS 20.0 (IBM Corp.) and Microsoft Excel. One-Way ANOVA, discriminant analysis and the Mantel test were performed. Maximal-minimal values, arithmetical means and standard deviations were calculated and analyzed for all populations. One-way analysis of variance (ANOVA) was used to determine significant differences between the means of variables. Bivariate Correlation-analysis was used to detect relationship between each seed variable and also, between seed variables and germination associate parameters. We applied discriminant-analysis at first only with seed morphological dataset, then with both morphological and germination parameters to predict a categorical dependent variable and to determine whether a set of variables is effective in predicting category membership. The analysis was performed by stepwise method. We sorted the studied populations according to the type of their habitat (peat bog, rocky surface, and mixed forest) to detect any grouping by traits, which are or not suited to the specific environment. Mantel test (Mantel 1967) was performed like in case of the morphoanatomical variables. Euclidean distances and geographical distances between populations were used for the evaluation using GenAlEx 6.5 (Peakall and Smouse 2012) software.

For editing and visual organization of the sequences as well as for the analysis of SNPs and indels (insertions/deletions) within the gels, the sequences were edited and analysed using BioEdit Sequence Alignment Editor version 7.0.9.0 (Hall 1999). Number of polymorphic sites, number of haplotypes, haplotype diversity, the variance and standard deviation of haplotype diversity, nucleotide diversity and the average number of nucleotide differences were calculated using DNA Sequence Polimorphism v6.10.01 (Rozas and Rozas 1995).

3. RESULTS

With pre-formed groups according to geographical distribution one-way MANOVA test on cone morphological traits yielded significant differences $(F(20;820) = 8.25; p < 0.001)$ with significant between-subject effects $(F(4;251) > 6.1; p < 0.001)$. We have found that in the case of five variables, populations from the Northern Carpathians and the Pannonian Basin form one group, while cones from the Eastern Carpathians are separated by four variables, being in one group with Central Island Mountains based on two variables (**Figure 2**).
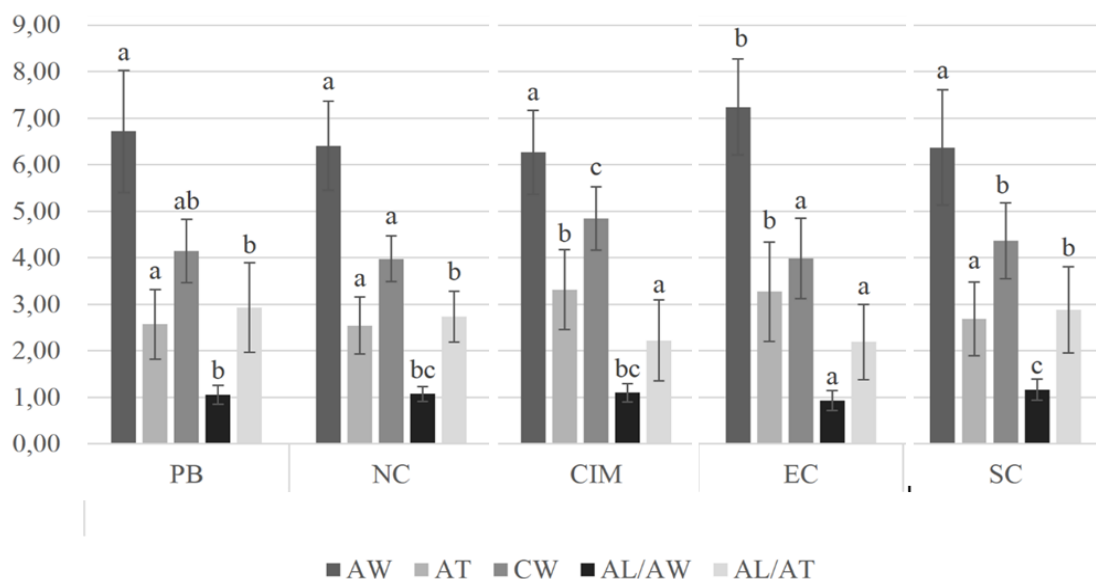

Figure 2. Statistically significant morphometric variables of Scots pine between the studied geographical regions. The PB abbreviation means Pannonian basin (HKO, HFE, HVE, HZA, HOR), NC: Northern Carpathians (SKV, STU, SLI, CHR), CIM: Central-Island Mountains (Apuseni) (RBI, RBE, RML, RMH), EC: Eastern Carpathians (RPO, RMO, RFE) and SC: Southern Carpathians (RPA) respectively. For population abbreviation see the Abbreviation list of the sampled populations.

Performing the MANOVA test on needle anatomical traits with pre-formed geographical groups, significant differences were found based on four variables and three ratios: $(F(16;758) = 10.35; p < 0.001; F(126;659) = 12.62; p < 0.001)$ with significant between-subjects effects $(F(4;251) > 14.90; p < 0.001; F(4;251) > 9.94; p < 0.001)$. Populations from the Pannonian Basin were significantly separated from the Northern Carpathians by four variables and three ratios (**Figure 3**).
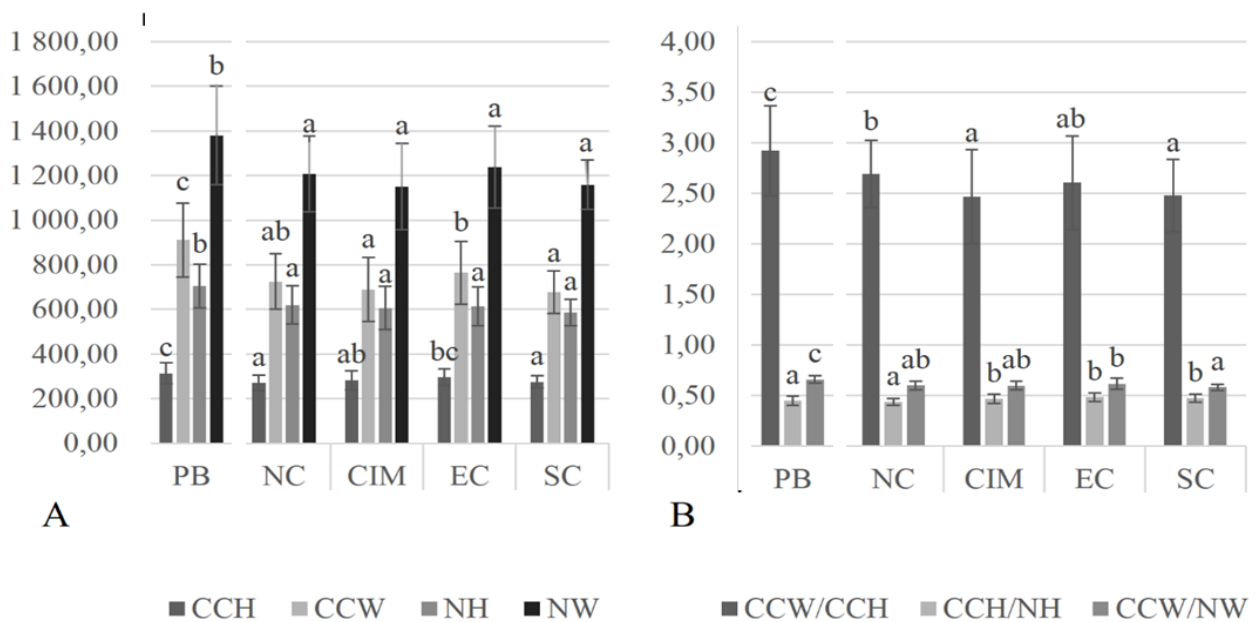
Figure 3. Statistically significant anatomical variables of Scots pine needles between the studied geographical regions. The PB abbreviation means Pannonian basin (HKO, HFE, HVE, HZA, HOR), NC: Northern Carpathians (SKV, STU, SLI, CHR), CIM: Central-Island Mountains (Apuseni) (RBI, RBE, RML, RMH), EC: Eastern Carpathians (RPO, RMO, RFE) and SC: Southern Carpathians (RPA) respectively. For population abbreviation see the Abbreviation list of the sampled populations.

According to the discriminant function analysis, on the basis of the first variable (Function 1), which was responsible for 60.6 % of the variation, the centroids of the populations were split into two distinct groups: one comprises the Central Island Mts. (3), the Northern Carpathians (2), and the Southern Carpathians (5), and a second, separate group is formed by populations from the Pannonian Basin (1) and the Eastern Carpathians (4) (**Figure 3A**). The second variable (Function 2), which was responsible for 20.7 % of the total variation, differentiated the populations from the Pannonian Basin (1) from populations in the Eastern Carpathians (4), but the Northern (2) and Southern Carpathians (5) still were clustered in one group. By carrying out discriminant function analysis on needle anatomical data set, by the first variable (Function 1), which was responsible for 41.3 % of the variation, populations formed two distinctive groups (**Figure 3B**): the Pannonian Basin (1), the Central Island Mts. (3), the Eastern Carpathian (4) vs. Northern (4) and Southern Carpathians (5). The second function (Function 2), which was responsible for 32.2 % of the total variation, made evident the separation of the Pannonian Basin (1) from the Central Island Mts. (3) and Eastern Carpathians.
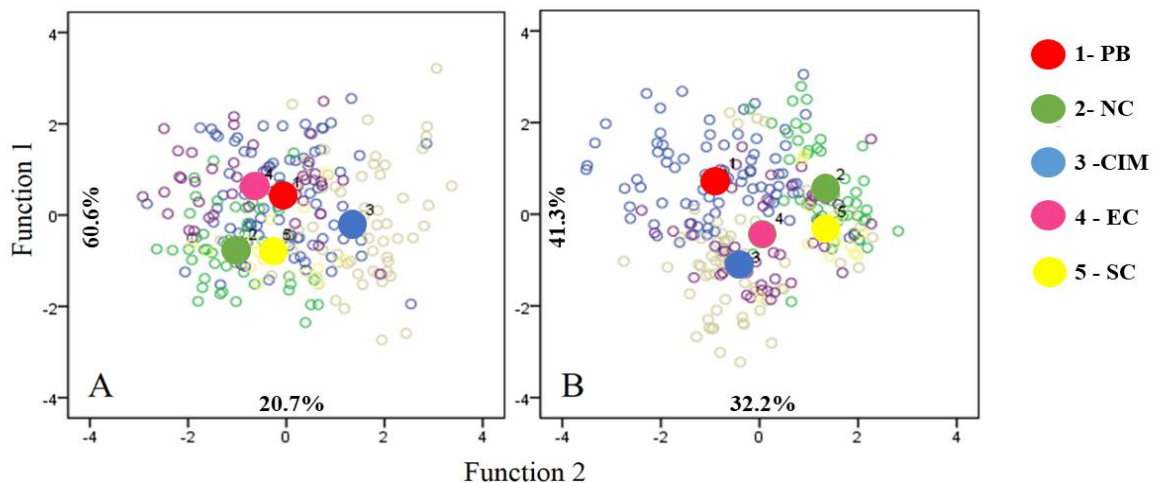
10

Figure 3. Canonical discriminant function analysis result of *Pinus sylvestris* populations by geographical distribution based on cone morphological characters (A) and needle anatomical variables (B). The PB abbreviation means Pannonian basin (HKO, HFE, HVE, HZA, HOR), NC: Northern Carpathians (SKV, STU, SLI, CHR), CIM: Central-Island Mountains (Apuseni) (RBI, RBE, RML, RMH), EC: Eastern Carpathians (RPO, RMO, RFE) and SC: Southern Carpathians (RPA) respectively. For population abbreviation see the Abbreviation list of the sampled populations.

The Mantel correlation test was not significant ($R2 = 0.017$, $p < 0.05$) neither on cone morphological, nor on needle anatomical data set.

By carrying out a one-way MANOVA test with pre-formed groups according to the habitat type using the morphological cone dataset, significant differences were detected ($F(10;498) = 13.69$; $p < 0.01$) by revealing significant between-subjects effects ($F(2;253) > 3.46$; $p < .05$). Populations of peatbog provenience differed significantly from those of rocky outcrops and mixed forests in the case of two, from dry rocky outcrops according to three variables (**Figure 4A**). On needle anatomical data, the one-way MANOVA test with pre-formed groups detected significant differences ($F(8;500) = 12.64$; $p < 0.001$) with significant between-subjects effects ($F(2;253) > 8.00$; $p < 0.001$). Significantly less resin ducts were detected in peat bogs and populations and on rocky surfaces. In the case of three proportions, populations from mixed forest differed significantly by higher values at two and lower values at one ratios (**Figure 4B**).
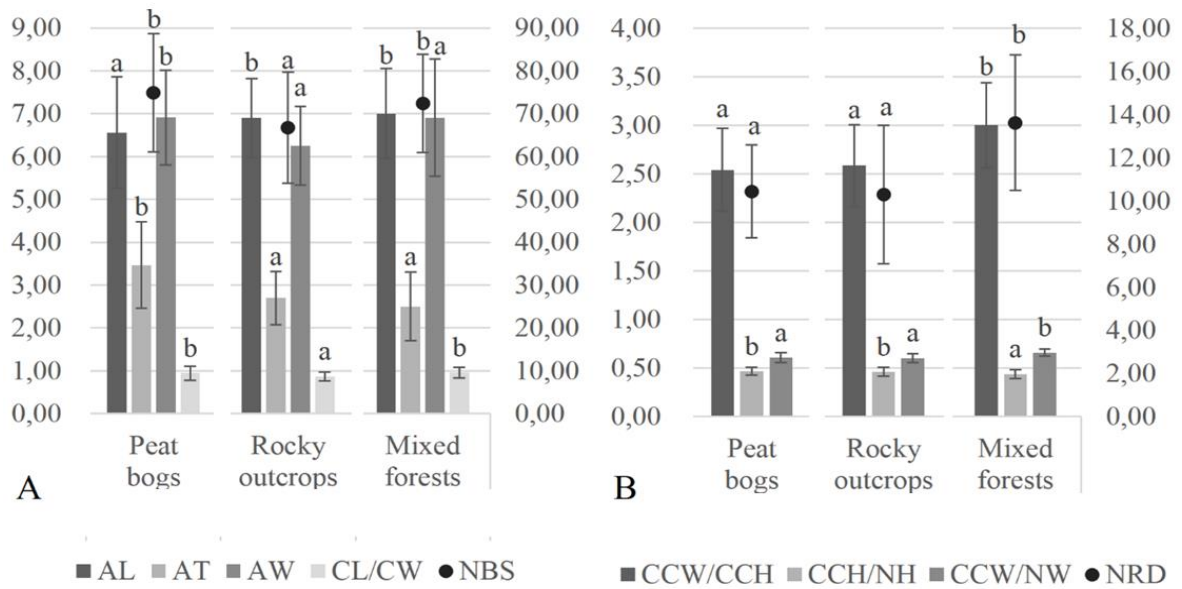
Figure 4. Statistically significant morphometric and anatomical variables of Scots pine individuals within the studied habitat types (peat bogs, rocky outcrops and mixed forests). For population abbreviation and population affiliation to the specific habitat type see the Abbreviation list of the sampled populations.

Discrimination analysis based on the measured cone morphological datasets, with the first variable (Function 1) responsible for 69.6 % of the variation and the second variable (Function 2) responsible for 30.4 % of the variation, revealed a slight pattern of populations by separating them into two groups: mixed forests (3) with rocky surfaces (2) vs. peat bogs (1)(**Figure 5A**). The same analysis with anatomical needle variables after a previous sorting of populations by habitat type showed that the first variable (Function 1) was responsible for 59.2 % of the variation and the second variable (Function 2) was responsible for 40.8 % of the variation. Three groups were identified: mixed forests (3), rocky surfaces (2), and peat bogs (**Figure 5B**).
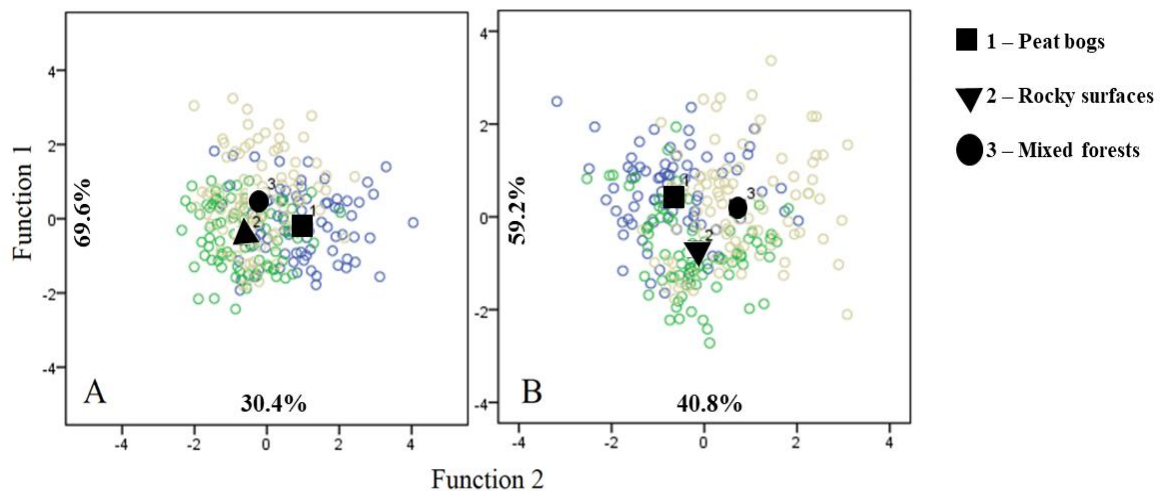

Figure 5. Canonical discriminant function analysis result of *Pinus sylvestris* populations by habitat type detected on morphological cone characters (A) and on needle anatomic variables (B): 1- peat bogs, 2- rocky surfaces, 3- mixed forests. For population abbreviation and population affiliation to the specific habitat type see the Abbreviation list of the sampled populations.

Discrimination analysis based on the measured seed morphological datasets with three groups pre-formed according to habitat type, with the first variable (Function 1) responsible for 87.3 % of the variation and the second (Function 2) responsible for 12.7 % of the variation revealed a slight pattern of populations by separating into two groups: mixed forests (3) with rocky substrate (2) vs. peat bog (1) (**Figure 6A**). With germination associated parameters also included in the analysis, with the first variable (Function 1) responsible for 93.3 % of the variation and the second (Function 2) responsible for 6.7 % of the variation revealed likewise in case of the result based only on morphological data, the separation of populations into two groups: mixed forests (3) with rocky substrate (2) vs. peat bog (1), but with a more stronger pattern of differentiation between the two groups (**Figure 6B**).
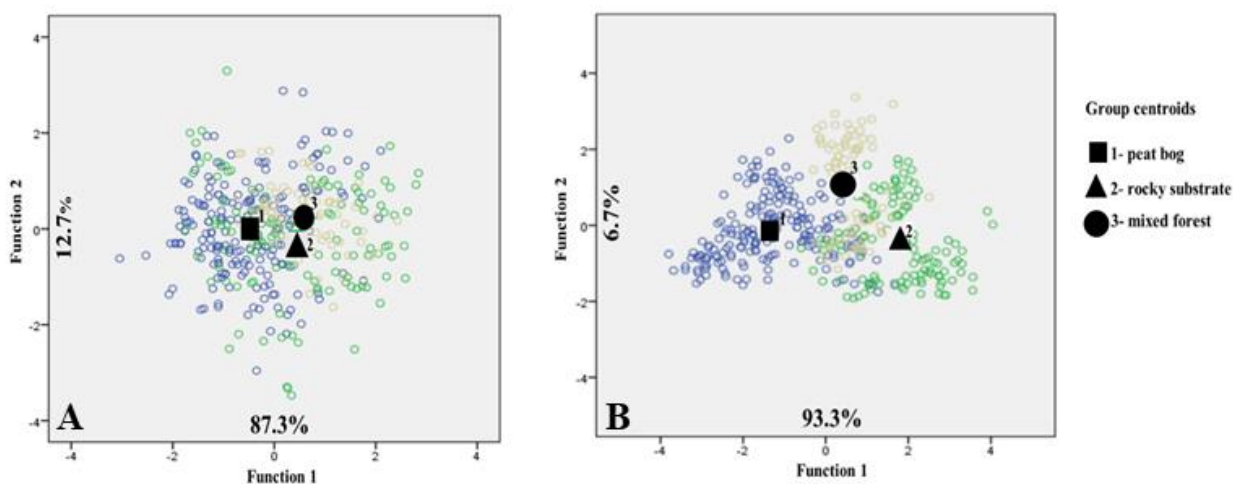

Figure 6. Differentiation of *Pinus sylvestris* L. populations by habitat type detected by discriminant function analysis, on the basis of morphological seed characteristics (A) and of morphological seed characteristics with germination associated parameters (B) : 1 – peat bogs, 2 – rocky substrates, 3 – mixed forests. For population abbreviation and population affiliation to the specific habitat type see the Abbreviation list of the sampled populations.

The dendrogram resulted from hierarchical cluster analysis, shows that samples from RMO and RPS with similar type of habitat (oligotrophic peat bog) form one subcluster, and also RCO and CHR samples (rocky substrate) form another subcluster. HZA population (mixed forest) is the closest to these two subclusters. Another subcluster comprises RBE and RVR (rocky substrate and mixed forest) samples and a special highlight is needed on SME peatbog population, characterised by a completely distant position (**Figure 7**).
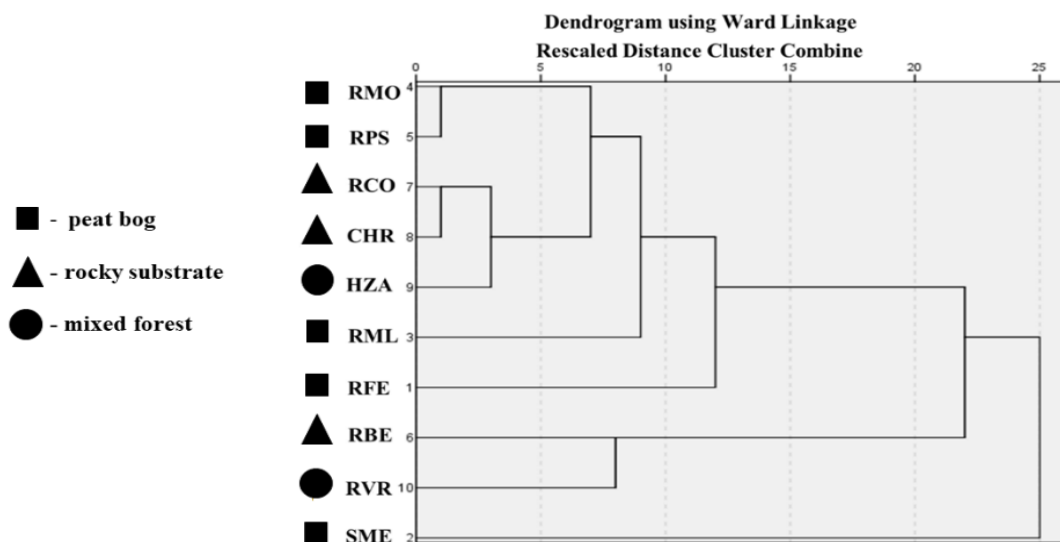
Dendrogram using Ward Linkage
Rescaled Distance Cluster Combine

Figure 7. Dendrogram generated with IBM SPSS 20.0 using Ward Linkage among 10 *Pinus sylvestris* L. populations from different habitats (peat bog, rocky substrate, mixed forest). For population abbreviation and population affiliation to the specific habitat type see the Abbreviation list of the sampled populations.

The results of Bivariate Correlation-analysis to detect relationship between seed morphological variables and defining correlation as significant at the 0.01 and 0.05 level, showed strong correlation between most of the variables. We performed the same analysis to detect relationship between seed morphology and germination associated data and define correlation as significant at the same level, correlation between Germination speed with seed length/wing length and seed width/wing width ratios. Likewise we calculated Germination percentage compared to seed length, seed width and seed weight.

Marker development and the cross species transfer test, after screening the contigs assembled from the three *Pinus cembra* transcriptomes using BLASTN and BLASTX toolkits against the BLAST database, from the MultiBLAST resulted 399 sequences that could have been selected, annotated and renamed by the coded protein. The designed primers successfully amplified 164 sequences on *Pinus cembra* samples. 68 PCR products appeared as a single band in the electrophoresis gel. 68 markers were tested on *Pinus sylvestris* DNA samples, and 53 provided PCR products. Of these 53, we detected 25 PCR products that occurred as a single band in the electrophoresis gels. After sequencing, we obtained in case of 20 sequences conclusive results on the amplified products of all three samples. The putative similarity of the sequences was estimated according to the best BLAST hit. In case of three genes no significant similarity was found. Accordingly, these sequences were excluded from further analyses. In total, 61 SNPs were found. Nine SNPs were found in transcription factor coding genes, but only in one particular sample. Indels were found in case of seven out of the 22 candidate

genes. Synonymous single nucleotide polymorphisms were found in case of six sequences, ranging from one to eight. Non-synonymous single base mutations were detected ranging from one to 14, in 11 of the candidate genes.

4. CONCLUSIONS

The detectable level of differences found in our study in case of peripheral populations of Scots pine are in line with the earlier literature results on the high level of variation. Differentiation among populations was reported based on needle (Boratynska and Hinca 2003, Jasińska *et al.* 2014, 2010, Pardos *et al.* 1990) and cone characteristics (Marcysiak 2006, Staszkiewicz 1993). Urbaniak *et al.* (2003) also found differentiation among populations, detected on the basis of morphological character expression, influenced by both the edaphic conditions and the distinct genetic structure. Morphological and anatomical differences among populations are also listed as distinguishing characteristics among populations in the work of Bobowicz and Korczyk (2000).

Our analyses showed in the first instance a grouping of populations by their geographical position. Populations from the Northern Carpathians and the Pannonian Basin formed a well-distinguished group on the basis of five cone and two needle characteristics. These findings are generally congruent with previous molecular studies (Bernhardsson *et al.* 2016, Cheddadi *et al.* 2006, Tóth *et al.* 2017, Naydenov *et al.* 2005), macrofossil and pollen data analyses (Damblon 1997, Haesaerts *et al.* 1996, Jankovska and Pokorný 2008, Richardson and Rundel 1998, Rudner *et al.* 1995, Rudner and Sümegi 2001, Stieber 1967, Willis and Van Andel 2004). According to these it is presumed that one main recolonization route in Europe existed for Scots pine, that presumably originated from around the Eastern Alps and the surroundings of the Danube plain. There are evidences of refugial locations in the Eastern Alps and East-Central Europe, e.g. the Hungarian plain *(*Bernhardsson *et al.* 2016, Cheddadi *et al.* 2006, Naydenov *et al.* 2005).

Our findings are in line with the literature. Based on needle anatomical characteristics, populations from the Pannonian Basin are significantly differentiated according to six anatomical variables. On the other hand, on the basis of four morphometrical variables in the MANOVA test and discriminant function analysis, the Eastern Carpathian region proved to be distant from the rest of the populations. The study by Magyari *et al.* (2014), on pollen based reconstruction of the LGM vegetation in the Eastern Carpathians between ∼22,870 BP and 19,150 BP, supported the persistence of *Pinus sylvestris* in the Eastern Carpathians.

Discriminant function analysis based on the eight measured cone characteristics revealed clearly discernible groups of populations by the habitat type. Nevertheless, cone size and weight can be influenced by tree age, general health of the trees, and the macro- and micro- habitat of the parent trees (Dangasuk and Panetsos 2004). The high levels of variation observed in cone morphology might

be explained by the long-term adaptation of populations to diverse and changing environmental conditions, and they can be due also to the lack of competition with other pines (Gil *et al.* 2002).

The results of one-way MANOVA test on the cone datasets revealed significant differentiation among populations growing in peat bogs and on rocky surfaces. Taking into consideration the fact that cones from several geographically different peat bog locations were not significantly larger than those from other habitat types, it can be concluded that in peat bogs cone structure, and not cone size, represents a difference that can be evaluated as a sign of local adaptation. Lack of geographic correspondence among populations with similar phenotypes was also observed in *P. canariensis* (Gil *et al.* 2002), *P. radiata* (Forde 1964), and *P. tecunumanii* (Eguiluz 1984).

One-way MANOVA test applied to needle anatomical data showed significantly less resin ducts in populations from peat bogs and rocky surfaces or significantly more in populations of mixed forest provenience. Though not well understood, pine resin may play a role in water regulation (Bell 2010, Farrell *et al.* 1991). All our samples from mixed forests were sampled from areas in the Pannonian Basin with the lowest altitudinal gradient. Accordingly, this could be regarded as a sign of altitudinal adaptation, or it might be due to common geographical origin. However, similar findings were described in *Pinus brutia* by Dangasuk and Panetsos (2004), who have reported the number of resin canals as useful trait for identifying altitudinal and longitudinal adaptative variations within and among populations. Considering that our mixed forest samples were collected from areas in the Pannonian Basin with the lowest altitudinal gradient, our data is in agreement with the findings of Wahid *et al.* (2006), who have found that needle width correlates negatively with altitude in maritime pine (*P. pinaster*).

The level of differences in seed morphology found in our study can be treated as being congruent with several earlier described variations regarded to specific morphological adaptations to different environments (Alía *et al.* 2001, Bilgen and Kaya 2007, Dzialuk *et al.* 2009, Jasińska *et al.* 2014, Kinloch *et al.* 1986, Köbölkuti *et al.* 2017, Labra *et al.* 2006, Prus-Glowacki *et al.* 2003, Pyhäjärvi *et al.* 2007, Semiz *et al.* 2007, Turna 2003). The lower seed weight, seed length/seed width, seed length/wing length and seed width/wing width traits of seeds from RML peat bog population may be an inevitable consequence of resource constraints that limits the ability of the parent plant to control individual seed size (Vaughton and Ramsey 1998). Definitely a peat bog can be characterized by specific edaphic conditions. Our significantly low values for wing width variable in case of SME, RMO and RPS peat bog populations are supported by the findings of (McGinley *et al.* 1990) in lodgepole pine. Our dendrogram showed that populations are not homogeneous regarding seed and

wings morphology when grouping by the habitat type. The reason of this differences existed among the studied population in terms of the morphological characters and type of habitat could be explain by different origin of populations. The distant position of SME population most probably is the result of the introgressive hybridization from the area within hybrid swarm populations of *Pinus sylvestris* and *P. mugo* formerly reported by Christensen and Dar (1997) and Wachowiak and Prus-Glowacki (2008).

Our coefficient correlation analysis result on each morphological seed variable were in agreement with the general tendency of relationships between most of the characters (Cervantes *et al.* 2016, Chambers and Macmahon 1994, Ehrenberg *et al.* 1955, Greene and Johnson 1993).

The Mantel correlation matrices showed no linear relationship between the morphological and geographic distances. This result may reflect that seed's phenotypic variation is determined by covariance between the genetic and environmental effects (Rehfeldt 1991), or could be due to the fact, that our studied Carpathian populations represent only a small geographic range from the species' large distribution area.

By carrying out the discriminant function analysis with both seed morphological and germination associated parameters with pre-formed three groups according to the habitat type, our results revealed a strong pattern of differentiation. Peat bog populations situated between the values of populations from mixed forests and rocky substrate were defined by lower values of percentage of germination, seed length and wing length for Function 1 and for Function 2 by values of Germination Speed, seed weight, seed width, wing width and Germination Value. The size and the number of the seeds produced by the plant are determined by the nutrient status of the mother plant at the time of flower bud initiation, since much of the nutrient content of the seeds must be translocated from the vegetative tissues. As peat bogs develop under ombotrophic or oligotrophic conditions, our lower germination percentage and smaller seed size from the peat bog communities can be explained by the specific conditions characterizing this type of habitat. The results of our Bivariate Correlation analysis was significant at 0.05 level between Percentage of Germination with seed length, seed width and seed weight likewise between Germination Speed with seed length/wing length and seed width/wing width variables. Earlier studies have shown that seedlings germinating from large seeds have higher seedling establishment, growth and survival.

Our results concerning the marker development and the cross species transfer test suggest that such primers can work across a wider range of *Pinus* species. The persistence of homologous sequences in the genomic DNA of these two species such as Swiss stone pine and Scots pine provides

further support for the use of *de novo* primers targeting such regions across different species from the *Pinus* genus. These results seem therefore to indicate that coding sequences functionally annotated can be amplified and utilized as genetic markers in relatively distant species from the same genus, using heterologous primers. By the detection of polymorphic sites, these primers can now be more extensively tested for polymorphisms that are more or less frequent in certain habitat type. It was shown that 25 primers worked in both species tested and several of the amplified fragments showed high levels of polymorphisms. The 22 candidate genes analyzed have shown a range of one to 14 polymorphic sites. The polymorphisms of these sites represents non-conserved positions. Therefore, the mutations may result in changes in above mentioned transcription factors and enzymes secondary structure and consequently, their function (Fu 1995). These observations also suggest that the identification of other coding DNA fragments from the genome may be very useful. The SNPs detected in our study may have important applications in further case studies of adaptation, as protein coding loci could be directly linked with adaptive genetic variation. On the other hand, sequence variation of transcription factors may also be dependent on DNA in 5' UTR regions. Altogether, these genetic differences might strongly determine population fitness. As the ability of species to adapt to changes of the environment depends on polymorphic populations (Krutovsky and Neale 2005), adaptive genetic variation in relevant genes is essential for the long term adaptation process.

5. SUMMARY

In our study we have focused on detecting the level of phenotypic differentiation based on cone morphology and needle anatomy in marginal populations of *Pinus sylvestris* in the Pannonian Basin and the Carpathian Mountains. Our results in concordance with paleo-botanical data indicate a common origin of the populations from the Northern Carpathians and the Pannonian Basin. High levels of variation was observed in cone morphology. Discriminant function analysis based on the eight cone characteristics revealed clearly discernible groups of populations and indicated significant differentiation among populations growing in peat bogs and on rocky surfaces. Significant differences among populations from different habitats were also revealed by comparing needle anatomical variables. The phenotypic differentiation by habitat type based on the measured characters might be evaluated as a sign of local adaptation with detectable phenotypic pattern.

The second part of the experiment was an attempt on quantifying variation in seed traits and germination power among and within the marginal populations, considering the type of habitat. Discriminant function analysis showed significant differentiation of populations growing in peat bogs. Seed length, wing length and germination rate were the most useful traits to identify seeds of peat bog origin, most probably adapted to that specific environment.

In final, as result of the development of new molecular markers on *Pinus cembra*, parts of 25 different candidate genes previously annotated with possible role in adaptation were tested and investigated on *Pinus sylvestris* samples. 22 candidate genes were analyzed considering the polymorphic sites, haplotype diversity, nucleotide diversity, insertions/deletions as well as synonymous and non-synonymous SNPs. Polymorphic sites' values ranged from one to 14, several sequences being considered of special interest in this aspect. The study was able to detect numerous SNPs and indels, potentially involved in adaptation processes to different environmental conditions. Some of the non-synonymous SNPs detected are of special interest because they might have an influence on the protein structure and function. The variation found in these sequences provide new SNP markers with importance in the study of adaptation. As adaptive genetic variation in these relevant genes are essential for the long term adaptation to stressful conditions, the characterisation of SNPs in these markers may contribute to the investigation of the genetic basis of adaptive variation.

6. NEW SCIENTIFIC ACHIEVEMENTS

I. **Common origin of the Scots pine populations from the Northern Carpathians and the Pannonian Basin was revealed based on the performed cone morphometrical and needle anatomical analyses. These findings are in accordance with former phylogeographical evidences.**

II. **Based on cone morphology and needle anatomy, significant differentiation was revealed among populations growing on different extreme ecological sites, as peat bogs and rocky surfaces. These findings allow us to gain important insights into the ongoing processes and predict ecosystem responses to changes in the environment.**

III. **Evidences of local adaptation to different ecological sites based on the variation in seed morphology and germination have been identified. The correlation between seed morphology and germination might be helpful in the early evaluation for seed selection.**

IV. **New molecular markers on previously annotated candidate genes with possible role in adaptation were developed and tested on *Pinus cembra* and *P. sylvestris*. The variation found within these genes are important in further studies of adaptation to stressful conditions.**

7. 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.

Bell, G. (2010). Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1537), 87–97.

Bernhardsson, C., Floran, V., Ganea, S. L., & García-gil, M. R. (2016). Forest Ecology and Management 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.

Bilgen, B. B., & Kaya, N. (2007). Allozyme variations in six natural populations of scots pine (*Pinus sylvestris*) in Turkey. *Biologia*, 62(6), 697–703.

Bobowicz, M. A. and Korczyk, A. (2000). The variability of the oldest trees of Scots pine (*Pinus sylvestris* L.) from the Bialowieza Primeval Forest. II. Variability of needle morphology traits in old-growth Scots pine trees from the Bialowieza Primeval Forest as compared to the variability. *Biological Bulletin of Poznań*, *1*, 37.

Boratynska, K. & Hinca, M. (2003). Morphological characteristic of *Pinus sylvestris* L. in the southernmost, isolated locality in the Sierra de Baza (Spain) as expressed. *Dendrobiology*, 50, 3–9.

Bridle, J. R., & Vines, T. H. (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*. 22(3), 140-147.

Cervantes, E., Martín, J. J., & Saadaoui, E. (2016). Updated methods for seed shape analysis. *Scientifica*, 2016, 5691825.

Chambers, J. C., & Macmahon, J. A. (1994). A say in the life of a seed: movements and fates of seeds and their implications for natural and managed systems. *Annual Review of Ecology and Systematics*, 25(25), 263–92.

Cheddadi, R., Vendramin, G. G., Litt, T., François, L., Kageyama, M., Lorentz, S., … Lunt, D. (2006). Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography*, 15(3), 271–282.

Christensen, K. I., & Dar, G. H. (1997). A morphometric analysis of spontaneous and artificial hybrids of *Pinus mugo x sylvestris* (Pinaceae). *Nordic Journal of Botany*, 17(1), 77–86.

Damblon, F. (1997). Palaeobotanical study of representative upper palaeolithic sites in the central European plain: a contribution to the SC-004 Project. *Préhistoire Européenne*, 11, 245–254.

Dangasuk, O. G., & Panetsos, K. P. (2004). Altitudinal and longitudinal variations in *Pinus brutia* (Ten.) of Crete Island, Greece: some needle, cone and seed traits under natural habitats. *New Forests*, 27(Panetsos 1981), 269–284.

Dzialuk, A., Muchewicz, E., Boratyński, A., Montserrat, J. M., Boratyńska, K., & Burczyk, J. (2009). Genetic variation of *Pinus uncinata* (Pinaceae) in the Pyrenees determined with cpSSR markers. *Plant Systematics and Evolution*, 277(3–4), 197–205.

Eguiluz, P. (1984). Geographic variation in needles, cones and seeds of *Pinus tecunumanii* in

Guatemala. *Silvae Genetica*, 33(2/3), 72-79.

Ehrenberg, C., Gustafsson, Å., Forshell, C., Simak, M. (1955). Seed quality and the principles of forest genetics. *Hereditas*, 41(3–4), 291–366.

Farrell, B. D., Dussourd, D. E., & Mitter, C. (1991). Escalation of plant defense: do latex and resin canals spur plant diversification? *The American Naturalist*, 138(4), 881.

Forde, M. B. (1964). Variation in natural populations of *Pinus radiata* in California. *New Zealand Journal of Botany*, 2(4), 459–485.

Fu, Y. X. (1995). Statistical properties of segregating sites. *Theoretical population biology*, 48(2), 172-197.

Gil, L., Climent, J., Nanos, N., Mutke, S., Ortiz, I., & Schiller, G. (2002). Cone morphology variation in *Pinus canariensis* Sm. *Plant Systematics and Evolution*, 235, 35–51.

Greene, D. F., & Johnson, E. A. (1993). Seed mass and dispersal capacity in wind-dipersed diaspores. *Oikos*, 67(1), 69–74.

Haesaerts, P., Damblon, F., Bachner, M., & Trnka, G. (1996). Revised stratigraphy and chronology of the Willendorf II sequence, Lower Austria. *Archaeologia Austriaca*, 80, 25–42.

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

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, 8(5), 461–467.

International Seed Testing Association-ISTA. (1985). International rules for seed testing. *Seed Science and Technology*, 13, 299–513.

Iverson, L. R., Schwartz, M. W., & Prasad, A. M. (2004). How fast and far might tree species migrate under climate change in the eastern United States? *Global Ecology and Biogeography*, 13, 209–219.

Jankovska, V., & Pokorný, P. (2008). Forest vegetation of the last full-glacial period in the Western Carpathians (Slovakia and Czech Republic). *Preslia*, 80(3), 307–324.

Jasińska, A. K., Boratyńska, K., Dering, M., Sobierajska, K. I., Ok, T., Romo, A., & Boratyński, A. (2014). Distance between south-European and south-west Asiatic refugial areas involved morphological differentiation: *Pinus sylvestris* case study. *Plant Systematics and Evolution*, 300(6), 1487–1502.

Kinloch, B. B., Westfall, R. D., Forrest, G. I. (1986). Caledonian Scots Pine: origins and genetic structure. *New Phytologist*, 104(4), 703–729.

Köbölkuti, Z. A., Tóth, E. G., Ladányi, M., & Höhn, M. (2017). Morphological and anatomical differentiation in peripheral *Pinus sylvestris* L. populations from the Carpathian region. *Dendrobiology*, 77, 105–117.

Krutovsky, K. V., & Neale, D. B. (2005). Forest genomics and new molecular genetic approaches to measuring and conserving adaptive genetic diversity in forest trees. In *Conservation and management of forest genetic resources in Europe*. Zvolen: Arbora Publishers. pp. 369–390.

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: Morphology, Distribution, Functional Ecology of Plants*, 201(6), 468–476.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*. 17(4), 183-189.

Losos, J. (1996). Phylogenetic perspectives on community ecology. *Ecology*, 77(5), 1344–1354.

Magyari, E. K., Veres, D., Wennrich, V., Wagner, B., Braun, M., Jakab, G., ... & Rethemeyer, J. (2014b). Vegetation and environmental responses to climate forcing during the Last Glacial Maximum and deglaciation in the East Carpathians: attenuated response to maximum cooling and increased biomass burning. *Quaternary Science Reviews*, 106(278), 2.

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

Marcysiak, K. (2006). Scots pine (Pinaceae) from the Crimea compared to the species variation in Europe on the basis of cone traits. *Analysis*, 12(2), 203–208.

McGinley, M. A., Smith, C. C., Elliott, P. F., & Higgins, J. J. (1990). Morphological constraints on seed mass in lodgepole pine. *Functional Ecolog*y, 183-192.

Naydenov, K. D., Tremblay, F. M., Alexandrov, A., & Fenton, N. J. (2005). Structure of *Pinus sylvestris* L. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: provenance tests. *Biochemical systematics and ecology*, 33(12), 1226-1245.

Pardos, J. A. , Lange, W and Weißmann, G. (1990). Morphological and chemical aspects of *Pinus sylvestris* L . from Spain. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 44(2), 143–146.

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28(19), 2537–2539.

Prus-Glowacki, W., Stephan, B. R., Bujas, E., Alia, R., & Marciniak, A. (2003). Genetic differentiation of autochthonous populations of *Pinus sylvestris* (Pinaceae) from the Iberian peninsula. *Plant Systematics and Evolution*. 239(1-2), 55-66.

Pyhäjärvi, T., García-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.

Rehfeldt, G. E. (1991). A model of genetic variation for *Pinus ponderosa* in the Inland Northwest (USA): applications in gene resource management. *Canadian Journal of Forest Research*, 21(10), 1491–1500.

Richardson, D. M., & Rundel, P. W. (1998). Ecology and biogeography of *Pinus*: an introduction. In *Ecology and biogeography of Pinus,* Cambridge, Cambridge University Press. pp. 3–46.

Rozas, J., & Rozas, R. (1995). DnaSP, DNA sequence polymorphism: an interactive program for estimating population genetics parameters from DNA sequence data. *Computer Applications in the Biosciences*, 11(6), 621–625.

Rudner, Z. and, & Sümegi, P. (2001). Recurring Taiga forest-steppe habitats in the Carpathian Basin in the Upper Weichselian. *Quaternary International*, 76, 177-189.

Rudner, Z., Sümegi, P., Toth, I., Beszeda, I., & Hertelendi, E. (1995). The vegetation history of the Upper Weichselian in the Central and Southern Part of The Great Hungarian Plain. *In 7th European Ecological Congress.* Budapest. p. 165.

Semiz, G., Heijari, J., Isik, K., & Holopainen, J. K. (2007). Variation in needle terpenoids among *Pinus sylvestris* L. (Pinaceae) provenances from Turkey. *Biochemical Systematics and Ecology.* 35(10), 652-661.

Staszkiewicz, J. (1961). Biometric studies on the cones of Pinus silvestris L., growing in Hungary. *Acta Botanicorum Academy of Science Hungary*, 7, 451–455.

Staszkiewicz, J. (1993). Zmienność morfologiczna szpilek, szyszek i nasion. *Biologia Sosny Zwyczajnej. PAN Instytut Dendrologii, Poznań–Kórnik.* pp. 33–42.

Steinbeck, K. (1966). Site, height and mineral nutrient content relations of Scotch pine provenances. *Silvae Genetica,* 15(2), 42–50.

Stieber, J. (1967). A magyarországi felsöpleisztocén vegetációtörténete az. anthrakotómiai eredmények (1957-IG). *Földtani Közlöny*, 97(3), 308–317.

Tóth, E. G., 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 and Genomes*, 13(2), 47.

Turna, I. (2003). Variation of some morphological and electrophoretic characters of 11 populations of Scots pine in Turkey. *Israel Journal of Plant Sciences*, 51(3), 223–230.

Turna, I., & Güney, D. (2009). Altitudinal variation of some morphological characters of Scots pine (*Pinus sylvestris* L.) in Turkey. *African Journal of Biotechnology*, 8(2), 202–208.

Urbaniak, L, Karlinski, L, Popielarz, R. (2003). Variation of morphological needle characters of Scots pine (*Pinus sylvestris* L.) populations in different habitats. *Acta Societatis Botanicorum Poloniae*, 72(1), 37–44.

Vaughton, G., & Ramsey, M. (1998). Sources and consequences of seed mass variation in *Banksia marginata* (Proteaceae). *Journal of Ecology*, 86(4), 563-573.

Vera, J. C., Wheat, C. W., Fescemyer, H. W., Frilander, M. J., Crawford, D. L., Hanski, I., & Marden, J. H. (2008). Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology*, 17(7), 1636–1647.

Wachowiak, W., & Prus-Głowacki, W. (2008). Hybridisation processes in sympatric populations of pines *Pinus sylvestris* L., *P. mugo* Turra and *P. uliginosa* Neumann. *Plant Systematics and Evolution*, 271(1–2), 29–40.

Wahid, N., González-Martínez, S. C., El Hadrami, I., & Boulli, A. (2006). Variation of morphological traits in natural populations of maritime pine (*Pinus pinaster* Ait.) in Morocco. *Annals of Forest Science*, 63(1), 83-92.

Willis, K. J., & Van Andel, T. H. (2004). Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, 23(23-24), 2369-2387.

8. ABBREVIATION LIST OF THE SAMPLED POPULATIONS

HKO	Hungary	Kőszeg	Mixed broadleaf pine forest
HFE	Hungary	Fenyőfő	Mixed broadleaf pine forest
HVE	Hungary	Pethőhenye	Mixed broadleaf pine forest
HZA	Hungary	Szalafő	Mixed broadleaf pine forest
HOR	Hungary	Csörötnek	Mixed broadleaf pine forest
SKV	Slovakia	Kvacany	Rock surface
STU	Slovakia	Svarin	Rock surface
SLI	Slovakia	Liptovszky Hrádok	Rock surface
RFE	Romania	Fântâna Brazilor	Peat bog
RPS	Romania	Poiana Stampei	Peat bog
RMO	Romania	Băile Tuşnad	Peat bog
RPA	Romania	Voineasa	Rock surface
RBI	Romania	Roșia	Rock surface
RBE	Romania	Poșaga de sus	Rock surface
RML	Romania	Ponor	Peat bog
RMH	Romania	Călăţele	Peat bog

PUBLICATIONS CONNECTED TO THE DISSERTATION

Papers in impact factored journals:

Köbölkuti, **Z. A.**, Tóth, E. G., Ladányi, M., Höhn, M. (2017). Morphological and anatomical differentiation in peripheral *Pinus sylvestris* L. populations from the Carpathian region. *Dendrobiology*, 77, 105–117.

Tóth, E. G.,**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:

Köbölkuti, Z. A., Höhn, M. (2018). Habitat type differentiation in peripheral *Pinus sylvestris* L. populations based on seed traits and germination data. *Studia botanica hungarica.* In press.

Tóth, E. G., **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):

Köbölkuti, Z. A., Tóth, E. Gy., Höhn, M. (2018). Adaptív markerek fejlesztése és tesztelése erdeifenyőn (*Pinus sylvestris* L.). XII. Aktuális Flóra- és Vegetációkutatás a Kárpát-medencében. Debrecen, 2018. február 23-25.

Köbölkuti, Z. A. Tóth, E. G., Jahn, D., Höhn, M., Heinze, B. (2017). Comparative analysis of adaptive variation in *Pinus cembra* L. and *Pinus sylvestris* L. based on Swiss stone pine transcriptome developed molecular markers. Integrated methods to detect polygenic adaptation from genomic data. Symposium and Summer School, Zürich. 2017.

Köbölkuti, Z. A., Jahn, D., Heinze, B. (2017). Cirbolyafenyő génexpressziós szekvenciaanalízise. IV.Transylvanian Horticulture and Landscape Studies Conference. Románia. Marosvásárhely, 2017. május 5-7.

Köbölkuti, Z. A., Tóth, K., Tóth, E. Gy., Buczkó, K., Höhn, M. (2017). Tűlevelek mikromorfológiai változatossága peremhelyzetű erdeifenyő (*Pinus sylvestris* L.) populációkban (P-8). XV. Magyar Növényanatómiai Szimpózium 2017. szeptember 7.

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. Románia. Marosvásárhely, 2017. május 5-7.

Tóth, E. G., **Köbölkuti, Z. A.,** Bede-Fazekas, Á., Vendramin, G. G., Bagnoli, F., Höhn, M. (2017). Population demographic inferences of Scots pine along the Carpathians and the Pannonian Basin based on bioclimatic and molecular genetic data. Biogeography of the Carpathians (Second Interdisciplinary Symposium). Cluj-Napoca, Romania.

Köbölkuti, Z. A., Tóth, E. G., Zahoretz, A., Tóth, K., Höhn, M. (2015). Comparative needle anatomy and cone morphometry of natural *Pinus sylvestris* (L.) populations from Central-Eastern Europe. III. Transylvanian Horticulture and Landscape Studies Conference. Románia. Marosvásárhely, 2015. május 15-16.